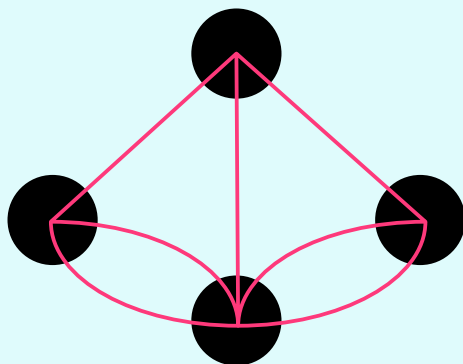
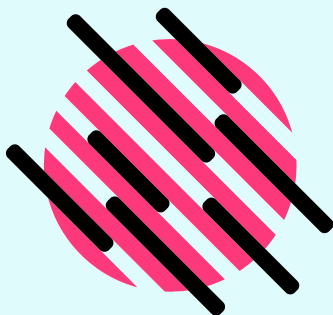


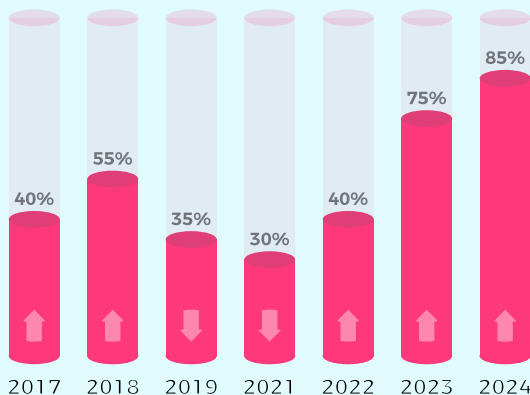
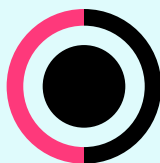


$$\sigma d = \sqrt{\left\{ \left[\frac{P1(1 - P1)}{n1} \right] + \left[\frac{P2(1 - P2)}{n2} \right] \right\}}$$

ESTATÍSTICA

PASSO A PASSO

Aliger dos Santos Pereira



EdUnB
Editora da Universidade do Estado da Bahia



Apoio:

ESTATÍSTICA PASSO A PASSO



Universidade do Estado da Bahia - UNEB

Adriana Marmori Lima
Reitora

Dayse Lago de Miranda
Vice-Reitora



Editora da Universidade do Estado da Bahia - EDUNEB

Diretora
Elisângela Santana dos Santos

Conselho Editorial

Titulares	Suplentes
Marluce Alves dos Santos	Angelita de Souza Leite
Josilda Batista Lima Mesquita Xavier	Francisco Hilder Magalhães e Silva
Astria Dias Ferrão Gonzales	Hilda Silva Ferreira
Emanuel Ernesto Fernandes Santos	Marcos Antônio Vanderlei Silva
Nilson Roberto da Silva Gimenes	Monica Matos Ribeiro
Alan da Silva Sampaio	Sandro dos Santos Correia
Raimundo Claudio Silva Xavier	Nilvo Luiz Cassol
José Ernane Carneiro Carvalho Filho	Marilecia Oliveira Santos
Carla Severiano de Carvalho	Lise Mary Arruda Dourado
Paulo César Souza García	Sílvio Roberto dos Santos Oliveira
Léa Costa Santana Dias	Cláudia Martins Moreira
Elzicleia Tavares dos Santos	Paulo Sérgio Dantas Vasconcelos
Marineuza Matos dos Anjos	Antonieta Miguel
Michael Daian Pacheco Ramos	Márcia Torres Neri Soares

ALIGER DOS SANTOS PEREIRA

ESTATÍSTICA PASSO A PASSO

Salvador
Eduneb
2025

© 2025 Autora

Direitos para esta edição cedidos à Editora da Universidade do Estado da Bahia.
Proibida a reprodução total ou parcial por qualquer meio de impressão, em forma idêntica,
resumida ou modificada, em Língua Portuguesa ou qualquer outro idioma.
Depósito Legal na Biblioteca Nacional.

Coordenação Editorial

Elisângela Santana dos Santos

Coordenação de Design

Sidney Silva

Capa e Diagramação

Lucas Lima - Tikinet

Revisão Textual e Normalização

Tikinet

Imagem de Capa

Freepik

Revisão Final

Textual - Tikinet

FICHA CATALOGRÁFICA

Bibliotecária: Fernanda de Jesus Cerqueira – CRB 162-5

Pereira, Aliger dos Santos

Estatística passo a passo / Aliger dos Santos Pereira. – Salvador:
EDUNEB, 2025.

E-Book: PDF, 300 p.; il.

ISBN nº 978-85-85813-45-1

1. Estatística. 2. Modelos Estatísticos. 3. Sustentabilidade. 4. Estudo e Ensino de
Estatística. 5. Estatística Aplicada. 6. Representações Estatísticas. 7. Interpretação
Estatística. 8. Modelos Estatísticos. I. Título.

CDD 310

Editora da Universidade do Estado da Bahia – EDUNEB

Rua Silveira Martins, 2555 – Cabula

41150-000 – Salvador – BA

editora@listas.uneb.br

portal.uneb.br

Esta Editora é filiada à



Associação Brasileira das
Editoras Universitárias

APRESENTAÇÃO DO *E-BOOK* PELA AUTORA



Olá a todos!



Hoje, a estatística é mais do que números e gráficos complicados. Ela permeia nosso cotidiano de maneiras surpreendentes e se encontra ao alcance de todos nós. Afinal, está presente nas nossas decisões pessoais e profissionais, contribuindo para uma melhor compreensão e otimização de processos e recursos da nossa vida e do planeta, como um indicativo de sustentabilidade.



Para se ter ideia, a estatística está presente:

- No nosso gasto de energia, água e internet, de forma a prever e avaliar o consumo diário desses serviços e gerenciar as melhores ofertas do mercado, inclusive a possível opção de modificar de empresa;
- No tráfego e no transporte, em que as agências e os aplicativos de trânsito utilizam dados estatísticos para monitorar o fluxo de veículos e calcular médias de velocidade nas estradas. Dessa forma, é possível identificar pontos para se elaborar planejamentos e melhorias na infraestrutura viária;
- Nas mídias sociais e redes sociais, que utilizam a estatística para analisar padrões de uso, como número de curtidas, compartilhamentos e interações, visando ajustar os algoritmos e as recomendações;
- Na saúde pública, na qual os órgãos de saúde utilizam a estatística para monitorar surtos de doenças, calcular taxas de vacinação e analisar dados demográficos para políticas de saúde preventiva;

-
- Nas finanças pessoais, em que as pessoas analisam seus gastos mensais, calculam médias de despesas e planejam orçamentos familiares;
 - Nos esportes, com os treinadores e analistas esportivos utilizando estatísticas para analisar o desempenho de atletas, como média de gols, percentual de acertos em arremessos e outras métricas para estratégias de jogo; e
 - Nas compras online, nas quais as empresas de comércio eletrônico utilizam estatísticas para analisar padrões de compra, comportamento do consumidor e recomendações personalizadas de produtos.

Este Livro tem o objetivo de mostrar de forma prática o uso de modelos estatísticos para a solução de problemas constates do dia a dia, da área empresarial, das atividades de pesquisa ou de outros itens que são muito questionados pelos alunos e pelas pessoas de uma forma geral nas oito áreas do conhecimento: Ciências Exatas e da Terra; Ciências Biológicas; Engenharias e Tecnologia; Ciências da Saúde; Ciências Agrárias; Ciências Humanas; Linguística, Letras e Artes; e Ciências Sociais Aplicadas. Afinal, a estatística pode e deve ser utilizada em todas as áreas citadas anteriormente, inclusive de forma interdisciplinar e transdisciplinar.

O nome deste *e-book* é *Estatística Passo a Passo*, pois os conteúdos dos capítulos estão inter-relacionados do início ao fim com o mesmo exemplo, a fim de resolvê-lo e interpretá-lo para uma tomada de decisão dentro de um contexto interdisciplinar e transdisciplinar.

Este livro é fruto também do meu aprendizado ministrando aulas de Estatística e Pesquisa Operacional na Universidade do Estado da Bahia (UNEB) e no Instituto Federal da Bahia (IFBA), onde percebi algumas dificuldades dos estudantes na compreensão dos referidos conteúdos.



O livro é estruturado em cinco partes essenciais:

1. Estatística e suas etapas aplicadas à pesquisa: Aqui, exploramos como aplicar metodologicamente a estatística em estudos e pesquisas;

2. Frequências e suas representações estatísticas: Nesta parte, nos aprofundamos nas diferentes formas de representar dados estatísticos através de frequências;

3. Medidas de Tendência Central, Separatrizes e Principais Desvios Estatísticos: Essenciais para se entender o comportamento dos dados, essas medidas são fundamentais na interpretação estatística;

4. Pesquisa Operacional e modelos estatísticos: Como a estatística pode otimizar processos e decisões na prática da pesquisa operacional;

5. Atividades com exercícios resolvidos: Para consolidar o aprendizado, proporcionamos atividades práticas com soluções detalhadas para os estudantes.



Para uma experiência de aprendizado completa, recomendamos seguir estes passos:

1. Leia o e-book e explore cada capítulo com atenção para compreender os conceitos fundamentais;

2. Veja as dicas em cada capítulo, elas são úteis e destacam pontos importantes para sua compreensão e interpretação;

3. Assista aos vídeos para complementar seu aprendizado, pois os tutoriais oferecem uma visão prática e visual dos conceitos. Eles podem ser apreciados pelo *link* da *playlist*: https://www.youtube.com/playlist?list=PLCBC7XIYk7dPEG8uNYbmqAZWArray_2IJJ



4. Aplique as planilhas de Excel, LINDO e Gephi. Ao utilizar essas ferramentas, há uma aplicação prática dos modelos estatísticos discutidos no livro de forma dinâmica;

5. Resolva as atividades, pois praticar leva à perfeição.

Ao praticar e realizar as atividades propostas, você estará preparado para utilizar a estatística de maneira contínua e integrada em diversas situações da vida pessoal e profissional e compreenderá que ela não é um “bicho de sete cabeças”, mas sim uma ferramenta poderosa ao alcance de todos.

Quero comprovar que o uso das estatísticas é muito simples e não pode ser fragmentado dentro do seu processo de aprendizado. Sua aplicação é fascinante como ferramenta, sendo importante para o levantamento a análise e a interpretação de dados em qualquer parte do conhecimento, pois esse não pode ser segregado em partes, e sim corresponder a um conjunto de saberes.



Esperamos que seja uma fonte valiosa de aprendizado e aplicação prática para vocês!



Aproveitem a jornada estatística, que é apaixonante
e tranquila! 

.....
Vídeo explicando os conteúdos deste livro de forma geral.



Link: <https://youtu.be/NWOu193-VG8>

.....

AGRADECIMENTOS

Gostaria de agradecer à Universidade do Estado da Bahia (UNEB), mais especificamente ao curso de Administração do Departamento de Ciências Humanas (DCH-1) e ao Instituto Federal da Bahia (IFBA), campus Camaçari, onde ministrou aulas para os cursos superiores de Ciências da Computação e de Matemática, além do curso integrado em Informática.

Também agradeço ao Grupo de Pesquisa Modelos e Estruturas Organizacionais a Nível Territorial para Ações Sustentáveis (Metas), vinculado à UNEB e ao DCH-1, bem como aos alunos Ana Beatriz Virgens Ribeiro, Arielton Evangelista da Silva, Mateus da Silva Guerreiro e Samuel Seabra Medeiros por dialogarem dentro desse grupo sobre o conteúdo deste *e-book* e sugerirem melhorias neste texto.

Desejo a todos uma boa leitura e um excelente aprendizado!

QUEM É A AUTORA?

Aliger dos Santos Pereira

Link do currículo Lattes: <http://lattes.cnpq.br/9514806025242255>



Link do meu site: <https://sites.google.com/view/aliger>

LinkedIn: www.linkedin.com/in/aliger-pereira-91b4aa15



Instagram: <https://www.instagram.com/aligerdossantos/>

Facebook: <https://www.facebook.com/aligerpereira>



Link do canal de YouTube: https://www.youtube.com/channel/UCEOfNp-fm71_rn4mZ2otd3pg

Email: aligerpereira@ifba.edu.br e aligerpereira@uneb.br

É graduada em Administração de Empresas (Universidade Salvador-UNIFACS) e tem licenciatura em Matemática (Universidade Cruzeiro do Sul-UNICSUL). Possui mestrado em Planejamento Territorial e Desenvolvimento Social (Universidade Católica do Salvador-UCSAL) e doutorado em Desenvolvimento Regional e Urbano (Universidade Salvador-UNIFACS).

Fez estágio de pós-doutoramento no Programa de Planejamento Territorial e Desenvolvimento Social (Universidade Católica do Salvador-UCSAL) e no Programa de Pós-Graduação em Difusão do Conhecimento (PPGDC). Escreveu *Clusters de veículo em Salvador: geoprocessamento e gestão de negócio para micro, pequenas e médias empresas (MPMEs)* e *Os Croatas?... presentes! e Inteligência Competitiva (IC): Micro, Pequenas e Médias Empresas (MPMES) em Salvador (Brasil) e Bragança (Portugal)*, o que favoreceu o uso da metodologia de Inteligência Competitiva (IC) para este livro. Ambos utilizaram o geoprocessamento e a estatística como principais instrumentos para análise e interpretações.

Atualmente, é docente do departamento de Ciências Humanas da Universidade do Estado da Bahia-UNEB (Campus I) e do Instituto Federal de Educação, Ciência e Tecnologia da Bahia-IFBA (Campus Camaçari).

Participa do Programa em Propriedade Intelectual e Transferência de Tecnologia para a Inovação (PROFNIT) ponto focal IFBA. Já foi professora colaboradora do Programa de Pós-Graduação Multi-Institucional em Difusão do Conhecimento (PPGDC). É integrante da Câmara Brasil-Croácia (BraCro). Atualmente, é Coordenadora de Pesquisa, Pós-Graduação de Inovação (COPPI-IFBA – Camaçari)

Além de líder do Grupo de Pesquisa Modelos e Estruturas Organizacionais a Nível Territorial para Ações Sustentáveis (METAS), também é editora e responsável pela revista *Scientia*. Foi coordenadora do curso de Administração do Centro Universitário de

Salvador (UNICEUSA). Coordenou o Curso de Turismo e Hotelaria e Administração da UNEB e a Tutoria do curso de Especialização em Gestão Pública (UNEB/ Universidade Aberta do Brasil [UAB]). Atua no Ensino a Distância (EAD) e presencial como professora, tutora e elaboradora de conteúdo.

PREFÁCIO

Este livro é uma jornada envolvente e esclarecedora através das complexidades da estatística aplicada, proporcionando uma compreensão profunda e prática de como os modelos estatísticos podem ser utilizados para extrair conhecimentos valiosos para a pesquisa aplicada e para a tomada de decisão a partir de dados.

A Estatística é uma ferramenta poderosa que permeia diversas áreas do conhecimento, da ciência à economia, da saúde à tecnologia. Neste livro, é explorada a aplicação de modelos estatísticos em uma variedade de contextos, abordando tanto os conceitos fundamentais quanto às técnicas avançadas, de modo a proporcionar uma base sólida para estudantes, pesquisadores e profissionais.

Cada capítulo é cuidadosamente elaborado para guiar o leitor desde os princípios básicos até aplicações práticas e desafios do mundo real. Traz dicas de filmes para complementar o aprendizado, sites para pesquisa exploratória com fontes secundárias, vídeos criados e disponibilizados pela autora em seu canal na internet para explicar ao leitor alguns tópicos, além dos exemplos atenciosamente escolhidos para ilustrarem a relevância e a versatilidade dos modelos estatísticos e os exercícios propostos, que oferecem a oportunidade de aprimoramento das habilidades analíticas.

O objetivo deste livro não é apenas transmitir conhecimento teórico, mas também capacitar o leitor a aplicá-los de forma eficaz em situações do cotidiano. A estatística é mais do que apenas números, é uma ferramenta essencial para a tomada de decisões informadas e a compreensão do mundo ao nosso redor, sendo importante para estudantes de graduação e pós-graduação e outros interessados no aprendizado das técnicas aqui apresentadas.

À medida que exploramos os diferentes tópicos, a autora nos convida a mergulhar nesse universo estimulante de forma leve, cuidadosa e precisa. Que este livro seja uma fonte de inspiração e conhecimento, capacitando-o a enfrentar os desafios estatísticos com confiança e discernimento.

Boa leitura e descoberta!

Maria Cristina Elyote Marques Santos

Professora Adjunta da Universidade do Estado da Bahia

**Licenciada em Matemática e Doutora em Educação
e Contemporaneidade.**

Link do currículo Lattes: <http://lattes.cnpq.br/5823568857777899>



DEPOIMENTOS DA IMPORTÂNCIA DA ESTATÍSTICA POR COORDENADORES DE CURSO

Com o objetivo de mostrar a importância dos conteúdos dos componentes de Estatística e Pesquisa Operacional para a formação dos estudantes de graduação, três coordenadores de distintos cursos deram seus depoimentos. São eles: Alexandre Boleira Lopo, Lúcio Marcos Silva dos Santos e Rosana Cruz.

1º Depoimento

É com grande entusiasmo que apresento este livro, fruto do notável aprendizado da professora Aliger dos Santos Pereira, resultante dos seus anos de experiência ministrando aulas de Estatística e Pesquisa Operacional na Universidade do Estado da Bahia (UNEB) e no Instituto Federal da Bahia (IFBA). Durante esse tempo, a autora percebeu as dificuldades dos estudantes na compreensão desses conteúdos, motivando a elaboração deste material.

Eu, docente e há mais de 20 anos formando professores e outras carreiras, como engenheiros, acredito que o estudo da Estatística incentiva o desenvolvimento do pensamento crítico. Professores capacitados na área são capazes de ensinar aos alunos como analisar e interpretar dados de maneira crítica, promovendo habilidades valiosas para a resolução de problemas na vida cotidiana.

Diante disso, o objetivo deste livro é proporcionar uma abordagem prática e descomplicada no uso de modelos estatísticos para solucionar problemas cotidianos, empresariais e de pesquisa. O livro está dividido em cinco capítulos e tem uma estrutura que oferece uma progressão natural e lógica, facilitando a compreensão e aplicação dos conceitos apresentados com vídeos da autora que acompanham cada capítulo e atuam como tutoriais a fim de

complementar os temas apresentados. Além disso, a utilização de programas e aplicativos práticos (Excel, LINDO e *Gephi*) facilitam o uso e aprendizagem da Estatística.

A Estatística é interdisciplinar pois se conecta a várias outras áreas, como ciências, economia, psicologia e sociologia. Sabe-se que muitas carreiras exigem habilidades estatísticas, e a formação na área fornece as bases necessárias para o mercado de trabalho, independentemente da área escolhida.

Além disso, os cidadãos (de forma geral) precisam ser capazes de entender e avaliar informações estatísticas para tomar decisões informadas em suas vidas diárias. Os professores de matemática e de qualquer área podem aproveitar esse livro com a intenção de preparem seus alunos para serem cidadãos críticos e bem-informados.

Entendo que o livro procura se alinhar com a Base Nacional Comum Curricular (BNCC), documento balizador da educação no Brasil, pois ela destaca a importância de interpretar dados, planejar e executar pesquisas amostrais e compreender medidas de tendência central, contribuindo para a formação integral dos alunos. Assim, este livro pode atender às expectativas e demandas educacionais, proporcionando oportunidades para os estudantes desenvolverem habilidades específicas relacionadas à estatística.

O documento da BNCC explica na página 275 que a incerteza e o tratamento de dados são estudados na unidade temática denominada Probabilidade e Estatística, que propõe uma abordagem de conceitos, fatos e procedimentos presentes em muitas situações-problema da vida cotidiana, das ciências e da tecnologia. O documento-base ressalta que todos os cidadãos precisam desenvolver habilidades para coletar, organizar, representar, interpretar e analisar dados em uma variedade de contextos, de maneira a fazer julgamentos bem fundamentados e tomar as decisões adequadas. Isso inclui raciocinar e utilizar conceitos, representações e índices estatísticos para descrever, explicar e prever fenômenos.

Em suma, esta obra é uma ferramenta valiosa para professores e alunos, incentivando uma abordagem dinâmica, prática e interdisciplinar do conhecimento estatístico. Desejo a todos uma leitura instigante e um aprendizado enriquecedor!

Alexandre Boleira Lopo, Doutor

Coordenador do Curso de Licenciatura em Matemática (IFBA)

Link do currículo Lattes: <http://lattes.cnpq.br/5108816345823701>



2º Depoimento

Em nosso mundo atual, existem alguns assuntos que são indissociáveis do nosso cotidiano. A matéria probabilidade trata das situações presentes em nosso dia a dia. Sua importância na compreensão e análise dos índices de custo de vida, da probabilidade de ganhos em seus investimentos, e das escolhas de amostras para suas pesquisas e tomadas de decisões a torna de suma importância em nossa formação. A probabilidade faz parte de um viés da matemática que trata da incerteza de dados, logo, com tantas incertezas de informações, nada melhor do que mapear seus dados usando as premissas básicas da probabilidade.

Na área de Ciências da Computação, os profissionais estão diretamente ligados à transformação de dados em informação, lidando em seu cotidiano com dados complexos e variados. Quando se estuda a matéria de probabilidade, o cientista da computação adiciona um entendimento da natureza de alguns dados,

de seus cálculos e de suas possíveis incertezas com um olhar fundamentado, identificando padrões e extraindo informações para uma melhor modelagem de seus sistemas.

Lúcio Marcos Silva dos Santos

**Coordenador do curso de Ciências da Computação e Licenciatura
em Computação (IFBA)**

Link currículo Lattes: <http://lattes.cnpq.br/5188567098810352>



3º Depoimento

Esta obra é fruto da experiência e do trabalho organizado ao longo do tempo pela professora doutora Aliger dos Santos Pereira, consolidando seus esforços em prol dos estudantes.

Ela apresenta conteúdos desenvolvidos para pessoas da área de Administração, Ciências Contábeis, Economia, Ciência da Computação, Matemática, dentre outros. Aborda também técnicas e aplicações a serem utilizadas nas pesquisas e subsidiando a tomada de decisões das diversas áreas.

A elaboração deste trabalho teve a intenção de colaborar com o avanço dos estudos dos Modelos Estatísticos e da Pesquisa Operacional. Consciente da importância e do valor desses temas para os estudantes, a autora apresenta um conteúdo bem didático, resultado de pesquisas e análises, com a finalidade de ajudar no aumento do conhecimento daqueles que o utilizarem.

Trata-se, portanto, de uma obra que irá contribuir muito para a aprendizagem dos estudantes.

Rosana Cruz Marques da Silva

Coordenadora do curso de Administração (UNEB-DCH-1)

Link currículo lattes: <http://lattes.cnpq.br/2954222111199312>



SUMÁRIO

1 A ESTATÍSTICA E SUAS ETAPAS APLICADAS	31
1.1 CONCEITO E IMPORTÂNCIA DA ESTATÍSTICA.....	32
1.2 ESTATÍSTICA COMO MÉTODO	39
1.3 TIPOS DE AMOSTRAGEM	51
1.4 INSTRUMENTOS DE PESQUISA.....	62
1.5 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS	74
2 FREQUÊNCIAS E SUAS REPRESENTAÇÕES ESTATÍSTICAS	79
2.1 DISTRIBUIÇÃO DE FREQUÊNCIA APÓS A COLETA DOS DADOS.....	80
2.2 DADOS NÃO AGRUPADOS	83
2.3 DADOS AGRUPADOS.....	86
2.4 DISTRIBUIÇÃO DE FREQUÊNCIA EM TABELA PARA INTERPRETAÇÕES.....	91
2.5 USO DO EXCEL PARA DISTRIBUIÇÃO DE FREQUÊNCIA EM GRÁFICOS	100
2.6 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS.....	110
3 MEDIDAS DE TENDÊNCIA CENTRAL, SEPARATRIZ E PRINCIPAIS DESVIOS ESTATÍSTICOS.....	116
3.1 MEDIDAS DE TENDÊNCIA CENTRAL OU DE POSIÇÃO	117
3.1.1 Medida de Tendência Central (Média Aritmética) e alguns desvios.....	118
3.1.2 Medida de Tendência Central (Mediana)	132
3.1.2.1 Separatriz	138
3.1.3 Medida de Tendência Central (Moda).....	148
3.2 MEDIDAS DE ASSIMETRIA E CURTOSE	151
3.3 CORRELAÇÃO E REGRESSÃO LINEAR	160
3.4 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS.....	169

4 A PESQUISA OPERACIONAL E OS SEUS MODELOS ESTATÍSTICOS	175
4.1 CONCEITO E IMPORTÂNCIA DA PESQUISA OPERACIONAL	176
4.2 PROGRAMAÇÃO LINEAR (PL) E SUA MODELAGEM	184
4.2.1 Modelo da Programação Linear	188
4.2.2 Cálculo da Programação Linear de forma Tabular pelo SIMPLEX	195
4.2.3 Cálculo da Programação Linear por gráfico.....	212
4.2.4 Cálculo da Programação Linear com o uso da Técnica do Transporte	219
4.2.5 Uso do software LINDO para Programação Linear	236
4.3 TRANSPORTE AVANÇADO E SUA RELAÇÃO COM PERT E CPM	248
4.4 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS.....	266
 5 EXERCÍCIOS DE ESTATÍSTICA COM SUAS PRINCIPAIS RESPOSTAS	 280
 REFERÊNCIAS.....	 297

ESTATÍSTICA PASSO A PASSO

RESUMO:

Este e-book traz conteúdos básicos de Estatística e de Pesquisa Operacional que podem ser utilizados nas áreas pessoais, profissionais e acadêmicas, ou seja, no dia a dia. Ele possui cinco capítulos, sendo eles: a Estatística e suas etapas aplicadas à Pesquisa; as Frequências e suas representações Estatísticas; as Medidas de Tendência Central, Separatriz e Principais Desvios Estatísticos; a Pesquisa Operacional e os seus modelos Estatísticos; e as atividades com exercícios já resolvidos para os estudantes praticarem. Neste último há vídeos entre cinco a 25 minutos. *Link* da playlist:

https://www.youtube.com/playlist?list=PLCBC7XIYk7dPE-G8uNYbmQAZWArY_2IJJ) que usam de forma prática três *softwares* (Excel, LINDO e Gephy) associado aos conteúdos dos assuntos já citados e atividades para serem resolvidas.



Palavras-chave: Estatística; Pesquisa Operacional.

ABSTRACT:

This E-book provides basic Statistics and Operational Research content that can be used in personal, professional and academic areas, that is, in everyday life. It has 5 chapters, they are: Statistics and its stages applied to Research; Frequencies and their Statistical representations; the Measures of Central Tendency, Separatrix and Main Statistical Deviations; Operational Research and its Statistical models; and activities with solved exercises for students to practice. There are videos between 5 and 25 minutes long (Playlist *link*:

https://www.youtube.com/playlist?list=PLCBC7XIYk7dPE-G8uNYbmQAZWArY_2IJJ) that practically use 3 software (Excel, LINDO and Gephy) associated with the contents of the subjects already mentioned and activities to be solved.



Keywords: Statistic; Operational Research.

LISTA DE FIGURAS

Figura 1	Etapas da pesquisa Estatística	33
Figura 2	Etapas da pesquisa Estatística	36
Figura 3	Relação entre o universo, população e amostra	50
Figura 4	Cálculo da amostra probabilística em três aplicativos <i>on-line</i> gratuitos	58
Figura 5	Resumo sintético do método estatístico	73
Figura 6	Elaboração da tabela de dados não agrupados (ROL-crescente) a partir de seu dado bruto	84
Figura 7	Apresentação de uma tabela conforme normas	86
Figura 8	Tipo de Intervalo para colocar na tabela em construção	88
Figura 9	Apresentação de um gráfico conforme normas	90
Figura 10	Resumo síntese sobre a coleta de dado e a distribuição de frequência	109
Figura 11a	Exemplo de cálculo da Média Aritmética Simples para Dados não agrupados	120
Figura 11b	Exemplo de cálculo da Média Aritmética Simples para dados agrupados	121
Figura 12	Passo a passo para o cálculo do desvio p do desvio relativo, do desvio absoluto, da variância, do desvio médio absoluto, do desvio padrão e do coeficiente de variância padrão para dados não agrupados	125
Figura 13	Passo a Passo para o cálculo do desvio Relativo, do Desvio Absoluto, da Variância, do Desvio Médio Absoluto, do Desvio Padrão e do Coeficiente de Variância para dados agrupados	129
Figura 14	Exemplo de cálculo da Mediana para dados não agrupados-par	134
Figura 15	Exemplo de cálculo da Mediana para dados não agrupados-ímpar	135
Figura 16	Exemplo de cálculo da Mediana para dados agrupados	137
Figura 17	Identificação dos quartis	139
Figura 18	Identificação dos decis	139
Figura 19	Exemplo de cálculo do Q1(25%) para dados agrupados	144
Figura 20	Exemplo de cálculo do C75(75%) para dados agrupados	146
Figura 21	Exemplo de cálculo da Moda através da fórmula de Czuber para dados agrupados	150

Figura 22	Tipos de curva de assimetria	153
Figura 23	Tipos de curva de curtose	158
Figura 24	Desenho dos cinco tipos de correlação de forma gráfica com a linha de tendência.	162
Figura 25	Resumo síntese de alguns cálculos que podem ser realizados na distribuição de frequência	168
Figura 26	Etapas da tomada de decisão na PO	180
Figura 27	Representação dos modelos da PL	185
Figura 28	Cálculo e representação gráfica da RT_1	186
Figura 29	Cálculo e representação gráfica da RT_2	214
Figura 30	Cálculo e representação gráfica da RT_3	214
Figura 31	Representação gráfica das 3 Restrições Técnicas	215
Figura 32	Representação gráfica das possíveis soluções	217
Figura 33	Representação do estudo de transporte	220
Figura 34	<i>Link</i> e local onde realizar download do LINDO	237
Figura 35	Aparência do LINDO em formato de documento em branco	237
Figura 36	Cálculo do LINDO do modelo de Maximização – Exemplo da Empresa Familiar	240
Figura 37	Salvar no LINDO do modelo de Maximização (Entrada) e a Resposta (Saída) - Exemplo da Empresa Familiar	242
Figura 38	Exemplo de transporte em rede da empresa familiar	251
Figura 39	Aparência do Gephi	259
Figura 40	Importação do Gephi primeira planilha	260
Figura 41	Importação do Gephi segunda planilha	261
Figura 42	Locais para clicar para fazer gráfico e cálculo estatístico	262
Figura 43	Desenho da rede no Gephi da Figura 39	265

LISTA DE GRÁFICOS

Gráfico 1	Diagrama de dispersão referente à relação entre o estudo por hora e notas do componente de estatística – Escola A-2024.1	161
-----------	--	-----

LISTA DE QUADROS

Quadro 1	Etapas da pesquisa Estatística	37
Quadro 2	Diferença entre o estudo qualitativo do quantitativo ao definir a abordagem	40
Quadro 3	Conceito e uso das três principais pesquisas para a área de estatística	42
Quadro 4	Conceito de universo, população, censo ou amostra	49
Quadro 5	Tipo de amostra probabilística e suas subdivisões com os respectivos conceitos para o uso na pesquisa estatística	52
Quadro 6	Tipo de amostra não probabilística e suas subdivisões com os respectivos conceitos para o uso na pesquisa estatística	60
Quadro 7	Algumas formas para coleta de dados	64
Quadro 8	Tipos de pergunta para questionário ou roteiro	69
Quadro 9	Dados brutos e uma coleta de dados referente à nota de 20 alunos	81
Quadro 10	Elaboração da tabela para dados agrupados a partir de dados brutos encontrados na pesquisa	90
Quadro 11	Principais tipo de gráfico com conceitos e exemplos	102
Quadro 12	Principais vantagens e desvantagens das três medidas de tendências centrais (média aritmética, mediana e moda)	151
Quadro 13	Principais técnicas de PO utilizadas para tomada de decisão	181
Quadro 14	Representação da função objetiva	191
Quadro 15	Representação da Primeira Restrição Técnica (RT_1)	192
Quadro 16	Representação da Segunda Restrição Técnica (RT_2)	192
Quadro 17	Representação da Terceira Restrição Técnica (RT_3)	193
Quadro 18	Representação da restrição de não negatividade	194
Quadro 19	Representação da saída	197
Quadro 20	Representação da entrada e da saída (elemento pivô)	199
Quadro 21	Cálculo da nova linha pivô	199

Quadro 22	Cálculo da nova 1.a linha	200
Quadro 23	Cálculo da nova 3.a linha	201
Quadro 24	Cálculo da nova 4.a linha	201
Quadro 25	Resultado matemático (não ótimo)	202
Quadro 26	Cálculo da nova entrada na função objetiva	203
Quadro 27	Cálculo da nova saída na restrição técnica	204
Quadro 28	Representação da nova entrada e da nova saída (elemento pivô)	205
Quadro 29	Cálculo da nova linha pivô	205
Quadro 30	Cálculo da nova 1.a linha	206
Quadro 31	Cálculo da nova 2.a linha	207
Quadro 32	Cálculo da nova 3.a linha	207
Quadro 33	Resultado matemático ótimo	208
Quadro 34	Entrega média diária e demanda diária de cada localidade	221
Quadro 35	Tempo para distribuição de mercadorias	221
Quadro 36	Somatório da demanda igual ao somatório da produção	222
Quadro 37	Cálculo utilizando o método do Custo Mínimo I	226
Quadro 38	Cálculo utilizando o método do Custo Mínimo II	227
Quadro 39	Cálculo utilizando o método do Custo Mínimo III	228
Quadro 40	Cálculo utilizando o método do Custo Mínimo IV	229
Quadro 41	Cálculo utilizando o método de Vogel I	231
Quadro 42	Cálculo utilizando o método de Vogel II	232
Quadro 43	Cálculo utilizando o método de Vogel III	233
Quadro 44	Cálculo utilizando o método de Vogel IV	234
Quadro 45	Modelo de maximização e minimização do exemplo do capítulo 4	238
Quadro 46	Modelo de maximização e minimização do exemplo do capítulo 4 no LINDO	239
Quadro 47	Nó do exemplo empresa familiar com seus respectivos números para colocar no LINDO	247
Quadro 48	Cálculo do caminho crítico	252
Quadro 49	Cálculo da folga dos 3 caminhos Quadro 48	252
Quadro 50	3 Durações das atividades	253

Quadro 51	Tempo de duração esperada da atividade e seu respectivo desvio padrão	255
Quadro 52	Cálculo da data mais cedo do início (DCI), data mais cedo do término e da folga	256
Quadro 53	Outra forma de calcular as 3 durações das atividades	256
Quadro 54	Valor da variância e do desvio da amostra	257
Quadro 55	Atributos da empresa de entrega de matéria prima a partir da Figura 39 para uso do Gephi (1ª planilha de Excel	258
Quadro 56	Identificação dos nós e arestas (caminho) para uso no Gephi (2ª planilha	259
Quadro 57	Principais explicações estatísticas do resultado do Gephi para o caso da empresa familiar	263

LISTA DE TABELAS

Tabela 1	Tabela de cálculo de amostra probabilística com 95% de confiabilidade e 5% de erro amostral (ajuste)	56
Tabela 2	Dados organizados em ROL- Turma A-2024.1	81
Tabela 3	Distribuição das frequências das notas de estudantes para futuras interpretações – Escola A-2024.1	98
Tabela 4	Dados organizados em ROL (Não Agrupado) das notas de estudantes – Escola A-2024.1	149
Tabela 5	Exemplo de cálculo de assimetria para dados não agrupados e agrupados-Escola A-2024.1	156
Tabela 6	Relação entre o estudo por hora e notas do componente de estatísticas- Escola A-2024.1	160
Tabela 7	Exemplo do uso da correlação de Pearson da relação entre o estudo por hora e notas da disciplina de estatística-Escola A-2024.1	165

1. A ESTATÍSTICA E SUAS ETAPAS APLICADAS

OBJETIVOS:

- Compreender os principais conceitos e a importância da estatística na atualidade e sua aplicação interdisciplinar e transdisciplinar nas diversas áreas de conhecimento.
- Entender as principais etapas da aplicação da estatística em projetos pessoais, acadêmicos, empresariais ou profissionais.

CONTEÚDO:

1.1 CONCEITO E IMPORTÂNCIA DA ESTATÍSTICA

1.2 ESTATÍSTICA COMO MÉTODO

1.3 TIPOS DE AMOSTRAGEM

1.4 INSTRUMENTOS DE PESQUISA

1.5 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS

Explicação, Objetivos e Conteúdos do Capítulo 1

Link: <https://youtu.be/6p4JVWkzhF4>



1.1 CONCEITO E IMPORTÂNCIA DA ESTATÍSTICA

A estatística é tão antiga quanto o primeiro ser humano, pois a necessidade de contar surgiu com as pessoas. Dessa forma, ela é uma parte da matemática aplicada que fornece métodos para a coleta, a organização, a descrição, a análise e a interpretação de dados e para a utilização dos mesmos na tomada de decisões em quaisquer (ou diversas) áreas do conhecimento. Os três primeiros métodos (coleta, organização e descrição) fazem parte da estatística descritiva, já os dois últimos (análise e interpretação) da estatística inferencial (Crespo, 2009, p. 3).

Tanto a Matemática como a Estatística têm origem semelhantes, pois surgem do convívio social, das trocas, da contagem, com caráter prático, utilitário, empírico, presente nas pessoas e na sociedade humana.

CURIOSIDADE – VOCÊ SABIA?

O termo estatística foi primeiramente empregado para designar o conjunto de dados referentes a assuntos do Estado, geralmente com a finalidade de controle fiscal ou de segurança nacional. Por esse motivo, o uso da palavra, segundo estudiosos, teria a sua origem na expressão latina *status*, que significa Estado.

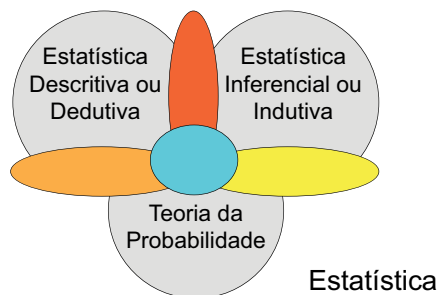
De acordo com a Associação Brasileira de Estatística (ABE) (2024):

A pesquisa científica é um processo interativo de acumulação de conhecimento e envolve a formulação de hipóteses, modelos e teorias, a observação de fenômenos e a verificação e rejeição de hipóteses sobre os mesmos. A estatística procura tornar esse processo o mais eficiente possível, através de suas técnicas de coletas de dados (amostragem e planejamento de experimentos); da apresentação de dados (análise exploratória e descrição: gráficos e tabelas); da modelagem (probabilidade e processos estocásticos);

da análise indutiva (inferência: testes de estimação) e da verificação (ajustamento, previsão e controle).

Dessa forma, a Estatística é um método que possibilita organizar a geração de dados e transformá-los em informação necessária para estudos das mais variadas áreas do conhecimento; e a melhor forma de representá-los estatisticamente é com o uso de tabelas e gráficos. Em outras palavras, a Estatística é um método ou um conjunto de técnicas que utiliza a coleta de dados e a sua classificação, apresentação ou representação, análise e interpretação visando utilizá-la dentro de um processo decisório, em que é possível envolver três ramos da Estatística que podem atuar de forma conjunta (pontos coloridos na Figura 1) ou separada individualmente (ponto cinza – Figura 1). Esses ramos (Figura 1) são: a Estatística Descritiva, a Estatística Inferencial e a Teoria da Probabilidade. A primeira procura descrever e analisar um certo grupo de observações denominado: amostra (veremos mais a frente), procurando expressar essas observações através de medidas e formas de representações (tabelas, gráficos e curvas, etc.). A segunda compreende um processo de generalização a partir da análise e interpretações de dados amostrais. A terceira busca métodos e técnicas apropriados para realizar um processo de generalizações que envolvem uma margem de risco ou incerteza. É possível um estudo que utilize os três ramos da Estatística (Figura 1 – Parte azul).

Figura 1 – Ramos da Estatística



Fonte: Elaboração própria, 2024.

Na atual Era do Conhecimento, quando estamos cercados de vários dados que são informações sistematicamente registradas dentro de um determinado contexto, as pessoas e empresas que souberem interpretá-los, ou seja, transformá-los em informação para uso pessoal, profissional ou empresarial, terão poder. Por isso que se tem relatado que informações são o ativo intangível mais importante do atual século.

Algumas áreas em que se usa a Estatística atualmente são:

- Indústria e empresas de forma geral: quando realizam testes de qualidade de produtos e serviços; na elaboração de previsão de vendas; como base para as pesquisas de opinião, que ajudam no lançamento de novos produtos ou aprimoramento dos que já estão no mercado; para estudos de ampliação ou expansão de empresas; em análises de amostras colhidas durante as várias etapas do processo produtivo; e para detectar erros e sugerir melhorias nos processos, produtos e serviços;
- Setor financeiro: ao realizar estudos de risco antes de conceder créditos ou personalizar seguros; para análise e verificação de aplicações financeiras;
- Agronegócio: quando identifica qual espécie produz melhor historicamente em determinado clima ou região; para realizar previsões sobre a situação da agricultura na economia;
- Bioestatística: em experimentos ou previsões das Ciências Médicas e Biológicas em relação a medicamentos existentes e novos, vacinas ou cosmético; em suas diversas etapas de testes ou mesmo quando já se encontra em uso no mercado;
- Tecnológica: em especial na área de Banco de Dados. As técnicas estatísticas são utilizadas para análise de dados, geração de modelos e previsões;
- Educação: para o controle no processo de gestão da sala de aula, constituindo-se em uma ferramenta ou uma forma

simples para o professor levantar dados diversos relacionados ao desempenho e ao comportamento das atividades de docência;

- Esporte: para identificar o desempenho dos atletas ou das equipes em uma partida e como esses fatores interferem no planejamento dos treinos futuros ou na contratação de atletas;
- Pesquisas eleitorais: fornecem dados para que os candidatos direcionem suas campanhas e para os cidadãos adquirirem informações que apoiem a decisão em quem irão votar;
- Pesquisa acadêmica, governamental e os institutos de pesquisa: na elaboração de relatórios para auxílio de políticas públicas.

Assim, a estatística retrata um fenômeno em um determinado tempo, e esses dados servirão para previsões e probabilidades futuras.

Dicas de filme: O homem que mudou o jogo (*Moneyball*)

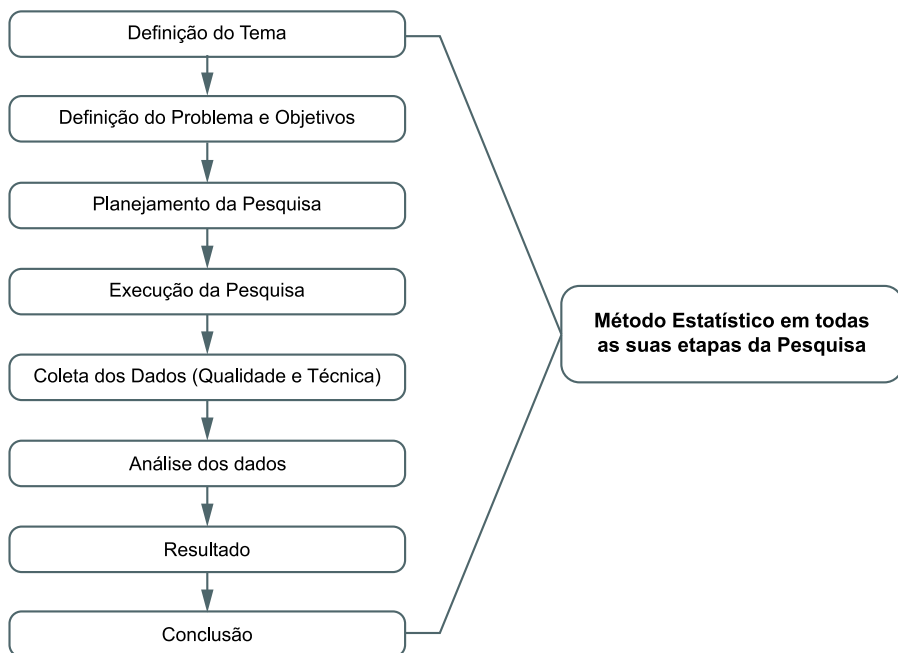


(Link do Trailer: <https://youtu.be/Buml-Yh0P1M>)

Um treinador visionário que fundamentou suas escolhas baseando-as em estatística e análise de dados para melhorar a performance do seu time. Ele contratou um cientista que estudou a porcentagem de acerto dos jogadores, substituindo assim o sexto sentido ancorado em expectativas dos olheiros.

Uma atitude bem simples do nosso dia a dia no uso da Estatística é quando desejamos comprar um produto ou serviço, e para que necessitamos ou desejamos comprá-lo. A partir daí, planejamos e pesquisamos onde podemos adquiri-lo conforme suas características e também seu preço, partindo de quanto podemos pagar. Esse simples fato do nosso cotidiano tem a ver com a Estatística, pois avaliamos as diversas alternativas e, através de uma análise processada em nossa mente, tomamos a decisão de comprá-lo ou não. Então, a primeira etapa é a definição do problema e dos objetivos, seguindo do planejamento e execução para a coleta dos dados; então analisamos as alternativas para chegarmos aos resultados e concluímos se o compramos ou não (Figura 2).

Figura 2 – Etapas da pesquisa Estatística



Fonte: Elaboração própria, 2024.

É bom salientar que a qualidade das informações analisadas e interpretadas para os resultados e conclusões dependerá de dois pontos: a qualidade do dado coletado e a técnica utilizada na coleta dos dados. Por isso é importante o uso do método estatístico nas diversas etapas da pesquisa. O Quadro 1 mostra de forma sintética os conceitos de cada item da Figura 2, presentes na pesquisa estatística.

Quadro 1 – Etapas da pesquisa estatística

Etapas	Conceitos	Pontos importantes
Tema	Defina o assunto que se deseja estudar e pesquisar.	Delimite bem o local e o período do estudo. O tema deve ser escolhido observando questões de viabilidade (de tempo, de informações e de recursos) por parte do pesquisador, importância (potencial de interessar e afetar muita gente e sustentabilidade) e originalidade (potencial de surpreender).
Problema	Necessário entender e explicar quais os problemas ou oportunidades relacionadas à pesquisa geram a necessidade de informação para a tomada de decisão (propósito da pesquisa).	Como o nome diz, deve ser uma pergunta que você responderá na sua conclusão. A resposta não pode ser sim ou não. Por isso é importante compreender bem o tema em questão e fazer uma boa revisão da literatura sobre ele. Faça perguntas usando, por exemplo: Qual (ou Quais), Como, Por que.
Objetivos	São determinados de maneira a trazer as informações que solucionem o problema de pesquisa. A indagação básica que o pesquisador deve fazer para elaboração correta dos objetivos é: Que informações são necessárias para resolver o problema de pesquisa?	Os objetivos devem ter sempre marcadores e verbos no infinitivo. Existem 2 tipos de Objetivos, o Geral e os Específicos. Como o nome diz, o Geral deve contemplar apenas um verbo. Já para os Específicos, sugere-se trabalhar entre dois até três verbos no infinitivo, pois devem ser respondidos nos seus resultados e conclusão.

continua...

...continuação

Etapas	Conceitos	Pontos importantes
Planejamento	Após definir o problema e os Objetivos (Geral e Específicos), é preciso determinar um processo para resolvê-los e uma forma de obter informações sobre a variável ou as variáveis em estudo. A variável corresponde ao conjunto de resultados possíveis de um fenômeno.	Aqui se deve: - Determinar os procedimentos necessários para resolver o problema e os objetivos. - Definir a metodologia: Tipo de abordagem da pesquisa, o tipo de pesquisa e os métodos que serão utilizados (veremos isso mais a frente). - Elaborar um cronograma das atividades e o custo da pesquisa.
Execução	Cumprir e realizar o que foi definido no planejamento para dar início a coleta de dados.	
Coleta de dados	Realizada após cuidadoso planejamento e a devida determinação das características mensuráveis do fenômeno coletivamente típico e sem qualquer tipo de tendência durante a coleta dos dados, que busca mensurar as informações com base em um método sistemático que busca responder o problema, testando hipóteses e avaliando resultados.	Pode ser: - Dados secundários ou indiretos: dados tabulados, coletados ou analisados prontos para consulta. Ex.: relatórios, livros, revistas, etc. - Fontes primárias ou diretas: dados brutos (que nunca foram tabulados, coletados ou analisados: Ex.: leitores, consumidores, análise de filmagem, fotos, entrevistas observações). Nesse caso, é necessário trabalhar ou com amostra ou com censo, através da avaliação de uma população (veremos isso mais a frente). Os dados podem ser coletados de forma contínua (todos os dias e de forma rotineira), periódica (em intervalos constantes de tempo), ou ocasional (a fim de atender a uma conjuntura ou a uma emergência, como no caso de epidemias e catástrofes).

continua...

conclusão

Etapas	Conceitos	Pontos importantes
Resultados	Trata dos resultados dos dados apurados e processados na coleta de dados de forma a trazer as interpretações e análises quantitativas ou qualitativas, e poderão ser representados estatisticamente através de tabelas e gráficos. Nesse caso, pode fazer parte tanto da Estatística Descritiva como da Inferencial. Assim, trata os dados numéricos do fenômeno e os interpreta conforme a medida estatística em estudo.	A apuração dos dados estatísticos pode ser feita manualmente ou com o uso de algum <i>software</i> ou aplicativo.
Conclusão	Responder o problema e os objetivos da pesquisa.	Após todas as etapas da pesquisa, basta divulgá-la.

Fonte: Elaboração própria, 2024.

No próximo tópico, explicaremos alguns termos estatísticos presentes no Quadro 1 e que fazem parte do método da Estatística.

1.2 ESTATÍSTICA COMO MÉTODO

Antes de iniciar os tipos de pesquisa Estatística, é necessário definir a Abordagem que será utilizada. A Abordagem é o grupo de atividades sistemáticas que permite contemplar os objetivos da pesquisa (Quadro 1), ou seja, corresponde ao conjunto de métodos que serão organizados de forma lógica para solucionarem o problema da pesquisa (Quadro 1).

Ao se escolher a Abordagem, existem duas: a Dedutiva e a Indutiva. Mas é somente possível escolher uma dessas. A Dedutiva é mais tradicional, que usa premissas gerais para uma conclusão específica; assim, utiliza-se a parte teórica e suas principais regras

para apenas depois serem apresentados os exemplos. Já a Indutiva sai de casos específicos para chegar em conclusões gerais, sendo uma abordagem de cunho de raciocínio e inferência lógica; assim, gera-se conclusões gerais a partir de observações específicas, existindo maior taxa de erros do que a dedutiva.

Após a escolha da Abordagem, deve-se escolher o Método, que pode ser quantitativo ou qualitativo. Essa escolha dependerá do que se deseja transmitir nos resultados e na conclusão, de acordo com o problema e os objetivos estipulados da pesquisa conforme o Quadro 2, que mostra a diferença entre os Métodos de cunho qualitativo do quantitativo. Também tem a ver com a maneira que o pesquisador pretende avaliar as variáveis de uso estatístico, que são denominadas de Qualitativas e Quantitativas.

Quadro 2 – Diferença entre o Estudo Qualitativo e Quantitativo ao se definir o Método

Qualitativa (categórica ou de atributo)	Quantitativa (numérica)
Descreve ou categoriza um elemento da população estudada.	Quantifica um elemento de uma população.
Trabalha amostras reduzidas.	Trabalha amostras amplas.
Os dados são analisados em seu conteúdo psicossocial.	Os dados são tratados estatisticamente.
Indicada quando se tem poucas informações sobre a população em estudo.	Mais indicado quando já se tem mais informações sobre a população em estudo.
Instrumentos de coleta de dados semiestruturados.	Instrumentos de coleta de dados mais estruturados.
Fornece análise mais profunda sobre a população em estudo (motivações, hábitos, atitudes e tendências de comportamento), embora nem sempre seja generalizável.	Fornece informações numéricas em relação ao consumo (quantidades, intensidades, etc.), que frequentemente são generalizáveis.

continua...

conclusão

Qualitativa (categórica ou de atributo)	Quantitativa (numérica)
Pode ser de duas formas durante os resultados e as conclusões da pesquisa: -Nominal: identifica as categorias; e -Ordinal: há uma ordenação das categorias.	Pode ser de duas formas durante os resultados e as conclusões da pesquisa: -Discreta: assume alguns valores; e -Contínua: assume diversos valores.
Amostra não probabilística.	Amostra probabilística.
Análise com o uso de conteúdos.	Análise com o uso da Estatística.

Fonte: Elaboração própria, 2024.

Observação: Alguns autores admitem a pesquisa quali quanti, que ocorre para interpretar as informações quantitativas por meio de símbolos numéricos e os dados qualitativos mediante a observação, fazendo um *mix* entre as duas (Rodrigues, 2021).

Após a escolha das variáveis do estudo presente na Abordagem, é necessário escolher o tipo de pesquisa. Existem várias classificações de pesquisas que resultam da utilização de diferentes variáveis de classificação pelos autores, mas para a área de estatística, optou-se por usar três tipos: a Exploratória ou *Desk Research*; a Descritiva ou *Ad Hoc*; e a Pesquisa Causal ou Experimental. O Quadro 3 traz o conceito e o uso de cada uma delas. Durante a pesquisa, deve-se usar apenas uma das três (Exploratória, Descritiva ou Experimental). Após a escolha de uma delas, pode-se utilizar mais de um modo de uso da mesma presente na coluna 3, Quadro 3.

Quadro 3 – Conceito e uso das três principais Pesquisas para a área de Estatística

Pesquisa	Conceito	Onde usar	Alguns verbos presentes nos seus objetivos da pesquisa (Quadro 1)	Pontos a serem avaliados ao escolher o tipo de Pesquisa
Exploratória ou <i>Desk Research</i>	É utilizada para elevar o conhecimento do pesquisador sobre um tema que ainda lhe é totalmente desconhecido, sendo muito útil para gerar questões e/ou hipóteses de pesquisa, e não deve ser utilizada para tomada de decisões; são informais, flexíveis e criativas.	- Levantamento de fontes secundárias: envolvem levantamentos bibliográficos, documentais, estatísticos e de pesquisas realizadas. -Levantamento de experiências (entrevistas ou observações informais de cunho individuais ou em grupo) -Estudo de caso.	- Identificar - Compreender - Citar	Vantagem: informações a baixo custo e coleta mais rápida do que as outras duas pesquisas Desvantagem: obsolescência das informações, por falta de atualização ou inexistência de dados fundamentais; e dificuldade de direcionar muitas vezes para o tema, o problema e os objetivos das pesquisas, pois os dados são secundários. Analisa dando uma noção do assunto, mas não aprofunda ou detalha, por isso o nome exploratória.

continua...

...continuação

Pesquisa	Conceito	Onde usar	Alguns verbos presentes nos seus objetivos da pesquisa (Quadro 1)	Pontos a serem avaliados ao escolher o tipo de Pesquisa
Descritiva ou Ad Hoc	Obter dados sobre características de grupos, estimar proporções de determinadas características e verificar a existência de relações entre variáveis. Dessa forma, visa descrever situações do mercado ou do ambiente a partir de dados primários, obtidos originalmente por meio de entrevistas pessoais ou discussões em grupo, relacionando e confirmando as hipóteses levantadas na definição do problema de pesquisa. Assim, constitui-se um estudo tanto quantitativo como qualitativo.	-Levantamento ou estudo de campo (dados representativos da população de interesse através de sumários estatísticos, como médias, modas, variâncias, decis, porcentagens e suas relações). -Estudos de campo (preocupa-se com algumas situações típicas, relativas as suas variáveis), com o uso de Entrevistas e observações formais. -Painéis (tipo de pesquisas evolutivas: informações coletadas e analisadas obtidas de forma periódica, levando em conta o acompanhamento do tempo; pode sofrer a influência dos dois itens citados a cima) - Dados primários. - <i>Omnibus</i> (análise da mesma população de forma constante).	-Descrever -Caracterizar -Atribuir -Detalhar	Vantagem: Informação direcionada para tema, problema e objetivos pois usa fontes primárias, promovendo detalhamento dos atributos. Desvantagem: Demora mais tempo e tem mais custo do que a exploratória.

continua...

Pesquisa	Conceito	Onde usar	Alguns verbos presentes nos seus objetivos da pesquisa (Quadro 1)	Pontos a serem avaliados ao escolher o tipo de Pesquisa
Pesquisa Causal ou Experimental	É um tipo de pesquisa na qual um pesquisador manipula e controla uma ou mais variáveis independentes e observa a oscilação na(s) variável(eis) dependentes concomitantemente à manipulação das variáveis independentes.	-Experimento de Campo, ocorre quando o pesquisador, em uma situação real, realiza manipulação de uma ou mais variáveis independentes, sob condições controladas quando a situação permitir, e é mensurado seu efeito sobre uma ou mais variáveis dependentes. -Experimento em laboratório, quando o pesquisador consegue criar uma situação exata do que deseja ter, controlando algumas e manipula outras variáveis. Também pode observar e medir os efeitos dessa manipulação das variáveis independentes sobre as variáveis dependentes. -Teste de Mercado, em que produtos e serviços são colocados em áreas selecionadas para se avaliar a potencial demanda, com alteração na variável de <i>marketing</i> (como propaganda, preço e promoções, embalagem, composição) em um período determinado.	-Comparar -Experimentar -Resultar -Causar	Vantagem: trabalha sempre com área controle e teste, que são pontos importantes para se abordar questões relacionadas com inovação, aprimoramento, melhoria e aprendizado. As relações de causa e efeito obtidas nesse tipo de pesquisa dificilmente serão obtidas nas pesquisas exploratórias e descritiva. Desvantagem: dentre as três pesquisas, é a mais cara de todas.

Fonte: Elaboração própria, 2024.

Explicação de alguns termos presentes no Quadro 3

- Variáveis dependentes são aquelas que variam concomitantemente a outras variáveis (chamadas independentes), cuja variação conjunta interessa ao pesquisador mensurar. Ex.: vendas, participação no mercado, experimentação, atitudes, imagem, etc.
- Variáveis independentes (ou tratamentos) são aquelas manipuladas e cujos efeitos sobre as variáveis dependentes se deseja medir. Ex.: preço, gastos em publicidade, tipos de promoções, meios de distribuição, componentes dos produtos, etc.
- Grupo experimental é a unidade de teste exposta ao tratamento experimental.
- Grupo controle são as unidades de teste não expostas ao tratamento experimental e que servem para medir os efeitos das outras variáveis a que tanto o grupo experimental quanto o de controle estão sujeitos.

Percebe-se, através do Quadro 3, que a pesquisa experimental deve sempre trabalhar com duas áreas, também denominadas de dois grupos: teste (novos dados) e controle (dados existentes). Assim, segue abaixo exemplos de como pode ocorrer a pesquisa experimental na área de administração (Samara; Barros, 2006):

- **Só depois e sem grupo de controle** (Exemplo): A veiculação de campanha publicitária em uma área restrita para observar o comportamento de determinado grupo de usuários ou cidadãos, já que só poderá avaliar após a introdução da variável comunicação. A comparação com grupo semelhante que não foi exposto ao teste e livre de interferência de outras variáveis é difícil de mensurar, por isso é denominado só depois e sem grupo de controle.
- **Antes e Depois sem grupo de controle** (Exemplo): Expectativas antes de conhecer um serviço e após conhecê-lo. Maior dificuldade deste estudo é o grupo selecionado ser influenciado pela

mediação antes de conhecer o produto ou serviço, assim tem tendência de a resposta posteriormente não estar de acordo com a realidade.

- **Só depois e com grupo de controle** (Exemplo): Um projeto em que há alteração da disposição dos produtos nas gôndolas de supermercados, considerando a proximidade com diversas marcas concorrentes e também as disposições usuais mantidas nos grupos de controle de uma empresa de distribuição de alimentos. Neste exemplo específico percebe-se que há unidades de testes que passarão pelos testes de alterações nas variáveis deste grupo de controle; e outras que servirão de grupo de controle e sem o procedimento de medições realizadas antes.
- **Antes e depois com grupo de controle** (Exemplo): Determinar os efeitos de uma mudança em uma característica da qualidade de um produto/serviço em termos de inovações tecnológicas. Seria realizada uma medição anterior sobre o consumo e opinião a respeito do produto em questão com as empresas selecionadas como unidades de testes e também com o grupo de controle. Nesse caso deve ser constituído por empresas similares que utilizam quantidades próximas do produto, e com o mesmo fim.

Após identificar o tipo de pesquisa, é possível calcular o tamanho da amostra a partir da população em estudo. Somente não se fará essa etapa se a pesquisa escolhida for a exploratória e a opção for o item fontes secundárias. Nesse caso específico, o pesquisador identifica os dados que têm interesse para iniciar a organização deles conforme parâmetros da sua pesquisa. Já se o item escolhido for outro da pesquisa exploratória, bem como da descritiva e da experimental, é possível calcular a amostra (se necessário) e o projeto da pesquisa deve passar pela aprovação do Comitê de Ética e Pesquisa (CEP) da universidade ou da empresa, antes de se iniciar.

**Dicas de Sites para fazer pesquisa
exploratória através de fontes secundárias**

AUDIÊNCIA IBOPE. Disponível em:
<https://www.ibope.com.br/>. Acesso em: 11 Fev. 2024.



CONSELHO REGIONAL ESTATÍSTICA – 3ª REGIÃO
(CONRE3). Banco de Dados. 2024. Disponível em:
<http://www.conre3.org.br/portal/banco-de-dados/>.
Acesso em: 11 fev. 2024.

DATA FOLHA. Disponível em: <https://datafolha.folha.uol.com.br/>. Acesso em: 11 Fev. 2024.



INDICADORES SOCIOECONÔMICOS DO BRASIL.
Disponível em: <https://www.ibge.gov.br/indicadores>.
Acesso em: 11 Fev. 2024.

MAPA DA VIOLÊNCIA. Disponível em:
<https://www.gov.br/pt-br>. Acesso em: 11 Fev. 2024.



BRASIL. *Instituto Brasileiro De Geografia E Estatística*
(IBGE). Brasília, DF, 2018. Disponível em: <https://www.ibge.gov.br/nossos-sites.html>. Acesso em: 11. Fev. 2024.

FGV IBRE. São Paulo, 2020. Disponível em:
<https://portalibre.fgv.br/>. Acesso em: 11.fev.2024





INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA). Brasília, DF, 2010. Disponível em: <https://www.ipea.gov.br/portal/>. Acesso em: 11 Fev. 2024.

PORTAL DE DADOS ABERTOS DO MINISTÉRIO DA EDUCAÇÃO. Brasília, DF, 2017. Disponível em: <https://dadosabertos.mec.gov.br/>. Acesso em: 11 fev. 2024.



PROGRAMA INTERNACIONAL DE AVALIAÇÃO DOS ESTUDANTES (PISA). Brasília, DF, 2020. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/pisa>. Acesso em: 11 fev. 2024.

SCIENTIFIC ELETRONIC LIBRARY ONLINE (SCIELO). São Paulo, 2000. Disponível em: <https://scielo.org/>. Acesso em: 11 fev. 2024.



SUPERINTENDÊNCIA DE ESTUDOS ECONÔMICOS E SOCIAIS DA BAHIA (SEI). Brasília, DF, 2020. Disponível em: <https://www.sei.ba.gov.br/>. Acesso em: 11 Fev. 2024.

UNESCO BRASÍLIA. Brasília, DF, 2022. Disponível em: <https://pt.unesco.org/fieldoffice/brasil>. Acesso em: 11 fev. 2024



Outro aspecto é saber o que irá pesquisar, por isso é necessário identificar inicialmente o universo, seguido da população e, por último, escolher se fará um censo ou amostragem (ato de tirar uma amostra). O Quadro 4 mostra os conceitos de cada item anterior.

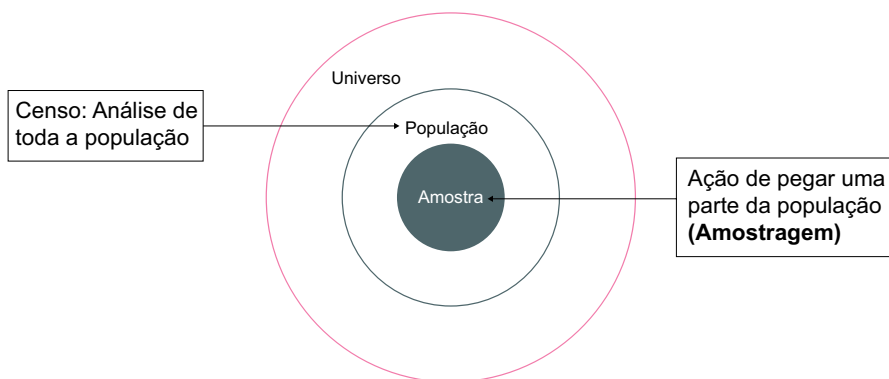
Quadro 4 – Conceito de universo, população, censo e amostra

Palavras estatísticas	Conceito	Tipos
Universo	É tudo aquilo que pode ser mensurado ou que é possível adquirir dados para futuras conclusões, como a totalidade de pessoas, animais, plantas ou objetos.	-
População	É o conjunto de todos os objetos que possam oferecer dados referentes ao estudo em questão.	- Finita: População cujos componentes podem ou poderão ser listados fisicamente, identificando-se por exemplo com nome ou números. - Infinita: População cujo número de componentes é ilimitado
Censo	É a análise de toda a população, não retirando um pedaço dela (amostra).	-
Amostra	É qualquer subconjunto da população.	- Probabilística: Aquela em que cada elemento da população tem uma chance conhecida e diferente de zero de ser selecionada para compor a amostra. Fornecem estimativas com precisão da população, pois utiliza a probabilidade de um evento, ou seja, é um número entre zero e um, que representa a possibilidade da ocorrência do evento-frequência de longo prazo. - Não Probabilística: Aquela em que a seleção dos elementos da população para compor a amostra depende ao menos em parte por critérios subjetivos de julgamento do pesquisador ou do entrevistado do campo, de acordo com sua experiência e com os objetivos do estudo; são utilizadas apenas para pesquisa exploratória. Os resultados desse tipo de amostra são tidos como preliminares e servem apenas para se ter uma ideia mais precisa de diversas possibilidades. Nela, faz-se a manipulação da população e se extraem somente os elementos com uma ou mais características. Assim, não garante a representatividade da população.

Fonte: Elaboração própria, 2024.

Para exemplificar o Quadro 4, foi feita a Figura 3. Percebe-se que dentro do universo existe inicialmente a população e depois a amostra, que representa a interseção entre os outros dois elementos. Assim, o pesquisador deve buscar uma amostra com representatividade tanto no universo quanto na população, ou seja, que esteja na parte azul da Figura 3.

Figura 3 – Relação entre o universo, população e amostra



Fonte: Elaboração própria, 2024.

Muitas vezes é melhor utilizar uma amostra ao invés do censo, pois geralmente é difícil completar o mesmo com todos os elementos do universo. Outro aspecto regularmente requer, a depender do tamanho da população em estudo, uma equipe de entrevistadores, necessitando de recursos e tempo. Dessa forma, na maioria dos estudos é melhor trabalhar com uma amostra, pois em um pequeno intervalo de tempo se pode gerar informações mais precisas e com representatividade, ao invés de se trabalhar com toda a população (censo) e faltarem dados.

Ao procurar uma amostra, deve-se sempre buscar a representatividade e evitar que sejam tendenciosas, pois caso haja uma escolha desse tipo, suas medidas características diferem dos atributos correspondentes da população representada, invalidando

assim suas conclusões. De modo geral, não há como remover a tendenciosidade após a extração da amostra e nenhuma informação pode ser retirada dela.

Para que a amostra tenha representatividade, deve-se escolher os elementos sem qualquer preconceito ou tendência. Seu tamanho deve ser grande ou o suficiente para minimizar os efeitos de ser considerada uma amostra atípica ou com resultados irrelevantes, que não são capazes de representar a população do estudo.

Para reduzir a tendenciosidade, é bom sempre fazer um pré-teste ou teste piloto do seu levantamento de dados e elaborar um relatório detalhado do seu método amostral, para ver se outras pessoas detectam alguma tendência onde não se espera encontrar nenhuma.

CURIOSIDADE – VOCÊ SABIA?

Pesquisas em que um grande grupo de voluntários é convidado a responder e todos eles são contados se inclinam a maior tendência e suas conclusões estão quase sempre em discordância com a população interessada. Esse tipo de pesquisa é muito usado em votações de *reality shows*, programas de rádio e televisão, e em pesquisa de opinião de internet. Esse tipo de amostra geralmente não tem valor.

Entretanto, quando trata de grupos voluntários para experimentos na área de saúde, são formas diferentes de pesquisa e têm grande valor por trabalharem com duas áreas ou grupos: o de controle e o de teste.

1.3 TIPOS DE AMOSTRAGEM

A amostragem, conforme o Quadro 4, pode ser Probabilística e não Probabilística. A primeira é subdividida em Aleatória,

Sistemática, Estratificada e Conglomerada (Quadro 5). Já a segunda se classifica em Conveniência, Intencional e Cotas (Quadro 6).

O pesquisador deve escolher apenas um tipo de amostra, ou probabilística ou não probabilística, para o cálculo a partir da população. Depois, precisa escolher também uma das subdivisões dessas.



O vídeo faz uma breve explicação da pesquisa probabilística



Link: <https://youtu.be/l0mLCmX4rZo>



O Quadro 5 indica os tipos de amostras probabilísticas com as suas subdivisões.

Quadro 5 – Tipo de amostra probabilística e suas subdivisões com os respectivos conceitos para o uso na pesquisa estatística

Subdivisão	Conceito	Como fazer
Aleatória ou simples	Há uma igual probabilidade, diferente de zero, de cada elemento da população ser escolhido por meio de um sorteio. É a escolha aleatória dos elementos que farão parte da amostra.	Para realizá-la, primeiro se deve listar a população; depois, determinar o tamanho da amostra, o índice de confiabilidade e o índice de precisão através de fórmulas. Após essa etapa, é necessário estabelecer o uso de número aleatórios, seja através de Tabela ou algoritmos computacionais.

continua...

...continuação

Subdivisão	Conceito	Como fazer
Sistemática	Fundada em uma lista de elementos já identificados e com o cálculo da amostra estabelecido a partir da quantidade da população. Escolhe-se os elementos em intervalos fixos e iguais. A principal vantagem é que é mais fácil detectar os elementos em grandes amostras. A desvantagem é em relação à posição dos elementos na lista e com a interferência de fenômenos sazonais e repetições durante a pesquisa.	Esse intervalo é obtido como divisão do número do universo, ou população, pelo número da amostra, como no seguinte exemplo: $I = N/n$, onde I = Intervalo, N = População, e n = amostra. Por exemplo: se a lista tiver 100 elementos, escolhe-se de forma aleatória o número 10. Dividindo $100/10 = 10$. Assim, a escolha de forma sistemática na lista será a cada 10 elementos.
Estratificada	Se pretendido trabalhar com atributos ou características de uma população que possui diversidade dos seus elementos. Consiste em dividir a população em subgrupos mais homogêneos (estratos) e retirar amostras aleatórias simples deles.	Primeiro há a identificação da população que se deseja estudar. Depois, dentro dessa haverá a definição dos estratos, que é feita por meio da retirada com o uso de amostragem simples e aleatória mediante listagem ou número aleatório. Então, junte isso numa só amostra a fim de se obter uma amostra de toda a população. A ideia principal é que se obtenha um grupo homogêneo que requer uma amostra menor que um grupo heterogêneo. Esse sorteio pode ser de forma Uniforme (sorteia-se igual número de elementos em cada estrato); Proporcional (o número de elementos em cada estrato é proporcional ao número de elementos existentes no estrato); e Ótima (quando se toma em cada estrato um número de elementos proporcional ao número de elementos do estrato e também à variação da variável de interesse no estrato, medida pelo seu desvio padrão).

continua...

conclusão

Subdivisão	Conceito	Como fazer
Conglomerada	Exige a utilização de mapas detalhados de regiões, estados, município e cidades, pois, para seleção da amostra, há subdivisões das áreas a serem pesquisadas por bairros, quarteirões e domicílios, que serão sorteados para composição dos elementos. A pesquisa será realizada de forma sistemática ou simples, para que não haja interferências de informações. Percebe-se que é um método muito utilizado por motivos de ordem prática e econômica: a população é dividida em pequenos grupos e se sorteia um número suficiente desses (conglomerados), que são a amostra.	Primeiro, define-se os elementos da amostra dentro do conglomerado através de mapas (usa-se o Google Maps ou outro aplicativo de mapa que tem a ver com a pesquisa). Então, extrai-se uma amostra aleatória simples, sistemática ou estratificada dos elementos da área, listando as localidades também de forma aleatória, sistemática ou estratificada, para depois iniciar a pesquisa dos elementos, seja com o uso de observação ou entrevista (por exemplo).

Fonte: Elaboração própria, 2024.

Através do Quadro 5, percebe-se que, nas quatro subdivisões da amostra probabilística, é necessário realizar o cálculo da amostra a partir da sua população existente de forma a ter seus graus de confiabilidade (intervalo definido por uma porcentagem fixa) e seu índice de precisão. É possível fazer isso usando fórmulas, tabelas e programas computacionais *on-line*.

A primeira fórmula é $n_0 = \frac{1}{\epsilon_0^2}$, onde n_0 é tamanho da amostra e E_0 é o erro amostral. O objetivo dessa primeira fórmula é definir o erro amostral, ou seja, seu índice de precisão. Se for definido como 0,05, usa-se a confiabilidade de 95% (delimitar o erro amostral em valores que variem entre 0,09 e 0,01, geralmente é utilizado o valor 0,05 e 95% de confiabilidade) pelo pesquisador, que colocando na fórmula, fica $n_0 = 1/(E_0)^2$ e se tem $n_0 = 1/(0,05)^2$, obtendo como resposta que $n_0 = 400$, ou seja, o tamanho da amostra não pode ultrapassar o valor de 400 (colocar o elemento em valor inteiro se cair em valor decimal). Depois, para saber o valor da amostra a partir da população delimitada na pesquisa, coloca-se na segunda fórmula $n = \frac{N \times n_0}{N + n_0}$, onde n é a amostra, N é a população, e n_0 é tamanho da amostra. Se N for 50.000 e n_0 for 400 (já calculado no primeiro passo), tem-se que $n = (N \times n_0)/(N + n_0)$ e então, que $n = (50.000 \times 400)/(50.000 + 400) = 396,8$ (aproximadamente 397), ou seja, teremos uma população para ser pesquisada de 397 pessoas dentro de uma de 50.000 e um erro de 0,05. O terceiro passo é confirmar o valor do erro amostral que foi definido no primeiro passo. A fórmula é $\epsilon_0 = \sqrt{\frac{N - n}{N \times n}}$, onde E_0 é o erro amostral; N é a população; e n é a amostra. Tem-se $E_0 = \sqrt{(N - n)/(N \times n)}$, então $E_0 = \sqrt{(50.000 - 397)/(50.000 \times 397)}$; $E_0 = 0,05$, aproximadamente 0,05 (valor confirmado no primeiro passo).

Caso deseje usar os dados da Tabela 1, que corresponde à tabela dos valores definidos no cálculo da amostra, constata-se que a coluna “N” corresponde à população e a coluna “A” são os valores da amostra. Os dados da Tabela 1 têm o valor do erro de 0,05 e 95% de confiabilidade e estão organizados hierarquicamente de

forma crescente a partir do valor da população. Se olharmos o valor de 50.000 na coluna N, tem-se que a amostra é de 381, e não de 397 como no cálculo anterior. Por isso é importante definir na pesquisa e deixar claro no texto para o leitor como chegou ao cálculo da amostra.

Tabela 1– Tabela de cálculo de amostra probabilística com 95% de confiabilidade e 5% de erro amostral (ajuste)

N	A	N	A	N	A
10	12	220	140	1200	291
15	14	230	144	1300	197
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
100	80	500	217	600	361
110	86	550	226	7000	364
120	92	600	234	8000	367
130	97	650	242	900	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	1000000	384

Observação:

N é o tamanho da população

A é o tamanho da amostra

Fonte: Krejcie e Morgan (1970, p. 608).

Se o cálculo for feito por programas *on-line*, posso sugerir três aplicativos gratuitos:

- O Cometto
(*link*: <https://comentto.com/calculadora-amostal/>);



- O Trujillo
(*Link* : <http://www.trujillo.com.br/calculadora.htm>);



- O Surveymonkey (*Link*: <https://pt.surveymonkey.com/mp/sample-size-calculator/>).



A Figura 4 mostra o cálculo da esquerda para direita de cada aplicativo citado anteriormente. Tem-se, respectivamente, que os valores foram 382, 381 e 382. Assim, os valores computacionais foram iguais ou com pequenas diferenças do obtido na Tabela 1. Para o valor dar 381 ou 382, o erro amostral no cálculo da fórmula manual deveria ser aproximadamente 0,06, e não 0,05.

Figura 4 – Cálculo da amostra probabilística em 3 aplicativos *on-line* gratuito

Programas <i>on-line</i> (gratuitos)	
Cometto	População: 50000 Erro amostral (%): 5 Nível de confiança: 95% Distribuição da população: Mais heterogênea (50/50) CALCULAR Resultado: 382
Trujillo	Calcule o tamanho da amostra Nível de confiança: ☒ 95% ☐ 99% Margem de erro: 5 Universo: 50000 Calcular Limpar Tamanho da amostra: 381
SurveyMonkey	Tamanho da população: 50000 Grav de confiança (%): 95 Margem de erro (%): 5 Tamanho da amostra: 382

Fonte: Elaboração própria, 2024.

.....

Vídeo cálculo da amostra probabilística por tabela e *softwares*



Link: https://youtu.be/UN1ON_qv0sA

.....

Observação: Amostra multiestágio é o esquema no qual os elementos do quadro de amostragem são subdivididos e a amostra combina vários métodos amostrais.

Dicas do uso de amostragem probabilística com o uso do Excel

•Vídeo amostra sem reposição probabilística no Excel:
(Link: https://youtu.be/sJpnHVJAg_k)



•Vídeo amostra com reposição probabilística no Excel:
(Link: <https://youtu.be/79p-54iN4xQ>)



Após a explicação da Amostra Probabilística (Quadro 5), iremos explicar a Amostra Não Probabilística (Quadro 6). A primeira é sempre superior a segunda, mas muitas vezes o pesquisador não tem escolha para seu uso. Entretanto, a Não Probabilística pode ser usada quando não se pretende generalizar os dados com o foco para a população, e isso ocorre em pesquisa exploratória quando se pretende ter conhecimento sobre um assunto, e não informações que deem representatividade a uma população. Também pode ocorrer quando não se segue de forma correta as instruções ao selecionar os entrevistados ou se faz isso através da observação, fazendo com que possa faltar informações pela escassez de elementos da população ou por razões de sigilo. Nesses casos, a amostra não probabilística é aceitável. Isso pode ocorrer, por exemplo, quando empresas se recusam a dar informações de produtos, processos ou serviços por questões sigilosas, ou quando profissionais da saúde se reúnem para dar pareceres iniciais sobre um medicamento ou um paciente. Por isso são estudos mais baratos que geram hipóteses para detectar oportunidades ou dificuldades, com o objetivo de gerar resultados e conclusões para trabalhos mais detalhados futuramente.

Quadro 6 – Tipo de Amostra Não Probabilística e suas subdivisões com os respectivos conceitos para o uso na pesquisa estatística

Subdivisão	Conceito
Conveniência ou Acidentais	Ocorre quando os elementos da amostra são selecionados de acordo com a conveniência deliberada do pesquisador. São as pessoas que estão ao alcance do pesquisador e dispostas a responder um questionário, serem observadas ou testarem um produto, muitas vezes de forma voluntária. O elemento é selecionado por estar no local ou no momento em que a pesquisa está sendo realizada. Nessa amostra, o foco é o ambiente de estudo, o local ou a área.

continua...

conclusão

Subdivisão	Conceito
Intencional ou Julgamento	Ocorre quando os elementos da amostra são selecionados segundo um critério de julgamento de forma deliberada pelo pesquisador, tendo como base o que se acredita que se possa fornecer ao estudo e exclui qualquer processo aleatório. Os critérios devem estar estipulados na pesquisa (tema, problema, objetivos e na metodologia). Por exemplo, se queremos verificar as razões e os motivos do uso (ou não) de determinada marca de produto, escolhemos dois grupos a serem entrevistados: os usuários e não usuários do produto. Esse método se caracteriza pelos casos selecionados e é pré-determinado e não depende de chances. Nessa amostra, escolhe-se geralmente as pessoas ou os cargos.
Cotas	O pesquisador procura uma amostra que se identifique em alguns aspectos com a população em estudo. Essa identificação pode estar relacionada ao sexo, idade, etc., e a quantidade a ser entrevistada é aleatória. Aqui, o valor da população deve ser conhecido, pelo menos de forma aproximada, mas não se pode fazer seleções aleatórias. Nessa amostra, o foco são as características ou os atributos da população.

Fonte: Elaboração própria, 2024.

O que é um questionário?

Um instrumento de coleta de informações que propiciará a tabulação, a análise e a medição para entendimento de eventos futuros.

Para a escolha do tipo do tipo de Amostragem, se Probabilística ou Não Probabilística, o pesquisador deve saber: o tipo de pesquisa que usará (Quadro 3); a acessibilidade dos elementos e das informações referentes à população pesquisada; a disponibilidade de ter os elementos organizados de forma crescente ou decrescente (Em ROL); a representatividade desejada ou necessária; o tempo e os recursos disponíveis; entre outros fatores. É uma decisão do pesquisador a partir do seu bom senso e experiência, mantendo as questões éticas e sigilosas se necessário. Outro aspecto importante é nunca divulgar resultados de amostras não probabilísticas como se fossem probabilísticas.

O que é representatividade para seleção de uma amostra?

A representatividade de uma amostra é a probabilidade de encontrar indivíduos com determinadas características nela, e deve ser a mesma de achar indivíduos com a mesma característica da população.

1.4 INSTRUMENTOS DE PESQUISA

Após o pesquisador definir se fará um censo ou utilizará um dos tipos de amostra (probabilística ou não probabilística), é necessário escolher a maneira de coleta dos dados (Quadro 7), que pode ser de forma natural (*face to face* ou cara a cara, apenas com o uso de papel e caneta) ou instrumental (usando alguns recursos), através do uso da comunicação ou a observação, ou a interação dos dois.

A comunicação (Quadro 7) de forma formal ou informal consiste no questionamento, verbal ou escrito, dos respondentes para obtenção do dado desejado, que será fornecido por declaração verbal ou escrita do próprio entrevistado. Já a observação (Quadro 7) corresponde ao registro de comportamento, fatos e ações relacionadas com o objetivo da pesquisa, e não envolve questionamento e respostas verbais ou escritas, sendo utilizada para levantar hipóteses preliminares sobre o comportamento de consumo. Tanto na comunicação como na observação, é sempre bom o pesquisador levar um questionário ou um roteiro para não se perder em relação aos objetivos da pesquisa. É bom relatar que também é possível fazer os dois simultaneamente.

A comunicação é mais versátil, rápida e tem menor custo que a observação. Entretanto, depende da boa vontade e sinceridade do respondente, dele dispor ou lembrar do dado solicitado e dos instrumentos ou a forma de coleta poder influenciar nas

respostas, sendo assim menos preciso do que a observação. Por trabalhar comportamento, a observação é sincera e real por parte do respondente, mas é necessário a análise de especialistas.

Tanto a comunicação como a observação podem ser direta ou indireta. Na direta o fato é coletado no momento de sua ocorrência (fontes primárias), já a indireta se refere aos registros deixados por comportamento ou fatos passados (fontes secundárias).

O instrumento de coleta de dados é o documento através do qual as perguntas e questões devem conter seu tema, problema e objetivos da pesquisa (Quadro 1). Essas serão apresentadas ou formuladas aos respondentes e então serão registradas as respostas e dados obtidos, para depois organizá-los e representá-los em forma de tabela ou gráficos estáticos e então fazer as interpretações para os resultados e conclusões. Dessa forma, o instrumento de coleta de dados mal elaborados, com questões dúbias, escalas incorretas e que apresentam dúvidas quanto ao procedimento para o seu preenchimento, constituem em uma das principais fontes de erros para as pesquisas amostrais probabilísticas ou não probabilística (Quadros 5 e 6).

O projeto e construção de um instrumento de coleta de dados compreendem os seguintes passos (Samara, Barros, 2006):

- Determinação dos dados a serem obtidos a partir do tema, problema e objetivos;
- Determinação do instrumento segundo o método escolhido;
- Redação de rascunho;
- Revisão e pré-teste (necessário fazer antes de aplicar para a amostra real); e
- Redação final.

Quadro 7 – Algumas formas para coleta de dados

Alguns métodos	Vantagens	Desvantagens
Comunicação e Observação		
Pessoal pode ser natural ou instrumental (com uso de celular e aplicativos como Google Meet, Microsoft Teams, Zoom, WhatsApp, entre outros).	• Controle da amostra. • Possibilidade de instruir o entrevistado. • Pode incluir a observação. • Possibilidade de entrevista e observação de longa duração.	• Alto custo. • Segmento de difícil acesso. • Treinamento de pessoal de campo. • Se utilizar instrumento para análise do comportamento através de observação, necessário adequar a câmara.
Celular (instrumento mais completo para pesquisa). Possui, por exemplo: • Câmeras fotográficas; • Câmera para filmagem; • Gravador de som; e • GPS.	A Câmera fotográfica analisa apenas o comportamento. O gravador de som, apenas sons para análise. A filmagem detecta comportamento e som para análise. O GPS identifica, através da posição de latitude e longitude, comportamento.	Se utilizar mais de instrumento, é preciso ter cuidado com a análise, pois muitas vezes necessitará de uma equipe polivalente e isso é caro.
Leitura ótica dos caixas das lojas.	Ótimo para analisar vendas.	Necessário que as informações estejam cruzadas estatisticamente com o banco de dados integrado (cliente, fornecedores, logística, dentre outros). Para a empresa, tem mais informações apenas de entrada e saída de mercadorias e o que mais vendeu.

continua...

...continuação

Alguns métodos	Vantagens	Desvantagens
Comunicação		
Telefone (presente também no celular).	Conversa com a pessoa e detecta elementos de voz.	• Desinteresse dos entrevistados de responder as perguntas. • Amostra não é tão criteriosa. • Custo alto em relação a e-mails, redes sociais e formulários eletrônicos.
E-mails, redes sociais, formulários eletrônicos (como do Google), <i>chat</i> (lembrando que os aplicativos de comunicação como Google Meet, Microsoft Teams, Zoom, WhatsApp, entre outros têm essa função).	• Possibilita grande abrangência da amostra. • Entrevistado tem mais tempo para responder. • Equipes menores para o projeto. • Redes sociais e WhatsApp trabalham com nichos específicos.	• Baixo retorno em relação ao custo de incentivos para cooperação. Se tiver um retorno de 5% até 8% do total da população, é considerado bom. • Há possibilidade de instrução e resolução de dúvidas. • Necessidade de manter os contatos atualizados.
Observação		
Psicogalvanômetro, ou detector de mentiras.	Mede a mudança na transpiração, através da qual se infere a reação emocional das pessoas sujeitas a determinado estímulo.	É um aparelho que trabalha muitas vezes com frequências cardíacas. A pessoa pode mentir se achar que algo é verdadeiro na sua mente, mesmo que o fato seja falso

continua...

...continuação

Alguns métodos	Vantagens	Desvantagens
Câmera do movimento dos olhos, ou Oftalmógrafo.	É utilizada para detectar como uma pessoa lê determinada propaganda impressa na revista ou jornal, embalagens, etc. As medidas são tomadas da sequência em que os movimentos de leitura são efetuados e do tempo despendido em cada ponto do objeto lido. Parte do pressuposto de que se começa a leitura pelos pontos que visualmente despertam a atenção e se dedica aos pontos que despertam o maior interesse no objeto lido. O Oftalmógrafo é o aparelho utilizado para a leitura do movimento oculares.	Uso de lentes prejudica seu uso.
Métodos interativos com o uso de Inteligência Artificial (IA) como apoio e consultoria na coleta dos dados.	Rapidez na coleta simultânea de dados e diminuição (ou eliminação) dos erros decorrentes da interferência humana.	• É imprescindível usar a IA correta, pois cada uma é especializada em áreas específicas como linguagem, desenho, figuras, filmagem, dentre outros. • O aprendizado da IA é similar aos dos seres humanos, então ela também pode possuir tendências e preconceitos durante a análise dos dados.

continua...

conclusão

Alguns métodos	Vantagens	Desvantagens
Pupilômetro.	Mede a mudança no diâmetro da abertura da pupila dos olhos, pois se sabe que as pessoas não têm controle sobre isso. O instrumento pressupõe que, à medida que as pessoas se interessam e têm reações favoráveis por determinado objeto, suas pupilas aumentam de diâmetro. A medição dele dará uma dimensão da favorabilidade da reação ao estímulo provocado.	• Observar se a pessoa não fez uso de algum medicamento ou droga que favoreça o dilatamento da pupila e se o aparelho foi bem instalado para não sofrer influência da iluminação externa.

Fonte: Elaboração própria, 2024.

Atualmente, o instrumento (Quadro 7) mais utilizado é o celular pelos alunos em sala de aula. Através dele, é possível elaborar formulários via Google Forms e aplicá-los diretamente com os entrevistados, desde que o aparelho tenha acesso à internet. Ao fazer isso, os dados já vêm tabulados e com gráficos, o que possibilita sua exportação para o Excel e sua transcrição, facilitando os cálculos estatísticos para futuras interpretações, desde que as perguntas sejam fechadas.

O Quadro 8 mostra os principais tipos de perguntas, que podem ser fechadas, abertas, semiabertas, dicotômica, encadeadas, na matriz de resposta, de ordem preferencial, na escala diferencial semântica, na Escala de Likert e na escala de 3 pontos.

As perguntas fechadas (Quadro 8) são mais fáceis de processar e facilitam o entendimento do entrevistado, mas limitam a espontaneidade do respondente, já que ele tem uma escolha forçada.

Nas perguntas abertas (Quadro 8), como o respondente expressa sua opinião de forma subjetiva ou emotiva, é possível ao pesquisador identificar dimensões antes não pensadas e não induzindo o entrevistado. Entretanto, demandam longo tempo para leitura, codificação e tabulação se forem feitas manualmente, tornam-se mais caras diante disso. Hoje há *softwares* que auxiliam o pesquisador na leitura, codificação e tabulação (como o Nvivo).

A diferença do pesquisador usar a escala de diferencial semântica, a Escala de Likert ou a escala de 3 pontos tem a ver com o aprofundamento da pesquisa, em que a primeira se aprofunda mais que a última. Outro aspecto é que essas escalas podem ser transformadas em nominal, ou seja, no caso da escala de 3 pontos, o número 1 significa péssimo, o 2 regular, e o 3 excelente. Pode-se usar também *emoji* com caras como atributos nominais, a depender do público que responderá o questionário para interagir mais com ele. Ao tabular as três escalas citadas anteriormente, o melhor é usar a média ponderada, que veremos mais à frente no momento das interpretações dos dados.

Quadro 8 – Tipos de pergunta para questionário ou roteiro

Tipo	Fechada
Explicação	Nesse tipo de perguntas, as respostas são fornecidas de forma que apenas uma delas seja possível.
Exemplo	Onde mora? () Centro () Lapa () Barra
Tipo	Aberta
Explicação	Nesse tipo de pergunta, o respondente expressa livremente no espaço reservado o que pensa, ou seja, sua opinião.
Exemplo	Onde você mora?.....
Tipo	Semi-aberta
Explicação	Mescla a pergunta fechada com uma aberta, dando a possibilidade de, se não tiver nos itens fechados, o entrevistado, colocar sua opinião.
Exemplo	Onde mora? () Centro () Lapa () Barra () Outros.....

continua...

...continuação

Tipo	Dicotômica										
Explicação	Admite apenas duas perguntas fechadas. Um questionário não deve ter muitas perguntas dicotômicas, pois não terá aprofundamento na interpretação.										
Exemplo	Sua moradia é própria? () Sim () Não										
Tipo	Encadeada ou Sequencial										
Explicação	Quando uma pergunta puxa a outra tendo relação, ou seja, a resposta de uma dependerá da que foi respondida anteriormente.										
Exemplo	Sua moradia é própria? () Sim () Não Se respondeu sim, antes de ser própria, houve financiamento? () Sim () Não										
Tipo	Matriz de resposta										
Explicação	Monta-se um quadro para facilitar a resposta do entrevistado, assim o respondente somente precisa marcar sua resposta dentro desse.										
Exemplo	<table><tr><td>Bairro</td><td>Mora</td><td>Morou</td></tr><tr><td>Centro</td><td>x</td><td></td></tr><tr><td>Lapa</td><td>x</td><td></td></tr></table>		Bairro	Mora	Morou	Centro	x		Lapa	x	
Bairro	Mora	Morou									
Centro	x										
Lapa	x										

continua...

...continuação

Tipo	Ordem preferencial
Explicação	É dada a possibilidade de escolher: o 1º, o 2º ou o 3º lugar.
Exemplo	Que bairro moraria por preferência? (2º) Centro (1º) Lapa (3º) Barra
Tipo	Diferencial Semântico (Osgood)
Explicação	Usa uma escala de 1 até 7 para dar nota e marcar um dos números.
Exemplo	Em relação à qualidade da sua casa, que nota você dá? Baixa qualidade (1234567) Alta qualidade
Tipo	Escala de Likert
Explicação	Usa uma escala de 1 até 5 para dar nota e marcar um dos números.
Exemplo	Em relação à qualidade da sua casa, que nota você dá? Baixa qualidade (12345) Alta qualidade

continua...

conclusão

Tipo	Escala de 3 pontos
Explicação	Usa uma escala de 1 até 3 para dar nota e marcar um dos números.
Exemplo	Em relação à qualidade da sua casa, que nota você dá? Baixa qualidade (123) Alta qualidade

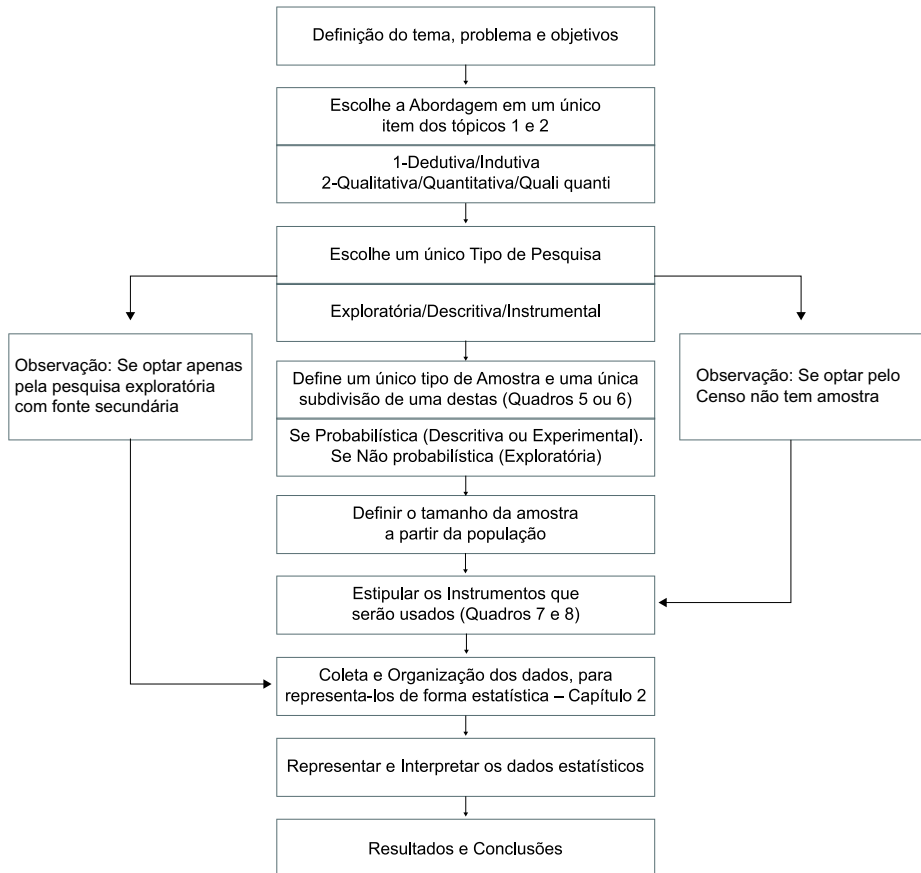
Fonte: Elaboração própria, 2024.

Observação: Lembrar que as perguntas devem ter o conteúdo do tema, problema e objetivos, e que o Quadro 8 tem apenas alguns tipos de perguntas. É possível elaborar outras em seu questionário. Também lembrar de fazer o teste piloto com um público similar a sua amostra antes de aplicar o questionário, para detectar qualquer falha e realizar as correções.

Um questionário deve ser o mais objetivo e claro possível para o entrevistado de forma que ele possa respondê-lo em até 90 segundos, para não se tornar cansativo.

A Figura 5 mostra as etapas que um pesquisador deve ter na elaboração do método estatístico de forma sintética. O capítulo 2 mostrará a organização e a representação dos dados através de tabelas e gráficos, em que a distribuição de frequência está presente com o objetivo de iniciar as interpretações iniciais dessas.

Figura 5 – Resumo sintético do método estatística



Fonte: Elaboração própria, 2024.

1.5 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS

Antes de realizar o exercício abaixo, revise os conteúdos nos 18 *flashcards*. Virar cada carta para ver a resposta, clicando em 📄. Se acertar, coloque 👍, e se errar, ❗. É importante clicar nos últimos dois *emojis* de acerto/erro para saber a quantidade de acertos e erros e melhorar seu aprendizado. Caso não queira seguir a ordem das cartas, pode misturá-las no símbolo 🔀.

Link dos 18s *flashcards*: <https://www.goconqr.com/pt-BR/flashcard/39649768/0-a-estatistica-e-suas-etapas-aplicadas-capitulo-1-estatistica-passo-a-passo-conceitos-basicos>



Veja o estudo de caso – Caso 1 - Potencial de mercado para *Software*

Uma empresa de informática que já está no mercado há 10 anos deseja relançar um jogo para o público maior de 18 anos. O produto apresenta como principais benefícios: interatividade, troca de dinheiro virtual, missões a serem realizadas pelo jogador, um excelente design gráfico na cidade, além de conteúdos sempre relacionados à conscientização dos 17 objetivos de desenvolvimento sustentável (ODS).

É bom lembrar que atos sexuais públicos são considerados crimes no *game*, podendo gerar a prisão do seu personagem. Outros atos ilegais, como usar substâncias ilícitas, servem como *power-ups* para o protagonista.

A empresa dispõe de dados sobre a demanda desse tipo de produto, pois se encontra no mercado há uma década. Então tem conhecimentos em seu banco de dados sobre os atributos do

público-alvo e suas preferências com relação às características do produto. Pergunta-se quais as possíveis:

- Abordagem(ns) que pode(m) ser usada(s);
- Tipos de pesquisa, e se precisa passar pelo CEP;
- Amostras;
- Instrumento(s) que pode(m) ser utilizado(s).

Resposta

Objetivo Geral: lançar no mercado um jogo para o público maior de 18 anos.

Segue abaixo algumas possíveis respostas:

Tipo de Pesquisa	Dados	Amostra	Instrumentos
Exploratória caso tenha como objetivos específicos avaliar o potencial de mercado e identificar alguns atributos do público para melhorar o produto). Não precisa passar pelo CEP se a pesquisa utilizar a opinião do público.	• Secundários. • Publicações na área de jogos, investigações de estudos já realizados sobre o perfil do usuário do jogo (+18) e aplicações de jogo +18, bem como pegar no banco de dados da empresa e em documentos que contemplem os objetivos.	Não probabilística por julgamento, com entrevistas informais especialistas e usuários do jogo (pode ser outra amostra não probabilística).	
Descritiva caso tenha como objetivos específicos avaliar o potencial de mercado, descrever os principais atributos do público para melhorar o produto) – Precisa do CEP	• Primários. • Levantamento ou estudo de campo (dados)	Amostra probabilística estratificada, em que se deve calcular, a partir da população, a amostra, confiabilidade e erro (pode ser outra amostra probabilística).	Comunicação e/ou observação com o uso de questionário. Usar o próprio jogo como instrumento de pesquisa.

continua...

conclusão

Tipo de Pesquisa	Dados	Amostra	Instrumentos
Experimental caso tenha como objetivos específicos avaliar o potencial de mercado, comparar os principais atributos do público em relação aos lançamentos desse jogo no passado (versão antiga ou jogo similares da empresa) para melhorar o produto. - Precisa do CEP	Antes e depois com grupo de controle. Exemplo: determinar os efeitos de uma mudança em uma característica do jogo em termos de inovações tecnológicas – usuários que já jogaram o jogo antes e que passam a jogar com as novas mudanças. Seria realizada uma medição anterior sobre a opinião a respeito do produto em questão com as empresas selecionadas como unidades de testes, e também com o grupo de controle. Nesse caso, deve ser constituído por empresas similares que utilizam quantidades próximas do produto, e com o mesmo fim.	Amostra probabilística estratificada, em que se deve calcular, a partir da população a amostra, confiabilidade e erro (pode ser outra amostra probabilística).	Comunicação e/ou observação com o uso de questionário. Depois dos novos dados coletados, confrontar/comparar antigos para gerar resultados e ver causas e efeitos. Usar o próprio jogo como instrumento de pesquisa.

Escolher apenas um tipo de pesquisa conforme recurso do pesquisador. Entretanto, é possível fazer inicialmente a exploratória para conhecer o mercado e o público. Depois, em outra etapa, fazer a descritiva. Finalizando, terá dados para a experimental. A abordagem para esse tudo pode ser qualitativa ou quali quanti, a depender da explicação do pesquisador em relação ao seu estudo.

2. FREQUÊNCIAS E SUAS REPRESENTAÇÕES ESTATÍSTICAS

OBJETIVOS:

- Organizar os dados coletados na pesquisa em forma de distribuição de frequência;
- Representar os dados da distribuição de frequência em gráficos e tabelas;
- Interpretar os dados das distribuições de frequência nas tabelas; e
- Aplicação do Excel para elaborar a distribuição de frequência e gráficos.

CONTEÚDO:

2.1 DISTRIBUIÇÃO DE FREQUÊNCIA APÓS A COLETA DOS DADOS

2.2 DADOS NÃO AGRUPADOS

2.3 DADOS AGRUPADOS

2.4 DISTRIBUIÇÃO DE FREQUÊNCIA EM TABELA PARA INTERPRETAÇÕES

2.5 USO DO EXCEL PARA DISTRIBUIÇÃO DE FREQUÊNCIA EM GRÁFICOS

2.6 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS

Explicação, Objetivos e Conteúdos do Capítulo 2

Link: https://youtu.be/LrusCE69_Yo



2.1 DISTRIBUIÇÃO DE FREQUÊNCIA APÓS A COLETA DOS DADOS

Como forma de recapitular o capítulo 1, vamos imaginar que o dado bruto ou a tabela primitiva (Quadro 9) referente às notas tiradas por 20 alunos podem ter sido coletadas através de quatro formas diferentes e possíveis:

- Da caderneta de uma sala de aula (fontes secundárias através de pesquisa exploratória);
- Através de uma amostra não probabilística por conveniência, pois o professor trabalha na Instituição de Educação e a amostra não busca tirar conclusões gerais, apenas identificar como anda o desempenho dos alunos daquela sala de aula;
- Através de uma amostra probabilística com o objetivo de descrever (pesquisa descritiva) quais atributos os estudantes mais aprenderam do assunto entre as séries do mesmo ano. Nesse caso, é possível o cálculo da amostra ou fazer um censo; ou
- Através de dados de antes de implantar uma metodologia ativa. Realizar uma comparação do aprendizado antes (área controle – notas antigas, mas da mesma unidade) e depois (área teste) da implantação na mesma sala de aula e na mesma unidade identifica as principais causas de melhoria dos aprendizados dos estudantes a partir da modificação do método de ensino.

Foram apresentadas essas quatro maneiras para mostrar que tudo dependerá da forma como foi delimitado o tema, o problema e os objetivos do pesquisador. Entretanto, o modo de cálculo e análise estatística será semelhante em todos os exemplos anteriores.

Quadro 9 – Dados brutos e uma coleta de dados referente à nota de 20 alunos

0	1	2	3	4	4	4	4	4	5
5	5	5	6	6	6	7	7	8	9

Fonte: Elaboração própria, 2024

O que são dados brutos?

Corresponde a uma lista de dados coletados sem ordenação.

Assim, o primeiro passo é a coleta dos dados brutos. A segunda etapa é transformá-los em Retorno de Aprendizado (ROL) em ordem crescente, decrescente ou em categorias (Tabela 2).

Tabela 2 – Dados organizados em ROL das notas de estudantes – Escola A -2024.1

ROL	Organização do ROL
A = {0;1;2;3;4;4;4;4;5;5;5;5;6;6;7;8;9}	Crescente
B = {9;8;7;7;6;6;6;5;5;5;5;4;4;4;4;3;2;1;0}	Decrescente

Fonte: Elaboração própria, 2024

Na terceira etapa, cabe ao pesquisador decidir como representar o dado de frequência dentro da tabela: se de forma agrupada ou não agrupada. Ambos estão corretos.

Vamos mostrar inicialmente a tabela para dados não agrupados e depois para dados agrupados, ambos de forma crescente e referente à Tabela 2. Mas também poderia ambos serem de maneira decrescente, tudo dependerá de como o pesquisador pretende apresentar os dados. O que importa é que as tabelas devem ter os dados em ROL.

A apresentação tabular é a forma não discursiva de apresentar informações, das quais o dado numérico se destaca como informação central. As tabelas resumem as informações e mostram ao leitor a síntese do fenômeno estudado.

Algumas explicações dos termos estatísticos antes de irmos para os próximos tópicos:

- Classe: corresponde à posição da linha da tabela. Se é dito classe 3, corresponde à linha 3 da tabela;
- Dados não agrupados: significa que o valores apresentados no ROL dentro da tabela não estão representados em formato de intervalo;
- Dados agrupados: significa que o valores apresentados no ROL dentro da tabela estão representados em formato de intervalo;
- Amplitude: é um desvio que corresponde ao maior menos o menor valor. É possível calcular a Amplitude do Intervalo referente a cada classe (linha) ou do dado da tabela como um todo (Tópico 2.3-3º Passo);
- Frequência (f_i): é o número de observações, fatos, acontecimentos ou eventos correspondentes numa classe na tabela;
- Simples: quer dizer que somente há um único resultado na estatística;
- Absoluta: quer dizer que o resultado é em número;
- Relativa: quer dizer que o resultado é em porcentagem ou em probabilidade;
- Acumulada: quer dizer que os valores se acumulam dentro de um intervalo;
- Desvio: Afastamento.

2.2 DADOS NÃO AGRUPADOS

Para realizar a explicação do assunto, foi feita a Figura 6. Do lado esquerdo estão os dados brutos e do lado direito a tabela de dados não agrupados. A explicação da Figura 6 será feita por etapa para facilitá-la.

- **1º Passo** – Elaborar a Tabela com sua Numeração, Título, e moldura sem fechamento do lado direito e esquerdo. Na primeira linha (classe) da tabela deve aparecer o significado de cada coluna (no caso específico, primeira coluna é Notas, segunda é contagem e valores de cada dado bruto, e a terceira coluna é frequência); a parte inferior, ou seja, seu rodapé, deve conter a fonte.
- **2º Passo** – Devido à opção de organizar os dados, tabular de forma crescente pelo pesquisador. Colocar na primeira coluna as notas organizadas de forma crescente do menor número (zero) até o maior (nove), a partir do que foi observado do lado esquerdo da Figura 6 (Item – Dados Brutos).
- **3º Passo** – Na coluna “Contagem dentro da Tabela 2 (Figura 6)”, é necessário identificar a quantidade de elementos existentes nos dados brutos e referentes a primeira coluna da tabela (denominada Notas) de maneira a facilitar a visualização e a quantidade de cada elemento encontrado no item “Dado Bruto”. Foi colocado com a mesma cor desse a quantidade de barras na coluna “Contagem – (Ver Dados Brutos – conforme cor)” e o número correspondente na coluna denominada frequência.
- **4º Passo** – A tabela de dados não agrupados com as notas dos estudantes de uma determinada série de uma escola A no período de 2024.1 já está representada através da Tabela 2, dentro da Figura 6. Na terceira coluna, chamada de “Frequência **Simples Absoluta** (f_i)”, é possível identificar a quantidade numérica de observações ou acontecimentos a partir dos dados brutos.

Figura 6 – Elaboração da tabela de dados não agrupados (ROL – crescente) a partir de seu dado bruto

Dados Brutos			Tabela para dados não agrupados		
0 4 4 5 5 6 6 7 6 8 9 3 4	1	2	Tabela 1 – Dados organizados em ROL (Não Agrupados) das notas de estudantes – Escola A – 2024.1		
	4	4			
	4	4			
	5 5	5			
	6 6	6			
	7 6	7			
	8	9 3 4			
			Tabela 1 – Dados organizados em ROL (Não Agrupados) das notas de estudantes – Escola A – 2024.1		

Fonte: Elaboração própria, 2024.

Caso não haja frequência (*fi*), deve-se colocar na tabela as seguintes expressões:

- um traço horizontal (—) quando o valor é zero, não só quanto à natureza das coisas como quanto ao resultado do inquérito;
- três pontos (...) quando não temos os dados;
- um ponto de interrogação (?) quando temos dúvida quanto à exatidão de determinado valor;
- zero (0) quando o valor é muito pequeno para ser expresso pela unidade utilizada; e
- se os valores são expressos em numerais decimais, precisamos acrescentar à parte decimal um número correspondente de zeros (0,0; 0,00; 0,000;...).

Qualquer tabela é diferente de quadro. Dentro da tabela, coloca-se informações de texto e números, e ela não é fechada do lado direito e esquerdo com linhas (Figura 7). O quadro apenas tem texto e é fechado do lado direito e esquerdo com linhas. A Figura 7 mostra como deve ser apresentada uma tabela. Através da Figura 7, percebe-se que a tabela deve ter os seguintes itens:

- Topo – Espaço superior de uma tabela destinado ao seu número e ao seu título;
- Centro – Espaço central de uma tabela destinado à moldura, aos dados numéricos e aos termos necessários à sua compreensão. Nele, identificam-se quatro espaços menores: o do cabeçalho, a coluna, a linha e a célula;
- Espaço do cabeçalho – Espaço superior do centro de uma tabela destinado a indicação do conteúdo de uma coluna.
- Coluna – Espaço vertical do centro de uma tabela que contém informações sobre o conteúdo das colunas ou dados numéricos de uma mesma direção vertical;
- Linha – Informações contidas em uma mesma direção horizontal;
- Célula – Espaço mínimo no centro de uma tabela, resultante do cruzamento de uma linha com uma coluna, destinado a um dado numérico ou a um sinal convencional;
- Rodapé – Espaço inferior de uma tabela destinado à fonte, à nota geral e à nota específica.

Figura 7 – Apresentação de uma tabela conforme normas

Coluna indicadora	Título		Coluna	← Topo da tabela
	SÉRIE	FREQUÊNCIA		← Cabeçalho
	xxx	yyy		← Célula
← Linha				← Centro
	SOMATÓRIO	VALOR DO SOMATÓRIO		
	Rodapé			← Rodapé

Fonte: Aksenen (2009, p. 39).

2.3 DADOS AGRUPADOS

Para realizar a explicação de como colocar os dados de forma agrupada (em intervalo) na tabela, optei por escrever passo a passo, facilitando cada etapa. Constitui em um procedimento mais formal, para que cada linha (classe) da tabela tenha um intervalo com o mesmo valor e, dessa forma, o leitor consiga ter uma melhor dimensão da distribuição das frequências ao visualizar a tabela. Quando há intervalos diferentes, os valores são concentrados em diversas classes e o leitor não consegue ter a real dimensão dos números distribuídos na frequência.

- **1º Passo:** Colocar os valores em ordem crescente ou decrescente (ROL) com o objetivo de organizar os dados para visualizar o maior e o menor valor. O pesquisador escolhe a forma que deseja organizar. Fiz a opção pelo ROL crescente.
- **2º Passo:** Identificar o maior (9) e o menor (0) valor da amostra da Tabela 2, organizados de forma crescente (ROL).

Menor Valor da amostra

Maior Valor da amostra

A = {0;1;2;3;4;4;4;4;4;5;5;5;5;6;6;6;7;7;8;9} – ROL notas dos alunos estudantes da Escola A – 2024.1

- **3º Passo:** Calcular a série a Amplitude Total (At) = maior valor – menor valor. No caso específico, para que haja substituição na fórmula anterior. Tem-se maior valor 9 (2º Passo) e o menor valor 0 (1º Passo). Assim, a fórmula da Amplitude Total (At) fica $At = 9 - 0 = 9$

Observação:

1) Toda vez que ver a palavra “amplitude” em conteúdos estatísticos, quer dizer: maior valor – menor valor.

2) No caso específico, para calcular a amplitude foi pego o valor da amostra. Mas também seria possível considerar o valor das notas dos alunos, levando em consideração que a avaliação teve como parâmetro a menor nota, que foi zero, e a maior, que foi dez; apesar de nenhum estudante da turma ter tirado. Ao fazer isso, o pesquisador analisará e fará a tabela com base no parâmetro Institucional Educacional que varia entre zero e dez, e não o parâmetro da amostra da sala. Ambas as análises estão corretas, mas tudo dependerá do que o pesquisador deseja mostrar, se o foco é o parâmetro institucional ou o da amostra da sala.

- **4º Passo:** Encontrar o número de linhas que terá a tabela. Para isso, aplica-se na Fórmula de Sturges $(K) = 1 + (3,3 \times \text{Log } N)$, onde:
 - N = número total de observações, frequência ou acontecimento. No caso é de 20 estudantes, corresponde ao Total ou o Somatório (representado pelo símbolo Σ).
 - 3,3 é um valor fixo da Fórmula de Sturges, alguns autores colocam 3,32 ou 3,33.
 - Log de N será $\text{Log } 20 = 1,301029996$.

Ao substituir na fórmula, tem-se que $K = 1 + (3,3 \times 1,301029996)$. Assim, $K = 1 + 4,2 = 5,2$. A tabela terá cinco linhas, pois se deve

considerar como a classe da tabela, ou seja, sua quantidade de linhas, um número inteiro.

- **5º Passo:** Encontrar o intervalo (I) fixo de cada linha ou classe da tabela. A fórmula é $I = At/K$, onde I é o valor do intervalo que se deseja encontrar. At é 9 (3º Passo) e K é 5 (número de linhas da Tabela – 4º Passo). Fazendo a substituição de fórmula, tem-se que $I = 9/5,2 = 1,73$, ou seja, **2** (é necessário transformar em número inteiro o resultado, mas a divisão deve ser feita com o número sem ser inteiro para se obter um valor o mais próximo possível da realidade).

A Figura 8 mostra os principais tipos de intervalos. Para o caso específico, por estar em ordem crescente e em ROL, o intervalo usado no 6º passo foi o fechado à esquerda, representado na segunda linha da Figura 8. Entretanto, para representação de intervalo em tabela, não usar aberto (4ª linha – Figura 8).

Figura 8 – Tipo de intervalo para colocar na tabela em construção

INTERVALOS	POSSÍVEIS REPRESENTAÇÕES	SIGNIFICADO
$x_1 \text{ — } x_2$	$[; [$ $[;)$	Inclui o extremo inferior (x_1) e exclui o extremo superior (x_2).
$x_1 \text{ — } x_2$	$] ;]$ $(;]$	Exclui o extremo inferior (x_1) e inclui extremo superior (x_2).
$x_1 \text{ — } x_2$	$] ; [$ $(;)$	Exclui o extremo inferior (x_1) e exclui o extremo superior (x_2).
$x_1 \text{ — } x_2$	$[;]$	Inclui o extremo inferior (x_1) e inclui o extremo superior (x_2).

Fonte: Aksenen (2009, p. 19).

Observação: Na estatística, não é possível usar o intervalo aberto, como $x_1 - x_2$, ou $] ; [$, ou $(;)$, ao colocar na tabela, pois não representa um intervalo. Estando totalmente aberto, vai para o infinito, e tal representação é válida apenas na matemática.

- **6º Passo:** Construir a tabela com cinco linhas (4º Passo) e intervalos de dois em dois (5º Passo) a partir do menor número da série, que nesse caso é o zero.
 - A primeira etapa para elaborar a tabela é iniciar com a primeira coluna denominada “Notas” da tabela do Quadro 10. Para isso, inicia-se com o menor valor da amostra ou do parâmetro. Por ter optado por um ROL crescente, começo do menor (zero), mas poderia ser do maior caso fosse decrescente (9). Em ambos os casos, o melhor intervalo para ser usado é o fechado à esquerda (+), conforme o 5º Passo. Assim, na primeira coluna denominada de “Notas”, colocar em todas as cinco linhas o intervalo fechado à esquerda (representado pelo símbolo +) entre os dois números de cada linha. Depois, soma-se o menor número com o valor fixo do intervalo, e então se tem $0 + 2$ (Valor do Intervalo) = **2** (Primeira classe ou linha da tabela-**cor negro em amarelo**). Como a primeira linha terminou com o número 2, repita-o na segunda classe justamente porque na primeira o valor é aberto do lado direito ($0 + 2$), ou seja, o zero faz parte e vai até o número 1,999....., mas nunca chegará no 2, pois não faz parte da contagem (lado direito aberto $0 + 2$) e somente será quantificado na próxima linha. Dessa forma, repete-se o 2, referente à primeira linha na segunda classe, e se soma mais o valor do intervalo, que é **2** ($2 + 2 = 4$). Então, tem-se agora na segunda classe o intervalo $2 + 4$ e faz isso de forma sucessiva até a quinta linha, parando no intervalo $8 + 10$.
 - A segunda etapa é fazer a contagem da coluna denominada “Dados Brutos” da Figura 9, com o objetivo de colocar nas duas últimas colunas da Tabela 2 da Figura 9. Para facilitar a visualização, foram usadas cores diferentes. Lembrando que se optou por fechado à esquerda e aberto à direita (+), assim o número da esquerda faz parte e contempla

todos os outros até chegar ao último número do intervalo, que não fará parte. Dessa forma, na primeira classe os números 0 e 1 fazem parte (cor vermelha nos dados brutos e na coluna de contagem dos dados [Quadro 10]), encontrando-se dois estudantes que tiraram entre 0 + 2. Mas o 2 não faz parte (+ sinal aberto à direita) e será contemplado na segunda classe, que é fechada à esquerda (2 + 4), e assim sucessivamente. Dessa forma, o intervalo de notas entre 2 + 4 possui dois alunos, no intervalo entre 4 + 6 há oito estudantes, entre 6 + 8 se tem seis e, por último, no intervalo entre 8 + 10 existem dois. Ao se somar todas as frequências (*f_i*), sempre deve dar o número total de elementos analisados na pesquisa, que corresponde à amostra ou ao censo. Ou seja, nesse caso o total do somatório foi 20.

Quadro 10 – Elaboração da Tabela para dados agrupados a partir de dados brutos encontrados na pesquisa

Dados Brutos		Tabela para Dados Agrupados				
0	1	2	Tabela 1 – Dados organizados em ROL (Agrupado) das notas de estudantes – Escola A-2024.1 Parou no último número 10, pois Tabela tem que ter 5 linhas 2 – soma-se sempre o intervalo com o número que esta no Limite Inferior (LI) Menor Número (ROL crescente)			
3	4					
4	4					
4	4					
5						
5	5					
6	6	6				
6	7					
7	8					
9						

Notas	Número de linha	Contagem (Ver Dados Brutos - conforme cor)	Frequência Simples Absoluta (<i>f_i</i>)
0 + 2 (0 + 2 = 2)	1	//	2
2 + 4 (2 + 2 = 4)	2	//	2
4 + 6 (4 + 2 = 6)	3		8
6 + 8 (6 + 2 = 8)	4		6
8 + 10 (8 + 2 = 10)	5	//	2
Total (Σ)			20

Fonte: Elaboração própria, 2024.

Fonte: Elaboração própria, 2024.

Observação:

- 1) Todos os números da primeira coluna que estão do lado esquerdo são chamados de Limite Inferior (LI), já os números do lado direito são chamados de Limite Superior (LS).
- 2) Se o pesquisador optar por fonte secundária oriunda de pesquisa exploratória e os dados estiverem brutos ou em formato tabelas não agrupadas, é possível também elaborar uma tabela para dados agrupados.

2.4 DISTRIBUIÇÃO DE FREQUÊNCIA EM TABELA PARA INTERPRETAÇÕES

A Tabela 3 foi formatada a partir das normas da Figura 7.

Os cálculos apresentados nas colunas da Tabela 3 podem ser usados tanto para dados agrupados como não agrupados, com exceção da última coluna, denominada Ponto Médio (X_i) e cujo cálculo está apenas presente nos dados agrupados, visto que os dados não agrupados correspondem às informações que foram coletadas na pesquisa e que correspondem a primeira coluna da Figura 6, denominada de Notas.

Inicialmente, mostraremos o cálculo mecânico de cada coluna e depois a explicação referente às respectivas. Hoje, qualquer programa faz o cálculo (como, por exemplo, o Excel). O mais importante para o pesquisador, que é o responsável pela tomada de decisão, é justamente as interpretações dos dados da tabela. Para facilitar a compreensão, manteremos a mesma tabela presente no Quadro 10.

Os cálculos presentes da Tabela 3 e suas explicações ocorrem nessa ordem: Frequência (f_i); Frequência Relativa, Frequência Acumulada Crescente-Abaixo (FACA); Frequência Acumulada

Crescente Relativa (FACR); Frequência Acumulada Decrescente-Acima (FADA); Frequência Acumulada Decrescente (FADR); e Xi (Ponto Médio).

A frequência denominada f_i na segunda coluna corresponde à Frequência Simples Absoluta. A palavra “frequência” na estatística quer dizer acontecimento, o termo “simples” significa que possui um único valor, enquanto a palavra “absoluta” significa que o valor encontrado deve ser numérico. A Frequência Simples Absoluta (f_i) já foi encontrada na pesquisa tanto no exemplo de dados não agrupados (Figura 6 – Tabela 1) como agrupados (Figura 9 – Tabela 1). Assim a segunda coluna da Tabela 1 (Figura 9, para dados agrupados) significa, dentro da interpretação, que dois alunos da Escola A no ano de 2024.1 tiraram notas entre os intervalos $0 \text{ - } 2$ e $2 \text{ - } 4$, oito alunos tiraram notas entre $4 \text{ - } 6$, seis alunos tiraram notas entre $6 \text{ - } 8$ e dois entre $8 \text{ - } 10$. A diferença entre a Tabela 1 – Figura 9 (dados agrupados) e a Tabela 1 – Figura 6 (dados não agrupados) é que, no primeiro caso, os dados das notas são apresentados em intervalo e não se sabe o valor numérico exato das notas dos alunos, apenas o intervalo dessas. Já na Tabela 1 – Figura 6 (dados não agrupados), sabe-se exatamente a nota do aluno associada à quantidade de alunos que tirou a respectiva nota, por exemplo, tem-se que um aluno tirou exatamente zero, um aluno tirou nota igual a 1, e assim sucessivamente. Ambas as tabelas estão corretas e dependerá do que o pesquisador deseja divulgar. Também em ambas as tabelas é importante observar que o Total de entrevistados (20 alunos) deve estar presente tanto nos dados agrupados (Tabela 1 – Figura 9) como nos dados não agrupados (Tabela 1 – Figura 6) no item denominado Total ou Somatório (Σ).

Aconselha-se sempre apresentar dados agrupados de forma tabular ou em gráfico quando a quantidade de informações é muito grande e se deseja sintetizá-las.

Já a Frequência Relativa (FR) significa que a resposta deve ser em porcentagem ou em probabilidade. Afinal, o termo

frequência significa acontecimento, observação ou ocorrência, e a palavra relativa significa porcentagem ou em probabilidade. Por isso que FR é a divisão dos fi presentes em cada classe (linha), dividido pelo total vezes 100 ($fi/\text{total} \times 100$). Caso não multiplique por 100, não transforma em porcentagem, e o que se terá é um índice. Outro aspecto importante é que o somatório de todas as porcentagens de fi deve dar 100% no final. Na primeira classe da coluna de FR, tem-se $2/20 \times 100 = 10$, na segunda $2/20 \times 100 = 10$, na terceira $8/20 \times 100 = 40$, e assim sucessivamente para as outras linhas. A interpretação da terceira coluna da Tabela 1 (Figura 9, para dados agrupados) significa que 10% dos alunos da Escola A no ano de 2024.1 tiraram notas entre os intervalos 0 + 2 e 2 + 4, respectivamente; 12% dos estudantes tiraram notas entre 4 + 6; 30% dos alunos tiraram notas entre 6 + 8; e 10% dos estudantes entre 8 + 10. A soma de todas as porcentagens ($10\% + 10\% + 40\% + 30\% + 10\%$) dá exatamente 100%.

Observação: Caso o somatório não dê 100%, é necessário arredondamento. Por exemplo, se o somatório der 99,9%, falta 0,1%, então esse valor deverá ser somado ao maior FR. Caso o somatório for 100,1%, há a mais 0,1% e, assim, deve-se subtrair do maior.

Em relação à Frequência Acumulada Crescente – Abaixo (FACA) – Tabela 3, significa que os dados são acumulados dentro de um intervalo e irão aumentando do menor para o maior (forma crescente). Já a palavra “abaixo” quer dizer que, durante a interpretação dos dados, deve-se olhar os números menores da referência de análise a partir do seu Limite Superior (LS). A partir dessa dica teórica, vamos iniciar os cálculos para uma melhor compreensão da FACA – Tabela 3. Primeiro, repete-se o primeiro número da Frequência Simples Absoluta (fi) da 1ª classe, no caso específico o número **2** na cor vermelha é levado para a FACA. Depois esse se soma ao 2 da segunda classe (linha) da Frequência

Simples Absoluta (f_i), e assim se tem $2 + 2$ (segunda linha de f_i da Tabela 3) = 4; ao achar o 4, coloca-o na coluna da FACA na segunda classe. Depois, soma-se o número encontrado, o 4 da FACA, com o elemento da terceira classe f_i , que no caso é 8. Então, tem-se $4 + 8 = 12$. Faz esse cálculo mecânico até a quinta linha da Tabela 3 (posição antes da classe do somatório), e o último valor da FACA deve ser sempre o valor do somatório, nesse caso 20 (Tabela 3). Na interpretação do caso, se quiser saber a quantidade de alunos que tiraram notas menores que 6, deve-se olhar o LS que tem o número 6 (primeira coluna Tabela 3) e as notas nessa coluna que têm valores menores que seis, afinal, a FACA quer dizer menor que seis. Assim, tem-se o intervalo que engloba e acumula os números de 0 a 6 e a resposta é 12 (terceira classe da FACA). Percebe-se que a FACA acumula os valores numéricos de f_i dos intervalos: de 0 a 2, que é 2; de 2 a 4, que é 2; e de 4 a 6, que é 8. Somando-se os valores, dará exatamente 12 ($2 + 2 + 8$), sendo essa outra forma de se chegar a esse valor. Outro fato a ser observado é que o primeiro valor da FACA será sempre igual ao primeiro valor numérico de f_i , que no caso é 2, e o último valor da FACA deve ser o total, que nesse caso é 20.

A Frequência Acumulada Crescente Relativa (FACR) nada mais é do que a $FACA / \text{total} \times 100$. Nesse caso específico, multiplica-se por 100, pois tem a palavra relativa no termo FACR, que significa que o valor deve ser em porcentagem. Pela fórmula apresentada anteriormente, a FACR tem origem na FACA, o que significa que seus valores em porcentagem também devem ir crescendo e que durante a sua interpretação, deve-se olhar os números menores à referência de análise a partir do seu Limite Superior (LS) da tabela. Assim, tem-se na primeira classe do FACR $2/2 \times 100 = 10\%$, na segunda classe o valor é $4/20 \times 100 = 20\%$, e assim se seguem os cálculos de forma sucessiva (Tabela 3). Na interpretação, caso queira saber a porcentagem de alunos que tiraram notas menores que 6, deve-se olhar o LS que tem o número 6 (primeira coluna Tabela 3) e as notas nessa coluna que têm os

valores percentuais menores que 6, afinal a FACR, como a FACA, quer dizer abaixo de 6. Assim, tem-se o intervalo que engloba e acumula os números de 0 + 6 e a resposta é 60% (terceira classe da FACR). Consta-se que o FACR acumula os valores percentuais de FR nos intervalos: de 0 + 2, que é 10%; de 2 + 4, que é 10%; e de 4 + 6, que é 40. A soma dos valores dará exatamente 60% (10% + 10% + 40%), sendo essa outra forma de se chegar ao valor de 60%. O primeiro valor do FACR será sempre igual ao primeiro valor percentual de FR, que no caso é **10%**, e o último valor da FACR deve ser o total de FR, ou seja, **100%**.

Em relação à Frequência Acumulada Decrescente – Acima (FADA) – Tabela 3, significa que os dados são acumulados dentro de um intervalo e diminuirão do maior (sempre o total ou somatório da tabela de f_i) para o menor (forma decrescente). Já a palavra “acima” quer dizer que, durante a interpretação dos dados, deve-se olhar os números maiores. A referência de análise é a partir do seu Limite Inferior (LI). A partir dessa dica teórica, vamos iniciar os cálculos para uma melhor compreensão da FADA – Tabela 3. Primeiro, repete-se o valor total ou o somatório da Frequência Simples Absoluta (f_i) da 1ª classe, no caso específico o número **20** na cor verde é levado para a primeira classe da FADA. Depois, o **20** que foi para FADA diminui do primeiro número da primeira classe da Frequência Simples Absoluta (f_i), que no caso é o **2**, e assim se tem **20 - 2** (primeira linha de f_i da Tabela 3) = 18. Ao achar o 18, coloca-o na coluna da FADA na segunda classe. Depois, subtrai o número encontrado, o 18 da FADA, com o elemento da segunda classe f_i , no caso 2. Então, tem-se $18 - 2 = 16$. Faz esse cálculo mecânico até a linha 5 da Tabela 3 (sempre a posição antes da classe do somatório). Na interpretação, caso quiser saber a quantidade de alunos que tiraram notas acima de 6, deve-se olhar o LI que tem o número 6 (primeira coluna Tabela 3) e as notas nessa coluna que têm valores maiores que 6, afinal a FADA quer dizer acima de 6. Assim, tem-se o intervalo que engloba e acumula os números de 6 + 10 e a resposta é 8 (quarta classe da FADA).

Percebe-se que a FADA acumula os valores numéricos de f_i dos intervalos de 6 + 8, que é 6, e de 8 + 10, que é 2. A soma dos valores dará exatamente 8 (6 + 2), sendo outra forma de se chegar a esse valor. Outro fato a ser observado é que o primeiro valor do FADA será sempre igual ao somatório, que é **20** nesse caso, e o último valor da FADA deve ser sempre igual ao valor de f_i anterior ao somatório, nesse caso **2** (Tabela 3).

A Frequência Acumulada Crescente Relativa (FADR) nada mais é do que a $FADA/total \times 100$. Nesse caso específico, multiplica-se por 100 pois tem a palavra relativa no termo FADR, significando que o valor deve ser em porcentagem. Pela fórmula apresentada anteriormente, a FADR tem origem na FADA, por isso seus valores em porcentagem também devem ir decrescendo e, durante a sua interpretação, deve-se olhar os números maiores. A referência de análise é a partir do seu Limite Superior (LI) da Tabela. Assim, tem-se na primeira classe do FADR $20/20 \times 100 = 100\%$, na segunda classe o valor é $18/20 \times 100 = 90\%$, e assim se seguem os cálculos de forma sucessiva (Tabela 3). Na interpretação, caso queira saber a porcentagem de alunos que tiraram notas acima de 6, deve-se olhar o LI que tem o número 6 (primeira coluna Tabela 3) e as notas nessa coluna que têm os valores percentuais maiores que 6, afinal a FADR, como FADA, quer dizer acima de 6. Assim, tem-se o intervalo que engloba e acumula as porcentagens de 6 + 10 e a resposta é 40% (quarta classe da FADR). Percebe-se que a FADR acumula os valores percentuais de FR dos intervalos de 6 + 8, que é 30%, e de 8 + 10, que é 10%. A soma dos valores dará exatamente 40 (30% + 10%), sendo outra forma de se chegar ao valor de 40%. Outro fato a ser observado é que o primeiro valor da FADA será sempre igual ao somatório de FR, que é **100%**, e o último valor da FADA deve ser sempre igual ao valor de FR anterior somatório, nesse caso **10%** (Tabela 3).

O Ponto Médio ($\bar{X}_i$), última coluna da Tabela 3, deve ser calculado somente para dados agrupados, pois nessa classe da

tabela há intervalos que possuem diversos números (com o objetivo de se ter um valor referência que dê significado para cada intervalo nas classe), e ele é calculado através da seguinte fórmula: $(LI + LS)/2$. Ao substituir essa fórmula na primeira classe de x_i (Tabela 3), tem-se $(0 + 2)/2 = 1$. Assim, a nota 1 representa as notas dos estudantes da Escola A do ano 2024 que estão presentes no intervalo de $0 \div 2$ (primeira classe). Já na segunda classe de x_i (Tabela 3), o cálculo é $(2 + 4)/2 = 3$, então a nota representa as notas dos estudantes da Escola A do ano 2024 que estão presentes no intervalo de $2 \div 4$ (segunda classe - x_2). Constata-se que a diferença entre o x_i da segunda com o da primeira classe (x_1) dará 2; ou seja, $3 - 1 = 2$ (valor calculado no item 2.3 – 4º Passo – fórmula de Sturges). A diferença do intervalo, que possui valor fixo e sem variação, constitui-se em outra forma de calcular o ponto médio a partir da consciência da variação do intervalo, basta somar com o valor da primeira classe. Ou seja, $1 + 2 = 3$ dará o mesmo valor encontrado na fórmula de X_i da segunda classe. O ponto médio (X_i) para dados não agrupados já possui a representatividade e corresponde aos itens da classe da própria primeira coluna da Tabela 1 – Figura 6, denominada notas.

Tabela 3 – Distribuição das frequências das notas de estudantes para futuras interpretações – Escola A-2024.1

Notas	Frequência Simples Absoluta (fi)	Frequência Relativa (FR) (%)	FACA (+)	FACR (%)	FADA (-)	FADR (%)	× i (Ponto Médio)
0 + 2	2	2/20 × 100 = 10	2	2/20 × 100 = 10	20	20/20 × 100 = 100	(0 + 2)/2 = 1
2 + 4	2	2/20 × 100 = 10	2 + 2 = 4	4/20 × 100 = 20	20 – 2 = 18	18/20 × 100 = 90	(2 + 4)/2 = 3
4 + 6	8	8/20 × 100 = 40	4 + 8 = 12	12/20 × 100 = 60	18 – 2 = 16	16/20 × 100 = 80	(4 + 6)/2 = 5
6 + 8	6	6/20 × 100 = 30	12 + 6 = 18	18/20 × 100 = 90	16 – 8 = 8	8/20 × 100 = 40	(6 + 8)/2 = 7
8 + 10	2	2/20 × 100 = 10	18 + 2 = 20	20/20 × 100 = 100	8 – 6 = 2	2/20 × 100 = 10	(8 + 10)/2 = 9
Total (Σ)	20	(10 + 10 + 40 + 30 + 10) = 100					

Fonte: Elaboração própria, 2024.

.....

Vídeo de como elaborar, organizar, tabular e calcular
frequências dos dados não agrupados no Excel



Link: <https://youtu.be/aVrQS97oBA0>

.....

Vídeo de como elaborar, organizar, tabular e calcular
frequências dos dados agrupados no Excel



Link: <https://youtu.be/KHapKASITzQ>

.....

Ver filme para fazer interpretações de Dados a partir
de distribuição de frequência:



Link: <https://youtu.be/SIZJqARRThM>

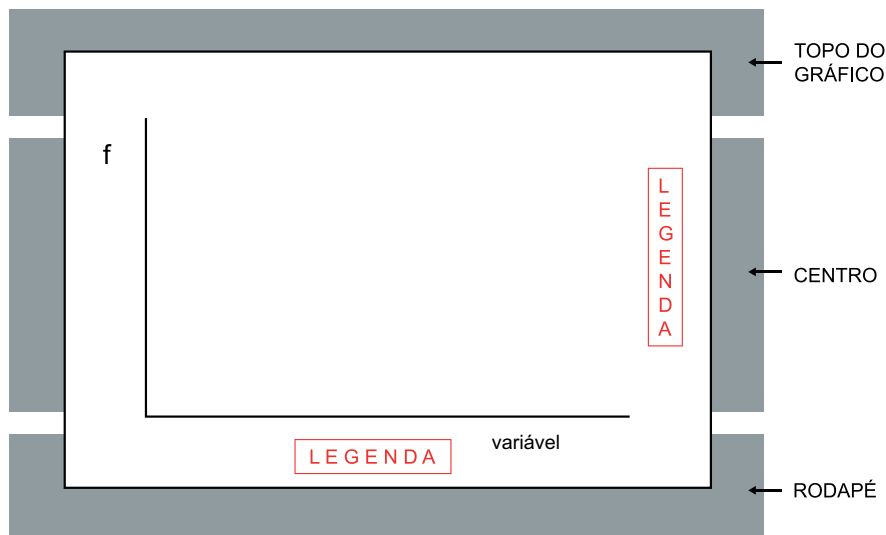
.....

2.5 USO DO EXCEL PARA DISTRIBUIÇÃO DE FREQUÊNCIA EM GRÁFICOS

Assim como as tabelas, os gráficos servem como representações estatísticas para resumir dados qualitativos e quantitativos, e seguem algumas regras para serem elaborados. Conforme a Figura 9, todo gráfico deve ter uma moldura, uma parte central, um sistema de coordenadas ou outro meio para representar os dados, uma legenda e uma fonte.

O topo do gráfico (Figura 9) é o local na parte superior destinado ao seu número e ao seu título (Tema-Local-Período). Já o centro do gráfico é o espaço destinado à base de exposição dos dados. Nesse centro estão identificados dois outros espaços menores: para exposição do sistema de coordenadas, onde serão mostrados os dados, e o espaço da legenda. Já o rodapé fica na parte inferior de um gráfico e é destinado à fonte.

Figura 9 – Apresentação de um gráfico conforme normas



Fonte: Elaboração própria, 2024.

O Quadro 11 mostra os principais tipos de gráficos com seus conceitos e usos, tendo como base a informação a ser representada em cada tabela ou dado bruto adquirida na pesquisa. O pesquisador deve selecionar que gráfico fará da representação e de sua interpretação, detectando se a informação é de cunho:

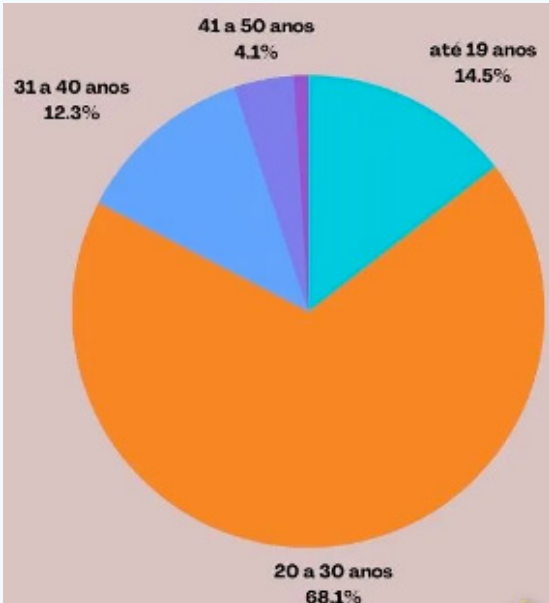
- Qualitativo, que poderá usar os gráficos: de Barra; de Setor ou Pizza; ou de Retangular de composição;
- Quantitativo, que tem como opções os gráficos de: Linha; *Dotplot* ou Diagrama de Pontos; Diagrama de Ramo de folhas; e Histograma; ou
- Quali quanti: Diagrama de Pareto.

Quadro 11 – Principais Tipo de Gráfico com conceitos e exemplos

Tipos	Conceito	Exemplo																																																																		
Informações Qualitativas																																																																				
Tipos	Conceito	Exemplo																																																																		
Barra	Mostra a quantidade de dados que pertence a cada categoria através de uma área retangular, seja de forma vertical ou horizontal, dimensionada proporcionalmente conforme valores. O Gráfico de Barra horizontal ao lado, 11.1, mostra a população do Brasil e de suas respectivas Regiões e Estados. O tamanho dos retângulos é proporcional aos dados numéricos. Assim, as Regiões e Estados que possuem os maiores valores têm os maiores retângulos horizontais.	Gráfico 11.1 – População por Estado, Regiões do País – Brasil – 2022 <table><tr><td>Brasil</td><td>203.062.512 habitantes</td></tr><tr><td>Sudeste</td><td>41,78%</td></tr><tr><td>Nordeste</td><td>26,91%</td></tr><tr><td>Sul</td><td>14,34%</td></tr><tr><td>Norte</td><td>8,54%</td></tr><tr><td>Centro-Oeste</td><td>8,02%</td></tr><tr><td>São Paulo</td><td>44.429.499</td></tr><tr><td>Minas Gerais</td><td>20.538.718</td></tr><tr><td>Rio de Janeiro</td><td>16.054.524</td></tr><tr><td>Bahia</td><td>14.136.417</td></tr><tr><td>Paraná</td><td>11.443.268</td></tr><tr><td>Rio Grande do Sul</td><td>10.888.506</td></tr><tr><td>Pernambuco</td><td>9.058.155</td></tr><tr><td>Ceará</td><td>8.791.688</td></tr><tr><td>Pará</td><td>8.116.132</td></tr><tr><td>Santa Catarina</td><td>7.609.601</td></tr><tr><td>Goiás</td><td>7.095.228</td></tr><tr><td>Maranhão</td><td>6.775.152</td></tr><tr><td>Paraíba</td><td>3.914.495</td></tr><tr><td>Amazonas</td><td>3.941.375</td></tr><tr><td>Espírito Santo</td><td>3.833.486</td></tr><tr><td>Mato Grosso</td><td>3.656.813</td></tr><tr><td>Rio Grande do Norte</td><td>3.302.406</td></tr><tr><td>Piauí</td><td>3.269.200</td></tr><tr><td>Alagoas</td><td>3.122.511</td></tr><tr><td>Distrito Federal</td><td>2.817.068</td></tr><tr><td>Mato Grosso do Sul</td><td>2.756.700</td></tr><tr><td>Sergipe</td><td>2.209.558</td></tr><tr><td>Rorônia</td><td>1.581.016</td></tr><tr><td>Tocantins</td><td>1.511.459</td></tr><tr><td>Acre</td><td>830.026</td></tr><tr><td>Amapá</td><td>703.508</td></tr><tr><td>Roraima</td><td>636.303</td></tr></table>	Brasil	203.062.512 habitantes	Sudeste	41,78%	Nordeste	26,91%	Sul	14,34%	Norte	8,54%	Centro-Oeste	8,02%	São Paulo	44.429.499	Minas Gerais	20.538.718	Rio de Janeiro	16.054.524	Bahia	14.136.417	Paraná	11.443.268	Rio Grande do Sul	10.888.506	Pernambuco	9.058.155	Ceará	8.791.688	Pará	8.116.132	Santa Catarina	7.609.601	Goiás	7.095.228	Maranhão	6.775.152	Paraíba	3.914.495	Amazonas	3.941.375	Espírito Santo	3.833.486	Mato Grosso	3.656.813	Rio Grande do Norte	3.302.406	Piauí	3.269.200	Alagoas	3.122.511	Distrito Federal	2.817.068	Mato Grosso do Sul	2.756.700	Sergipe	2.209.558	Rorônia	1.581.016	Tocantins	1.511.459	Acre	830.026	Amapá	703.508	Roraima	636.303
		Brasil	203.062.512 habitantes																																																																	
Sudeste	41,78%																																																																			
Nordeste	26,91%																																																																			
Sul	14,34%																																																																			
Norte	8,54%																																																																			
Centro-Oeste	8,02%																																																																			
São Paulo	44.429.499																																																																			
Minas Gerais	20.538.718																																																																			
Rio de Janeiro	16.054.524																																																																			
Bahia	14.136.417																																																																			
Paraná	11.443.268																																																																			
Rio Grande do Sul	10.888.506																																																																			
Pernambuco	9.058.155																																																																			
Ceará	8.791.688																																																																			
Pará	8.116.132																																																																			
Santa Catarina	7.609.601																																																																			
Goiás	7.095.228																																																																			
Maranhão	6.775.152																																																																			
Paraíba	3.914.495																																																																			
Amazonas	3.941.375																																																																			
Espírito Santo	3.833.486																																																																			
Mato Grosso	3.656.813																																																																			
Rio Grande do Norte	3.302.406																																																																			
Piauí	3.269.200																																																																			
Alagoas	3.122.511																																																																			
Distrito Federal	2.817.068																																																																			
Mato Grosso do Sul	2.756.700																																																																			
Sergipe	2.209.558																																																																			
Rorônia	1.581.016																																																																			
Tocantins	1.511.459																																																																			
Acre	830.026																																																																			
Amapá	703.508																																																																			
Roraima	636.303																																																																			
Fonte: IBGE (2024).																																																																				

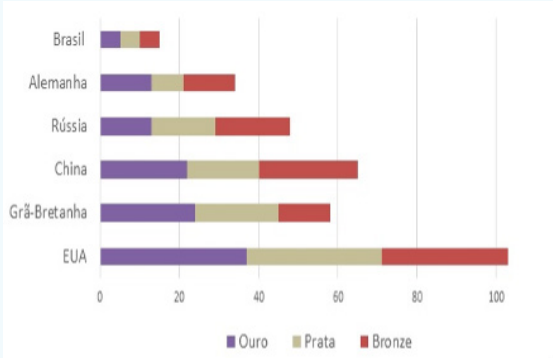
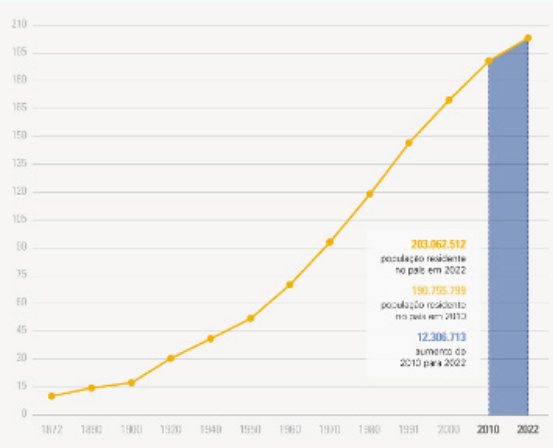
continua...

...continuação

Tipos	Conceito	Exemplo										
Setor ou Pizza	Indicado para mostrar como se divide um todo. Por exemplo, o Gráfico 11.2 mostra em porcentagem a idade dos candidatos que solicitaram Fundo de Financiamento ao Estudante do Ensino Superior (FIES) no ano de 2023.2 no Brasil, em que 68,1% tinham a idade entre 20 e 30 anos. Constata-se que o gráfico em formato de círculo tem porcentagem referente a um em cada parte, ou seja, sua FR e o somatório da porcentagem devem ser 100%.	Gráfico 11.2 – Idade dos candidatos Inscritos no Fundo de Financiamento ao Estudante do Ensino Superior (FIES) – Brasil – 2023.2  <table><tr><th>Idade</th><th>Porcentagem</th></tr><tr><td>até 19 anos</td><td>14.5%</td></tr><tr><td>20 a 30 anos</td><td>68.1%</td></tr><tr><td>31 a 40 anos</td><td>12.3%</td></tr><tr><td>41 a 50 anos</td><td>4.1%</td></tr></table> Fonte: FIES (2024).	Idade	Porcentagem	até 19 anos	14.5%	20 a 30 anos	68.1%	31 a 40 anos	12.3%	41 a 50 anos	4.1%
Idade	Porcentagem											
até 19 anos	14.5%											
20 a 30 anos	68.1%											
31 a 40 anos	12.3%											
41 a 50 anos	4.1%											

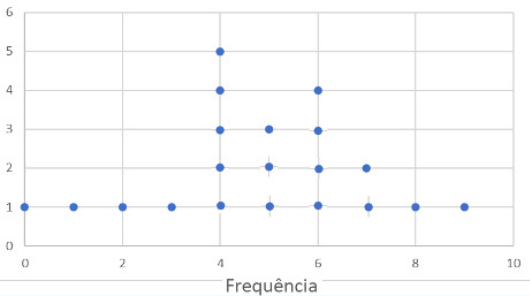
continua...

...continuação

Tipos	Conceito	Exemplo																												
Retangular de composição	É uma alternativa para gráfico de setor, pois também mostra o todo dividido em partes. É constituído por retângulos que podem estar em posição horizontal ou vertical com subdivisões, em que cada parte representa uma categoria de variável. O gráfico 11.3, por exemplo, mostra nas barras horizontais a quantidade de medalha de ouro, prata e bronze que cada país ganhou nas Olimpíadas de 2016.	Gráfico 11.3 – Tipo de Medalha ganho por seis países – Olimpíadas – 2016  <table border="1"><thead><tr><th>País</th><th>Ouro</th><th>Prata</th><th>Bronze</th></tr></thead><tbody><tr><td>Brasil</td><td>10</td><td>5</td><td>5</td></tr><tr><td>Alemanha</td><td>15</td><td>10</td><td>10</td></tr><tr><td>Rússia</td><td>15</td><td>15</td><td>15</td></tr><tr><td>China</td><td>25</td><td>20</td><td>25</td></tr><tr><td>Grã-Bretanha</td><td>25</td><td>20</td><td>15</td></tr><tr><td>EUA</td><td>35</td><td>35</td><td>30</td></tr></tbody></table> Fonte: Tribunal de Contas do Estado do Espírito Santo (2024).	País	Ouro	Prata	Bronze	Brasil	10	5	5	Alemanha	15	10	10	Rússia	15	15	15	China	25	20	25	Grã-Bretanha	25	20	15	EUA	35	35	30
	País	Ouro	Prata	Bronze																										
Brasil	10	5	5																											
Alemanha	15	10	10																											
Rússia	15	15	15																											
China	25	20	25																											
Grã-Bretanha	25	20	15																											
EUA	35	35	30																											
Informações Quantitativas																														
Linha	São ótimos para retratar uma grande quantidade de dados numéricos e contínuos durante um período específico. O Gráfico de linha 11.3a mostra o crescimento da população entre os anos de 1872 até 2022, período que ocorreu o último censo do Brasil.	Gráfico 11.3a – Crescimento da população (em milhões) no Brasil – entre os anos de 1872 até 2022  <table border="1"><thead><tr><th>Ano</th><th>População residente no país (em milhões)</th></tr></thead><tbody><tr><td>1872</td><td>~10</td></tr><tr><td>1890</td><td>~12</td></tr><tr><td>1900</td><td>~15</td></tr><tr><td>1950</td><td>~45</td></tr><tr><td>1960</td><td>~65</td></tr><tr><td>1970</td><td>~95</td></tr><tr><td>1980</td><td>~125</td></tr><tr><td>1990</td><td>~155</td></tr><tr><td>2000</td><td>~185</td></tr><tr><td>2010</td><td>190.755.799</td></tr><tr><td>2022</td><td>203.062.512</td></tr></tbody></table> 203.062.512 população residente no país em 2022 190.755.799 população residente no país em 2010 12.306.713 aumento de 2010 para 2022 Fonte: IBGE (2024).	Ano	População residente no país (em milhões)	1872	~10	1890	~12	1900	~15	1950	~45	1960	~65	1970	~95	1980	~125	1990	~155	2000	~185	2010	190.755.799	2022	203.062.512				
	Ano	População residente no país (em milhões)																												
1872	~10																													
1890	~12																													
1900	~15																													
1950	~45																													
1960	~65																													
1970	~95																													
1980	~125																													
1990	~155																													
2000	~185																													
2010	190.755.799																													
2022	203.062.512																													

continua...

...continuação

Tipos	Conceito	Exemplo
Dotplot ou Diagrama de Pontos	Exibe os dados de uma amostra representando cada valor por meio de um ponto posicionado ao longo de uma escala, que pode ser apresentada de forma horizontal ou vertical. Assim, é um gráfico para ser utilizado ao se iniciar a análise dos dados, pois gera uma imagem deles e classifica-os em ordem numérica. O Exemplo do Gráfico 11.4 mostra os dados não agrupados da Figura 6, sendo possível perceber que é um gráfico para mostrar poucas informações.	Gráfico 11.4 – Dados organizados em ROL (Não Agrupado) das notas de estudantes – Escola A-2024.1  Fonte: Elaboração própria (2024).

continua...

...continuação

Tipos	Conceito	Exemplo																
Diagrama de Ramo de folhas	Exibe os dados de uma amostra utilizando dígitos reais que formam os valores desses dados. Cada valor numérico é dividido em duas partes: o(s) dígito(s) inicial(is) forma(m) o(s) ramo(s), o(s) dígito(s) adicional(is) forma(m) a folha. Os ramos são dispostos ao longo do eixo principal, sendo inserida uma folha para cada valor de dados de forma que se exiba a distribuição dos dados em ROL. O objetivo é observar uma distribuição concentrada sobre um valor central, e em seguida dispensar de algumas maneiras em uma ou ambas as direções, analisando algo que o pesquisador pode ou não ter previsto. O Gráfico 11.5 mostra a concentração das notas 7 e 8, não existindo notas nas casas 3 e 6.	Notas de 10 alunos com casas decimais: 7,4;8,1;7,5;2,0;8,0;9,0;10,0;5,5;4,5;1,2 Agora colocando essas Notas no Gráfico de Ramo de Folhas, que tem a Árvore principal (lado esquerdo – cor verde – Gráfico 11.5) e nas folhas as casas decimais das dez notas (lado direito – cor azul – Gráfico 11.5). Gráfico 11.5 – Notas de 10 alunos – Escola H-2024 <table><tr><td>1</td><td>2</td></tr><tr><td>2</td><td>0</td></tr><tr><td>4</td><td>5</td></tr><tr><td>5</td><td>5</td></tr><tr><td>7</td><td>4 5</td></tr><tr><td>8</td><td>0 1</td></tr><tr><td>9</td><td>0</td></tr><tr><td>10</td><td>0</td></tr></table>	1	2	2	0	4	5	5	5	7	4 5	8	0 1	9	0	10	0
	1	2																
2	0																	
4	5																	
5	5																	
7	4 5																	
8	0 1																	
9	0																	
10	0																	

continua...

...continuação

Tipos	Conceito	Exemplo																																	
Histograma	Quando os dados são contínuos (podem assumir qualquer valor dos números reais). Caso a amostra seja grande, não se deve fazer gráficos de pontos, e sim de histograma, pois através das barras das frequências das observações de uma série de dados, é possível visualizar melhor a amostra. Representa a frequência simples ou relativa dos elementos agrupados ou não agrupados. No exemplo do gráfico 11.6, foram colocados os dados através de gráfico de linha a frequência acumulada da Figura 6.	Gráfico 11.6 – Dados organizados em ROL (Não Agrupado) das notas de estudantes – Escola A-2024.1 <table border="1"><thead><tr><th>Bloco</th><th>Frequência</th><th>% cumulativo</th></tr></thead><tbody><tr><td>0</td><td>1</td><td>16,67%</td></tr><tr><td>1</td><td>1</td><td>33,33%</td></tr><tr><td>2</td><td>1</td><td>50,00%</td></tr><tr><td>3</td><td>1</td><td>66,67%</td></tr><tr><td>4</td><td>5</td><td>83,33%</td></tr><tr><td>5</td><td>4</td><td>100,00%</td></tr><tr><td>6</td><td>3</td><td>100,00%</td></tr><tr><td>7</td><td>2</td><td>100,00%</td></tr><tr><td>8</td><td>1</td><td>100,00%</td></tr><tr><td>Mais</td><td>1</td><td>100,00%</td></tr></tbody></table> Fonte: Elaboração própria (2024).	Bloco	Frequência	% cumulativo	0	1	16,67%	1	1	33,33%	2	1	50,00%	3	1	66,67%	4	5	83,33%	5	4	100,00%	6	3	100,00%	7	2	100,00%	8	1	100,00%	Mais	1	100,00%
Bloco	Frequência	% cumulativo																																	
0	1	16,67%																																	
1	1	33,33%																																	
2	1	50,00%																																	
3	1	66,67%																																	
4	5	83,33%																																	
5	4	100,00%																																	
6	3	100,00%																																	
7	2	100,00%																																	
8	1	100,00%																																	
Mais	1	100,00%																																	

continua...

Tipos	Conceito	Exemplo																																	
Informações Quantitativas e qualitativas																																			
Diagrama de Pareto	Categoriza em um parâmetro aproximado de 80% em relação a atributos que influenciam mais (prioridades) e compara com cerca de mais ou menos 20% das características que terão interferência menor. Por isso se usa o gráfico de barras ordenadas, das mais altas para mais baixas. Junto há também um gráfico de linha que mostra as porcentagens e contagem cumulativas com relação às barras. No exemplo do gráfico 11.7, foi colocado os dados da Figura 6 que deixa claro que 80% dos estudantes tiraram as respectivas notas, na seguinte ordem: 4, 5, 6, 7, 0 e 1 (ver eixo horizontal Gráfico 11.7). Já 20% dos alunos tiraram as notas 2, 3, 8 e 9.	Gráfico 11.7 – Dados organizados em ROL (Não Agrupado) das notas de estudantes – Escola A-2024.1 <table border="1"><thead><tr><th>Nota</th><th>Frequência</th><th>% cumulativo</th></tr></thead><tbody><tr><td>4</td><td>5</td><td>20,00%</td></tr><tr><td>5</td><td>4</td><td>40,00%</td></tr><tr><td>6</td><td>3</td><td>60,00%</td></tr><tr><td>7</td><td>2</td><td>80,00%</td></tr><tr><td>0</td><td>1</td><td>90,00%</td></tr><tr><td>1</td><td>1</td><td>100,00%</td></tr><tr><td>2</td><td>1</td><td>110,00%</td></tr><tr><td>3</td><td>1</td><td>120,00%</td></tr><tr><td>8</td><td>1</td><td>130,00%</td></tr><tr><td>9</td><td>1</td><td>140,00%</td></tr></tbody></table> Fonte: Elaboração própria (2024).	Nota	Frequência	% cumulativo	4	5	20,00%	5	4	40,00%	6	3	60,00%	7	2	80,00%	0	1	90,00%	1	1	100,00%	2	1	110,00%	3	1	120,00%	8	1	130,00%	9	1	140,00%
Nota	Frequência	% cumulativo																																	
4	5	20,00%																																	
5	4	40,00%																																	
6	3	60,00%																																	
7	2	80,00%																																	
0	1	90,00%																																	
1	1	100,00%																																	
2	1	110,00%																																	
3	1	120,00%																																	
8	1	130,00%																																	
9	1	140,00%																																	

Fonte: Elaboração própria, 2024.

Vídeo de como elaborar gráficos com o uso do Excel

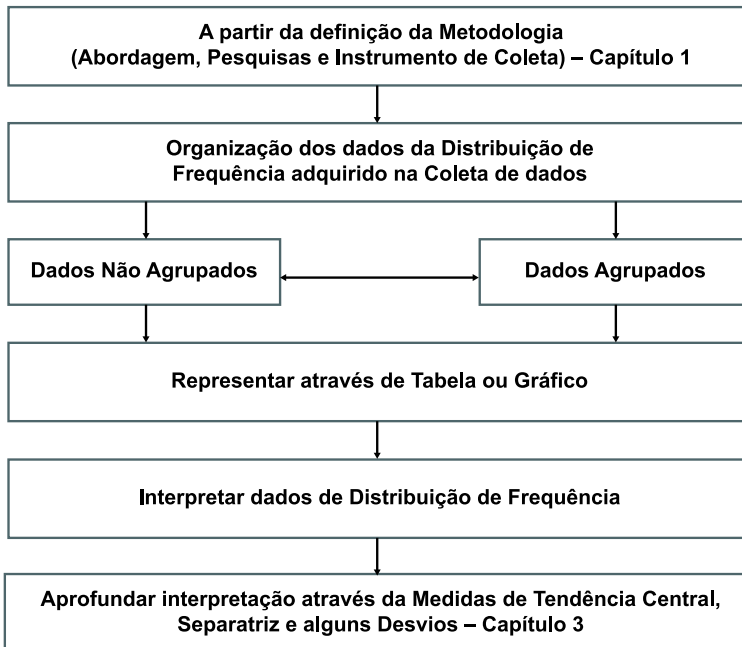


Link: https://youtu.be/5D_XuFu5kfs

A Figura 11 mostra de forma sintética que, após a coleta de dados, inicia-se a organização dos dados agrupados ou não agrupados, para poder representá-los em forma de tabela ou gráfico. Atualmente a melhor maneira de representar os dados estatísticos através das distribuições de frequências é com o uso de programas, como por exemplo o Excel. Nesse momento, já é possível fazer algumas interpretações oriundas da estatística descritiva

No capítulo 3 veremos as medidas de tendência centrais, separatriz e principais desvios estatísticos, que são medidas estatísticas que visam aprofundar mais as interpretações das informações encontradas nas distribuições de frequências e representadas através de gráficos e tabelas (Figura 11).

Figura 11 – Resumo síntese sobre a coleta de dado e a distribuição de frequência



Fonte: Elaboração própria, 2024.

2.6 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS

1 – A distribuição não agrupada abaixo indica o número de acidentes ocorridos com 70 motoristas de uma empresa de ônibus:

Tabela 1 – Distribuição do número de acidentes ocorridos com 70 motoristas – Empresa de ônibus-2024

N.º ACIDENTES	0	1	2	3	4	5	6	7
N.º MOTORISTAS	20	10	16	9	6	5	3	1

Fonte: Elaboração própria, 2024.

Determine, em relação aos dados da Tabela 1:

- a) o número de motoristas que não sofreram nenhum acidente;
- b) o número de motoristas que sofreram pelo menos 4 acidentes;
- c) o número de motoristas que sofreram menos de 3 acidentes;
- d) o número de motoristas que sofreram no mínimo 3 e no máximo 5 acidentes;
- e) a porcentagem de motoristas que sofreram no máximo 2 acidentes.

Faça um *quiz* das 5 perguntas acima no *link*: <https://www.goconqr.com/pt-BR/quiz/39647874/frequencias-e-suas-representacoes-estatisticas-capitulo-2-estatistica-passo-a-passo>



- a) Elabore as frequências da Tabela 1 do número de acidentes com motorista.

Resposta:

PERGUNTAS	RESPOSTAS
a) O número de motoristas que não sofreram acidente:	20
b) O número de motoristas que sofreram pelo menos 4 acidentes:	15
c) O número de motoristas que sofreram menos de 3 acidentes:	35
d) O número de motoristas que sofreram no mínimo 3 e no máximo 5 acidentes:	20
e) A porcentagem dos motoristas que sofreram no máximo 2 acidentes:	52%

Tabela 1 – Distribuição do número de acidentes ocorridos com 70 motoristas – Empresa de ônibus-2024

Acidentes	Motoristas	FR (%)	FACA (+)	FACR (%)	FADA (-)	FADR (%)
0	20	40%	20	40,00%	70	100,00%
1	10	20%	30	60,00%	50	71,43%
2	16	32%	46	92,00%	40	57,14%
3	9	18%	55	110,00%	24	34,29%
4	6	12%	61	122,00%	15	21,43%
5	5	10%	66	132,00%	9	12,86%
6	3	6%	69	138,00%	4	5,71%
7	1	2%	70	140,00%	1	1,43%
TOTAL (Σ)	50	140%				

Fonte: Elaboração própria, 2024.

2 – Uma empresa que vender um conjunto residencial com 400 lotes. Os lotes com tamanho acima de 900 m² serão mantidos de forma sustentável de forma a preservar a floresta nativa da região. Já nos lotes entre 800 e 900 m², serão feitas áreas de lazer como, por exemplo, quadras esportivas, piscina, academia, entre outros. Após realizar o dimensionamento dos lotes, os mesmos ficaram conforme tabela de distribuição agrupada abaixo:

Tabela 2 – Quantidade de Lotes (m²) – Condomínio Residencial-2024

Áreas (m ²)	Nº de lotes (fi)
300 400	14
400 500	46
500 600	58
600 700	76
700 800	68
800 900	62
900 1000	48
1000 1100	22
1100 1200	6
TOTAL (Σ)	400

Fonte: Elaboração própria, 2024.

Pergunta-se:

- a) a amplitude total;
- b) o limite superior da quinta classe;
- c) o limite inferior da oitava classe;
- d) a amplitude do intervalo da segunda classe;
- e) o ponto médio da sétima classe;

- f) a frequência da quarta classe;
- g) a frequência relativa da sexta classe, que ficará para área de lazer;
- h) a frequência acumulada da quinta classe;
- i) o número de lotes cuja área não atinge 700;
- j) o número de lotes cuja área atinge e ultrapassa 800;
- k) a porcentagem dos lotes cuja área não atinge 600;
- l) a porcentagem dos lotes cuja área seja maior ou igual a 900 m^2 e irão para a área sustentável do condomínio;
- m) a porcentagem dos lotes que estão entre 500 e 1000 m^2 , ou seja, 500 m^2 é o menor valor, mas não pode ultrapassar o valor de 1000 m^2 ;
- n) a classe do 72º lote;
- o) até que classe estão incluídos 60% dos lotes.

Faça um *quiz* das dez perguntas acima no *link*: <https://www.goconqr.com/pt-BR/quiz/39647804/2-frequencias-e-suas-representacoes-estatisticas-capitulo-2-estatistica-passo-a-passo>



- a) Elabore as frequências da Tabela 2

Resposta:

PERGUNTAS	RESPOSTAS
a) Amplitude: Amplitude total (At) = Maior valor – menor valor	$1200 - 300 = 900$
b) Limite superior da 5ª classe: $700 \mid 800$	Número máximo da linha = 800
c) Limite Inferior da 8ª classe: $1000 \mid 1100$	Menor número da classe/linha = 1000
d) Amplitude do intervalo da 2ª classe: $400 \mid 500$	$500 - 400 = 100$
e) Ponto médio da 7ª classe: $900 \mid 1000 \mid (Ls + Li/2)$	$1000 + 900 / 2 = 950$
f) A frequência da 4ª classe: fi da respectiva 4ª linha	76
g) A frequência relativa da 6ª classe: $800 \mid 900$	15,5%
h) A frequência acumulada da 5ª classe: FACA/FADA	FACA: 262 FADA: 206
i) O número de lotes cuja área não atinge 700 m²: (FACA 4ª CLASSE)	194
j) O número de lotes cuja área atinge e ultrapassa 800 m²: (FADA 6ª CLASSE)	136
k) A porcentagem dos lotes cuja área não atinge 600 m²: (FACR 3ª CLASSE)	29,5%
l) A porcentagem dos lotes cuja área seja superior a 900 m²: (FADR 7ª CLASSE)	19%
m) A porcentagem dos lotes cuja área é de 500 m², no mínimo, e inferior a 1000 m²:	78%
n) A classe do 72º lote: (Analisar FACA)	3º classe
o) Até que classe estão 60% dos lotes: (Avaliar FACR)	Até a 5ª classe

p) frequências

Tabela 2 – Quantidade de Lotes (m²) – Condomínio Residencial-2024

Áreas (m²)	Nº de lotes (fi)	FR (%)	FACA (+)	FACR (%)	FADA (-)	FADR (%)	XI
300 ┆ 400	14	3,50%	14	3,50%	400	100,00%	350
400 ┆ 500	46	11,50%	60	15,00%	386	96,50%	450
500 ┆ 600	58	14,50%	118	29,50%	340	85,00%	550
600 ┆ 700	76	19,00%	194	48,50%	282	70,50%	650
700 ┆ 800	68	17,00%	262	65,50%	206	51,50%	750
800 ┆ 900	62	15,50%	324	81,00%	138	34,50%	850
900 ┆ 1000	48	12,00%	372	93,00%	76	19,00%	950
1000 ┆ 1100	22	5,50%	394	98,50%	28	7,00%	1050
1100 ┆ 1200	6	1,50%	400	100,00%	6	1,50%	1150
TOTAL (Σ)	400	100%					

Fonte: Elaboração própria, 2024.

3. MEDIDAS DE TENDÊNCIA CENTRAL, SEPARATRIZ E PRINCIPAIS DESVIOS ESTATÍSTICOS

OBJETIVOS:

- Calcular e interpretar o uso das Medidas de Tendência Central e Separatriz;
- Compreender o uso da assimetria e curtose;
- Calcular e interpretar o uso de alguns desvios; e
- Calcular e interpretar o uso da correlação.

CONTEÚDO:

3.1 MEDIDAS DE TENDÊNCIA CENTRAL OU DE POSIÇÃO

3.1.1 Medida de Tendência Central (Média Aritmética) e alguns desvios

3.1.2 Medida de Tendência Central (Mediana)

3.1.2.1 Separatriz

3.1.3 Medida de Tendência Central (Moda)

3.2 MEDIDAS DE ASSIMETRIA E CURTOSE

3.3 CORRELAÇÃO E REGRESSÃO LINEAR

3.4 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS

Explicação, Objetivos e Conteúdos do Capítulo 3

Link: <https://youtu.be/5bVPidku4Fs>



3.1 MEDIDA DE TENDÊNCIA CENTRAL OU DE POSIÇÃO

No capítulo 2, fizemos a organização e as representações das distribuições de frequências, em que os dados foram adquiridos através da definição da amostra, censo ou fontes secundárias (capítulo 1).

Neste capítulo 3, após a representação dos dados de frequência em tabela ou gráficos, iremos utilizar os elementos típicos da distribuição, que são:

- As medidas de tendência central;
- As medidas de variabilidade ou dispersão;
- As medidas de assimetria; e
- As medidas de curtose.

Vamos iniciar com as medidas de tendência central, também chamadas de posição, para aprofundarmos o conhecimento e as explicações presentes nas distribuições de frequências presentes nas tabelas e gráficos.

Há três medidas de tendência central ou de posição, e são elas:

- Média Aritmética;
- Mediana; e
- Moda.

Essas três medidas citadas acima têm esse nome pois os dados observados tendem a se agrupar em torno dos valores centrais distribuídos no eixo horizontal (eixo das abscissas), onde é possível observar e identificar o tipo de assimetria

Já as separatrizes englobam:

- a própria mediana;
- os decis; e
- os percentis.

Começaremos nossa explicação com as medidas de tendência central e depois com as separatrizes. Ambas devem sempre estar em ROL antes de se iniciarem os cálculos.

3.1.1 Medida de Tendência Central (Média Aritmética) e alguns desvios

Há diversos tipos de Média (como por exemplo a média móvel, a geométrica, a harmônica, etc.), mas a mais utilizada é a Média Aritmética, que corresponde ao valor mais representativo e é representada por $\bar{x}$.

O que é a média móvel?

Link: <https://g1.globo.com/ciencia-e-saude/video/entenda-como-funciona-a-media-movel-em-casos-do-novo-coronavirus-8716049.ghtml>



Como calcular a média móvel no Excel e elaborar o gráfico em linha do dado real e da previsão?

Link: <https://youtu.be/9UyQLTm-GMw>



A Média Aritmética corresponde ao quociente da divisão da soma dos valores da variável pelo número deles, ou seja, $\bar{x} = \frac{\sum x_i}{n}$; onde $\bar{x}$ é a Média Aritmética; x_i os valores da variável; e n o número de valores. Primeiro, vamos calcular a média para dados não agrupados (Figura 6 – Capítulo 2), e depois agrupados (Figura 9 – Capítulo 2).

A Média Aritmética para dados não agrupados é denominada Simples. A fórmula de Média Aritmética Simples ($\bar{x}$) é igual ao numerador (parte superior) dividido pelo denominador (parte inferior). O numerador corresponde ao somatório de cada elemento da amostra, ou seja, " $x_1, x_2, \dots, x_n$ ". Já o denominador é igual a n , que significa o total ou o número de elementos da amostra ou da observação dos fenômenos, conforme a Fórmula: $\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$. Veja o exemplo da Figura 11a, no lado esquerdo a Tabela 1, que representa uma distribuição não agrupada das notas de alunos da Escola A no ano de 2024.1 e que também poderia estar representada por $A = \{0;1;2;3;4;4;4;4;5;5;5;5;6;6;6;7;7;8;9\}$. Enquanto na direita se tem o cálculo da Média Aritmética Simples para dados não agrupados, ao fazer o cálculo, tem-se que a nota mais representativa dessa sala é de 4,8, aproximadamente 5,0, já que todos os números de A são inteiros.

Vídeo com o cálculo no Excel da Medida de Tendência Central, Separatriz e Desvio para dados primitivos



Link: <https://youtu.be/ymUcWHcsHqc>

Figura 11a – Exemplo de cálculo da Média Aritmética Simples para dados não agrupados

Tabela 1 – Dados organizados em ROL (Não Agrupado) das notas de estudantes – Escola A-2024.1		Cálculo da Média Aritmética Simples para dados não agrupados através da fórmula:
Notas (Xi)	Frequência Simples Absoluta (fi)	$\bar{X} = \frac{x_1 + x_2 + x_3 + ... + x_n}{n}$
0	1	Então, tem-se que: $\bar{x} = (0 + 1 + 2 + 3 + 4 + 4 + 4 + 4 + 4 + 5 + 5 + 5 + 5 + 6 + 6 + 6 + 7 + 7 + 8 + 9)/20$. Tem-se como resposta 4,8 (Média Aritmética das notas dos estudantes) .
1	1	
2	1	
3	1	
4	5	
5	3	
6	4	
7	2	
8	1	
9	1	
Total (Σ)	20	
Fonte: Elaboração própria, 2024.		

Fonte: Elaboração própria, 2024.

A Média Aritmética para dados agrupados é denominada Ponderada. A fórmula de Média Aritmética Ponderada ($\bar{x}$) é igual ao numerador (parte superior) dividido pelo denominador (parte inferior). O numerador corresponde ao somatório dos produtos dos valores pelos pesos de cada um, dividido pelo somatório dos pesos onde n corresponde aos números: “ $x_1, x_2, …, x_n$ ”, aos quais são atribuídos fatores de ponderação (pesos ou frequências do valor em um ROL). Já o denominador é igual $P_1, P_2, P_3, …, P_n$ ou $f_1, f_2, f_3, …, f_n$, sendo P para pesos e f para frequência, respectivamente, conforme a fórmula:

$$\bar{x} = \frac{x_1P_1 + x_2P_2 + ... + x_nP_n}{P_1 + P_2 + ... + P_n} \text{ ou } \bar{X} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i}$$

Veja o exemplo da Figura 11b, a Tabela 1, que representa uma distribuição agrupada das notas de alunos da Escola A no ano de 2024.1 (Exemplo do Quadro 10 – Capítulo 2). Enquanto na parte superior (Figura 11b) se tem o cálculo da Média Aritmética Ponderada para dados Agrupados, ao fazer o Cálculo (Figura 11b parte inferior), tem-se que a nota mais representativa dessa sala é de 5,4.

Figura 11b – Exemplo de cálculo da Média Aritmética Simples para dados agrupados

Tabela 1 – Dados organizados em ROL (Agrupado) das notas de estudantes – Escola A-2024.1		
Notas	Frequência Simples Absoluta (fi)	Ponto Médio (Xi) 1º Passo
0 + 2	2	1
2 + 4	2	3
4 + 6	8	5
6 + 8	6	7
8 + 10	2	9
Total (Σ)	20	
Fonte: Elaboração própria, 2024		
Cálculo da Média Aritmética ponderada para dados agrupados, fiz o passo a passo para facilitar a interpretação.		
1º Passo: Calcular o Ponto Médio (LI + LS)/2		
2º Passo: Multiplicar o valor do Ponto Médio pela Frequência Simples Absoluta (fi) de cada classe. Dessa forma, o Ponto Médio representa todos os valores do seu respectivo intervalo, enquanto a Frequência Simples Absoluta (fi) corresponde ao valor do peso presente na fórmula da Média Aritmética Ponderada. Então, tem-se: (1 × 2) + (3 × 2) + (5 × 8) + (7 × 6) +(9 × 2) = 108		
3º Passo: Soma-se os valores da Frequência Simples Absoluta (fi). Então, tem-se: 2 + 2 + 8 + 6 + 2 = 20 (o somatório do peso sempre deve dar o total de observações pesquisadas).		
4º Passo: Divide o resultado do 2º Passo pelo 3º Passo, ao fazer isso estamos substituindo na fórmula de Média Aritmética Ponderada $\bar{x} = \frac{x_1P_1 + x_2P_2 + ...+ x_nP_n}{P_1 + P_2 + ...+ P_n}$, onde o 2º passo é o numerador da fórmula e o 3º passo é o denominador. Assim, 108/20 = 5,4 (valor representativo das notas dos estudantes da Escola A no ano de 2024.1).		
Fonte: Elaboração própria, 2024.		

É possível fazer o cálculo da Figura 11a através da fórmula da Média Aritmética Ponderada, pois há uma frequência de números se repetindo. Assim, tem-se:

1º Passo: Calcular o Ponto Médio $(LI + LS)/2$. Nesse caso não é preciso calcular, pois cada classe já possui uma representatividade que corresponde à coluna denominada “Notas”.

2º Passo: Multiplicar valor da coluna “Notas-Figura 11a” pela Frequência Simples Absoluta (f_i) de cada classe da Tabela 1 presente na Figura 11a. O valor do Peso é a coluna da “Frequência Simples Absoluta (f_i)”. Então, tem-se: $(0 \times 1) + (1 \times 1) + (2 \times 1) + (3 \times 1) + (4 \times 5) + (5 \times 3) + (6 \times 4) + (7 \times 2) + (8 \times 1) + (9 \times 1) = 96$

3º Passo: Somar os valores da Frequência Simples Absoluta (f_i). Então, tem-se: $1 + 1 + 1 + 1 + 1 + 5 + 3 + 4 + 2 + 1 = 20$ (o somatório do peso sempre deve dar o total de observações pesquisadas).

4º Passo: Dividir o resultado do 2º Passo pelo 3º Passo. Ao fazer isso, estamos substituindo na fórmula de Média Aritmética Ponderada $\bar{x} = \frac{x_1 P_1 + x_2 P_2 + \dots + x_n P_n}{P_1 + P_2 + \dots + P_n}$, onde o 2º passo é o numerador da fórmula e o 3º passo é o denominador. Assim, $96/20 = 4,8$ aproximadamente 5, o mesmo valor achado na Figura 11a.

Como a estatística é oriunda da matemática, é possível chegar ao mesmo valor com operações diferentes. O cálculo do Figura 11a é mais usado para distribuições menores, já o da Figura 11b para distribuições maiores. Percebe-se que a resposta do Figura 11b teve um valor diferente de 5,4, apesar da amostra de dados brutos ter sido igual para a elaboração da Tabela 1 presente na Figura 11a e da Tabela 1 da Figura 11b. Isso ocorre pois se usou o ponto médio para representar os intervalos ao invés do dado real da nota da tabela. Quero deixar claro que ambas as respostas podem ser utilizadas para a interpretação.

Muitas vezes o valor da Média Aritmética é considerado atípico, ou seja, não representa a realidade. Isso ocorre quando há

números muito extremos que puxam o valor da Média Aritmética. Por exemplo: num setor há três funcionários que recebem de salário, respectivamente, R\$ 3.000,00; R\$ 3.000,00; e R\$ 25.000,00. Ao fazer o cálculo da Média Aritmética de forma simples, tem-se: $(3.000 + 3.000 + 25.000)/3$ R\$ 10.333,33. Ou seja, apenas um funcionário ganha acima de R\$ 10.333,33 no setor, sendo R\$ 25.000,00. Esse dado puxou o valor da Média Aritmética Simples para cima. Nesse caso, considero a Média Aritmética atípica. Se isso ocorre, essa medida de tendência central não é a melhor para a interpretação, e sim a Mediana.

Outro aspecto é que, no exemplo anterior, posso realizar o cálculo através da Média Aritmética ponderada, pois há valores de referência que se repetem mais de uma vez, dando pesos. Assim, tem-se: $(3.000 \times 2) + (25.000 \times 1)/(2 + 1) =$ R\$ 10.333,33. Dessa maneira, tanto a Média Aritmética simples como a ponderada darão o mesmo valor.

Dica:

- 1) Todas as vezes que existir um valor em que a frequência se repetir, use a Média Aritmética ponderada para agilizar o cálculo.
- 2) Se a Média Aritmética for atípica, use para a interpretação a Mediana, pois o valor é mais fiel à realidade.
- 3) Toda vez que ver a palavra desvio na estatística, interprete como afastamento de um determinado valor.

O valor da média serve também como referência de comparação entre os valores de distribuição presentes na amostra. Por exemplo, vou escolher de forma aleatória dois alunos, um que tirou a nota 9 e outro que tirou a nota 3. Se fizermos dois cálculos distintos em relação à nota deles, tem-se que $9 - 4,8 = 4,2$. Assim o aluno que tirou 9 tem 4,2 pontos acima da nota que representa a

média da sala, que é 4,2. Já com o estudante que tirou 3, foi feita a seguinte conta: $3 - 4,8 = -1,8$. Isso significa que esse aluno tirou uma nota 1,8 abaixo da nota que representa a média da sala, que é 4,2. Aqui não estou avaliando se a média achada é alta ou baixa em relação aos parâmetros educacionais, isso é algo que o pesquisador fará depois observando a realidade dos dados (exigência numérica de qualidade educacional, programas do curso, etc.).

Esse cálculo que fizemos no parágrafo anterior se denomina desvio em relação à média, ou seja, é a diferença entre cada elemento de um conjunto de valores e a Média Aritmética. Ele também é chamado Desvio Relativo. Sua fórmula é: $dr = (X_i - \bar{x})$, onde X_i é o ponto médio, ou valor representativo daquela classe, e $\bar{x}$ é a Média Aritmética.

As medidas de dispersão são importantes para a análise de dados em relação à aproximação ou ao afastamento destes quanto à média. Assim, servem para pesquisar ou verificar como se dispersam em relação à média aritmética e/ou mediana (veremos o cálculo da Mediana no Tópico 3.1.3), que representam os valores centrais encontrados na amostra, para depois tomar uma decisão. As principais medidas de dispersão são:

1. Amplitude: diferença entre o maior e o menor valor (visto no capítulo 2);
2. Desvio Relativo (Dr): que é igual à $(X_i - \bar{x})$ e corresponde ao valor da média de cada classe menos a média aritmética ponderada das classes;
3. Desvio absoluto (Da): o módulo do respectivo desvio relativo referente a cada classe;
4. Variância: somatório dos desvios relativos elevado ao quadrado, multiplicado pelas respectivas frequências e dividido pelo somatório das frequências;
5. Desvio médio absoluto (Dma): somatório dos módulos dos desvios absolutos, multiplicado pelas respectivas frequências e dividido pelo somatório delas;

6. Desvio padrão: raiz quadrada da variância. Ele pode ser amostral, se o total de f_i for menor ou igual a 30 acontecimentos (representado pela letra S), ou populacional, se for maior que 30 (representado pela letra “ σ ” – pronuncia-se “sigma”). A média aritmética populacional é representada pela letra “ μ ”, que se fala “Mi”.

Estou mostrando essas seis medidas de dispersão pois é necessário utilizar do item 1 até 6 para se chegar no desvio padrão, seguindo a ordem citada anteriormente como passo a passo. Além disso, o desvio padrão é super útil para a interpretação da Média Aritmética. Para tornar a explicação mais clara, utilizaremos os mesmos exemplos anteriores, iniciando com dados não agrupados e depois com os agrupados. Assim, a Figura 12 mostrará as etapas para o cálculo do desvio relativo, da variância, do desvio padrão, do erro padrão e do coeficiente de variância.

Figura 12 – Passo a passo para o cálculo do desvio Relativo, do Desvio absoluto, da Variância, do Desvio Médio Absoluto, do Desvio Padrão e do Coeficiente de Variância para Dados não Agrupados

Tabela 1 – Dados organizados em ROL (Não Agrupado) das notas de estudantes – Escola A-2024.1					
Notas (X_i) 1º Passo	Frequência Simples Absoluta (f_i)	$Dr = (X_i - \bar{x})$ 3º Passo	$Da = (X_i - \bar{x}) $ 4º Passo	$Dma = (X_i - \bar{x}) ^2$ 5º Passo (Variância)	Variância $x f_i$ 6º Passo
0	1	$0 - 4,8 = -4,8$	$ -4,8 = 4,8$	$(4,8)^2 = 23,04$	$23,04 \times 1 = 23,04$
1	1	$1 - 4,8 = -3,8$	$ -3,8 = 3,8$	$(3,8)^2 = 14,44$	$14,44 \times 1 = 14,44$
2	1	$2 - 4,8 = -2,8$	$ -2,8 = 2,8$	$(2,8)^2 = 7,84$	$7,84 \times 1 = 7,84$
3	1	$3 - 4,8 = -1,8$	$ -1,8 = 1,8$	$(1,8)^2 = 3,24$	$3,24 \times 1 = 3,24$
4	5	$4 - 4,8 = -0,8$	$ -0,8 = 0,8$	$(0,8)^2 = 0,64$	$0,64 \times 5 = 3,2$
5	3	$5 - 4,8 = 0,2$	$ 0,2 = 0,2$	$(0,2)^2 = 0,04$	$0,04 \times 3 = 0,12$
6	4	$6 - 4,8 = 1,2$	$ 1,2 = 1,2$	$(1,2)^2 = 1,44$	$1,44 \times 4 = 5,76$
7	2	$7 - 4,8 = 2,2$	$ 2,2 = 2,2$	$(2,2)^2 = 4,84$	$4,84 \times 2 = 9,68$
8	1	$8 - 1,4,8 = 3,2$	$ 3,2 = 3,2$	$(3,2)^2 = 10,24$	$10,24 \times 1 = 10,24$
9	1	$9 - 4,8 = 4,2$	$ 4,2 = 4,2$	$(4,2)^2 = 17,64$	$17,64 \times 1 = 17,64$
Total (Σ)	20		25	83,4	95,2

Fonte: Elaboração própria, 2024.

continua...

Figura 12 – Passo a passo para o cálculo do Desvio relativo, da Variância, do Desvio Padrão, do erro padrão e do Coeficiente de Variância para Dados não Agrupados

1º Passo: Calcular o Ponto Médio ($\bar{X}$). No caso de dado não agrupado quantitativo, as próprias notas dos alunos referentes a primeira coluna já correspondem aos valores. Não é necessário calcular.

2º Passo: Calcular a Média Aritmética. Ver Figura 12 do cálculo da Média Aritmética, que nesse caso específico teve resultado 4,8.

3º Passo: Calcular o Desvio Relativo, que é $(X_i - \bar{x})$. Assim, para o aluno que tirou zero, tem uma diferença negativa de 4,8 em relação à nota representativa tirada pela média da sala, que foi de 4,8. Caso a análise seja para o aluno que tirou 1, tem uma diferença negativa de 3,8 em relação à nota representativa tirada pela média da sala. O mesmo raciocínio serve para os resultados encontrados na Tabela 1, referentes à 3ª coluna.

4º Passo: Coloca os valores do Desvio Relativo em Módulo, com o objetivo de ter uma noção da variabilidade do conjunto dos dados referente a cada classe. Caso desejássemos calcular o desvio absoluto médio, basta somar os valores $D_a = |(X_i - \bar{x})|$ da quarta coluna, encontrando 25, e depois dividir pelo total da Frequência Simples Absoluta (f_i) – 2ª coluna, assim, tem-se $25/20 = 1,25$. Dessa forma, cada nota ficou em torno de 1,25 “pontos” distante da média, que é 4,8.

5º Passo: Após ter os dados do “Da” na Tabela, a próxima etapa é calcular a variância. O objetivo de se calcular a variância é tomar os desvios dos valores X_i em relação à média aritmética, ou seja, $(X_i - \bar{x})$. Mas a soma desses desvios é sempre igual a zero. Para evitar que isso aconteça (e que não interfira nos cálculos), cada desvio deve ser elevado ao quadrado, pois o quadrado de qualquer número pertencente aos reais é sempre positivo. Nessa fórmula, iremos calcular a variância amostral (Fórmula da Variância amostral: $s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$) e não populacional (fórmula da Variância populacional: $s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$), pois o somatório total é 20, que é um número menor/igual a 30. Se o somatório fosse maior que 30, seria a variância populacional. A diferença entre ambas é que o denominador da amostral sempre subtrai o total de 1 (Total ou $n - 1$), e a populacional sempre manterá o valor Total (n) da amostra sem qualquer subtração. Assim, para se calcular a variância, coloca-se o valor do “Da” de cada classe ao quadrado, depois se faz o somatório de cada valor ao quadrado ($23,04 + 14,44 + 7,84 + 3,24 + 0,64 + 1,44 + 4,84 + 10,24 + 17,64$) e se encontra 83,4. Após isso, coloca na fórmula da variância amostral devido às explicações já dadas, e se tem $83,4/(20 - 1) = 3,76$. Como o valor não foi zero, ela não é uma amostra homogênea, mas seria necessário calcular as notas de alunos de uma outra sala levando em consideração o mesmo período que foi colhida essa amostra, bem como atributos similares. Se, calculando-se da outra sala, a variância fosse 2, por ser um número menor do que 3,76, os valores das notas dos alunos dessa outra sala seriam menos dispersos do que as notas não agrupadas da Figura 12. Entretanto, se o valor da variância for maior que 3,76, por exemplo 4, significa que as notas não agrupadas da Figura 12 são mais dispersas do que as da sala que teve o valor das notas referentes à variância maior.

continua...

6º Passo: Corresponde ao cálculo do desvio padrão, em que teremos que multiplicar o valor da Variância pela Frequência Simples Absoluta (f_i) de cada classe. Depois, faz-se o somatório dos resultados dos valores encontrados em cada classe, que no caso da Tabela 1 – Figura 13, fica $23,04 + 14,44 + 7,84 + 3,24 + 3,2 + 0,12 + 5,76 + 9,68 + 10,24 + 17,64$, o que dará como resultado 95,2; então, divide-se $95,2 / (\text{total} - 1)$. Por se tratar de uma variância amostral, o denominador da fórmula fica 19 ao invés de 20, e a divisão tem como resultado 5,01. Esse último resultado encontrado é colocado na raiz e então se tem $\sqrt{5,01}$, que é igual a **2,23**.

7º Passo: O desvio padrão corresponde ao conjunto de dados obtidos por amostragem de uma população denominada “normal”. É possível verificar que quanto maior é a amostra, menor é o desvio padrão, ou seja, pode-se chegar ao valor mais próximo à realidade. Após achar o valor do desvio padrão, deve-se realizar os seguintes cálculos: $(\bar{x} - S)$ e $(\bar{x} + S)$, que inclui aproximadamente 68,26% dos resultados obtidos na pesquisa; e o intervalo compreendido entre $(\bar{x} - 2.S)$ e $(\bar{x} + 2.S)$, que inclui aproximadamente 95,44% desses resultados. Agora, se for considerado o intervalo compreendido entre $(\bar{x} - 3.S)$ e $(\bar{x} + 3.S)$, teremos aproximadamente 99,71% dos resultados, isto é, quase a totalidade das observações efetuadas. Cabe ao pesquisador definir qual dos intervalos citados anteriormente irá utilizar. Assim, optei pelo de $(\bar{x} - S)$ e $(\bar{x} + S)$, onde se tem $(4,8 - 2,23 = 2,56)$ e $(4,8 + 2,23 = 7,03)$. Então, o intervalo em que estão 68% das notas está entre [2,56;7,03].

8º Passo: Calcular o erro padrão. Anteriormente (7º Passo), vimos que o desvio padrão mede a dispersão dos valores em um conjunto de dados em relação à média. Enquanto agora veremos que o erro padrão mede a precisão da média da amostra em relação à média da amostra ou da população. Para calculá-lo, por se tratar de uma amostra menor que 30 (pois possui 20 estudantes), basta dividir o valor do desvio padrão pela $\sqrt{\text{total} - 1}$, por se tratar de erro padrão amostral e não populacional. Diante disso, tem-se: $2,23 / (\sqrt{20 - 1})$, que é igual a **0,51**. Ou seja, dentro do intervalo encontrado no 7º Passo, que foi de [2,56 - **0,51** = 2,04; 7,03 + **0,51** = 7,54], é possível colocar a precisão estimada de um intervalo de notas dos estudantes da Escola Ano 2024.1 estimada dentro de uma precisão de 68%, entre [2,04;7,54]

9º Passo: Calcular o Coeficiente de Variância (CV), que corresponde a uma medida estatística que expressa a variabilidade relativa (porcentagem) dos dados em relação à média. Sua fórmula é $CV = (\text{desvio padrão} / \text{média}) \times 100$. Quanto maior o valor do CV, maior a variabilidade relativa dos dados em relação à média. Por outro lado, um CV menor indica uma variabilidade relativa menor. Nesse exemplo, tem-se $CV = (2,23 / 4,8) \times 100$, com o resultado 46,6%, que pode ser considerada uma porcentagem alta. Entretanto, seu uso é melhor, por exemplo, através de comparações de CV, que poderiam ser feitas com notas de outra turma ou através do CV da presença dos estudantes em aula, observando-se sempre o mesmo local e o mesmo período.

Fonte: Elaboração própria, 2024.

.....

Vídeo com o cálculo no Excel da Medida de Tendência Central (Média Aritmética), Separatriz e Desvio para dados não agrupados



Link: <https://youtu.be/Bs9yEYHoxPo>

.....

Observação:

1) A variância é uma medida que tem pouca utilidade como estatística descritiva, porém é extremamente importante na inferência estatística e em combinações de amostras. É usada para comparar amostras que possuem a mesma unidade em relação à dispersão significativa de quanto os conjuntos de dados são heterogêneos, mas isso não é suficiente para se tomar decisões de desvios sobre médias. Se a variância for zero, a amostra é homogênea, e quanto mais longe de zero, mais dispersa e heterogênea é a amostra.

2) Um desvio padrão grande significa que os valores amostrais estão bem distribuídos em torno da média, enquanto um desvio padrão pequeno indica que eles estão condensados próximos a ela. Em poucas palavras, quanto menor o desvio padrão, mais homogênea é a amostra.

Vídeo desvio padrão para dados não agrupados



Link: <https://youtu.be/qRwKrVmQdeE>

Agora veremos, através da Figura 13, o passo a passo para o cálculo do Desvio Relativo, da Variância, do Desvio Padrão, do erro padrão e do Coeficiente de Variância para dados agrupados.

Figura 13 – Passo a Passo para o cálculo do Desvio Padrão para dados agrupados

Tabela 1 – Dados organizados em ROL (Agrupado) das notas de estudantes – Escola A – 2024.1						
Notas Limite Inferior (LI) + Limite Superior (LS)	Frequência Simples Absoluta (fi)	Ponto Médio (Xi) 1º Passo	Dr = (Xi - x̄) 3º Passo	Da = (Xi - x̄) 4º Passo	Dma = (Xi - x̄) ² 5º Passo (Variância)	Variância x fi 6º Passo
0 + 2	2	(0 + 2)/2 = 1	(1 - 5,4) = -4,4	-4,4 = 4,4	(4,4)² = 19,36	19,36 × 2 = 38,72
2 + 4	2	(2 + 4)/2 = 3	(3 - 5,4) = -2,4	-2,4 = 2,4	(2,4)² = 5,76	5,76 × 2 = 11,52
4 + 6	8	(4 + 6)/2 = 5	(5 - 5,4) = -0,4	-0,4 = 0,4	(0,4)² = 0,16	0,1 × 8 = 1,28
6 + 8	6	(6 + 8)/2 = 7	(7 - 5,4) = 1,6	1,6 = 1,6	(1,6)² = 2,56	2,56 × 6 = 15,36
8 + 10	2	(8 + 10)/2 = 9	(9 - 5,4) = 3,6	3,6 = 3,6	(3,6)² = 12,96	12,96 × 2 = 25,92
Total (Σ)	20			12,4	40,8	92,8

Figura 13 – Passo a Passo para o cálculo do Desvio Relativo, da Variância, do Desvio Padrão, do erro padrão e do Coeficiente de Variância do desvio Relativo, do Desvio Absoluto, da Variância, do Desvio Médio Absoluto, do Desvio Padrão e do Coeficiente de Variância para dados agrupados

1º Passo: Calcular o Ponto Médio (Xi). Por ser dado agrupado, corresponde ao (Limite Inferior - LS + Limite Superior - LS)/2), e é necessário a realização do cálculo de cada classe. Para relembrar esse cálculo, ver Capítulo 2 – Tabela 2 – Tópico 2.4.

2º Passo: Calcular a Média Aritmética (ver Figura 11b do cálculo da Média Aritmética). Nesse caso específico, o resultado foi 5,4.

continua...

...continuação

3º Passo: Calcular o Desvio Relativo, que é $(X_i - \bar{x})$. Assim, para o aluno que tirou entre 0 e 2, tem uma diferença negativa de 4,4 em relação à nota representativa tirada pela média da sala, que foi de 5,4. Caso a análise seja para os alunos que tiraram entre 2 e 4, existe uma diferença negativa de 2,4 em relação à nota representativa tirada pela média. O mesmo raciocínio serve para os outros resultados encontrados na Tabela 1, referentes a 4ª coluna.

4º Passo: Colocar os valores do Desvio relativo em Módulo, com o objetivo de ter uma noção da variabilidade do conjunto dos dados referente a cada classe. Caso desejássemos calcular o desvio absoluto médio, basta somar os valores $D_a = |(X_i - \bar{x})|$ da quinta coluna, encontrando 12,4. Depois, é preciso dividir pelo total da Frequência Simples Absoluta (f_i) – 2ª coluna, e assim, tem-se $12,4/20 = 0,62$. Dessa forma, cada nota ficou em torno de 0,62 “pontos” distante da média, que é 5,4.

5º Passo: Após ter os dados do “Da” na tabela, a próxima etapa é calcular a variância. O objetivo é de tomar os desvios dos valores X_i em relação à média aritmética, ou seja, $(X_i - \bar{x})$, mas a soma desses desvios é sempre igual a zero. Para evitar que isso aconteça (e que não interfira nos cálculos), cada desvio deve ser elevado ao quadrado, pois o quadrado de qualquer número pertencente aos reais é sempre positivo. Nessa fórmula, iremos calcular a variância amostral (Fórmula da Variância Amostral: $s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$) e não populacional (fórmula da Variância Populacional: $s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$) pois o somatório total é 20, que é um número menor igual a 30. Se o somatório fosse maior que 30, seria a variância populacional. A diferença entre ambas é que o denominador da amostral sempre subtrai o total de 1 (Total ou $n - 1$), e a populacional sempre manterá o valor Total (n) da amostra, sem qualquer subtração. Assim, para se calcular a variância, coloca-se o valor do “Da” de cada classe ao quadrado e depois se faz o somatório dos valores de cada valor ao quadrado ($19,36 + 5,76 + 0,16 + 2,56 + 12,96$), encontrando-se 40,8. Depois, coloca na fórmula da variância amostral devido às explicações já dadas, em que se tem $40,8/(20 - 1) = 2,15$. Como valor não foi zero (Figura 12/14 – Variância), ela não é uma amostra homogênea, mas seria necessário calcular as notas de alunos de uma outra sala para fazer uma comparação. Por isso, optou-se pelos dados das notas da Figura 11b como parâmetro comparativo com a Figura 12, pois são amostras do mesmo período e da mesma avaliação, só que representadas de forma diferente (dados não agrupados e dados agrupados). Ambas possuem atributos similares: na Figura 13, o valor da variância foi de 2,15, e por ser um número menor do que 3,76 (Figura 12), significa que os valores das notas dos alunos da Figura 14 estão menos dispersos do que as notas não agrupadas da Figura 13.

continua...

6º Passo: Corresponde ao cálculo do desvio padrão, em que teremos que multiplicar o valor da Variância pela Frequência Simples Absoluta (f_i) de cada classe. Depois, faz-se o somatório dos resultados dos valores encontrados em cada classe. No caso da Tabela 1 – Figura 13, fica $38,72 + 11,52 + 1,28 + 15,36 + 25,92$, o que dará como resultado 92,8. Então, divide-se $92,8/(\text{total} - 1)$, por se tratar de uma variância amostral. Assim, o denominador da fórmula fica 19 ao invés de 20, e a divisão tem como resultado 4,88. Esse último resultado encontrado é colocado na raiz, e então, tem-se $\sqrt{4,88}$, que é igual a **2,21**.

7º Passo: O desvio padrão corresponde ao conjunto de dados obtidos por amostragem de uma população denominada “normal”. É possível verificar que quanto maior é a amostra, menor é o desvio padrão, ou seja, pode-se chegar ao valor mais próximo à realidade. Após achar o valor do desvio padrão, deve-se realizar os seguinte cálculo: $(\bar{x} - S)$ e $(\bar{x} + S)$, que inclui aproximadamente 68,26% dos resultados obtidos na pesquisa; e o intervalo compreendido entre $(\bar{x} - 2 \cdot S)$ e $(\bar{x} + 2 \cdot S)$, que inclui aproximadamente 95,44% desses resultados. Agora, se for considerado o intervalo compreendido entre $(\bar{x} - 3 \cdot S)$ e $(\bar{x} + 3 \cdot S)$, teremos aproximadamente 99,71% dos resultados, isto é, quase a totalidade das observações efetuadas. Cabe ao pesquisador definir qual dos intervalos citados anteriormente irá utilizar. Assim, optei pelo de $(\bar{x} - S)$ e $(\bar{x} + S)$, onde se tem $(5,4 - 2,21 = 3,19)$ e $(5,4 + 2,21 = 7,6)$. Então, o intervalo que estão 68% das notas dos estudantes da Escola Ano 2024.1 compreende as notas que vão de [3,19;7,61].

8º Passo: Calcular o erro padrão. Anteriormente (7º Passo), vimos que o desvio padrão mede a dispersão dos valores em um conjunto de dados em relação à média. Enquanto agora veremos que o erro padrão mede a precisão da média da amostra em relação à média da amostra ou da população. Para calculá-lo, por se tratar de uma amostra menor que 30 (pois possui 20 estudantes), basta dividir o valor do desvio padrão pela $\sqrt{\text{total} - 1}$, por se tratar de erro padrão amostral e não populacional. Diante disso, tem-se: $2,21/(\sqrt{20} - 1)$, que é igual a **0,51**. Ou seja, dentro do intervalo encontrado no 7º Passo, que foi de [3,19 – **0,51** = **2,68**; 7,61 + **0,51** = 8,12], é possível colocar a precisão estimada de um intervalo de notas dos estudantes da Escola Ano 2024.1 dentro de 68% entre [2,68; 8,12].

9º Passo: Calcular o Coeficiente de Variância (CV), que corresponde a uma medida estatística que expressa a variabilidade relativa (porcentagem) dos dados em relação à média. Sua fórmula é $CV = (\text{desvio padrão}/\text{média}) \times 100$. Quanto maior o valor do CV, maior a variabilidade relativa dos dados em relação à média. Por outro lado, um CV menor indica uma variabilidade relativa menor. Nesse exemplo, tem-se $CV = (2,21/5,4) \times 100$, com resultado 41%, que pode ser considerado uma porcentagem alta. Entretanto, seu uso é melhor, por exemplo, através de comparações de CV, assim se comparando com o CV da Figura 12, que foi 46,6%. Dessa forma, a Figura 13 tem um CV que, por ser menor, é melhor do que o da Figura 12, pois os dados estão mais próximos da média.

Fonte: Elaboração própria, 2024.

Vídeo com o cálculo no Excel da Medida de Tendência Central
(Média Aritmética), Separatriz e Desvio para Agrupados



Link: https://youtu.be/ZiFA3bStH_w

3.1.2 Medida de Tendência Central (Mediana)

A mediana corresponde a outra Medida de Tendência Central responsável por identificar, dentro da distribuição de frequência em ROL, o ponto que define a metade (50%), ou seja, o valor situado de tal forma no conjunto que separa em dois subconjuntos de mesmos números de elementos. Assim como a média aritmética, o cálculo é feito para dados agrupados e não agrupados. Ela é representada pela sigla “Md”.

Para dados não agrupados, é necessário observar a quantidade de elementos organizados sempre em ROL se tem um número par ou ímpar. Assim, mostraremos como realizar esse cálculo através de conjunto e depois em tabela, a partir da quantidade de elementos, se for par ou ímpar.

Se for par e com a apresentação em conjunto, utiliza-se a fórmula $P_{Md} = \frac{n+1}{2}$, onde P_{Md} é a posição do elemento mediano e n é o total de *fi*. Vamos fazer o exemplo já utilizado nos capítulos 2 e 3, que podem ser representados de duas formas: em conjunto ou em tabela.

Primeiro, iremos ver em forma de conjunto, em que se tem $A = \{0;1;2;3;4;4;4;4;4;5;5;5;5;6;6;6;7;7;8;9\}$. O conjunto “A” possui 20 elementos, o que corresponde a 20 estudantes, o que é um número par. O conjunto “A” é o mesmo exemplo utilizado nos capítulos 2 e 3 e já está em ROL.

O primeiro passo é colocar na fórmula: $P_{Md} = (20 + 1)/2 = 10,5$ devido ao 20 ser o número par, e então achamos a posição do elemento 10,5.

O segundo passo é sinalizar no conjunto “A” a posição do 10,5, que está entre dois elementos marcados na cor verde ($A = \{0;1;2;3;4;4;4;4;4;5;5;5;5;6;6;6;7;7;8;9\}$). No caso, são os dois números 5, pois a posição é a 10,5ª, está entre a 10ª e a 11ª posição.

O terceiro passo é colocar em uma segunda fórmula, que tira a média aritmética simples entre os dois termos centrais do ROL identificados pela cor verde, sendo nesse caso os elementos da 10ª e da 11ª posição. Então, teremos: $\bar{x} = (5 + 5)/2 = 5$ e a Mediana (Md) = 5. Fazendo a interpretação, temos que a nota 5 corresponde ao ponto que separa 50% das notas para cima e 50% para baixo dos 20 estudantes da Escola A no ano de 2024.1.

Se a apresentação estiver em tabela, faremos a partir do passo a passo presente na Figura 14.

Figura 14 – Exemplo de cálculo da Mediana para dados não agrupados – par

Tabela 1 – Dados organizados em ROL (Não Agrupado) das notas de estudantes – Escola A-2024.1		
Notas	Frequência Simples Absoluta (fi)	FACA 3º Passo
0	1	1
1	1	2
2	1	3
3	1	4
4	5	9
5	3	12
6	4	16
7	2	18
8	1	19
9	1	20
Total (Σ)	20	

4º Passo:
FACA da 10,5 Posição
está aqui-5ª classe

Fonte: Elaboração própria, 2024.

1º Passo: Colocar notas em ROL. No caso específico, já está de 0 até 9.

2º Passo: Colocar na fórmula: $P_{Md} = \frac{n+1}{2}$. Substituindo, tem-se $P_{Md} = (20 + 1)/2 = 10,5$, devido ao 20 ser o número par. E então, achamos que o elemento está na 10,5ª posição.

3º Passo: Fazer a Frequência Acumulada Crescente – Abaixo (FACA), para ver a posição correspondente a 10,5. Ver como faz cálculo da FACA no Capítulo 2 – Tabela 2

4º Passo: Olhar na FACA onde está a posição da 10,5ª. Percebe-se que está na sexta classe, pois a quarta classe vai até o 9º elemento.

5º Passo: Tirar a média aritmética simples entre os dois termos centrais. No caso estão na sexta classe, possuindo três números 5 (pois fi = 3) que são, respectivamente, o 10º, a 11º e a 12º, já que a quinta classe termina na nona posição conforme FACA igual a 9 e nota igual a 4. Então, teremos $\bar{x} = (5 + 5)/2 = 5$, e a Mediana (Md) = 5. Fazendo a interpretação, obtemos que a nota 5 corresponde ao ponto que separa 50% das notas para cima e 50% para baixo dos 20 estudantes da Escola A no ano de 2024.1.

Observação: Caso tenha dificuldade de fazer por tabela os dados não agrupados, monte em forma de conjunto, conforme explicação anterior para número par.

Fonte: Elaboração própria, 2024.

Caso a quantidade de elementos organizados em ROL seja ímpar e apresentado em forma de conjunto para dados não agrupados, utilizaremos a fórmula da mediana: $P_{Md} = \frac{n+1}{2}$. Vamos pegar um conjunto hipotético denominado B = {5, 15, 12, 2, 18, 15, 6, 16 e 9}.

O primeiro passo, como o conjunto B não está em ROL, é organizá-lo dessa forma. Então, temos $B = \{2, 5, 6, 9, 12, 15, 15, 16, 18\}$.

O segundo passo é colocar na fórmula: $P_{Md} = \frac{n+1}{2}$. já que B possui de nove elementos, um número ímpar. Ao substituir na fórmula, tem-se: $P_{Md} = (9 + 1)/2 = 5^{\text{a}}$ posição.

O terceiro passo é identificar o elemento na 5ª posição dentro do conjunto “B”. Assim, tem-se o número 12, sinalizado com a cor verde: $B = \{2, 5, 6, 9, 12, 15, 15, 16, 18\}$. A Mediana (Md) é o valor 12, e ao fazer a interpretação, temos que o número 12 corresponde ao ponto que separa 50% dos valores para cima e 50% para baixo de nove estudantes de uma Escola A no ano de 2024.1.

Se a apresentação estiver em tabela, faremos a partir do passo a passo presente na Figura 15.

Figura 15 – Exemplo de cálculo da Mediana para dados não agrupados – ímpar

Tabela 1 – Dados organizados em ROL (Não Agrupado) das Idades de estudantes – Escola A-2024.1

Idade	Frequência Simples Absoluta (fi)	FACA 3º Passo
2	1	1
5	1	2
6	1	3
9	1	4
12	1	5
15	2	7
16	1	8
18	1	9
Total (Σ)	9	

4º Passo:
FACA da 5ª Posição
está aqui-5ª classe

Fonte: Elaboração própria, 2024.

continua...

1º Passo: Colocar notas em ROL, no caso específico já está de 2 até 18.

2º Passo: Colocar na fórmula $P_{Md} = \frac{n+1}{2}$. Substituindo, tem-se $P_{Md} = (9 + 1)/2 = 5$, devido ao 9 número ímpar. E então, achamos que a posição do elemento é a 5ª.

3º Passo: Fazer a Frequência Acumulada Crescente – Abaixo (FACA) para ver a posição correspondente a 5ª. Ver como fazer cálculo da FACA no Capítulo 2 – Tabela 2

4º Passo: Identificar diretamente na tabela (coluna da FACA) o elemento da 5ª posição e ver que corresponde ao número 12, presente na primeira coluna da tabela. Então, a Mediana (Md) = 12. Fazendo a interpretação, temos que a idade 12 anos corresponde ao ponto que separa 50% das idades para cima e 50% para baixo dos nove estudantes de uma Escola A no ano de 2024.1.

Observação: Caso tenha dificuldade de fazer por tabela os dados não agrupados, monte em forma de conjunto, conforme explicação anterior para número ímpar.

Fonte: Elaboração própria, 2024.

Para calcular a mediana para dados agrupados, usaremos duas fórmulas. A primeira é para identificar a posição da Mediana e é $P_{Md} = \frac{n}{2}$, onde P_{MD} é a posição da Mediana e n é o somatório total de elementos. Divide-se por 2 pois corresponde à Mediana, ou seja, ao ponto de 50% que divide a distribuição de frequência em duas partes iguais. Já a segunda fórmula é para achar o valor numérico da mediana, e é $Md = L_i + \frac{\left[\frac{n}{2} - \sum f_{ant}\right]}{f_{Md}} \cdot A$, onde:

- L_i = limite inferior da classe que contém a mediana;
- n = tamanho da amostra ou população que está sendo pesquisada, ou seja, o total de observações do acontecimento. Lembrando que o $n/2$. Então na hora de substituir, basta pegar o resultado da fórmula $P_{Md} = \frac{n}{2}$
- $\sum f_{ant}$ = valor da faca anterior à classe em que se identificou a posição da Mediana;
- A = amplitude da classe que contém a mediana;
- f_{Md} = frequência da classe que contém a mediana.

A Figura 16 irá mostrar o passo a passo para o cálculo da Mediana para dados agrupados.

Figura 16 – Exemplo de cálculo da Mediana para dados agrupados

Tabela 1 – Dados organizados em ROL (Agrupado) das notas de estudantes – Escola A-2024.1		
Notas	Frequência Simples Absoluta (f_i)	FACA 2º Passo
0 - 2	2	2
2 - 4	2	4
4 - 6	8	12
6 - 8	6	18
8 - 10	2	20
Total (Σ)	20	

3º Passo:

FACA da 10ª Posição está aqui-3ª classe

Fonte: Elaboração própria, 2024

1º Passo: Usaremos a fórmula $P_{Md} = \frac{n}{2}$, onde n é somatório (Σ) de f_i (Segunda Coluna). Não importa se o número é par ou ímpar, mas nesse caso $n = 20$. Colocando na fórmula, temos: $P_{Md} = 20/2 = 10$. Assim, a mediana se encontra na **10ª posição**.

2º Passo: Fazer a Frequência Acumulada Crescente-Abaixo (FACA) para ver a posição correspondente a **10 (Encontrado no 1º Passo)**. Ver como faz cálculo da FACA no Capítulo 2 – Tabela 2.

3º Passo: Identificar diretamente, na tabela na coluna da FACA, o elemento da **10ª posição** (Encontrado no 1º Passo) e ver que corresponde às notas entre 4 - 6 presentes na primeira coluna da tabela. Afinal, ao olhar a FACA, identifica-se que a primeira classe incorpora a 1ª e a 2ª posição, a segunda classe inicia da 3ª e vai até a 4ª posição e a terceira classe inicia da 5ª e vai até a 12ª posição.

4º Passo: Substituir na segunda fórmula da Mediana, que é $Md = L_i + \left[\frac{\frac{n}{2} - \Sigma f_{ant}}{f_{Md}} \right] \cdot A$, a partir de todos os dados referentes a 3ª Classe. Dessa forma, temos:

- L_i = limite inferior da classe que contém a mediana, que é igual a **4 (cor roxa)** pois o Intervalo da Mediana está na 3ª classe 4 - 6;
- n = tamanho da amostra ou população que está sendo pesquisada, ou seja, o total de observações do acontecimento. Então, basta substituir o resultado da fórmula: $P_{Md} = \frac{n}{2}$ (ver 1º Passo) = **10ª Posição**;
- Σf_{ant} = valor da FACA anterior à classe em que se identificou a posição da Mediana. Assim, tem-se que a FACA da posição da mediana referente a 3ª classe é 12 e a anterior é 4;
- A = amplitude da classe que contém a mediana, é $6 - 4 = 2$;
- f_{Md} = frequência da classe que contém a mediana (ver segunda coluna em que $f_i = 8$).

Substituindo os dados na fórmula, tem-se:

$Md = 4 + (10 - 4)/8 \times 2 = 5,5$

continua...

conclusão

5º Passo: Verificar se o número encontrado está dentro da classe identificada. No caso específico é a 3ª Classe, no intervalo entre 4 e 6. Se a resposta não estiver coerente, rever cálculos. Como o 5,5 está dentro da 3ª classe, pode-se iniciar a interpretação.

6º Passo: 5,5 corresponde ao ponto que separa 50% das notas para cima e 50% para baixo dos 20 estudantes de uma Escola A no ano de 2024.1.

Fonte: Elaboração própria, 2024.

No próximo subtópico falaremos de Separatriz, pois o cálculo é similar ao da Mediana e tem total relação com essa Medida de Tendência Central. Depois desse subtópico, falaremos de Moda, que corresponde à última medida de tendência central.

3.1.2.1 Separatriz

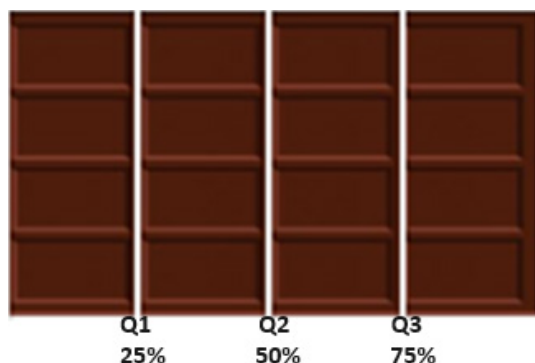
As separatrizes correspondem a um conjunto de parâmetros que subdividem uma série ordenada em várias partes iguais, apresentando o mesmo número de valores. Elas são divididas em: quartil, decil e centil.

Nos quartis, os valores de uma série são divididos em quatro partes iguais e cada parte tem um valor igual a 25%, representado pela letra “Q”. O primeiro quartil (Q1) ocorre quando uma quarta parte (25%) dos dados na série é menor que ele e as três quartas partes restantes (75%) são maiores. Já o segundo quartil (Q2) coincide com a mediana ($Q2 = Md$) e seu valor corresponde a 50%. O terceiro quartil (Q3) se encontra na posição em que as três quartas partes (75%) dos termos são menores que ele e uma quarta parte (25%) é maior.

Para visualização, foi feita a Figura 17 a fim de se mostrar a divisão de cada quartil em 25% e sua respectiva posição dentro de uma barra de chocolate que deve ser dividida de forma igual para quatro pessoas. A primeira pessoa ficaria com a porcentagem até 25%(Q1); a segunda de 25%, local em que o chocolate

foi quebrado até o 50%; a terceira do 50% até o 75%; e a quarta pessoa do 75% até a posição de 100%. O mesmo raciocínio de divisão ocorre com o decil e centil, a única diferença é o tamanho da porcentagem que será dividida.

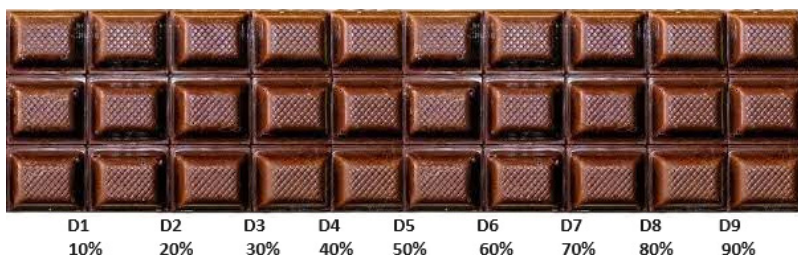
Figura 17 – Identificação dos Quartis



Fonte: Elaboração própria, 2024.

No decil, os valores de uma série são divididos em 10 partes iguais. Cada uma tem um valor igual a 10% e é representado pela letra “D”. Para visualizar o decil foi feita a Figura 18, mostrando a divisão de cada quartil em 10% e sua respectiva posição dentro de uma barra de chocolate a ser repartida para 10 indivíduos diferentes.

Figura 18 – Identificação dos Decis



Fonte: Elaboração própria, 2024.

Já o centil também tem o nome de percentil e corresponde aos valores que dividem um conjunto de dados ordenados em 100 partes iguais. Cada uma tem valor igual a 1% e é representado pela letra “C” ou “P”.

Os pontos que o quartil, o decil e o centil têm iguais é o de 50% (MD) e o de 100%, por isso que a separatriz possui relação com a Mediana (Md – ponto de 50%). Eles são usados quando queremos fazer divisões exatas para identificar dentro da distribuição de frequência uma determinada frequência relativa (porcentagem). Entretanto, seu uso tem a ver com a porcentagem que o pesquisador deseja usar: se a localização que tenha variação de probabilidade ocorrer a cada 25% será o quartil; se for a cada 10%, o decil; e caso seja a cada 1%, o percentil.

Vamos calcular agora o quartil, o decil e o centil para a distribuição não agrupada das notas de alunos da Escola A no ano de 2024.1, que também poderia estar representada em $A = \{0;1;2;3;4;4;4;4;4;5;5;5;5;6;6;6;7;7;8;9\}$. Assim, escolhi de forma aleatória as porcentagens de 25%, 20% e 30%, e quero saber as notas que estão na posição relativa citada anteriormente. A fórmula que encontra a posição do elemento e que irei usar em todas as separatrizes será:

**(Posição do numerador referente à fração da separatriz que se deseja/
número de partes da separatriz) $\times$ Total + (1/2)**

Lembrando que para uso da fórmula, os elementos da distribuição de frequência devem estar sempre em ROL.

Para 25%, irei usar o Q1 que corresponde à representação fracionária de $\frac{1}{4}$. Então seu numerador será 1, pois a divisão de $\frac{1}{4} \times 100 = 25\%$. O quartil divide em quatro partes iguais e o resultado será o total de estudantes, que no caso é 20. O valor de $\frac{1}{2}$ na fórmula é sempre fixo pois é o ponto da Mediana, local onde as três separatrizes (quartil, decil e centil) se encontram. Ao substituir na fórmula, temos:

(Posição do numerador referente à fração da separatriz que se deseja/ número de partes da separatriz) $\times$ Total + (1/2):

Tem-se que $(\frac{1}{4}) \times 20 + (1/2) = (0,25 \times 20) + 0,5 =$

$0,5 + 0,5 = 1^{\text{a}}$ Posição.

Então, a nota que representa 25% é zero, pois é o elemento que se encontra na primeira posição no conjunto $A = \{0; 1; 2; 3; 4; 4; 4; 4; 4; 5; 5; 5; 5; 6; 6; 6; 7; 8; 9\}$

1ª Posição

Para 20% irei usar o D2, que corresponde à representação fracionária de $2/10$. Então seu numerador será 2, pois a divisão de $2/10 \times 100 = 20\%$. O decil divide em 10 partes iguais e o resultado será o total de estudantes, que no caso é 20. O valor de $\frac{1}{2}$ na fórmula é sempre fixo pois é o ponto da Mediana, local onde as três separatrizes (quartil, decil e centil) se encontram. Ao substituir na fórmula temos:

(Posição do numerador referente à fração da separatriz que se deseja/ número de partes da separatriz) $\times$ Total + (1/2);

Tem-se que $(2/10) \times 20 + (1/2) = (0,25 \times 20) + 0,5 =$

$2 + 0,5 = 2,5^{\text{a}}$ Posição. Quando não dá número inteiro na posição, como é o caso desse exemplo, é necessário identificar no Conjunto “A” a posição do 2,5, que está entre dois elementos marcados na cor verde, ou seja, entre o que está na segunda posição (que é o 1) e o da terceira posição (que é o 2), pois o número 2,5 se encontra entre a posição dois e a três ($A = \{0; 1; 2; 3; 4; 4; 4; 4; 4; 5; 5; 5; 5; 6; 6; 6; 7; 8; 9\}$). Depois disso, coloca-se em uma segunda fórmula, que irá tirar a média aritmética simples entre os dois termos centrais do ROL identificados pela cor verde, nesse caso os elementos da 2ª e da 3ª posição. Então, teremos: $\bar{x} = (1 + 2)/2 = 1,5$, com o D2 = 1,5. Fazendo a interpretação, temos que a nota 1,5 corresponde ao ponto que é exatamente 20% das notas de 20 estudantes da Escola A no ano de 2024.1.

Para 30% irei usar o C30, que corresponde à representação fracionária de 30/100. Então seu numerador será 30, pois a divisão de $30/100 \times 100 = 30\%$. O centil divide em 100 partes iguais e o resultado será o total de estudantes, que no caso é 20. O valor de $\frac{1}{2}$ na fórmula é sempre fixo pois é o ponto da Mediana, local onde as três separatrizes (quartil, decil e centil) se encontram. Ao substituir na fórmula, temos:

(Posição do numerador referente à fração da separatriz que se deseja/ número de partes da separatriz) $\times$ Total + $(1/2)$;

Tem-se que $(30/100) \times 20 + (1/2) = (0,3 \times 20) + 0,5 =$

$6 + 0,5 = 6,5^{\text{a}}$ **Posição**. Quando não dá número inteiro na posição, como é o caso desse exemplo, é necessário identificar no Conjunto “A” a posição do 6,5, que está entre dois elementos marcados na cor verde, ou seja, entre o que está na sexta posição (que é o 4) e o da sétima posição (que também é o 4), pois o número 6,5 se encontra entre a posição seis e a sete ($A = \{0;1;2;3;4;4;4;5;5;5;5;6;6;6;7;7;8;9\}$). Depois, coloca-se em uma segunda fórmula, que irá tirar a média aritmética simples entre os dois termos centrais do ROL identificados pela cor verde, nesse caso os elementos da 6ª e da 7ª posição. Então, teremos: $\bar{x} = (4 + 4)/2 = 4$, com o C30 = 4. Fazendo a interpretação, temos que a nota 4 corresponde ao ponto que é exatamente 30% das notas de 20 estudantes da Escola A no ano de 2024.1.

Gostaria de esclarecer que os três exemplos acima foram calculados pelas separatrizes (quartil, decil e centil) diferentes para dados não agrupados, mas poderiam ser todos calculados pelo centil. Por exemplo, o Q1 poderia ser calculado pelo C25 e o D2 pelo C20, assim como o C30 pelo D3. Todos os cálculos chegariam ao mesmo resultado.

O quartil, o decil e o centil usados para dados agrupados também utilizam as duas fórmulas da Mediana, e os cálculos do passo a passo são similares ao da Mediana para dados agrupa-

dos. Para encontrar a posição do elemento do quartil, do decil e do centil, a fórmula é $P = (\text{Total} \times \text{Posição do numerador referente à fração da separatriz que se deseja})/(\text{número de partes})$, onde:

- P é a posição do Quartil, do Decil e do Centil que se deseja encontrar;
- Total (também denominado de n) é o somatório total de elementos;
- Posição do numerador é referente à fração da separatriz que se deseja; e
- Divide-se pelo número de partes de cada separatriz: se Quartil, em quatro partes; caso for Decil, em 10 partes; e se for Centil, 100 partes.
- Já a segunda fórmula é para achar o valor numérico de qualquer uma das três separatrizes (Quartil, Decil e do Centil), sendo $s = L_i + \frac{P - \sum f_{\text{ant}}}{f_s} \cdot A$, onde:
- S = Valor da separatriz (Quartil, Decil e do Centil) que se deseja encontrar;
- L_i = limite inferior da classe que contém a separatriz (Quartil, Decil e do Centil);
- P = tamanho da amostra ou população que está sendo pesquisada, ou seja, o total de observações do acontecimento. É possível de substituir na fórmula, bastando pegar o resultado da fórmula de P calculada anteriormente;
- $\sum f_{\text{ant}}$ = valor da FACA anterior à classe em que se identificou a posição da Separatriz (do Quartil, do Decil e do Centil);
- A = amplitude da classe que contém a separatriz (do Quartil, do Decil e do Centil);

- f_s = frequência da classe que contém a separatriz (do Quartil, do Decil e do Centil).

A Figura 19 irá mostrar passo a passo do cálculo do Q1(25%) para dados agrupados.

Figura 19 – Exemplo de cálculo do Q1(25%) para dados agrupados

Tabela 1 – Dados organizados em ROL (Agrupado) das notas de estudantes – Escola A-2024.1		
Notas	Frequência Simples Absoluta (f_i)	FACA 2º Passo
0 - 2	2	2
2 - 4	2	4
4 - 6	8	12
6 - 8	6	18
8 - 10	2	20
Total (Σ)	20	

3º Passo :
FACA da
5ª Posição está
aqui-3ª classe

Fonte: Elaboração própria, 2024

1º Passo: Usaremos a fórmula: $P = (\text{Total} \times \text{Posição do numerador referente à fração da separatriz que se deseja})/(\text{número de partes})$, onde o Total é o somatório (Σ) de f_i (Segunda Coluna), nesse caso $n = 20$. A Posição do numerador referente à fração da separatriz que se deseja é 1 pois é oriunda de $\frac{1}{4} = 0,25 \times 100 = 25\%$, já o número de partes (por ser o quartil) é igual a quatro, cada uma igual a 25%. Ao colocar na fórmula, temos: $P = (20 \times 1)/4 = 5$. Assim, o Q1 se encontra na **5ª posição**.

2º Passo: Fazer a Frequência Acumulada Crescente-Abaixo (FACA) para ver a posição correspondente a **5ª** (Encontrado no 1º Passo). Ver como faz cálculo da FACA no Capítulo 2 – Tabela 2.

3º Passo: Identificar diretamente na Tabela, na coluna da FACA, o elemento da **5ª posição** (Encontrado no 1º Passo) e ver que corresponde às notas entre 4 - 6 presentes na primeira coluna da tabela. Afinal, ao olhar a FACA, identifica-se que a primeira classe vai incorporar a 1ª e a 2ª Posição, a segunda classe inicia da 3ª e vai até a 4ª posição, e a terceira classe inicia da 5ª e vai até a 12ª posição. Então, o elemento da 5ª posição se encontra nessa classe.

continua...

conclusão

4º Passo: Substituir na segunda fórmula da separatriz para no caso o Q1, que é $s = L_i + \frac{P - \sum f_{ant}}{f_s} \cdot A$, a partir de todos os dados referentes a 3ª Classe. Assim, temos:

- L_i = limite inferior da classe que contém o Q1. É igual a 4 (cor roxa), pois o Intervalo da Mediana está na 3ª classe 4 - 6.
- n = tamanho da amostra ou população que está sendo pesquisada, ou seja, o total de observações do acontecimento. Então, basta substituir na fórmula e pegar o resultado da fórmula do 1º Passo, que corresponde a 5ª posição.
- $\sum f_{ant}$ = valor da FACA anterior à classe, que identificou a posição do Q1. Assim, tem-se que a FACA da posição do Q1 referente a 3ª classe é 12, e a anterior é 4.
- A = amplitude da classe que contém a mediana, que é 6 - 4 = 2.
- f_s = frequência da classe que contém a mediana (ver segunda coluna, em que $f_5 = 8$).

Substituindo os dados na fórmula, tem-se:

$$Q1 = 4 + (5 - 4)/8 \times 2 = 4,25$$

5º Passo: Verificar se o número encontrado no 4º Passo, que no caso foi 4,25, está dentro 3ª Classe, no intervalo entre 4 - 6. **Se a resposta não tiver coerente, rever cálculos. Como o 4,25 está dentro da 3ª classe, pode-se iniciar a interpretação.**

6º Passo: 4,25 corresponde ao ponto de 25% das notas de 20 estudantes de uma Escola A no ano de 2024.1.

Fonte: Elaboração própria, 2024

Já a Figura 20 irá mostrar o cálculo referente a 75%. O mesmo pode ser realizado pelo Q3 ou pelo C75, como na Figura 19. Como já foi realizado ao cálculo d Quartil(Q1), farei pelo Centil para que o leitor veja que é possível usar esse ao invés do outro e chegar ao mesmo resultado.

Figura 20 – Exemplo de cálculo do C75 (75%) para dados agrupados

Tabela 1 – Dados organizados em ROL (Agrupado) das notas de estudantes – Escola A-2024.1		
Notas	Frequência Simples Absoluta (fi)	FACA 2º Passo
0 - 2	2	2
2 - 4	2	4
4 - 6	8	12
6 - 8	6	18
8 - 10	2	20
Total (Σ)	20	

3º Passo : FACA da 15ª Posição está aqui-4ª classe

Fonte: Elaboração própria, 2024

1º Passo: Usaremos a fórmula: $P = (\text{Total} \times \text{Posição do numerador referente à fração da separatriz que se deseja})/(\text{número de partes})$, onde o Total é o somatório (Σ) de f_i (Segunda Coluna), nesse caso $n = 20$. A Posição do numerador referente à fração da separatriz que se deseja é 1 pois é oriunda de $75/100 = 0,75 \times 100 = 75\%$. Já o número de partes (por ser o centil) é igual a 100, cada uma igual a 1%. Ao colocar na fórmula, temos: $P = (20 \times 75)/100 = 15$. Assim, o C75 se encontra na **15ª posição**.

2º Passo: Fazer a Frequência Acumulada Crescente-Abaixo (FACA) para ver a posição correspondente a **15ª** (Encontrado no 1º Passo). Ver como fazer cálculo da FACA no Capítulo 2 – Tabela 2.

3º Passo: Identificar diretamente na Tabela, na coluna da FACA, o elemento da **15ª posição** (Encontrado no 1º Passo) e ver que corresponde às notas entre 6 - 8 presentes na primeira coluna da tabela. Afinal, ao olhar a FACA, identifica-se que a primeira classe vai incorporar a 1ª e a 2ª Posição, a segunda classe inicia da 3ª e vai até a 4ª posição, a terceira classe inicia da 5ª e vai até a 12ª posição, e já 4ª classe vai da 13ª posição até 18ª. Então, o elemento da 15ª posição se encontra nessa classe.

4º Passo: Substituir na segunda fórmula da separatriz o C75, que é: $s = L_i + \left[\frac{P - \sum f_{\text{ant}}}{f_s} \right] \cdot A$, a partir de todos os dados referentes a 3ª Classe. Assim temos:

- L_i = limite inferior da classe que contém o C75. É igual a **6 (cor roxa)** pois o Intervalo da Mediana está na 3ª classe 4 - 6;
- n = tamanho da amostra ou população que está sendo pesquisada, ou seja, o total de observações do acontecimento. Então, basta substituir na fórmula e pegar o resultado da fórmula do 1º Passo, que corresponde a 5ª posição;
- $\sum f_{\text{ant}}$ = valor da FACA anterior à classe, que identificou a posição do C75. Assim, tem-se que a FACA da posição do C75 referente a 4ª classe é 18, e a anterior é 12;
- A = amplitude da classe que contém a mediana, que é $8 - 2 = 6$;
- f_s = frequência da classe que contém a mediana (ver segunda coluna $f_5 = 6$).

Substituindo os dados na fórmula, tem-se:

$C75 = 6 + (15 - 12)/6 \times 2 = 7$

continua...

conclusão

5º Passo: Verificar se o número encontrado no 4º Passo, no caso 7, está dentro 4ª Classe, no intervalo entre 6 e 8. Se a resposta não estiver coerente, deve-se rever os cálculos. Como o 7 está dentro da 4ª classe, pode-se iniciar a interpretação.

6º Passo: 7 corresponde ao ponto que 75% das notas de 20 estudantes de uma Escola A no ano de 2024.1.

Fonte: Elaboração própria, 2024.

No tópico 3.1.1 falamos que, quando a média aritmética for atípica, deve-se usar a mediana. Assim, a média aritmética deve ser sempre somada e subtraída do desvio padrão e a mediana deve ser somada e subtraída do desvio quartil. Entretanto, se a média aritmética for atípica, deve-se usar na interpretação a mediana, pois corresponde ao ponto de 50%, e assim há uma interpretação melhor da realidade dos dados.

O desvio quartil (representado por Dq) corresponde ao intervalo definido pelos limites ($Md - Dq$) e ($Md + Dq$). Nesse intervalo são encontrados 50% dos elementos da série, aproximadamente. No caso de a distribuição ser simétrica, essa porcentagem será exata. Outra grande importância do desvio quartil é o fato de ser uma medida não afetada pelos valores extremos da série como ocorre (por exemplo) com a média aritmética atípica, portanto será útil quando esse fato ocorre. Sua fórmula é $Dq = (Q3 - Q1)/2$, onde Dq é o desvio quartil, Q1 é o valor do primeiro quartil (correspondente a 25%), e Q3 é o valor do terceiro quartil (correspondente a 75%). O desvio quartil também é chamado de amplitude semi-interquartilica.

Ao calcularmos o desvio quartil dos dados das Figura 17 (que tem o valor do Q1 como 4,25) e da Figura 18 (que tem o valor do Q3 como 7, lembrando que o valor do $C75 = Q3$ e que ambos correspondem a 75%), temos a seguinte substituição de fórmula:

- $Dq = (Q3 - Q1)/2$
- $Dq = (7 - 4,25)/2 = 1,375$

Depois de achar o Dq , é necessário encontrar o intervalo definido pelos limites $(Md - Dq)$ e $(Md + Dq)$. Então, temos que a Md é 5,5 (conforme Figura 15) ao acharmos que o intervalo de $Md - Dq = 5,5 - 1,375 = 4,125$, e que o intervalo de $Md + Dq = 5,5 + 1,375 = 6,875$. Dessa forma, 50% das notas que compreendem aos estudantes da Escola Ano 2024.1 estão entre $[4,125;6,875]$. Já ao fazer essa comparação com o intervalo da $(\bar{x} - S)$ e $(\bar{x} + S)$ – Figura 12, constata-se que 68% das notas dos estudantes da Escola no Ano 2024.1 estão entre $[3,19;7,61]$.

3.1.3 Medida de Tendência Central (Moda)

A Moda (identificada pela sigla “Mo”) é uma outra medida de Tendência Central. Ela ocorre quando há com maior frequência em uma série de valores. Pode ser calculada tanto para dados não agrupados como agrupados.

Vamos começar a explicação através dos dados não agrupados, em que se tem o seguinte conjunto: $A = \{0;1;2;3;4;4;4;4;4;5;5;5;5;6;6;6;7;7;8;9\}$. Assim, a moda é o número 4, pois ele se repete cinco vezes. Essa moda é denominada unimodal, pois existe apenas um único valor dela. Assim, a nota 4 foi a nota mais tirada pelos estudantes da Escola no Ano 2024.1, com cinco alunos obtendo-a. A distribuição de frequência do conjunto A anterior pode ser representada também de forma tabular (conforme Tabela abaixo), e quando isso ocorre, observaremos o valor que mais se repete na coluna de fi (no caso, o número 5). Ao verificar a coluna de notas, percebe-se que o número 5 está relacionado com o valor 4. Assim, a resposta para a representação em conjunto e tabular será idêntica para esses dados não agrupados.

A nota 4 foi a nota mais tirada pelos estudantes da Escola no Ano 2024.1 e cinco alunos tiraram a nota 4. Assim, a tabela tem uma única Moda – Unimodal, presente na 5ª classe

Tabela 4 – Dados organizados em ROL (Não Agrupado) das notas de estudantes – Escola A-2024.1

Notas	Frequência Simples Absoluta (fi)
0	1
1	1
2	1
3	1
4	5
5	3
6	4
7	2
8	1
9	1
Total (Σ)	20



Fonte: Elaboração própria, 2024.

Entretanto, podemos encontrar na pesquisa séries de distribuição em ROL nas quais não exista um valor que mais se repete, isto é, nenhum valor aparece mais vezes do que outros. É o caso do seguinte conjunto: $C = \{4, 5, 9, 11, 12, 14\}$. Quando isso ocorre, dizemos que é amodal, ou seja, não há moda.

Pode haver dois valores de concentração que se repetem na mesma quantidade, por exemplo no conjunto $D = \{1, 2, \underline{4}, \underline{4}, \underline{4}, 5, 6, \underline{7}, \underline{7}, \underline{7}, 9, 9\}$. Nesse caso temos duas modas: o número 4 e o 7. Cada um se repetiu em igual frequência, ou seja, três vezes, e assim, relatamos que a moda é bimodal por serem dois números. Caso tenha mais de dois valores se repetindo, consideraremos como multimodal.

Para dados agrupados em intervalos de classes, você pode calcular a moda por meio do método de Czuber, que se baseia na influência das classes adjacentes na moda se deslocando no sentido da classe de maior frequência. A fórmula é: $Mo = l_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot h$, onde:

- li é o limite inferior da classe de maior frequência;
- h é o valor do intervalo de classe;
- Δ_1 é a diferença entre a maior frequência e a frequência simples anterior; e
- Δ_2 é a diferença entre a maior frequência e a frequência simples posterior.

Vamos explicar passo a passo a Moda através da fórmula de Czuber para dados agrupados utilizando a Figura 21.

Figura 21 – Exemplo de cálculo da Moda através da fórmula de Czuber para dados agrupados

Tabela 1 – Dados organizados em ROL (Agrupado) das notas de estudantes – Escola A-2024.1															
	<table><tr><th>Notas</th><th>Frequência Simples Absoluta (f_i)</th></tr><tr><td>0 ÷ 2</td><td>2</td></tr><tr><td>2 ÷ 4</td><td>2</td></tr><tr><td>4 ÷ 6</td><td>8</td></tr><tr><td>6 ÷ 8</td><td>6</td></tr><tr><td>8 ÷ 10</td><td>2</td></tr><tr><td>Total (Σ)</td><td>20</td></tr></table>	Notas	Frequência Simples Absoluta (f_i)	0 ÷ 2	2	2 ÷ 4	2	4 ÷ 6	8	6 ÷ 8	6	8 ÷ 10	2	Total (Σ)	20
Notas	Frequência Simples Absoluta (f_i)														
0 ÷ 2	2														
2 ÷ 4	2														
4 ÷ 6	8														
6 ÷ 8	6														
8 ÷ 10	2														
Total (Σ)	20														
Passo 1: Maior valor de f_i é o número 8, e está no intervalo das notas entre 4 ÷ 6 na 3ª classe															
Fonte: Elaboração própria, 2024															
Passo 1: Identificamos a maior frequência na segunda coluna de f_i. No caso específico, é o número 8, que corresponde ao intervalo de 4 ÷ 6 (3ª Classe) Passo 2: Colocar na fórmula de Czuber $Mo = l_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} . h$, onde: • li é o limite inferior da classe de maior frequência é 4; • h é o valor do intervalo de classe (6 - 4 = 2); • Δ_1 é a diferença entre a maior frequência e a frequência simples anterior (8 - 2 = 6); • Δ_2 é a diferença entre a maior frequência e a frequência simples posterior (8 - 6 = 2). Após identificar cada valor, substituir na fórmula: $Mo = 4 + (6/6 + 2) \times 2 = 5,5$ Passo 3: Constatar se o número encontrado no Passo 2, no caso 5,5, está no intervalo da 3ª classe (4 ÷ 6). Estando, pode-se fazer a interpretação; se não, rever cálculos. Passo 4: É a Interpretação dos dados da moda. Assim, tem-se que o valor de 5,5 corresponde à nota mais tirada pelos alunos da Escola no Ano 2024.1															

Fonte: Elaboração própria, 2024.

O Quadro 12 mostra as principais vantagens e desvantagens das três Medidas de Tendências Centrais (Média Aritmética, Mediana e Moda).

Quadro 12 – Principais vantagens e desvantagens das três Medidas de Tendências Centrais (Média Aritmética, Mediana e Moda)

Medidas de Tendências Central	Vantagem	Desvantagem
Média Aritmética	• Fácil de compreender e usar • Usa todos os dados disponíveis • Fácil de incluir em equações matemáticas	• É afetada pelos valores extremos • É necessário conhecer todos os valores da distribuição
Mediana	• Fácil de determinar • Não é afetada pelos valores extremos • Parece ser uma medida correta, pois divide a série em duas partes iguais a 50%	• Difícil de incluir em equações matemáticas • Não usa os dados disponíveis
Moda	• Fácil de calcular • Não é afetada pelos valores extremos	• Pode estar afastada do centro das observações • Difícil de incluir em equações matemáticas • A distribuição pode ter mais de uma moda • Não usa todos os dados disponíveis

Fonte: Bruni (2008, p. 55).

3.2 MEDIDAS DE ASSIMETRIA E CURTOSE

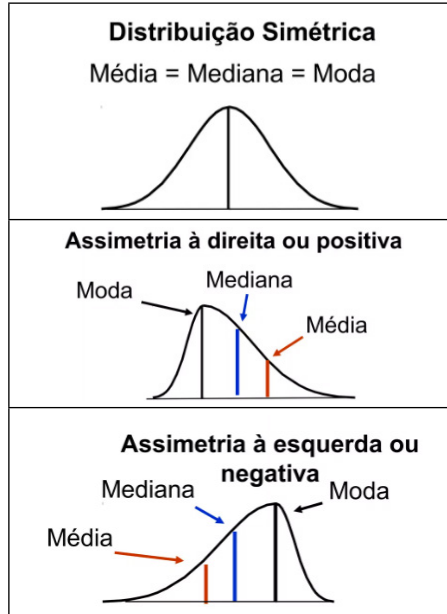
O nome assimetria significa desvio ou afastamento da simetria, ou seja, é o grau de formação da curva de frequências. Permite analisar uma distribuição de acordo com as relações entre suas medidas de moda, média e mediana quando observadas graficamente ou apenas por meio da análise dos valores. A partir

dos valores das três medidas de tendência central, a curva da distribuição de frequência pode ser:

- Simétrica ou nula, quando apresenta o mesmo valor para a média, a mediana e a moda, $\bar{x} = Md = Mo$ (Figura 22);
- Assimétrica positiva ou à direita, quando a cauda da curva da distribuição declina para direita. Assim, o valor da média é maior que o valor da moda e se tem a seguinte relação: $Mo \leq Md \leq \bar{x}$ (Figura 22);
- Assimétrica negativa ou à esquerda, quando a cauda da curva da distribuição declina para esquerda, então o valor da média é menor que o valor da moda. Tem-se a seguinte relação: $\bar{x} \leq Md \leq Mo$ (Figura 22).

O estudo dessa curva de distribuição de frequência, que pode ser simétrica ou assimétrica, será representado pela letra “ A_s ”. Uma distribuição é assimétrica quando a igualdade dos valores das medidas de tendências centrais não ocorre. A Figura 22 mostra os tipos de curva da distribuição de frequência citados anteriormente de cima para baixo, respectivamente.

Figura 22 – Tipos de curva de assimetria



Fonte: Elaboração própria, 2024.

Outra forma de mensurar a assimetria é através da distância entre a média e a moda: quanto maior a distância entre essas duas medidas (seja assimétrica positiva ou negativa), maior será a assimetria da distribuição. Outra forma de identificar a curva de distribuição de frequência, além da comparação entre as três medidas de tendência central, é através da fórmula do primeiro coeficiente de Pearson $A_s = \frac{\bar{X} - Mo}{S}$ ou $A_s = \frac{\bar{X} - Mo}{\sigma}$, onde:

- A_s = coeficiente de assimetria;
- $\bar{x}$ = média aritmética dos dados;
- Mo = moda;
- S = desvio padrão amostral ou σ = desvio padrão populacional.

Então, tem-se como identificação do tipo de distribuição de frequência:

- Se a resposta de $A_s = 0$, a distribuição é simétrica;
- Se a resposta de $A_s > 0$, a distribuição é assimétrica positiva;
- Se a resposta de $A_s < 0$, a distribuição é assimétrica negativa.

Também existe uma segunda fórmula, denominada segundo coeficiente de assimetria de Pearson, na qual é possível encontrar a distribuição de frequência através da seguinte Fórmula:

$$A_s = \frac{3 \cdot (\bar{X} - Md)}{S}, \text{ onde}$$

- A_s = coeficiente de assimetria;
- $\bar{x}$ = média aritmética dos dados;
- Md = Mediana;
- S = desvio padrão amostral, que pode ser substituído também pelo σ = desvio padrão populacional.

Então, tem-se como identificação do tipo de distribuição de frequência:

- Se $-0,15 < |A_s| < 0,15$, consideramos a curva pouco assimétrica ou simétrica;
- Se $0,15 < A_s < 1$, temos uma assimetria moderada à direita;
- Se $A_s > 1$, temos uma assimetria forte à direita;
- Se $-1 < A_s < -0,15$, temos uma assimetria moderada à esquerda; e
- Se $-1 < A_s$, temos uma assimetria forte à esquerda.

E a fórmula do segundo coeficiente de assimetria do Coeficiente Quartílico:

$$A_s = \frac{Q_1 + Q_3 - 2M_d}{Q_3 - Q_1}, \text{ onde}$$

- A_s = coeficiente de assimetria;
- Q_1 = Quartil 1(25%);
- Q_3 = Quartil 3(75%);
- M_d = Mediana;
- A_s = coeficiente de assimetria;
- $\bar{x}$ = média aritmética dos dados;
- M_d = Mediana; e
- S = desvio padrão amostral, que pode ser substituído também pelo σ = desvio padrão populacional.

A fórmula para calcular a distribuição de frequência para ver a posição de sua assimetria dependerá da facilidade de encontrar os números para os cálculos estáticos.

A Tabela 5 mostra como identificar e realizar o cálculo de assimetria dos exemplos das notas dos estudantes de uma escola no ano de 2024.1. O valor da Média aritmética, Mediana, Moda e Desvio Padrão foram feitos anteriormente neste capítulo, assim como as outras informações sintetizadas presentes na Tabela 5.

Através dos dados da Tabela 5, percebe-se que ambas as tabelas foram oriundas do mesmo dado bruto (capítulo 2) e possuem o mesmo somatório total de 20 estudantes. Entretanto, ao colocar os dados agrupados em intervalos, há uma maior concentração dos dados pois diminuem os números de linhas das tabelas (classe), fazendo com que os valores tenham como parâmetro o ponto médio dos intervalos de cada classe, e não o real valor das notas nos dados coletados da amostra. Todos os dados presentes na Tabela 5 estão corretos. Cabe ao pesquisador, dentro dos recursos que possui (tempo, quantidade de informações, objetivo da pesquisa, dentre outros), ver a melhor forma de representá-los.

Outro aspecto observado na Tabela 5 é que, dentro dos dados agrupados e não agrupados, há diferenças de assimetria. O dado não agrupado teve duas assimetrias positivas e duas negativas. Já o

agrupado teve três negativas e uma positiva. Isso ocorre pois, ao usar como parâmetro a Mo em relação à média aritmética, se $Mo < \bar{x}$ ela será negativa, mas se $Mo > \bar{x}$ ela será positiva. Então, a Mo influencia na interpretação da assimetria quando confronta apenas com o valor das três medidas de tendência central, assim como na primeira fórmula do primeiro coeficiente de Pearson. Entretanto, a fórmula do segundo coeficiente de Pearson e do Coeficiente Quartílico à Md tendem a influenciar no cálculo da assimetria. Assim, todos os cálculos estão corretos, mas depende dos dados disponíveis ou o que se deseja mostrar na pesquisa para a utilização da fórmula.

Tabela 5 – Exemplo de cálculo assimetria para dados não agrupados e agrupados. Escola A-2024.1

Medidas	Dado Não Agrupado	Dado Agrupado
Amostra	20	20
$\bar{x}$	4,8, aproximadamente 5	5,4
Da	1,25	0,62
Variância	3,76	2,15
Desvio padrão-68%	2,23 - [2,56;7,03].	2,21 - [3,19;7,61].
Erro	0,51 - [2,04;7,54]	0,51 - [2,68;8,12]
CV	46,6%,	41%
Mediana	5	5,5
Q1	0	4,25
Q3	6	7
Desvio quartil	3 - [2;8]	1,375 - [3,19;7,61].
Moda	4	5,5
Assimetria confrontando as três medidas de tendências centrais	$Mo(4) < \bar{x}(4,8) < Md(5)$ assimétrica positiva $Mo(4) < Md(5) = \bar{x}(5)$ Se usar a $\bar{x}$ aproximada, tem-se o número 5 assimétrica positiva	$\bar{x}(5,4) < Md(5,5) = Mo(5,5)$ assimétrica negativa

continua...

conclusão

Medidas	Dado Não Agrupado	Dado Agrupado
Primeiro coeficiente de Pearson	$(4,8 - 4)/2,23 = 0,8/2,23 = 0,35$ $A_s > 0$, assimétrica positiva	$(5,4 - 5,5)/2,21 = -0,1/2,21 = -0,044$ $A_s < 0$, assimétrica negativa
Segundo coeficiente de Pearson	$3 \times (4,8 - 5)/2,23 = -0,6/2,23 = -0,27$ (assimétrica moderada negativa e moderada)	$3 \times (5,4 - 5,5)/2,21 = -0,3/2,21 = -0,14$ (assimétrica moderada negativa)
Coeficiente Quartílico	$(0+6) - (2 \times 5)/6 - 0 = 6 - 10/6 = -0,67$ (assimétrica moderada negativa e moderada)	$(4,25+7) - 2 \times (5,5)/(7 - 4,25) = 11,25 - 11/3,25 = 0,077$ (assimétrica moderada positiva)

Fonte: Elaboração própria, 2024.

Já a curtose é a característica que se refere ao grau de achatamento da curva de distribuição de frequência em relação a uma curva normal I (curva correspondente a uma distribuição teórica de probabilidade) de referência. De acordo com ela, a distribuição é classificada da seguinte forma:

- Mesocúrtica ocorre quando a curva de frequências apresenta um grau de achatamento equivalente ao da curva normal;
- Platicúrtica ocorre quando a curva de frequências se apresenta mais aberta (ou mais achatada na parte superior) que a curva normal;
- Leptocúrtica ocorre quando a curva de frequências se apresenta mais fechada (ou mais afilada na parte superior) que a curva normal.

A fórmula para a medida da curtose é

$$C = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}, \text{ onde}$$

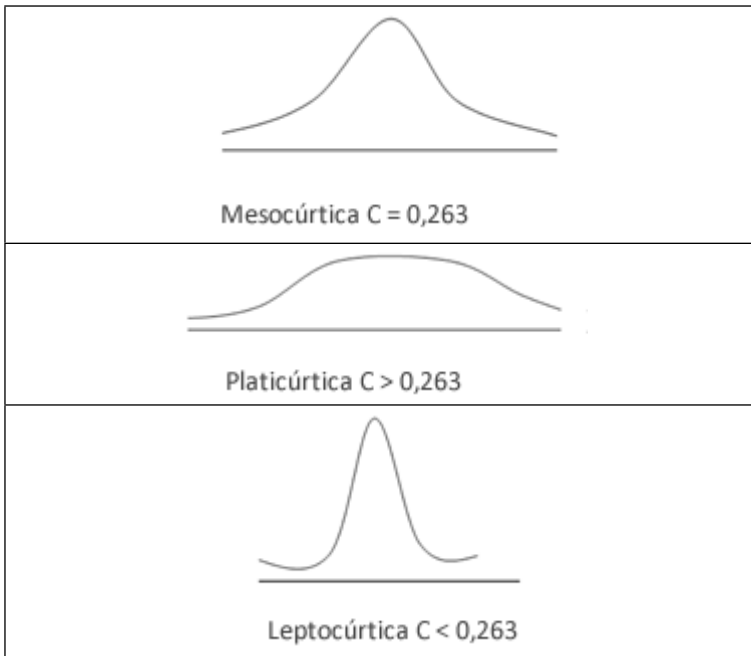
- C é o valor do coeficiente percentílico de curtose;
- Q_3 é quartil 3(75%);
- Q_1 é o Quartil 1(25%);
- P_{90} é o Percentil 90(90%); e
- P_{10} é o Percentil 10(10%).

Então, tem-se como identificação do tipo de curtose:

- $C = 0,263 \rightarrow$ curva mesocúrtica;
- $C > 0,263 \rightarrow$ curva platicúrtica;
- $C < 0,263 \rightarrow$ curva leptocúrtica.

A Figura 23 mostra da esquerda para direita as três curvas de curtose relatada anteriormente de forma respectiva

Figura 23 – Tipos de curva de Curtose



Fonte: Elaboração própria, 2024.

Para saber o valor da curtose e consequentemente o tipo de curva, fazer os cálculos inicialmente com os dados não agrupados e depois com os dados agrupados. Em relação ao mesmo exemplo dos capítulos 2 e 3, em que existe uma amostra com 20 notas de estudantes de uma Escola A no período de 2024.1, corresponde ao conjunto $A = \{0;1;2;3;4;4;4;4;5;5;5;5;6;6;7;7;8;9\}$

$$C = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}, \text{ onde}$$

- Q_3 é Quartil 3(75%) dos dados não agrupados = 6;
- Q_1 é o Quartil 1(25%) dos dados não agrupados = 0;
- P_{90} é o Percentil 90(90%) dos dados não agrupados = 7,5;
- P_{10} é o Percentil 10(10%) dos dados não agrupados = 2,5.

Observação : Para calcular os itens da separatriz presente na fórmula de curtose, verificar item 3.1.2.1 sobre Separatriz

A partir dos dados, basta substituir na fórmula do coeficiente percentílico de curtose, onde teremos $(6 - 0)/2 \times (7,5 - 2,5) = 6/2 \times 5 = 0,6$. Então $0,6 > 0,263$ - curva platicúrtica.

Agora, vamos fazer o cálculo para dados agrupados, colocando-o na fórmula do coeficiente percentílico de curtose. Assim, a fórmula é $C = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}$, onde

- Q_3 é o Quartil 3(75%) dos dados não agrupados = 7;
- Q_1 é o Quartil 1(25%) dos dados não agrupados = 4,25;
- P_{90} é o Percentil 90(90%) dos dados não agrupados = 8; e
- P_{10} é o Percentil 10(10%) dos dados não agrupados = 2.

Realizando a substituição de fórmula, tem-se $:(7 - 1)/2 \times (8 - 2) = 6/2 = 0,5$. Então, $0,5 > 0,263$ curva platicúrtica.

3.3 CORRELAÇÃO E REGRESSÃO LINEAR

A correlação é usada em pesquisa para verificar se há ou não relação entre as variáveis estudadas e qual o grau dessa. Pode-se destacar a relação da $\text{Peso} \times \text{Idade}$, $\text{Consumo} \times \text{Renda}$, $\text{Demanda} \times \text{Preço}$, $\text{Consumo de cigarro} \times \text{Incidência de câncer}$, $\text{Nota do aluno} \times \text{frequência}$, dentre outras.

Para realizar o estudo da correlação entre as variáveis, usa-se o Diagrama de Dispersão. Nele, os dados são organizados em um sistema coordenado cartesiano ortogonal em pares ordenados (x, y) . A Tabela 6 mostra um exemplo hipotético e que foram escolhidos 10 estudantes de 20 através de um sorteio probabilístico aleatório, sem reposição. Nessa Tabela 6, há a relação de duas variáveis: o tempo em horas de estudo dos alunos para Estatística (X) e a sua respectiva nota na disciplina (Y)

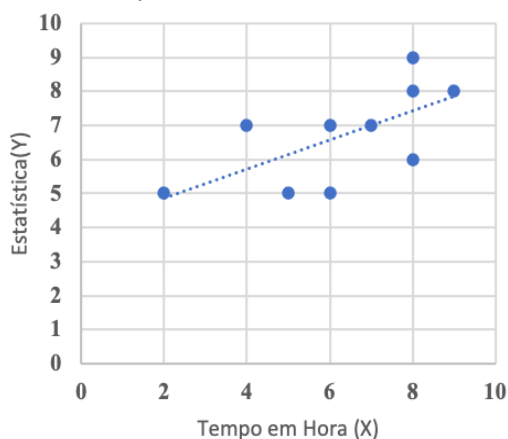
Tabela 6 – Relação entre o estudo por hora e notas do componente de Estatística- Escola A-2024.1

Sorteio de Amostra de alunos através de uma lista Probabilística Simples	Relação entre duas variáveis	
	Tempo em hora praticando Estatística (X)	Notas de Estatística (Y)
2	6	7
7	8	9
18	7	7
24	8	8
33	6	5
57	4	7
60	9	8
73	2	5
80	8	6
88	5	5

Fonte: Elaboração própria, 2024.

O Gráfico 1 mostra o Diagrama de Dispersão e uma linha de tendência pontilhada. Para que haja sempre correlação, é necessário que os pontos sigam ou fiquem próximo da linha de tendência, seja subindo ou descendo; se isso não ocorre, não haverá correlação. Nesse caso, existe correlação pois os pontos seguem a linha de tendência para direção superior, estando próxima da mesma.

Gráfico 1 – Diagrama de Dispersão referente à Relação entre o estudo por hora e notas do componente de Estatística- Escola A-2024.1



Fonte: Elaboração própria, 2024.

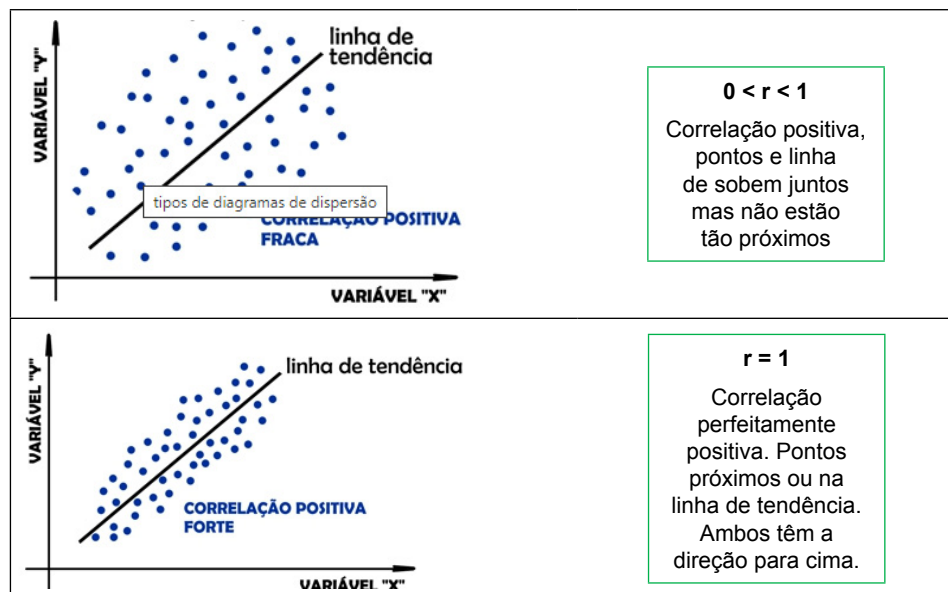
Há cinco tipos de correlação. Ao analisar em gráfico através da matriz de correlação a partir do coeficiente de Pearson(r) com o nível de confiança de 5%, classificamos da seguinte forma se:

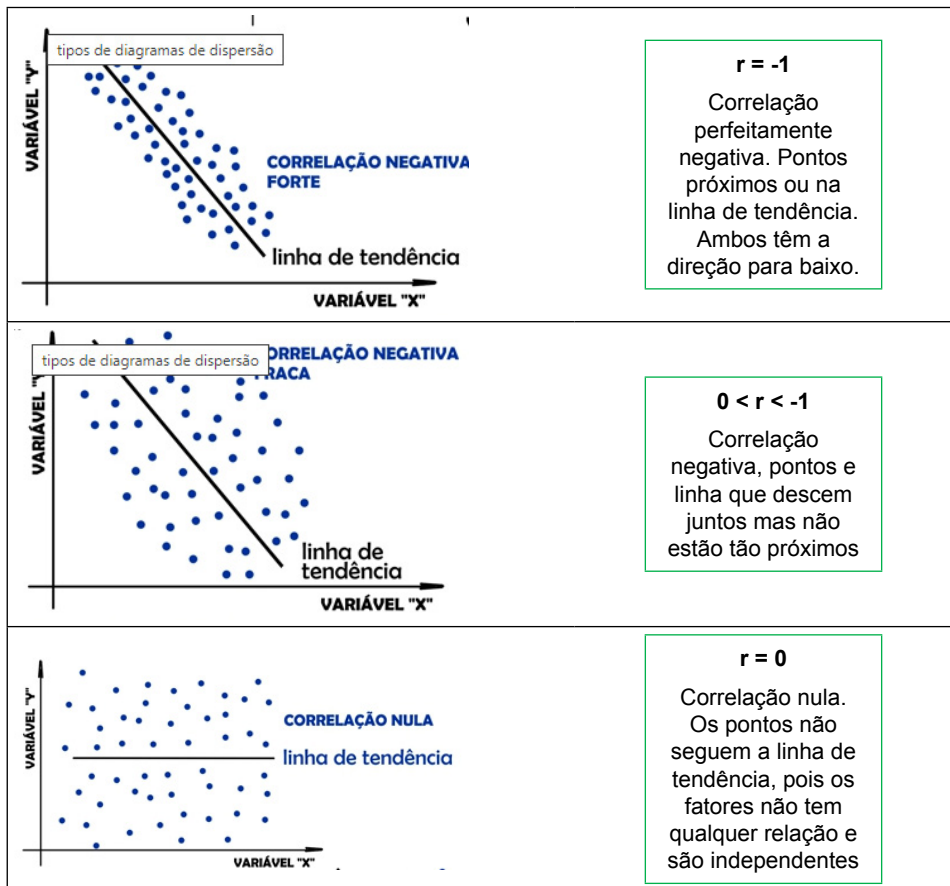
- $0 < r < 1$ é uma correlação positiva. Nesse caso, quanto mais próximo r estiver de + 1, mais próximos estarão os pontos de ajustes, ou seja, quando Y cresce, X cresce também; quando Y decresce, X decresce também;
- $r = 1$ é uma correlação perfeitamente positiva. Nesse caso, ocorre quando X e Y estão perfeitamente alinhados;

- $-1 < r < 0$ é uma correlação linear negativa. Quando isso ocorre, Y cresce, X decresce e vice-versa;
- $r = -1$ é uma correlação perfeitamente negativa. Ocorre quando os valores de X e Y estiverem perfeitamente alinhados, mas em sentido contrário; assim, estarão mais próximos os pontos de ajustes integrais a uma reta decrescente;
- $r = 0$ ocorre se não houver relação entre X e Y. Assim, as variáveis ocorrem independentemente (pereira, 2019, p. 169-170).

Para que haja uma visualização dos cinco tipos de correlação, foi elaborada a Figura 24 para que o leitor possa identificá-la ao fazer o seu Gráfico de Dispersão.

Figura 24 – Desenho dos cinco Tipos de Correlação de forma Gráfica com a Linha de Tendência





Fonte: Elaboração própria, 2024.

O Gráfico 1 elaborado a partir dos dados da Tabela 6 deixa claro pela Figura 24 que é uma correlação positiva. A fórmula da Correlação de Pearson é dada por:

$$r = \frac{n \sum x_i y_i - (\sum x_i) (\sum y_i)}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2] [n \sum y_i^2 - (\sum y_i)^2]}} , \text{ onde } n \text{ é o número de observações.}$$

Para maior compreensão dessa fórmula, fizemos o cálculo passo a passo do exemplo da Tabela 6 para encontramos a Correlação de Pearson da Relação entre o estudo por hora e notas da disciplina de Estatística. Por isso elaboramos a Tabela 7.

1º Passo: Fazer o somatório da ordenada X (Ver Tabela 7).

2º Passo: Fazer o somatório da ordenada Y (Ver Tabela 7).

3º Passo: Abrir uma coluna colocando X^2 (Ver Tabela 7).

4º Passo: Fazer os somatórios da coluna de X^2 (Ver Tabela 7).

5º Passo: Abrir uma coluna colocando Y^2 (Ver Tabela 7).

6º Passo: Fazer os somatórios da coluna de Y^2 (Ver Tabela 7).

7º Passo: Abrir uma coluna colocando XY (Ver Tabela 7).

8º Passo: Fazer os somatórios da coluna de XY (Ver Tabela 7).

9º Passo: Substituir na Fórmula a partir dos dados encontrados.

$$r = \frac{n \sum x_i y_i - (\sum x_i) (\sum y_i)}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2] [n \sum y_i^2 - (\sum y_i)^2]}}, \text{ onde:}$$

$$10 \times (440) - (63 \times 67) / \sqrt{[(10 \times 439) - (63)^2] \times [(10 \times 467) - (67)^2]} \\ = 179 / \sqrt{76.201}$$

$179 / 276,05 = \mathbf{0,6484}$ (Valor entre $0 < r < 1$ - Correlação positiva).
Entretanto, r está entre $0,6 < |r| < 0,1$, então há correlação entre o fato de estudar estatística e tirar boas notas nessa disciplina, sendo possível tirar conclusões.

Observação:

Se tiver entre $0,6 \leq |r| \leq 1$ é possível tirar conclusões significativas sobre o comportamento das variáveis.

Se estiver entre $0,3 \leq |r| < 0,6$, há uma correlação relativamente fraca entre as variáveis. Se estiver entre $0 < |r| < 0,3$, a correlação é muito fraca e, praticamente, nada podemos concluir sobre a relação entre as variáveis em estudo.

Tabela 7 – Exemplo do uso da correlação de Pearson da Relação entre o estudo por hora e notas da disciplina de Estatística-Escola A-2024.1

Sorteio de Amostra de alunos através de uma lista Probabilística Estratificada	Tempo em hora praticando Estatística (X)	Notas de Estatística (Y)	3º Passo: X²	5º Passo: Y²	7º Passo: XY
2	6	7	36	49	42
7	8	9	64	81	72
18	7	7	49	49	49
24	8	8	64	64	64
33	6	5	36	25	30
57	4	7	16	49	28
60	9	8	81	64	72
73	2	5	4	25	10
80	8	6	64	36	48
88	5	5	25	25	25
Total (Σ)	63	67	439	467	440
	1º Passo	2º Passo	4º Passo	6º Passo	8º Passo

Fonte: Elaboração própria, 2024.

Vídeo do uso da correlação



Link: <https://youtu.be/FodXGyal-ds>

Vídeo do Cálculo da correlação no Excel e a elaboração do seu gráfico e dispersão com a linha de tendência



Link: <https://youtu.be/mJbzSrhi7pc>

Já a análise de regressão tem por objetivo descrever, através de um modelo matemático, a relação entre duas variáveis, partindo de n observações das mesmas.

Uma vez caracterizada a relação entre as variáveis, procura-se descrevê-la através de uma função matemática. A variável sobre a qual se deseja fazer uma estimativa recebe o nome de variável dependente ou explicada, e a outra recebe o nome de variável independente ou explicativa. Vamos supor que X é uma variável independente e Y a dependente, procura-se então determinar, por meio de uma função, uma reta que melhor represente a relação verdadeira entre essas variáveis através da fórmula: $Y = aX + b$. Então, deve-se calcular o valor de a e b . O primeiro é calculado através da fórmula: $a = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$. Já b é através da Fórmula $b = \bar{y} - a\bar{x}$, onde:

- n é o número de observações;
- $\bar{x}$ é a média dos valores x_i , que corresponde a $\left(\bar{x} = \frac{\sum x_i}{n}\right)$; e
- $\bar{y}$ é a média dos valores y_i , que corresponde a $\left(\bar{y} = \frac{\sum y_i}{n}\right)$.

Por estar sendo usada uma amostra para se obter os valores dos parâmetros, como resultado se tem uma estimativa da verdadeira equação de regressão, então:

$$\hat{Y} = aX + b - \text{Sendo } \hat{Y} \text{ o } Y \text{ estimado.}$$

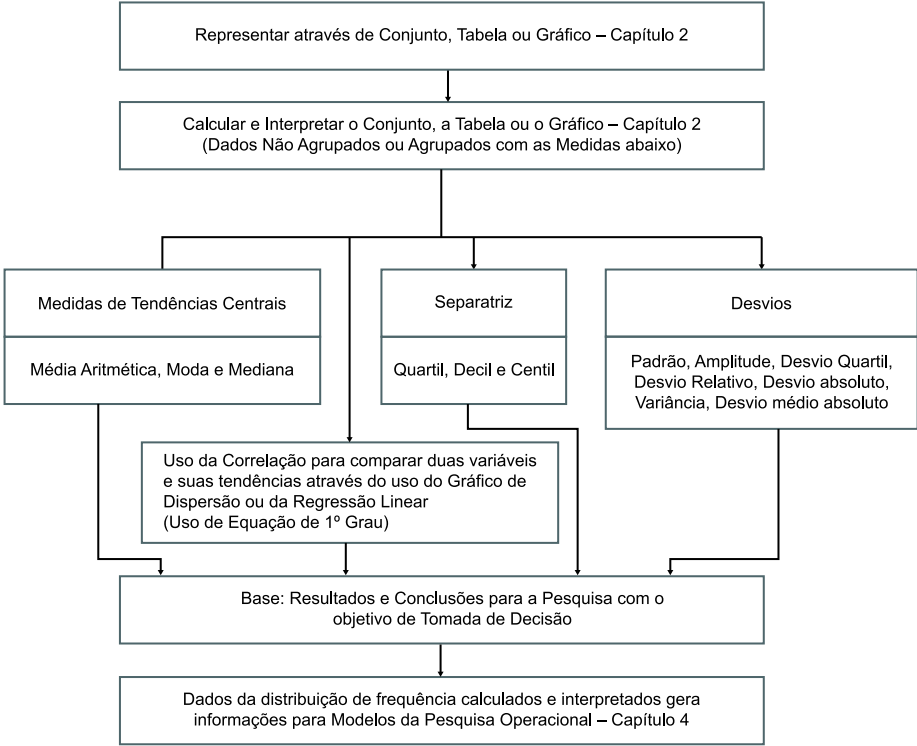
Vamos utilizar as informações da Tabela 7, fazendo a substituição na fórmula: $a = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$. Assim, tem-se $(10 \times 440) - (63 \times 67)/(10 \times 439) - (63)^2 = 179/421 = 0,4252$

Como $\left(\bar{x} = \frac{\sum x_i}{n}\right) = 63/10 = 6,3$ e $\left(\bar{y} = \frac{\sum y_i}{n}\right) = 67/10 = 6,7$. Depois, substitui-se nessa fórmula $b = \bar{y} - a\bar{x}$. E então, tem-se $b = 6,7 - (0,4252 \times 6,3) = 4,02$. Assim, temos $a = 0,4252$ e $b = 4,02$. Dispomos da seguinte equação $b = \bar{y} - a\bar{x}$, onde teremos $b = 6,7 - 0,4252X$. Essa equação é importante para traçamos pontos futuros dentro do próprio gráfico 1, simulando pontos de X . Por exemplo, se $X = 0$, teremos $b = 6,7 - (0,4252 \times 0)$. Então, pontos que poderiam ser colocados no Gráfico 1 seriam $(0;6,7)$, e é possível realizar outras simulações para interpretar as notas dos alunos.

No capítulo 4, falaremos da Pesquisa Operacional, que corresponde a uma das áreas em que se pode aplicar a Estatística a partir do uso de equações do 1º Grau ou Matrizes para a tomada de decisão.

A Figura 25 mostra de forma sintética que, após estarmos com os gráficos e as tabelas da pesquisa de distribuição de frequência prontas, é necessário aprofundar os dados para as interpretações e tomadas de decisões. Assim, é possível usar as Medidas de Tendências Centrais, Separatrizes e alguns desvios.

Figura 25 – Resumo síntese de alguns cálculos que podem ser realizados na Distribuição de frequência



Fonte: Elaboração própria, 2024.

3.4 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS

1 – Um Condomínio tem 80 apartamentos e necessita economizar energia para ser mais sustentável no uso de seus recursos. Por isso a companhia de energia elétrica, junto com você que é o síndico, necessita avaliar:

- a) Média aritmética
- b) Moda
- c) Mediana
- d) Assimetria
- e) Quartil 1
- f) Quartil 3
- g) Desvio quartil
- h) Desvio padrão

Tabela 1 – Consumo de Energia em (Kw/h) – Condomínio Paraíso-2024

Consumo (Kw/h)	Número de usuário (fi)
5 25	4
25 45	6
45 65	14
65 85	26
85 105	14
105 125	8
125 145	6
145 165	2
Total(Σ)	80

Fonte: Companhia de Energia Elétrica, 2024,

Faça um **quiz** das seis perguntas acima no **link**: <https://www.goconqr.com/pt-BR/quiz/39648798/3-medidas-de-tendencia-central-separatriz-e-principais-desvios-estatisticos-capitulo-3-estatistica-passo-a-passo>



Resposta:

Consumo (Kw/h)	Número de usuários (fi)	FACA	LI	LS	xi(LS + LI/2)	Xi × fi
5 ┤ 25	4	4	5	25	15	60
25 ┤ 45	6	10	25	45	35	210
45 ┤ 65	14	24	45	65	55	770
65 ┤ 85MD	26	50	65	85	75	1950
85 ┤ 105	14	64	85	105	95	1330
105 ┤ 125	8	72	105	125	115	920
125 ┤ 145	6	78	125	145	135	810
145 ┤ 165	2	80	145	165	155	310
Total	80					6360
Valor da Média aritmética					6360/80	79,5

• Média aritmética = **79,5 kw/h**

• Cálculo Mediana:

$Emd = 80/2 = 40$

$Md = 65 + [20(40 - 24)/26] = 77,3 \text{ kw/h}$

• Cálculo da Moda:

$Mo = 65 + 20[12/(12 + 12)]$

$1 = 26 - 14 = 12$

$2 = 26 - 14 = 12$

Mo = 75 Kw/h

• Assimetria = Média aritmética > Mediana > Moda, ou seja, $79,5 > 77,3 > 75$
(Assimetria positiva)

Consumo (Kw/h)	Número de usuários (fi)	FACA
5 25	4	4
25 45	6	10
45 65	14	24
65 85MD	26	50
85 105	14	64
105 125	8	72
125 145	6	78
145 165	2	80
Total	80	

- Quartil 1 = $E_q = 80/4 = 20$
- $Q_1 = 45 + 20(20 - 10)/14 = 59,28 \text{ Kw/h}$
- Quartil 3 = $(80 \times 3)/4 = 60$
- $Q_3 = 85 + 20(60 - 50)/14 = 99,29$
- Desvio quartil

$$D_q = (Q_3 - Q_1)/2 = (99,29 - 59,28)/2 = 20,005$$

$$\text{Mediana } \pm d_q =$$

$$77,3 - 20 = 57,3$$

$$77,3 + 20 = 97,3$$

Intervalo que 50% dos condôminos usam de energia é [57,3 kw/h; 97,3 kw/h]

- Desvio padrão

Consumo (Kw/h)	Número de usuário (fi)	LI	LS	xi(LS + LI/2)	(Xi - Média) ² .fi
5 ┤ 25	4	5	25	15	1664,1
25 ┤ 45	6	25	45	35	881,5
45 ┤ 65	14	45	65	55	8403,5
65 ┤ 85MD	26	65	85	75	526,5
85 ┤ 105	14	85	105	95	3363,5
105 ┤ 125	8	105	125	115	10082
125 ┤ 145	6	125	145	135	18481,5
145 ┤ 165	2	145	165	155	11400
Total	80				
Valor da Média aritmética					80780

Desvio padrão = 31,98 kw/h

Média aritmética +/- desvio padrão = [47,52 kw/h;31,98 kw/h]

2 – Uma empresa que vender um conjunto residencial com 400 lotes. Os lotes com tamanho acima de 900 m² serão mantidos de forma sustentável de forma a manter a floresta nativa da região. Já nos lotes entre 800 e 900 m², serão feitas áreas de lazer como, por exemplo, quadra esportiva, piscina, academia, entre outros. Após realizar o dimensionamento dos lotes, os mesmos ficaram conforme tabela de distribuição agrupada abaixo:

Tabela 2 – Quantidade de Lotes (m²) – Condomínio Residencial-2024

Áreas (m ²)	Nº de lotes (fi)
300 400	14
400 500	46
500 600	58
600 700	76
700 800	68
800 900	62
900 1000	48
1000 1100	22
1100 1200	6
TOTAL (Σ)	400

Fonte: Elaboração própria, 2024.

Pergunta-se:

- O valor da média aritmética;
- O valor da variância;
- O valor do desvio padrão;
- O valor do erro padrão.

Faça um quiz das quatro perguntas acima no link: <https://www.goconqr.com/pt-BR/quiz/39648798/3-medidas-de-tendencia-central-separatriz-e-principais-desvios-estatisticos-capitulo-3-estatistica-passo-a-passo>



Resposta:

Áreas (m²)	Nº de lotes (fi)	XI	XI*FI	XI - X	DESVIO²	DESVIO² * FI
300 ┆ 400	14	350	4900	-365,50	133590,25	1870263,50
400 ┆ 500	46	450	20700	-265,50	70490,25	3242551,50
500 ┆ 600	58	550	31900	-165,50	27390,25	1588634,50
600 ┆ 700	76	650	49400	-65,50	4290,25	326059
700 ┆ 800	68	750	51000	34,50	1190,25	80937
800 ┆ 900	62	850	52700	134,50	18090,25	1121595,50
900 ┆ 1000	48	950	45600	234,50	54990,25	2639532
1000 ┆ 1100	22	1050	23100	334,50	111890,25	2461585,50
1100 ┆ 1200	6	1150	6900	434,50	188790,25	1132741,50
TOTAL (Σ)	400		286200		610712,25	14463900

Média de m² Lote	715,50	Xi * FI / N	
Variância	1526,78	Desvio²/ FI	
Desvio Padrão	39,07	Raiz da Variância	
erro padrão	4,10	Dp/Raiz N	

4. A PESQUISA OPERACIONAL E OS SEUS MODELOS ESTATÍSTICOS

OBJETIVOS:

- Compreender o conceito e a importância da Pesquisa Operacional para a tomada de decisão com o uso de modelos.
- Identificar alguns instrumentos presentes na Pesquisa Operacional.
- Elaborar a modelagem estatística de maximização e de minimização da Programação Linear (um dos instrumentos da Pesquisa Operacional) e sua aplicação de forma prática.
- Mostrar cálculo da Programação Linear de maximização e de minimização de forma manual e com o uso do *Software* LINDO.
- Mostrar a relação do modelo de Programação Linear de Minimização com PERT e CPM; e o uso no programa *Gephi*.

CONTEÚDO:

4 A PESQUISA OPERACIONAL E OS SEUS MODELOS ESTATÍSTICOS

4.1 CONCEITO E IMPORTÂNCIA DA PESQUISA OPERACIONAL

4.2 PROGRAMAÇÃO LINEAR (PL) E SUA MODELAGEM

4.2.1 Modelo da Programação Linear

4.2.2 Cálculo da Programação Linear de forma Tabular pelo SIMPLEX

4.2.3 Cálculo da Programação Linear por gráfico

4.2.4 Cálculo da Programação Linear com o uso da Técnica do Transporte

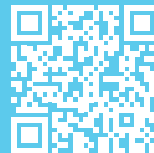
4.2.5 Uso do *software* LINDO para Programação Linear

4.3 TRANSPORTE AVANÇADO E SUA RELAÇÃO COM PERT E CPM

4.4 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS

Explicação Objetivos e Conteúdos do Capítulo 4

Link: <https://youtu.be/ZqC2xoHym2g>



4.1 CONCEITO E IMPORTÂNCIA DA PESQUISA OPERACIONAL

A Pesquisa Operacional (PO) é um ramo da ciência administrativa, científica e tecnológica que tem como objetivo tomar decisões baseadas em modelos matemáticos e estatísticos, a partir de análises quantitativas e qualitativas das informações e do conhecimento existente com o uso de equipes interdisciplinares. Hoje, a Pesquisa Operacional está relacionada a modelos que buscam otimizar sistemas numéricos, com o uso de *softwares* e *hardwares* que apoiam o pesquisador na coleta e organização dos dados e nos cálculos e simulações ótimas dentro dos padrões da realidade, para que se chegue a melhor interpretação e tomada de decisão.

A PO teve seu maior crescimento no final da Segunda Guerra Mundial, quando equipes de pesquisadores (principalmente ingleses e americanos) procuraram desenvolver métodos para

resolver determinados problemas de operações militares. Antes, ela tinha um enfoque quantitativo.

Os estudos ingleses focavam em casos ou problemas específicos. Já os trabalhos americanos abordaram modelos e métodos matemáticos em diversos temas como, por exemplo, teoria de estoques, transporte, substituições de equipamentos, teoria de filas, programação (*scheduling*) de tarefas, teoria de jogos, fluxos em redes e otimização linear.

Somente a partir de 1947 o uso da PO passou a ter um enfoque quantitativo e qualitativo, graças ao seu uso por uma equipe de especialistas em diversas áreas, chefiada pelo físico Blackett, para analisar os usos de radares. A partir daí houve a expansão e o uso da PO em diversos países, empresas públicas e privadas, assim como em vários setores.

Os sucessos das aplicações da PO ocorreram devido às habilidades das equipes interdisciplinares em relação à análise e avaliação de dados, ao estabelecimento das relações quantitativas e à constituição de hipóteses testáveis de forma a construir a ligação entre a teoria e a prática, observando sempre os recursos existentes.

As primeiras pesquisas brasileiras sobre PO ocorreram no ano de 1968 no Instituto Tecnológico de Aeronáutica (ITA), em São José dos Campos.

Filme sobre o uso e importância da Pesquisa Operacional



Link: <https://youtu.be/Kq9gJ1vcGtA>

Curiosidade sobre a Pesquisa Operacional

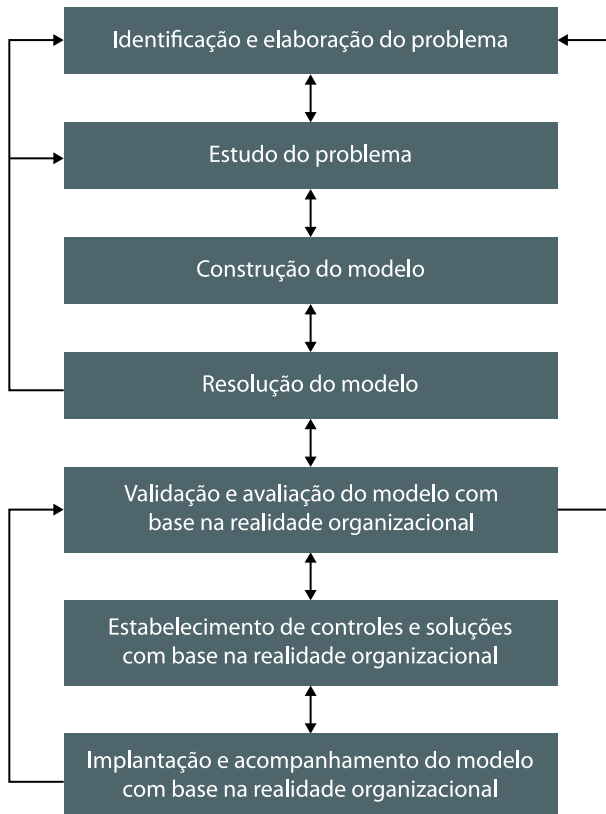
O termo pesquisa operacional é uma tradução (brasileira) direta do termo em inglês *Operational Research*, que em Portugal foi traduzido por investigação operacional e nos países de língua hispânica, por investigação operativa. O surgimento deste termo está ligado à invenção do radar na Inglaterra em 1934. Dois anos depois, o ministério britânico da Aviação criou a Estação de Pesquisa Manor Bawdsey, em Suffolk, para estudar como a tecnologia do radar poderia ser utilizada para interceptar aviões inimigos. O termo pesquisa operacional é atribuído ao superintendente da estação AP Rowe, que, em 1938, coordenou equipes para examinar a eficiência de técnicas de operações advindas de experimentos com interceptação de radar. Em 1941 foi inaugurada a Seção de Pesquisa Operacional do Comando da Força Aérea de Combate, com equipes envolvidas em problemas de operações de guerra, como manutenção e inspeção de aviões, escolha do tipo de avião para uma missão e melhoria na probabilidade de destruição de submarinos. Outros problemas incluíam controle de artilharia antiaérea e dimensionamento de comboios de frota. A análise científica do uso operacional de recursos militares de maneira sistemática foi iniciada na Segunda Guerra Mundial. Após o final da guerra, uma pesquisa operacional evoluiu rapidamente na Inglaterra e nos Estados Unidos. Em 1947, foi implantado o projeto SCOOP (*Scientific Computation of Optimal Programs*) no Pentágono, com o objetivo de apoiar decisões de operações na força aérea americana. O projeto continha um grupo de pesquisa coordenado pelo economista Marshall Wood e pelo matemático George Dantzig. Durante esse projeto, Dantzig desenvolveu, formalizou e testou o método SIMPLEX para resolver problemas de programação linear (otimização linear). Esses desenvolvimentos se basearam em trabalhos precursores, notadamente do matemático russo Leonid Kantorovich, e a programação termo linear foi sugerida a Dantzig pelo economista TC Koopmans (tanto Koopmans quanto Kantorovich ganharam posteriormente o prêmio Nobel de economia de 1975, pelas contribuições à teoria de alocação ótima de recursos). Em 1952, foi fundada a sociedade científica americana de pesquisa operacional (ORSA – *Operations Research Society of America*) e, em 1953, a sociedade inglesa de pesquisa operacional (ORS – *Operational Research Society*) e a americana de ciências de administração (TIMS – *The Institute de Ciências de Gestão*). Em 1957 foi realizada a primeira conferência internacional, como informa, por exemplo, Kirby (2003). O trabalho de 1939 de Kantorovich (intitulado “Métodos matemáticos na organização e no planejamento de produção”) é considerado um dos precursores da pesquisa operacional, ao apontar que uma das maneiras de se aumentar a eficiência em uma empresa seria por meio de melhorias na organização e no planejamento da produção (Arenales, 2007, p. 1-3).

Dessa forma, os modelos da Pesquisa Operacional mais antigos têm enfoque quantitativo, com os modernos de cunho quali quanti tendo sete etapas (Figura 26). A primeira é a Identificação do Problema (já visto no capítulo 1), em que se irá definir os objetivos e os caminhos a serem alcançados para que o mesmo seja resolvido ou minimizado. Depois vem o Estudo do Problema, que corresponde à etapa em que há coleta de informações relacionadas ao problema com o máximo de precisão. O terceiro passo é a Construção do Modelo: aqui, há a elaboração da representação dos elementos e dados capazes de descrever os aspectos mais importantes da situação e das interações que o modelo estatístico expressa, devendo ser fidedigno à realidade. A quarta etapa é resolver os modelos elaborados na etapa anterior através de cálculos, e se não ocorre a possibilidade de resolução do modelo ou da sua solução, será considerado não viável e então é necessário retornar à etapa 2 ou à etapa 1, conforme setas da Figura 26. Caso a etapa 4 tenha resultado, é necessário validá-lo, ou seja, deve-se testar o resultado do modelo para verificar se os resultados oriundos dos dados e das informações são adequados à realidade da empresa ou ao contexto da pesquisa. Se a solução não for viável, será necessário definir e/ou construir um novo modelo.

Durante a elaboração e validação do modelo (Figura 26), é necessário estabelecer controles e experimentações com o objetivo de identificar parâmetros fundamentais para a solução do problema. Qualquer mudança nesses parâmetros deverá ser controlada para garantir sua validade. Caso alguns desses parâmetros sofram desvios além do permitido, poderá ser necessário o cálculo de um novo controle ou a reformulação do modelo.

Estando todas as etapas cumpridas da Figura 26, pode-se iniciar a implantação e acompanhamento do modelo com base na realidade organizacional e no contexto pesquisado.

Figura 26 – Etapas da tomada de decisão na PO



Fonte: Pereira (2010, p. 21).

O Quadro 13 mostra o conceito e o uso de algumas técnicas inerentes à PO. Podemos citar (Quadro 13) a Programação Linear, a Programação Inteira, a Programação dinâmica, a Teoria dos Grafos, a Teoria da Fila, e a Simulação e Teoria dos Jogos.

A Pesquisa Operacional pode ser usada, por exemplo, em:

- Empresas que trabalham com entrega de produtos para decidir os melhores itinerários a serem percorridos por carteiros, entregadores e transportadoras, principalmente

com o aumento do comércio eletrônico e a criação de empresas que não utilizam espaços físicos, e sim virtuais;

- Aplicativos de entrega *delivery*;
- Investimentos financeiros;
- Mistura de matéria-prima para elaborar um produto;
- Cálculo de filas;
- Estudo de Localização de empresas, seja da área industrial (proximidade fornecedor e matéria prima), agropecuária (infraestrutura para escoamento de produção) ou de serviço (proximidade ao cliente);
- Alocação de mão de obra e turnos de trabalho na empresa.

Quadro 13 – Principais técnicas da PO utilizadas para na tomada de decisão

Técnica	Conceito	Alguns exemplos do uso da Pesquisa Operacional
Programação linear	É uma técnica de programação matemática em que se procura maximizar ou minimizar uma função, sujeitando-a a certas limitações ou restrições de recursos presentes na empresa.	Localização de empreendimentos, logística interna e externa, composição de produtos a serem fabricados com base nas limitações do mercado e dos equipamentos disponíveis, avaliação de desempenho de funcionários, avaliação de carteira de investimentos.
Programação inteira	É a programação matemática em que uma ou mais variáveis de decisão têm valores inteiros (programação inteira total) e outras variáveis de escolha podem ter valores contínuos (programação inteira mista).	Trabalha-se com modelos de aquisição de equipamentos, nos quais o resultado do modelo somente pode ser em números inteiros.
Programação dinâmica	Ao contrário da programação linear, em que as soluções são estáticas, a programação dinâmica procura verificar a trajetória temporal das variáveis dentro de determinado período.	Alocar recursos escassos entre diversos empreendimentos, efetuar um planejamento otimizado em diversos períodos, com restrições específicas para cada um e com disponibilidade de recursos para os globais. É muito utilizado para fornecer uma maneira de projetar algoritmos que sistematicamente exploram todas as possibilidades (corretude). Armazena resultados a fim de evitar computação redundante (eficiência).

continua...

conclusão

Técnica	Conceito	Alguns exemplos do uso da Pesquisa Operacional
Teoria dos grafos	É um ramo da matemática que estuda as relações entre os objetos de um determinado conjunto.	Planejamento de projetos em que são identificadas diversas ações que devem ser feitas dentro de um tempo considerado, além de analisar caminhos para percursos na área logística. Utilizada para otimizar problemas de fluxo máximo (encontrar o maior percurso) e achar o caminho mais curto (cálculo do menor percurso), seja para veículos ou pessoas. Muito utilizado em aplicativos de entrega de mercadoria ou deslocamento de pessoas.
Teoria da fila	É um setor da Pesquisa Operacional que utiliza conceitos básicos de processos estocásticos e de matemática aplicada para analisar o fenômeno de formação de filas e suas características.	Cálculo do tempo em filas de banco, supermercados ou qualquer outro estabelecimento ou espera em chamadas telefônicas. Foi desenvolvida com a finalidade de prever o comportamento das filas de modo a permitir o dimensionamento adequado de instalações, equipamentos e sua infraestrutura. Muito usado também para problemas de congestionamento de sistema, com o objetivo de diminuição do tempo de espera, fila em serviços e em transporte público.
Simulação	É a técnica de solução de um problema pela análise de um modelo que descreve o comportamento do sistema. Para isso, podem ser usados computadores digitais.	O aumento da produção de uma empresa, na qual será testado (simulado) o funcionamento de algumas novas máquinas, com o objetivo de verificar a necessidade de outros equipamentos ou recursos. O teste de medicamentos ou vacina em grupo de pessoas para comparar os sintomas entre um grupo de controle e o grupo teste.
Teoria dos jogos	É um problema de otimização em que as partes envolvidas usam estratégias para cada uma convencer o seu oponente.	Negociação econômica e política internacional.

Fonte: Pereira (2010, p. 23-24).

Dica de filme em que se usa a Pesquisa Operacional:

Estrelas além do tempo: O período é a corrida espacial. Estados Unidos e Rússia vivem o momento decisivo para saber qual das duas nações finalmente vai decolar em direção ao espaço. Um grupo de mulheres afro-americanas se mostrou essencial na busca por desvendar os cálculos e equações que tornaram possível a vitória norte-americana nessa corrida.



Link Trailer: <https://youtu.be/wx3PVtrU-Os>

Uma mente brilhante: Esse filme é baseado na vida de John Forbes Nash Jr., matemático que trabalhou com a teoria dos jogos e geometria diferencial, entre outros conceitos, e se tornou reconhecido na área. Em 1994, o matemático compartilhou o Prêmio de Ciências Econômicas em Memória de Alfred Nobel com Reinhard Selten e John Harsanyi.



Link Trailer: <https://youtu.be/q8vUMD1f0ss?t=2>

O homem que viu o infinito: Traz a história do matemático indiano Srinivasa Ramanujan, um autodidata em técnicas de PO em simulações e que tinha dificuldades em ingressar na universidade. Isso porque ele era obcecado pelo estudo da matemática. Aos 25 anos, conhece o professor Hardy, que acredita no jovem e faz de tudo para que ele vá estudar em Cambridge, onde passam a trabalhar com afinco pelo reconhecimento das ideias inovadoras do jovem.



Link trailer: <https://youtu.be/mfyAyfGrA4A>

Como o estudo na área de Pesquisa Operacional é multidisciplinar e há várias técnicas, optamos em trabalhar nos próximos tópicos com Programação linear e Teoria dos Grafos.

4.2 PROGRAMAÇÃO LINEAR (PL) E SUA MODELAGEM

A PL é um conjunto de técnicas matemáticas e estatísticas com as quais pode ser determinada uma solução ótima para problemas que apresentam várias soluções possíveis, e indica um método iterativo que determina a melhor combinação de valores que as variáveis do modelo devem assumir a fim de otimizar ela, com cada variável obedecendo certas restrições. Qualquer modelagem de PL deve ter três partes, que podem estar no formato de equação, de inequação ou de matriz (Figura 27).

- Na Função objetiva (FO) – (Figura 27) estará sempre, em um dos termos, o de maximizar (Max) ou de minimizar (Min). Essa parte da PL encontrará às variáveis de decisão, que são aquelas que medem a quantidade dos distintos recursos da empresa ou do evento que está sendo analisado, como funcionários, dinheiro, número de clientes, capacidade de produção, dentre outros. A FO é responsável por determinar a melhor forma de utilizar os recursos, que são limitados durante a análise empresarial.
- As Restrições técnicas (RT) – (Figura 27) estão relacionadas aos recursos que a empresa tem à sua disposição. Dessa forma, esses são fatores limitados e finitos ao serem definidos dentro da Restrição Técnica. Esse recurso pode ser quantidade de clientes, espaço de estocagem, quantidade de funcionários, dentre outros.
- As Restrição de não negatividade (RN) – (Figura 27) relata que os valores das variáveis não podem ser nunca negativos, já que as variáveis estão presentes na FO e devem ser todas identificadas dentro da Restrição de Não negatividade.

Assim, todos os resultados das variáveis encontrados em FO não podem ser negativos.

Figura 27 – Representação dos modelos da PL

PARTES DA PL	PADRÃO (EQUAÇÃO)	CANÔNICO (INEQUAÇÃO)	MATRICIAL (MATRIZ)
Função objetiva	Otimizar (máx./mín.) = $C \cdot X$ $\text{máx. (mín.)} \cdot Z = C_1 X_1 + C_2 X_2 + \dots + C_n X_n$	Otimizar (máx./mín.) = $C \cdot X$ $\text{máx. (mín.)} \cdot Z = C_1 X_1 + C_2 X_2 + \dots + C_n X_n$	Otimizar (máx./mín.) = $C \cdot X$ $[C_1 \ C_2 \ \dots \ C_n] \cdot \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$
Restrições técnicas (limitações dos recursos empresariais)	$A \cdot X = B$ $A_{11} X_1 + A_{12} X_2 + \dots + A_{1n} X_n = B_1$ $A_{21} X_1 + A_{22} X_2 + \dots + A_{2n} X_n = B_2$ $\dots$ $A_{m1} X_1 + A_{m2} X_2 + \dots + A_{mn} X_n = B_m$	$A \cdot X \leq/\geq B$ $A_{11} X_1 + A_{12} X_2 + \dots + A_{1n} X_n \leq/\geq B_1$ $A_{21} X_1 + A_{22} X_2 + \dots + A_{2n} X_n \leq/\geq B_2$ $\dots$ $A_{m1} X_1 + A_{m2} X_2 + \dots + A_{mn} X_n \leq/\geq B_m$	$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{bmatrix} \cdot \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \leq/\geq/ = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix}$
Restrições de não negatividade	$X \geq 0; B \geq 0$	$X \geq 0; B \geq 0$	$X \geq 0; B \geq 0$

$X \in \mathbb{R}$

X = variáveis de decisão

A = coeficientes das variáveis

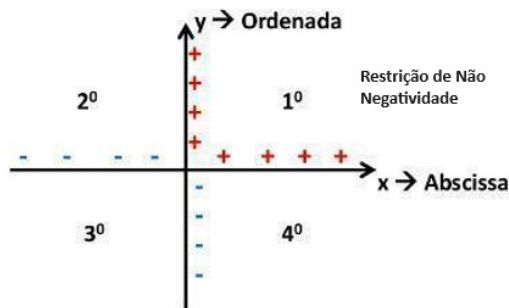
B = termo independente que representa os recursos disponíveis

C = valor relacionado à variável de decisão

Fonte: Pereira (2010, p. 33).

Devido à Restrição de Não Negatividade, as respostas das variáveis da Função Objetiva devem estar dentro lado direito e superior da Figura 28, ou seja, presente no primeiro Quadrante, onde valor da abscissa (X) e da Ordenada (Y) são positivos.

Figura 28 – Quadrante dos resultados das variáveis de não negatividade



Fonte: Elaboração própria, 2024.

Algumas palavras chave da área de Pesquisa Operacional que veremos nesse capítulo:

- **Programação linear:** Técnica matemática destinada a determinar a melhor utilização de recursos limitados, de forma que uma função-objetivo seja otimizada e sejam satisfeitas determinadas condições estabelecidas.
- **Modelagem:** É a atividade de construir modelos que expliquem as características ou o comportamento de uma organização.
- **Modelagem Padrão da PL:** Utilizado quando a modelagem possui equação. Nesse modelo, há a presença do sinal de igualdade (=).
- **Modelagem Canônico da PL:** Utilizado quando a modelagem possui inequações. Nesse modelo, há a presença dos sinais de maior ou igual ($\leq$) e de menor ou igual ($\geq$).
- **Modelagem Matricial da PL:** Utilizado quando a modelagem possui matrizes. Nesse modelo, há a presença dos sinais (=), ($\leq$) e ($\geq$).
- **Variáveis de decisão:** Variáveis que correspondem às decisões a serem tomadas, visando encontrar a solução para o problema em análise.
- **Função-objetivo:** Expressão matemática por meio da qual se relacionam as variáveis de decisão e o objetivo a ser atingido.

- **Método SIMPLEX:** Método matemático que combina conceitos de álgebra matricial com um conjunto de regras básicas que levam à solução dos problemas de PL.
- **Restrições:** Limitações impostas sobre possíveis valores que podem ser assumidos pelas variáveis de decisão.
- **Preço sobra (ou sobra de preço):** Campo de relatório de sensibilidade que indica quanto se deixa de ganhar por não se dispor de mais uma unidade de determinada variável restritiva.
- **Reduzido custo:** Informação contida no relatório de sensibilidade, que pode ser entendida como “custo de oportunidade”, indicando quanto se deixa de ganhar (ou perder) ao se adotar uma alternativa diferente da solução ótima indicada.
- **Análise de sensibilidade:** Análise que permite avaliar os reflexos sobre a solução ótima de eventuais alterações nos dados do problema, dentro de intervalos definidos.
- **Variáveis de folga:** Variáveis que representam eventuais folgas entre os limites fixados pelas restrições e o montante efetivamente definido na solução do problema. Introduzidas nas inequações representativas das restrições, transformam-nas em equações.
- **Método SIMPLEX:** Método matemático que combina conceitos de álgebra matricial com um conjunto de regras básicas que levam à solução dos problemas de PL.
- **Ambiente de risco:** Ambiente em que o tomador de decisão é capaz de atribuir probabilidade de ocorrência dos cenários (ocorrência futuro).
- **Solução:** Qualquer especificação de valores, dentro do domínio da função objetiva, para as variáveis de decisão, independentemente de se tratar de uma escolha desejável ou permissível.

- **Solução viável:** Uma solução em que todas as restrições são satisfeitas.
- **Solução ótima:** Uma solução viável que tem o valor mais favorável da função-objetivo, o $f(X)$. Isto é, maximiza ou minimiza a função-objetivo, podendo ser única ou não.

Após toda a explicação teórica acima, vamos iniciar a Modelagem da Programação Linear e depois realizar os cálculos através do SIMPLEX tabular, com o uso de equação ou inequação.

4.2.1 Modelo da Programação Linear

Para elaborar o Modelo e depois realizar seu cálculo através do SIMPLEX, optamos em apresentar um estudo de caso, no qual o pesquisador adquiriu os dados através de fontes secundárias ou utilizando como instrumento a observação para montar a sua modelagem.

O texto seguinte é apresentado em cores diferentes para facilitar a compreensão da explicação dos dados associados à elaboração da Modelagem da PL, que corresponde ao primeiro passo antes de se iniciar o cálculo usando a técnica do SIMPLEX.

Calça com bolso antifurto de celular é sucesso de inovação e de vendas

Adaptado – 19/02/2017 – Pequenas Empresas & Grande Negócios

Usar um martelo na fabricação de uma calça jeans foi a inovação de uma empresa que está há mais de 30 anos no mercado. A fábrica de calças de São Paulo investiu em um modelo que tem um bolso quase invisível para guardar o celular.

Não foi fácil lançar a calça com bolso para celular. O pai, que comanda a fábrica desde 1986, era contra. E os filhos tiveram que tocar o projeto sem ele saber. “Fizemos dez modelos meio escondido. Em uma semana vendemos seis mil peças, e aí ele acreditou”, conta o filho do empresário.

Uma calça comum na empresa tem um lucro unitário de R\$ 120,00, já a calça com o bolso invisível tem um lucro unitário de R\$ 150,00.

A maior dificuldade da empresa é que o custo da mão de obra da calça de bolso invisível é maior do que a calça comum, pois a calça comum utiliza apenas dois funcionários em seu processo produtivo e a com o bolso invisível utiliza quatro empregados, não podendo ultrapassar 100 funcionários. A empresa atualmente tem um espaço de estoque para a calça comum de 3 m² e a calça com bolso invisível tem um espaço de 2 m², sendo que o espaço total para ambas as calças é de até 90 m² da loja das vendas e da fábrica onde são produzidas as calças. A calça com bolso invisível utiliza durante o processo de fabricação como ferramenta até três equipamentos, enquanto a calça comum utiliza cinco, não podendo ultrapassar 120.

A ideia patenteou o projeto da calça de bolso invisível e aproveitaram a mesma estrutura de fabricação da indústria têxtil da família. “A calça é um sucesso de vendas”. Qual é o lucro máximo que a calça comum e a invisível poderão dar à empresa familiar por dia de trabalho?

Vamos acompanhar agora a resolução desse estudo empresarial, etapa a etapa, conforme sugere Pereira (2010).

1º passo – Identifique as variáveis de decisão, que sempre estarão relacionadas à palavra “máximo” ou “mínimo”. Nesse caso, é possível identificá-las na última frase do texto, em cor azul: “[...] o lucro máximo que a calça comum e a invisível poderão dar a empresa familiar?”. Para esse caso da empresa familiar, a decisão está relacionada ao lucro e conseqüentemente à Maximização da venda da calça comum e da calça do bolso invisível. Sendo assim, ambos são as variáveis de decisão e chamaremos de X_1 a calça comum e de X_2 a calça do bolso invisível. Observe a palavra “máximo” ou “mínimo” estarão sempre ligadas às variáveis de decisões. Então, as variáveis de decisão serão denominadas X_1 e X_2 :

- X_1 = Calça Comum
- X_2 = Calça do Bolso Invisível

2º passo – Após estarem definidas as variáveis de decisão (X_1 e X_2), é hora de identificar a função objetiva. Ela sempre está associada às variáveis de decisão e à palavra “máximo” ou “mínimo”. Podemos chegar à função objetiva por meio das informações presentes em dois trechos na cor azul no texto:

Texto 1: “[...] Uma calça comum na empresa tem um lucro unitário de R\$ 120,00 já a calça com o bolso invisível tem um lucro unitário de R\$ 150,00”.

Texto 2: “[...] Qual é o lucro máximo que a calça comum e a invisível poderão dar a empresa familiar?”

Dessa forma, a função objetiva, que corresponde a essa etapa do SIMPLEX a ser montada no modelo, visa maximizar o lucro unitário das calças comuns e invisível.

Quadro 14 – Representação da função objetiva

Variáveis do Processo de Decisão	X_1 (Calça Comum)	X_2 (Calça do Bolso Invisível)
Lucro unitário	R\$ 120,00	R\$ 150,00

Fonte: Elaboração própria, 2024.

A função objetiva para o caso da empresa familiar é, portanto, o lucro máximo a ser obtido com a venda de calça comum e calça do bolso invisível, representado por “máx(L)”. A representação do modelo matemático da Função Objetiva (FO) é: $\text{máx}(L) = 120X_1 + 150X_2$.

3º passo – Identifique os recursos finitos que a empresa tem à disposição. Esses são importantes para que você chegue à restrição técnica da PL. Nesse caso, a limitação dos recursos, ou seja, as Restrições Técnicas da Programação Linear, estão sendo identificadas no texto em três cores distintas: Verde, Vermelho e Roxo, respectivamente. Para facilitar a visualização delas, que correspondem ao limite de recursos dentro da empresa familiar, iremos montar três Quadros (15,16 e 17) com cada Restrição Técnica e com a identificação dos seus limites de recursos, que chamaremos de variável independente e representada pela letra “ b_i ”. Somente quem tem b_i são as restrições Técnicas e de não negatividade; a Função objetiva não terá b_i (limite de recursos disponíveis).

- Primeira Restrição Técnica (RT_1) é referente à mão de obra para produção das calças. O Texto da Primeira Restrição Técnica (RT_1) é: “[...]A maior dificuldade da empresa é que o custo da mão de obra da calça de bolso invisível é maior do que a calça comum, pois a calça comum utiliza apenas dois funcionários em seu processo produtivo e a calça com o bolso invisível utiliza 4 empregados, não podendo ultrapassar 100 funcionários.”

Quadro 15 – Representação da Segunda Restrição Técnica (RT_1)

	X_1 (Calça Comum)	X_2 (Calça do Bolso Invisível)	b_i
RT_1	2	4	≤ 100 (quantidade de funcionários que não pode ser ultrapassada)

Fonte: Elaboração própria, 2024.

A representação do Modelo da Restrição Técnica (RT_1) – Quadro 16 referente à quantidade de funcionários que trabalham para a fabricação das calças comum e invisível (nesse caso, a segunda restrição técnica) é: $2X_1 + 4X_2 \leq 100$.

- Segunda Restrição Técnica (RT_2) é referente ao m^2 para estocar as calças: O Texto da Segunda Restrição Técnica (RT_2) é: “[...] A empresa atualmente tem um espaço de estoque para a calça comum de 3 m^2 e a calça com bolso invisível tem um espaço de 2 m^2 , sendo que o espaço total para ambas as calças é de até 90 m^2 da loja das vendas e da fábrica onde são produzidas as calças”.

Quadro 16 – Representação da Primeira Restrição Técnica (RT_2).

	X_1 (Calça Comum)	X_2 (Calça do Bolso Invisível)	b_i
RT_2	3	2	≤ 90 (área que não pode ser ultrapassada)

Fonte: Elaboração própria, 2024.

A representação do Modelo da Restrição Técnica (RT_2) – Quadro 15 referente ao espaço para armazenagem das calças comum como invisível (nesse caso, a primeira restrição técnica) é: $3X_1 + 2X_2 \leq 90$.

- Terceira Restrição Técnica é sobre a quantidade de ferramentas durante o processo de fabricação das calças. O Texto

da Terceira Restrição Técnica (RT_3) é : “[...]A calça com bolso invisível utiliza durante o processo de fabricação como ferramenta até três equipamentos, enquanto a calça comum utiliza cinco, não podendo ultrapassar 120.”

Quadro 17 – Representação da Terceira Restrição Técnica (RT_3)

	X_1 (Calça Comum)	X_2 (Calça do Bolso Invisível)	b_i
RT_3	5	3	≤ 120 (quantidade de equipamentos que não pode ser ultrapassada)

Fonte: Elaboração própria, 2024.

A representação do Modelo da Restrição Técnica (RT_3) – Quadro 17 referente à quantidade de equipamentos que trabalham para a fabricação das calças comum e invisível (nesse caso, a segunda restrição técnica) é: $5X_1 + 3X_2 \leq 120$.

Observação: Ao organizar e montar cada Restrição Técnica, é necessário utilizar a mesma unidade de medida para todas. Caso não esteja, é necessário transformar. Além disso, em cada Restrição todos devem estar com a mesma unidade. Veja que, no caso da empresa familiar, todas as unidades formam coletadas a partir de um dia de trabalho. Outros aspectos em que cada restrição deve ter o padrão da mesma unidade são, por exemplo, a RT_1 trabalha com m^2 , a RT_2 com quantidade de equipamentos em unidade e a RT_3 com quantidade de funcionários em unidade.

4º passo – Represente a restrição técnica de não negatividade. Ela significa que as variáveis de decisão sempre devem ser maiores ou iguais a zero, não podendo ser negativas. Caso contrário, X_1 (Calça Comum) e X_2 (Calça do Bolso Invisível) teriam uma quantificação negativa, o que, com a PL, não é possível, pois

essa é uma técnica que visa à otimização. A restrição técnica de não negatividade é representada da seguinte forma, nesse exemplo:

Quadro 18 – Representação da restrição de não negatividade

Variáveis do processo de decisão	b_i
X_1 (Calça Comum)	≥ 0
X_2 (Calça do Bolso Invisível)	≥ 0

Fonte: Elaboração própria, 2024.

Com base no Quadro 18, tem-se que a representação matemática da restrição técnica de não negatividade é: $X_1 \geq 0$ e $X_2 \geq 0$.

5º passo – Organize e represente o modelo de forma clara.

- Variáveis de decisões:
 - X_1 = calça comum
 - X_2 = Calça do Bolso Invisível
- Função objetiva: $Máx(L) = 120X_1 + 150X_2$
- Restrição Técnica (recursos limitados)
 - RT_1 (quantidade e funcionários em unidade):
 $2X_1 + 4X_2 \leq 100$
 - RT_2 (espaço em m^2): $3X_1 + 2X_2 \leq 90$
 - RT_3 (quantidade de Equipamentos em unidade):
 $5X_1 + 3X_2 \leq 120$
- Restrição de não negatividade: $X_1 \geq 0$ e $X_2 \geq 0$.

A partir da elaboração do modelo acima, que teve como problema “Qual é o lucro máximo que a calça comum e a invisível poderão dar a empresa familiar?”, iniciaremos o cálculo do SIMPLEX pelo método Tabular e mostraremos passo a passo no tópico 4.2.2.

Curiosidade

O SIMPLEX foi desenvolvido por Dantzig no ano de 1947, e seu objetivo era encontrar resultados ótimos para problemas militares. Atualmente, de acordo com Passos (2008), o SIMPLEX

“[...] é um método iterativo que, utilizando um algoritmo, procura maximizar ou minimizar (otimizar) uma função objetiva de um modelo matemático (modelo linear), considerando-se algumas restrições lineares. A cada iteração realizada, a função objetiva melhora a solução do problema, passando de um ponto extremo da região de soluções a outro, até chegar à solução ótima.”.

Filme explicando o uso do PL



Link: <https://youtu.be/zFSwyKMAiHI>

4.2.2 Cálculo da Programação Linear de forma Tabular pelo SIMPLEX

1º passo – Identifique, no modelo, a função objetiva pela palavra “máximo” ou “mínimo”. No caso da concessionária, buscamos calcular o máximo lucro: $\text{Max}(L) = 120X_1 + 150X_2$

2º passo – Passar as variáveis de decisão, com o sinal trocado, para o lado do $\text{Max}(L)$. Assim, tem-se: $\text{Max}(L) - 120X_1 - 150X_2$

Observação:

Se o caso fosse de minimizar (se o caso em estudo tratasse de custo em vez de lucro), seria necessário transformarmos a expressão em máximo. Multiplicaríamos toda a expressão por -1 . Então, ela ficaria: $\text{Máx}(Z) = -120X_1 - 150X_2$. Ao passarmos as variáveis para o outro lado, ficaria: $\text{Máx}(Z) + 120X_1 + 150X_2$.

3º passo – Identifique em FO o número relacionado à variável de decisão que possui o maior valor absoluto (maior valor, não importando o sinal), com o objetivo de detectar a variável de decisão que será responsável pela entrada na função objetiva.

Entrada



$$\text{FO: Max}(L) - 120X_1 - 150X_2$$

A entrada é em X_2 por ser o maior valor absoluto do X_1 . Então, ela será em $X_2 = 150$; ou seja, $150 > 120$. Nesse caso, apenas temos duas variáveis de decisões em FO, mas poderíamos ter mais. Caso isso ocorra, é necessário comparar todas e escolher o maior valor absoluto em relação às variáveis de decisões presentes apenas em FO.

4º passo – Identifique, nas restrições técnicas, o local onde haverá a **SAÍDA**, que é pelo menor valor não negativo. Para identificá-la, divida o b_i (variável independente) de cada restrição técnica pelos valores da variável que entra da restrição respectiva. O b_i é o valor presente na Restrição Técnica após o sinal ($\leq$, $\geq$, $=$, $<$ ou $>$) – Quadro 20.

Quadro 19 – Representação da saída

Restrição Técnica (RT)	Modelo de RT	b_i	X_2 (Entrada- Variável de decisão -3º passo)	b_i de cada restrição técnica/variável que entra de cada restrição técnica
RT_1 (Quantidade e funcionários em unidade):	$2X_1 + 4X_2 \leq 100$	100	4	$100/4 = 25 \rightarrow$ SAÍDA (o menor valor não negativo)
RT_2 (espaço em m²)	$3X_1 + 2X_2 \leq 90$	90	2	$90/2 = 45$
RT_3 (Quantidade de Equipamentos em unidade)	$5X_1 + 3X_2 \leq 120$	120	3	$120/3 = 40$

Fonte: Elaboração própria, 2024.

5º passo – Acrescente a variável de folga, para isso, avalie o sinal de cada restrição técnica:

- Se o sinal for o de menor ou igual ($\leq$), coloque variável de folga positiva (+ XF).
- Se o sinal for o de maior ou igual ($\geq$), coloque variável de folga negativa (– XF).
- Entretanto, se o sinal for de igualdade (=), não haverá variável de folga (XF), apenas variável auxiliar (variável colocada para auxiliar nos cálculos, representada pela letra “A”). Nesse caso, as inequações precisam ser transformadas em equações. Para tanto, basta trocar os sinais de maior ou igual ($\geq$) ou o de menor ou igual ($\leq$) por igual (=); Caso contrário, não será possível colocar as variáveis de folga (XF).
- É preciso transformar as inequações em equações trocando os sinais de $\geq/\leq$ por igual (=). Se isso não for feito, não será possível colocar as variáveis de folga.

No estudo acima sobre a Empresa Familiar que vende dois tipos de calça, a comum e a de bolso invisível, percebe-se que

todas as Restrições Técnicas tenham sinal de ($\leq$). Assim, em todas serão colocadas variável de folga positiva (+XF), que terão sua numeração em relação à posição da sua restrição técnica. Então, a RT_1 terá XF_1 , e assim sucessivamente nesse exemplo, ficando:

$$2X_1 + 4X_2 + XF_1 = 100 \quad (RT_1)$$

$$3X_1 + 2X_2 + XF_2 = 90 \quad (RT_2)$$

$$5X_1 + 3X_2 + XF_3 = 120 \quad (RT_3)$$

Reafirmamos que, no caso apresentado, a variável de folga (XF) foi positiva pois todas as restrições técnicas têm sinal de $\leq$. Observe também que cada variável de folga (XF) é representada com um número específico, com o objetivo de identificar a origem de cada uma delas no cálculo (à qual restrição técnica ela está relacionada). Assim, a primeira restrição técnica terá a variável de folga denominada de XF_1 ; a de XF_2 ; e a terceira, de XF_3 . Perceba que o número tem como referência a posição da variável no sistema ou modelo.

6º passo – Organize o sistema, destacando a entrada e a saída. No local em que a entrada e a saída são iguais, está o elemento pivô (no ponto comum entre elas). A linha em que esse elemento estiver será a linha pivô. A denominação “pivô” ocorre porque todos os cálculos futuros têm essa linha como base, então ela será “o pivô” de tudo. No quadro a seguir, Z é a variável responsável pelo valor da função objetiva.

Nesse caso, o elemento pivô é o número 4, pois é ele que está no ponto em que a entrada é igual à saída (ponto de encontro entre entrada e saída). A Linha que o Elemento pivô está é a linha que será pega para o 7º Passo seguinte.

Quadro 20 – Representação da entrada e da saída (elemento pivô)

Partes da PL	Z	X_1	X_2 (Entrada)	XF_1	XF_2	XF_3	b_i
FO	1	-120	-150	0	0	0	0
RT_1	0	2	4	1	0	0	100 (SAÍDA)
RT_2	0	3	2	0	1	0	90
RT_3	0	5	3	0		1	120

Fonte: Elaboração própria, 2024.

Observação: Ao organizar o sistema, perceba que a primeira linha será sempre a função objetiva (FO), e o Z correspondente a ela será sempre igual a 1. As outras linhas serão sempre as restrições técnicas (RT) e devem ser sempre zero na coluna de Z.

7º passo – Faça uma tabela começando pela linha pivô, que nesse caso é a linha correspondente a RT_1 . Em seguida, divida cada um dos valores da linha pivô pelo valor do elemento pivô. Com os resultados dessas divisões, você obterá uma nova linha pivô unitária. Nesse caso, ela é a 2ª linha (do quadro 20). Divida cada um de seus elementos pelo valor do elemento pivô (nesse caso, 4) e assim se obterá a nova linha pivô.

Quadro 21 – Cálculo da nova linha pivô

Partes da PL	Z	X_1	X_2 (Entrada)	XF_1	XF_2	XF_3	b_i
RT_1 (Linha Pivô – Quadro 20)	0	2	4	1	0	0	100
Divisão por 4 (elemento pivô), para chegarmos à nova linha pivô.	0	0,5	1	0,25	0	0	25 (SAÍDA)

Fonte: Elaboração própria, 2024.

Observação: Esse valor de b_i deve ser igual ao valor da saída obtido no quarto passo.

8º passo – Calcule as novas linhas com base na nova linha pivô. Para chegar a elas, siga esses procedimentos:

a) Multiplique cada elemento da nova linha pivô pelo valor da variável de entrada (em verde, no Quadro 20) de cada linha: para chegar à 1ª nova linha, multiplique pelo valor da entrada da 1ª linha; para chegar à 2ª nova linha, multiplique pelo valor de entrada da 2ª linha; e assim por diante. Todavia, o valor da entrada deve ter o sinal trocado.

b) Some os elementos encontrados no procedimento a), termo a termo, com os elementos da 1ª linha (do Quadro 20).

c) Repita esses procedimentos (a e b) com as outras linhas do quadro 12, exceto com a) em que está o elemento pivô (nesse caso, a 4ª linha).

Acompanhe o desenvolvimento desses procedimentos por linha:

1ª linha

Quadro 22 – Cálculo da nova 1ª linha

	Z	X ₁	X ₂	XF ₁	XF ₂	XF ₃	b _i
Nova linha pivô	0	0,5	1	0,25	0	0	25
Procedimento a: multiplique cada um dos elementos por 150 (valor da entrada da 1ª linha com o sinal trocado)-FO.	0	75	150	37,5	0	0	3.750
1ª linha (do quadro 20)-FO	1	-120	-150	0	0	0	0
Procedimento b: some os elementos da linha multiplicada pelo valor da entrada, com sinal trocado, aos elementos da 1ª linha do quadro 20-FO, para obter a nova 1ª linha-FO.	1	-45	0	37,5	0	0	3.750

Fonte: Elaboração própria, 2024.

3ª linha

Quadro 23 – Cálculo da nova 3ª linha

	Z	X_1	X_2	XF_1	XF_2	XF_3	b_i
Nova linha pivô	0	0,5	1	0,25	0	0	25
Procedimento a: multiplique cada um dos elementos por (-2) (valor da entrada da 3ª linha com o sinal trocado) – RT_2 .	0	-1	-2	-0,5	0	0	-50
3ª linha (do quadro 20) – RT_2	0	3	2	0	1	0	90
Procedimento b: some os elementos da linha multiplicada pelo valor da entrada, com sinal trocado, aos elementos da 1ª linha do quadro 20 – RT_2 , para obter a nova 3ª linha – RT_2 .	0	2	0	-0,5	1	0	40

Fonte: Elaboração própria, 2024.

4ª linha

Quadro 24 – Cálculo da nova 4ª linha

	Z	X_1	X_2	XF_1	XF_2	XF_3	b_i
Nova linha pivô	0	0,5	1	0,25	0	0	25
Procedimento a: multiplique cada um dos elementos por (-3) (valor da entrada da 4ª linha com o sinal trocado) – RT_3 .	0	-0,5	-3	-0,75	0	0	-75
4ª linha (do quadro 20) – RT_3	0	5	3	0	0	1	120
Procedimento b: some os elementos da linha multiplicada pelo valor da entrada, com sinal trocado, aos elementos da 1ª linha do quadro 20 – RT_3 , para obter a nova 4ª linha – RT_3 .	0	3,5	0	0,75	0	1	45

Fonte: Elaboração própria, 2024.

9º passo – Elabore uma tabela com as novas linhas obtidas. Para saber a resposta da questão, identifique nessa nova tabela as colunas que possuem o número 1 (visualize-os, no quadro 25, cada um com uma cor rosa), e depois associe os valores (1) aos números presentes na coluna de b_i (variáveis independentes), com suas

respectivas cores. Veja que cada número 1 possui uma cor e está associado ao seu b_i correspondente. O Z sempre representará a função objetiva e sempre terá um valor para responder ao problema inicial de maximização ou minimização. Já as restrições técnicas serão identificadas e subdivididas em variáveis básicas (todas as que tiverem valor 1) e não básicas (as em que não há o valor 1 em nenhuma linha).

Quadro 25 – Resultado matemático (não ótimo)

Partes da PL	Z	X_1	X_2 (Entrada)	XF_1	XF_2	XF_3	b_i
FO (1ª Linha)	1	-45	0	37,5	0	0	3.750
RT ₁ (2ª Linha)	0	0,5	1	0,25	0	0	25
RT ₂ (3ª Linha)	0	2	0	-0,5	1	0	40
RT ₃ (4ª Linha)	0	3,5	0	-0,75	0	1	45

Fonte: Elaboração própria, 2024.

Assim, nesse caso, Z , X_1 , X_2 , XF_1 , XF_2 e XF_3 assumem o valor do b_i da linha em que estiver o valor 1 no Quadro 25. Chegamos, então, à seguinte conclusão matemática (Quadro 25):

- Variáveis básicas (que têm o valor 1):
 - $Z = 3.750$ (valor da FO em relação à Maximização de Lucro da empresa e em relação à venda das duas calças: a normal e a com bolso invisível);
 - $X_2 = 25$ (pois o b_i é 25 na linha em que está o valor 1);
 - $XF_2 = 40$ (pois o b_i é 40 na linha em que está o valor 1);
 - $XF_3 = 45$ (pois o b_i é 45 na linha em que está o valor 1).

- Variáveis não básicas (que não têm o valor 1): X_1 e XF_1 .

Observação: Se as colunas das variáveis de decisão da tabela obtida no 9º passo contiverem números negativos, a solução encontrada não será ótima em razão da restrição de não negatividade. Então, será necessário refazê-las. Se as variáveis de decisão obtidas com o 9º passo apresentarem apenas valores positivos, teremos uma função ótima e não precisaremos refazê-la.

Nesse caso, a função ainda não é ótima, pois há um valor negativo na coluna X_1 (-45), apesar de X_2 ter apenas valores positivos. Lembrando que X_1 e X_2 são as variáveis de decisões desse exemplo específico, sendo necessário refazê-lo. À medida que a função é refeita, o valor de Z deve ir aumentando, caso seja de maximização; se o valor de Z estiver relacionado com minimização, à medida que a função vai sendo refeita, o valor de Z deve diminuir, pois busca a solução ótima.

10º passo – Identifique a função objetiva no modelo e o maior valor absoluto presente nas variáveis de decisão. Entrada é em X_1 por ser o maior valor absoluto, nesse caso, $X_1 = -45$ (valor absoluto = 45). Ao refazer, estamos retornando para o **3º passo**. O valor absoluto de X_1 ($4,5$) > X_2 (0).

Quadro 26 – Cálculo da nova entrada na função objetiva

Entrada em X_1 , maior valor absoluto das variáveis de decisões							
Partes da PL	Z	X_1	X_2 (Entrada)	XF_1	XF_2	XF_3	b_i
FO	1	-45	0	37,5	0	0	3.750

Fonte: Elaboração própria, 2024.

11º passo – Nas restrições técnicas, identifique o local em que haverá a saída, tendo como base o 9º passo. Para isso, divida o valor de b_i (variável independente) pelo valor da variável que entra da restrição correspondente (da mesma linha).

Quadro 27 – Cálculo da nova saída na restrição técnica

Restrição Técnica (RT)	Modelo de RT	b_i	X_1 (Entrada – Variável de decisão – 10º passo)	bi de cada restrição técnica/variável que entra de cada restrição técnica
RT ₁ (quantidade e funcionários em unidade):	$0,5X_1 + X_2 + 0,21XF_1 = 25$	25	0,5	$25/0,5 = 50$
RT ₂ (espaço em m²)	$2X_1 - 0,5XF_1 + XF_2 = 40$	40	2	$40/2 = 20$
RT ₃ (quantidade de equipamentos em unidade)	$3,5X_1 - 0,75XF_1 + XF_3 = 45$	45	3,5	$45/3,5 = 12,86$ (SAÍDA)

Fonte: Elaboração própria, 2024.

Como já foram encontradas as variáveis de folga, organize o sistema, identificando a entrada e a saída, e tendo como base o quadro obtido no 9º passo. No local em que o valor da entrada e o da saída forem iguais, estará o elemento pivô (no ponto comum). A linha do elemento pivô corresponderá à linha pivô (mesmo procedimento descrito no 6º passo). Assim, temos: a Entrada agora em X_1 (10º passo) e a Saída em RT₃ (11º passo); o local onde as duas linhas se encontram é o elemento Pivô; e a linha onde está esse elemento é a Linha Pivô.

Quadro 28 – Representação da nova entrada e da nova saída (elemento pivô)

Nesse caso, o elemento pivô é o número 3,5, pois é ele que está no ponto em que a entrada é igual à saída (ponto de encontro entre entrada e saída). A Linha que o Elemento pivô está é a que será pega para o 12º Passo seguinte.

Partes da PL	Z	X_1	X_2 (Entrada)	XF_1	XF_2	XF_3	b_i
FO (1ª Linha)	1	-45	0	37,5	0	0	3.750
RT_1 (2ª Linha)	0	0,5	1	0,25	0	0	25
RT_2 (3ª Linha)	0	2	0	-0,5	1	0	40
RT_3 (4ª Linha)	0	3,5	0	-0,75	0	1	45

Fonte: Elaboração própria, 2024.

12º passo – Divida os elementos da linha pivô pelo valor do elemento pivô, obtendo uma nova linha pivô unitária (mesmo procedimento descrito no 7º passo). Nesse caso, a linha pivô é a 4ª linha (do quadro 28) e será a base para os outros cálculos.

Quadro 29 – Cálculo da nova linha pivô

Partes da PL	Z	X_1	X_2 (Entrada)	XF_1	XF_2	XF_3	b_i
RT_3 (Linha Pivô – Quadro 28)	0	3,5	0	-0,75	0	1	45
Divisão por 3,5 (elemento pivô) para chegarmos à nova linha pivô	0	1	0	-0,21	0	0,28	12,86 (SAÍDA)

Fonte: Elaboração própria, 2024.

Observação: Esse valor de b_i deve ser igual ao valor da saída obtido no 11º passo.

13º passo – Calcule as novas linhas, com base na nova linha pivô. Para obtê-las, siga esses procedimentos (como feito no 8º passo):

- a) Multiplique cada elemento da nova linha pivô pelo valor da entrada obtido no 11º passo (em verde, no quadro 28), porém com o sinal trocado;
- b) Some os elementos encontrados no procedimento a), termo a termo, com os elementos da 1ª linha (do quadro 28); e
- c) Repita os procedimentos a) e b) nas outras linhas do quadro 28, exceto na qual está o elemento pivô (nesse caso, RT₃-Quadro 28).

Quadro 30 – Cálculo da nova 1ª linha

1º linha

	Z	X ₁	X ₂	XF ₁	XF ₂	XF ₃	b _i
Nova linha pivô	0	1	0	-0,21	0	0,28	12,86
Procedimento a: multiplique cada um dos elementos por 45 (valor da entrada da 1ª linha com o sinal trocado) – FO.	0	45	0	-9,64	0	12,6	578,57
1ª linha (do quadro 28) – FO	1	-45	0	37,5	0	0	3.750
Procedimento b: some os elementos da linha multiplicada pelo valor da entrada, com sinal trocado, aos elementos da 1ª linha do quadro 28 – FO, para obter a nova 1ª linha – FO.	1	0	0	27,86	0	0	4.328,57

Fonte: Elaboração própria, 2024.

Quadro 31 – Cálculo da nova 2ª linha

2ª linha

	Z	X_1	X_2	XF_1	XF_2	XF_3	b_i
Nova linha pivô	0	1	0	-0,21	0	0,28	12,86
Procedimento a: multiplique cada um dos elementos por -0,5 (valor da entrada da 2ª linha com o sinal trocado) – RT_1 .	0	-0,5	0	-0,105	0	-0,14	-6,43
2ª linha (do quadro 28) – RT_1	0	0,5	1	0,25	0	0	25
Procedimento b: some os elementos da linha multiplicada pelo valor da entrada, com sinal trocado, aos elementos da 2ª linha do quadro 28 – FO, para obter a nova 2ª linha – RT_1 .	0	0	1	-0,145	0	-0,14	18,57

Fonte: Elaboração própria, 2024.

Quadro 32 – Cálculo da nova 2ª linha

3ª linha

	Z	X_1	X_2	XF_1	XF_2	XF_3	b_i
Nova linha pivô	0	1	0	-0,21	0	0,28	12,86
Procedimento a: multiplique cada um dos elementos por -2 (valor da entrada da 3ª linha com o sinal trocado) – RT_3 .	0	-2	0	0,42	0	0,56	-25,71
3ª linha (do quadro 28) – RT_3	0	2	0	-0,5	1	0	40
Procedimento b: some os elementos da linha multiplicada pelo valor da entrada, com sinal trocado, aos elementos da 3ª linha do quadro 28 – FO, para obter a nova 3ª linha – RT_3 .	0	0	0	-0,08	1	0,56	14,29

Fonte: Elaboração própria, 2024.

14º passo – Elabore uma nova tabela com os resultados obtidos. Para saber a resposta da questão, identifique as variáveis que têm valor 1 e associe-as ao valor de b_i respectivo. O Z que representa a função objetiva sempre terá um valor para responder

ao problema inicial de maximização ou minimização. No caso da questão da concessionária, o objetivo é maximizar. Já nas restrições técnicas, serão identificadas as variáveis básicas (variáveis que têm o valor 1-cor rosa no quadro 33) e não básicas (variáveis que não têm o valor 1).

Quadro 33 – Resultado matemático ótimo

Partes da PL	Z	X_1	X_2 (Entrada)	XF_1	XF_2	XF_3	b_i
FO (1ª Linha)	1	0	0	27,86	0	0	4.328,57
RT_1 (2ª Linha)	0	0	1	-0,145	0	-0,14	18,57
RT_2 (3ª Linha)	0	0	0	-0,08	1	0,56	14,29
RT_3 (4ª Linha)	0	1	0	-0,21	0	0,28	12,86

Fonte: Elaboração própria, 2024.

As variáveis básicas (que apresentam valor 1) assumem o valor de b_i de suas respectivas linhas. Chegamos, então, à seguinte conclusão matemática:

- Variáveis básicas:
 - $Z = 4.328,57143$ (valor de máximo da FO para otimizar o sistema, correspondendo ao lucro máximo obtido na venda das duas calças na Empresa familiar).
 - $X_1 = 12,8571429$ (necessário vender 12,86 calças comuns. Cabe aproximar para 13, pois não se pode vender 12,85 de uma calça. Outro aspecto é que $X_1 \geq 0$, atendendo à restrição de não negatividade, então o resultado está dentro do que foi estabelecido como realidade).

- $X_2 = 18,5714286$ (necessário vender 18,57 calças com bolso invisível. Cabe aproximar para 19, pois não se pode vender 18,57 de uma calça. Outro aspecto é que $X_1 \geq 0$, atendendo à restrição de não negatividade, então o resultado está dentro do que foi estabelecido coma realidade).
- $XF_2 = 14,2857142$ Referente a RT_2 (espaço em m^2). Assim, tem-se que sobrou aproximadamente 12,29 m^2 .
- Variáveis não básicas:
 - XF_1 (Referente a RT_1 [Quantidade e funcionários em unidade]). Assim, não sobrou nada de funcionários, todos os onze foram usados.
 - XF_3 (Referente RT_3 [Quantidade de Equipamentos em unidade]). Também não sobrou nada dos equipamentos, todos foram usados.

Essa função é ótima pois os números presentes nas colunas das variáveis de decisão (X_1 e X_2) são positivos. Assim, não é necessário refazer as linhas, mas se o resultado fosse negativo seria necessário até as colunas das variáveis de decisão terem apenas números positivos.

15º passo – Interprete os dados obtidos com o estudo de caso da Empresa familiar (Quadro 33) para a realidade empresarial. Vamos fazer a substituição dos valores no modelo de PL apresentado no 5º Passo do item 4.2.1 do capítulo 4, começando pela Função objetiva e depois as restrições, onde $X_1 = 12,8571429$ e $X_2 = 18,5714286$.

Relembrando que a PL presente no 5º Passo do item 4.2.1 do capítulo 4 é:

- Variáveis de decisões:
 - X_1 = Calça Comum
 - X_2 = Calça do Bolso Invisível
- Função objetiva: Máx (L) = $120X_1 + 150X_2$
- Restrição Técnica (recursos limitados)
 - RT_1 (Quantidade e funcionários em unidade): $2X_1 + 4X_2 \leq 100$
 - RT_2 (espaço em m^2): $3X_1 + 2X_2 \leq 90$
 - RT_3 (Quantidade de Equipamentos em unidade): $5X_1 + 3X_2 \leq 120$
- Restrição de não negatividade: $X_1 \geq 0$ e $X_2 \geq 0$.

Assim, tem-se ao fazer a substituição:

- O Lucro Máximo e ótimo diário da empresa familiar é de \$ 4.328,57 em relação à venda de calças comuns e com bolso invisível. Afinal, $Z = 4.328,57$ (Quadro 33 – FO – **cor azul**). Para confirmar os dados, basta substituir os valores de $X_1 = 12,8571429$ e $X_2 = 18,57$ (Aproximadamente 19) na fórmula de Função objetiva: $\text{Max}(L) = 120X_1 + 150X_2$. Tanto o valor encontrado em $X_1 \geq 0$ como em $X_2 \geq 0$ são verdadeiros, pois atendem à restrição de não negatividade.

Já ao fazer a substituição na:

- RT_1 (Quantidade e funcionários em unidade), tem-se: $2X_1 + 4X_2 \leq 100 \rightarrow 2(12,8571429) + 4(18,5714286) \rightarrow 25,7142858 + 74,285944 = 100,00023100 = 100$ (Verdadeiro). Por isso, a Variável de Folga (XF_1) oriunda de RT_1 é não básica, pois utilizou todos os funcionários, ou seja, 100, para a produção das calças comuns e com bolso invisível.

- RT_2 (espaço em m^2), tem-se:
 $3X_1 + 2X_2 \leq 90 \rightarrow 3(12,8571429) + 2(18,5714286) \rightarrow$
 $38,5714287 + 37,1428572 = 75,7142859 \rightarrow 75,71 < 90$
 (Verdadeiro). Por isso, a Variável de Folga (XF_2) oriunda de RT_2 é básica e teve como va-lor $XF_2 = 14,2857142$, pois $90 - 75,7142859 = 14,2857142$. O nome é variável de folga porque é a sobra de recurso referente a RT_2 , que é o espaço de armazenagem para calças comuns e com bolso invisível em m^2 , o valor de aproximadamente $14,29 m^2$.
- RT_3 (Quantidade de Equipamentos em unidades), tem-se:
 $5X_1 + 3X_2 \leq 120 \rightarrow$
 $5(12,8571429) + 3(18,5714286) \rightarrow$
 $64,2857145 + 55,7142864 = 120,000001 \rightarrow 120 = 120$
 (Verdadeiro). Por isso, a Variável de Folga (XF_3) oriunda de RT_3 é não básica, pois utilizou todos os equipamentos, ou seja, 120, para a produção das calças comuns e com bolso invisível.

Essa questão também é possível de resolver de forma gráfica por ter apenas duas variáveis (X_1 e X_2).

Passo a Passo do cálculo do SIMPLEX (capítulo 2 – Pereira, 2010)



Link: https://youtu.be/1jxD_FiJUra

4.2.3 Cálculo da Programação Linear por gráfico

A representação gráfica consiste em apresentar um sistema de eixos cartesianos – um horizontal (X_1) e outro vertical (X_2). Os dois devem obedecer às limitações dos recursos, ou seja, as restrições técnicas e de não negatividade impostas pelo modelo de PL. O gráfico apresenta, além da solução ótima, todos os possíveis pontos viáveis para a tomada de decisão. Para fazermos o gráfico, inicialmente é necessário termos a modelagem do sistema. Usaremos o modelo já trabalhado anteriormente no 5º Passo do item 4.2.1 do capítulo 4 e mostraremos passo a passo a elaboração do gráfico conforme Pereira (2010).

1º passo – Identifique as restrições técnicas e zere os valores das variáveis de decisão, encontrando assim os pontos de delimitação do gráfico para cada variável de decisão (resolva a equação, considerando cada uma das variáveis de decisão como zero). No caso da empresa familiar, existem três restrições técnicas, que nas fórmulas são apresentadas como inequações. Para realizarmos o cálculo desse item, devemos transformar a inequação em uma equação. Vamos iniciar os cálculos das Restrições técnicas:

- RT_1 (Quantidade e funcionários em unidade) tem-se:

$$2X_1 + 4X_2 \leq 100$$

Figura 28 – Cálculo e Representação gráfica da RT₁

Para $X_1 = 0$ $2X_1 + 4X_2 \leq 100$ $2X_1 + 4X_2 = 100$ $2(0) + 4X_2 = 100$ $0 + 4X_2 = 100$ $X_2 = 100 / 4$ $X_2 = 25$ Ponto A₁ (0;25) Se $X_1 = 0$, então $X_2 = 25$. Ponto A₁ (0;25). - + 25	Para $X_2 = 0$ $2X_1 + 4X_2 \leq 100$ $2X_1 + 4X_2 = 100$ $2X_1 + 4(0) = 100$ $2X_1 + (0) = 100$ $X_1 = 100 / 2$ $X_1 = 50$ Ponto A₂ (50;0) Se $X_2 = 0$, então $X_1 = 50$. Ponto A₂ (50;0). - + 50
--	--

Fonte: Elaboração própria, 2024.

Observação

- Se X for positivo e o sinal for de $\leq$, marque o lado à esquerda;
- Se X for positivo e o sinal for de $\geq$, marque o lado à direita;
- Se X for negativo e sinal for de $\leq$, marque o lado à direita;
- Se X for negativo e o sinal for de $\geq$, marque o lado à esquerda.

- Para RT₂ (espaço em m²), tem-se: $3X_1 + 2X_2 \leq 90$

Figura 29 – Cálculo e Representação gráfica da RT_2

Para $X_1 = 0$ $3X_1 + 2X_2 \leq 90$ $3X_1 + 2X_2 = 90$ $3(0) + 2X_2 = 90$ $X_2 = 90/2$ $X_2 = 45$ Ponto $B_1(0;45)$ Se $X_1 = 0$, então $X_2 = 45$. Ponto $B_1(0;45)$. - + 45	Para $X_2 = 0$ $3X_1 + 2X_2 \leq 90$ $3X_1 + 2X_2 = 90$ $3X_1 + 2(0) = 90$ $X_1 = 90/3$ $X_1 = 30$ Ponto $B_2(30;0)$ Se $X_2 = 0$, então $X_1 = 30$. Ponto $B_2(30;0)$. - + 30
---	---

Fonte: Elaboração própria, 2024.

- Para RT_3 (Quantidade de Equipamentos em unidade), tem-se: $5X_1 + 3X_2 \leq 120$

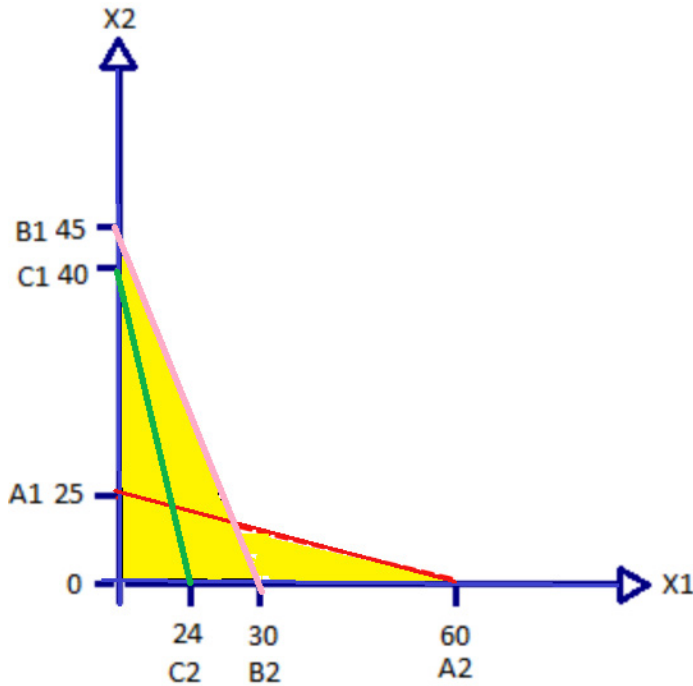
Figura 30 – Cálculo e Representação gráfica da RT_3 .

Para $X_1 = 0$ $5X_1 + 3X_2 \leq 120$ $5X_1 + 3X_2 = 120$ $(0) + 3X_2 = 120$ $X_2 = 120/3$ $X_2 = 40$ Ponto $C_1(0;40)$. Se $X_1 = 0$, então $X_2 = 40$. Ponto $C_1(0;40)$. - + 40	Para $X_2 = 0$ $5X_1 + 3X_2 \leq 120$ $5X_1 + 3X_2 = 120$ $5X_1 + 3(0) = 120$ $X_1 = 120/5$ $X_1 = 24$ Ponto $C_2(24;0)$. Se $X_2 = 0$, então $X_1 = 24$. Ponto $C_2(24;0)$. - + 24
---	--

Fonte: Elaboração própria, 2024.

2º passo – Elabore o gráfico de inequações das restrições técnicas e de não negatividade, inserindo os pontos encontrados. Lembrando que a resposta deve estar apenas no 1º Quadrante, conforme explicação da Figura 28 item 4.2. A RT_1 está representada pela linha vermelha, a RT_2 na linha rosa e a RT_3 na linha verde da Figura 31. Já as restrições de não negatividade estão no eixo x_1 e x_2 até $+\infty$.

Figura 31 – Representação gráfica das três Restrições Técnicas

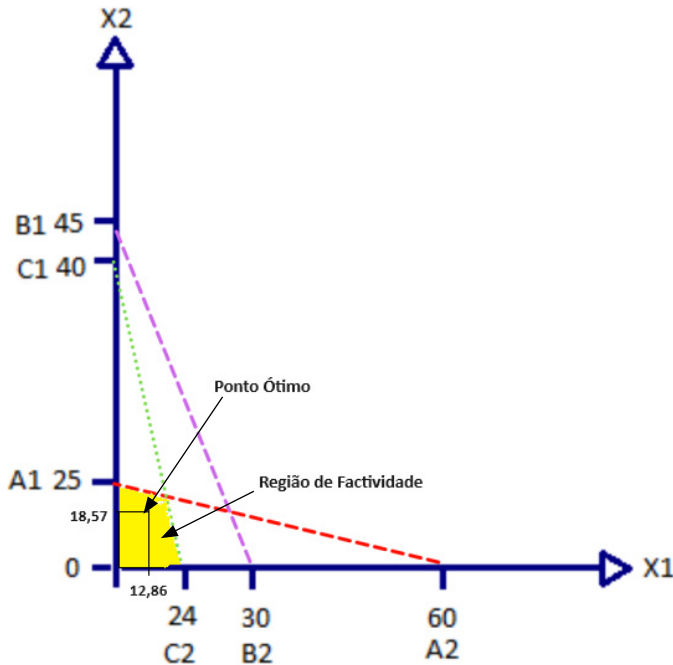


Fonte: Elaboração própria, 2024.

3º passo – Com base no jogo do sinal das RT_1 , RT_2 , RT_3 e nas restrições de não negatividade ($x_1 \geq 0$ e $x_2 \geq 0$), encontrar os pontos de interseção entre todos os itens citados anteriormente.

Na Figura 32, foram traçadas todas as restrições técnicas e de não negatividade. Dessa forma, foi encontrado o espaço solução, que é o conjunto de todos os pontos entre os quais está aquele que pode vir a ser considerado como o ponto ótimo (Figura 32 – cor amarela), ou seja, os que “obedecem” a todas as restrições do modelo. No gráfico, o espaço solução é o polígono amarelo. Qualquer ponto na região de factibilidade (Figura 32 – região amarelo) é solução para o problema, mas a solução ótima será encontrada em um dos vértices desse polígono. Assim, a Figura 32 identifica todos os possíveis pontos para a solução da questão do estudo da empresa familiar, e X_1 pode variar de (0;24), e X_2 de (0;25). Perceba que, na Figura 32, estão todos os pontos possíveis para X_1 e X_2 e as relações existentes entre X_1 e X_2 , que podem ser substituídos na função objetiva. No SIMPLEX, cada iteração apresenta o valor de um vértice do polígono. A Figura 32 (Ponto ótimo) indica a decisão que dará o maior lucro para a empresa, mas para isso é necessário traçar retas paralelas: no local onde a reta tocar o ponto mais alto do polígono (pois nesse caso o problema é de maximização) está a solução ótima (nesse caso, $X_1 = 12,86$ e $X_2 = 18,57$). Nem sempre o maior lucro é a melhor solução para o pesquisador. Essa decisão dependerá da circunstância ambiental em que a empresa está inserida. Por exemplo: se considerarmos uma greve geral nas empresas que fabricam tecido para calças e não houver outras com material similar, a empresa terá que se restringir à venda de calças. Nesse caso, $X_1 = 12,8571429$ (seria necessário vender 12,86 calças comuns, cabendo aproximar para 13 pois não se pode vender 12,85 de uma calça) e $X_2 = 18,5714286$ (seria necessário vender 18,57 calças com bolso invisível, cabe aproximar para 19). Ambos os números são positivos, atendem à condição da restrição de não negatividade e estão presentes na parte amarela do Figura 32.

Figura 32 – Representação gráfica das possíveis soluções



Fonte: Elaboração própria, 2024.

4º passo. O pesquisador pode pegar outros pontos dentro da área amarela da Figura 32 para fazer simulações, já que através do SIMPLEX ele já sabe o valor do ponto ótimo (maior lucro possível a partir dos recursos que possui é de R\$ 4.328,57 por dia). Na realidade de um empreendimento de forma rotineira, o ponto do SIMPLEX nem sempre será alcançado. Assim, é possível fazer uma simulação para não se perder um valor muito alto no lucro. O empresário pode optar, de forma hipotética, trabalhar apenas com calça de bolso invisível devido ao aumento de furtos/roubos de celular, contemplando esse nicho de mercado. Os pontos $(0;25)$ são exemplo, mas poderia ser qualquer um dentro da área de facticidade da Figura 32.

Se fizer isso, é necessário realizar a substituição primeiro em todas as restrições técnicas, para que todas sejam verdadeiras, não sendo possível admitir nenhuma falsa. Se todas deram verdade, faz-se a substituição em FO. Assim, os pontos são $X_1 = 0$ e $X_2 = 25$, tem-se:

- na RT_1 (quantidade e funcionários em unidade):
 $2X_1 + 4X_2 \leq 100 \rightarrow$
 $2(0) + (4 \cdot 25) \leq 100 \rightarrow 100 = 100$ (VERDADEIRO). Não há sobra, usará todos os funcionários.
- RT_2 (espaço em m^2): $3X_1 + 2X_2 \leq 90 \rightarrow 3(0) + 2(25) \leq 90 \rightarrow$
 $50 \leq 90$ (VERDADEIRO). Sobra 40 m^2 de espaço ($90 - 50 = 40$), que o empresário pode alugar ou usar para outra atividade dentro da própria empresa caso queira.
- RT_3 (Quantidade de Equipamentos em unidade):
 $5X_1 + 3X_2 \leq 120 \rightarrow$
 $5(0) + 3(25) \leq 120 \rightarrow 75 \leq 120$ (VERDADEIRO). Sobram 45 equipamentos ($120 - 75 = 45$), que o empresário pode alugar ou usar para outra atividade dentro da própria empresa caso queira.

5º passo. Se todas as restrições forem atendidas, é possível substituímos esses valores na função objetiva (se houver uma restrição falsa, a substituição não será possível). Então, temos: Função objetiva: $\text{Max}(L) = 120X_1 + 150X_2 \rightarrow$
 $120(0) + (150 \cdot 25) = \text{R\$ } 3.750$. Se fizer essa opção, perderá por dia de trabalho $\text{R\$ } 578,57$, mas talvez consiga recuperar o dinheiro com novas atividades oriundas das sobras de recursos de RT_2 e RT_3 .

Veja casos práticos e reais do uso da PL nos livros abaixo:

HOSPITAL SANTA IZABEL. *Relatório de Gestão Hoteleira com o Uso de Modelos Matemáticos*. Salvador: UNEB, 2018.



Disponível em: https://issuu.com/fvianaoliveira/docs/livro_santa_izabel.
Acesso em: 11 fev. 2024.

PEREIRA, Aliger dos Santos. *Organização, sistemas e métodos: estudo de caso micro e pequenas empresas de Salvador*. Salvador: UNEB, 2020.



Disponível em: https://issuu.com/ibesvoceadm/docs/livro_adm_uneb_2020.
Acesso em: 11 fev. 2024

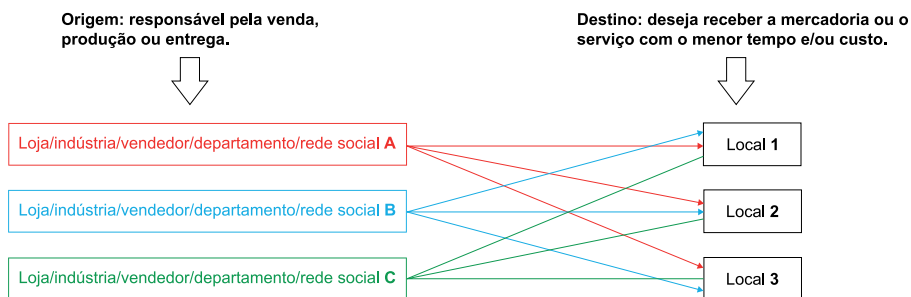
Vimos aplicação da PL em exemplos com Maximização que também podem ser usados para Minimização. Entretanto, queremos mostrar o uso da Minimização com a Técnica de Transporte, nesse caso específico somente podendo ser usada visando minimizar a PL.

4.2.4 Cálculo da Programação Linear com o uso da Técnica do Transporte

Apesar do nome ser transporte, seria melhor nomear como alocação de recursos de um local (origem) para um destino.

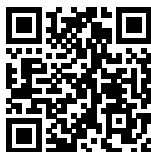
O problema de transporte pode ser representado, esquematicamente, pelas origens e destinos (conforme Figura 33), considerando que **A**, **B** e **C** devem entregar seus produtos para os locais **1**, **2** e **3**. Note que, na origem, está o produtor/vendedor/entregador (como fornecedores, lojas, indústrias ou mesmo um departamento dentro de uma organização). Do outro lado do processo, há o destino, que pode ser o cliente interno ou externo (clientes internos são os funcionários/colaboradores da empresa, e cliente externo é o mercado consumidor), o local ou a região que demanda um produto ou serviço, cujas necessidades devem ser atendidas. Em nosso exemplo, a quantidade de origens é igual à quantidade de destinos. Com exceção do método da designação, para o qual é obrigatório que o número de origens e o de destinos seja igual, na resolução pelos outros métodos não há essa obrigatoriedade. Hoje, é possível utilizá-lo também para fluxo de informações entre pessoas ou equipamentos na gestão de conhecimento com rede sociais.

Figura 33 – Representação do estudo de transporte



Fonte: Adaptado de Pereira (2010).

Ver filme sobre o uso do Transporte



Link: https://youtu.be/_mZY-yLsnrg

Imagine que a empresa familiar das calças tem três filiais para entregar em três locais distintos as mercadorias em um menor Tempo. Dessa forma, o foco é na minimização do Tempo.

Quadro 34 – Entrega média diária e demanda diária de cada localidade

Média diária de entregas das filiais	Demanda diária de cada localidade
Filial A = 25 (loja A)	Local 1 = 30
Filial B = 60 (loja B)	Local 2 = 40
Filial C = 55 (loja C)	Local 3 = 70

Fonte: Adaptado de Passos (2008).

O Quadro 35 mostra o tempo de distribuição dos materiais (em minutos) para as localidades:

Quadro 35 – Tempo para distribuição de mercadorias

Para De	Local 1	Local 2	Local 3
Filial A	36	41	25
Filial B	28	25	31
Filial C	26	30	41

Fonte: Adaptado de Passos (2008).

Antes de realizarmos a representação matemática (PL), precisamos fazer uma única representação física das atividades presentes nos Quadros 34 e 35 em um único (Quadro 36). No Quadro 36, constata-se que o somatório dos valores da demanda deve ser igual ao somatório dos valores da produção (quadro 36). O somatório da produção ($25 + 60 + 55$) é igual ao somatório da demanda ($30 + 40 + 70$), ou seja, 140 minutos. Se não forem iguais, será necessário acrescentarmos uma linha (no caso da produção) ou uma coluna (no caso da demanda) com a quantidade que falta, e a resolução ocorrerá da mesma forma.

Quadro 36 – Somatório da demanda igual ao somatório da produção

Para De	Local 1	Local 2	Local 3	Produção
Filial A	36	41	25	25
Filial B	28	25	31	60
Filial C	26	30	41	55
Demanda	30	40	70	140

Somatório da produção
(25 + 60 + 55)

Somatório da demanda (30 + 40 + 70)

Demanda = Produção

Fonte: Elaboração própria, 2024.

A partir do Quadro 36, pode-se montar o seguinte modelo de PL para a minimização do Tempo (em minutos) de entrega de todas as mercadorias das três lojas em um dia:

- **Função objetiva** – Busca minimizar o tempo, portanto, considera os valores relacionados a ele, associando a origem ao destino:
 - Na cor vermelha, estão os tempos de entrega da mercadoria referentes à Filial **A**: 36 minutos para que ele chegue ao Local 1; 41 minutos para a entrega da mercadoria no Local 2; e 25 minutos para levar a mercadoria ao Local 3.

- Na cor azul, estão os tempos de entrega da mercadoria referentes à Filial **B**: 28 minutos para que ele chegue ao Local 1; 25 minutos para a entrega da mercadoria no Local 2; e 31 minutos para levar a mercadoria ao Local 3.
- Na cor verde, estão os tempos de entrega da mercadoria referentes à Filial **C**: 26 minutos para que ele chegue ao Local 1; 30 minutos para a entrega da mercadoria no Local 2; e 41 minutos para levar a mercadoria ao Local 3.

$$\text{Min}(T) = 36X_{A1} + 41X_{A2} + 25X_{A3} + 28X_{B1} + 25X_{B2} + 31X_{B3} + 26X_{C1} + 30X_{C2} + 41X_{C3}$$

Observe que a função objetiva do estudo de caso foi elaborada com base nos tempos de distribuição do produto em minutos, sem considerarmos nenhum valor de produção ou de demanda. Perceba também que as letras de cada variável (X) estão associadas à origem (primeira letra subscrita) e ao destino (número da localidade subscrita).

- **Restrição técnica** – Corresponde ao limite de recursos que a empresa possui. Nesse caso, as restrições estão relacionadas à produção e à demanda:
 - Produção da Filial **A** (valores da segunda linha do quadro 36):

$$X_{A1} + X_{A2} + X_{A3} = 25$$
 - Produção da Filial **B** (valores da terceira linha do quadro 36):

$$X_{B1} + X_{B2} + X_{B3} = 60$$
 - Produção da Filial **D** (valores da quarta linha do quadro 36):

$$X_{C1} + X_{C2} + X_{C3} = 55$$

- Demanda do Local 1 (valores da segunda coluna do quadro 36):
 - $X_{A1} + X_{B1} + X_{C1} = 30$
- Demanda do Local 2 (valores da terceira coluna do quadro 36):
 - $X_{A2} + X_{B2} + X_{C2} = 40$
- Demanda do Local 3 (valores da quarta coluna do quadro 36):
 - $X_{A3} + X_{B3} + X_{C3} = 70$
- **Restrição de não negatividade** – Todas as variáveis presentes no sistema (modelo) e, conseqüentemente existentes também na função objetiva, devem ser consideradas, pois são as variáveis de tomada de decisão da PL (capítulo 2). Essa restrição diz respeito à exigência de que todas as variáveis de decisão sejam maiores ou iguais a zero:
 - $X_{A1}; X_{A2}; X_{A3}; X_{B1}; X_{B2}; X_{B3}; X_{C1}; X_{C2}; X_{C3} \geq 0$

Percebemos que o método do transporte é oriundo da PL. Entretanto, como há muitas variáveis de decisão, assim como restrições técnicas, é mais rápido resolvê-lo manualmente, utilizando as técnicas de transporte, do que pela PL. Indicamos a utilização de programas computacionais para tornar a resolução mais prática e dinâmica, mas de qualquer forma vamos mostrar manualmente a técnica do Custo Mínimo e de Vogel passo a passo (Pereira, 2010).

Observação: Por meio do Modelo de PL acima, podemos resolvê-lo através: do SIMPLEX; de algumas técnicas de transporte, como por exemplo Custo mínimo, canto noroeste, Vogel e Designação; ou do uso de *software* (por exemplo Excel através do Solver, OMP, COM, QSB, LINDO ou LINGO). Como esse modelo tem mais de duas variáveis, é difícil resolver manualmente utilizando gráficos. Outros aspectos é que as técnicas de Custo mínimo, canto noroeste e Vogel podem dar o mesmo resultado ou com pequenas diferenças. Ao resolver no *software* sempre dará o menor valor, não importando a técnica. Neste livro, mostrarei apenas a do custo Mínimo, as outras três técnicas podem ser vistas nos livros de Pereira (2010), Passos (2008) e Lachtermacher, (2016) caso queira mais aprofundamento no respectivo assunto.

O nome do método é Custo Mínimo porque os cálculos sempre começarão pelos menores valores de tempo ou custo. Assim, caracteriza-se pelo fato de que os valores das cargas (representadas pela cor marrom no Quadro 37), ou seja, o tempo de deslocamento em minutos para se levar os produtos, devem ser inseridos em uma célula durante a resolução, considerando o menor tempo dentro da capacidade de cada linha (disponibilidade, oferta ou produção) ou de cada coluna (necessidade de demanda) e seguindo em ordem crescente.

O Quadro 37 mostra os dados referentes ao estudo de caso “Entrega de Calças de uma Empresa Familiar para três localidades”. Em marrom estão os valores de tempo presentes nesse estudo de caso, cujo objetivo é minimizá-los. Se o problema priorizasse o custo, os valores considerados seriam referentes a esse.

Acompanhe essas instruções para a realização do cálculo pelo método do Custo Mínimo:

1º passo – Como indica o nome do método, Custo Mínimo, comece identificando o menor valor de tempo ou de custo. Nesse

caso, o menor valor é 25 minutos e corresponde ao tempo que a Filial **A** leva para chegar o **Local3**, ou seja, **A3**, mas poderia ser também o 25 de **B2**. Em seguida, encontre o menor valor de demanda ou produção, que nessa situação é 25 e corresponde à produção média diária da Filial **A**. Depois, insira o valor da menor produção na mesma célula em que está o menor valor de tempo ou custo (no caso 25) e zere o valor da produção nessa coluna ($25 - 25 = 0$). Ao fazer isso, zerará a linha da **Filial A**.

Como esse valor de produção (25) é menor do que a demanda do **Local 3** (70), subtraia o valor da demanda desse **Local 3** do valor da produção da **Filial A**. Assim, o valor da demanda passa a ser a diferença entre a Demanda **Local 3** pela Produção **Filial A** ($70 - 25 = 45$), e o da produção fica 0 ($25 - 25 = 0$).

Quadro 37 – Cálculo utilizando o método do Custo Mínimo I

Para De	Local 1	Local 2	Local 3	Produção
Filial A	36 -	41 -	25 25	25 0
Filial B	28	25	31	60
Filial C	26	30	41	55
Demanda	30	40	70 45	140

Fonte: Elaboração própria, 2024.

2º passo – No quadro 38, identifique o menor valor de tempo ou de custo entre os valores que não sofreram nenhuma alocação de recursos, ou seja, entre os itens das linhas referentes à **Filial B** e à **Filial C** que estão em marrom e que ainda não foram alterados (os números que não foram alterados nesse quadro são aqueles em cuja célula não foi inserido um número na cor rosa). Nesse caso é o 25, valor que corresponde ao tempo necessário para a **Filial B** fazer a entrega no **Local 2**, ou seja, **B2** corresponde ao menor valor não alterado em comparação aos outros: $B1 = 28$; $B2 = 25$;

B3 = 31; C1 = 26; C2 = 30; e C3 = 41. Depois, encontre o menor valor correspondente, se é o da produção ou se é o da demanda. Nesse caso é o da produção, que é 40, menor que o da produção, que é 60. Insira o menor valor entre o da demanda e o da produção na célula **B2** (que aqui é o valor demanda do **Local 2**). Em seguida, zere a coluna referente ao **Local 2** (pois fica $40 - 40 = 0$), porque esse valor foi alocado na célula **B2**, que corresponde à demanda dessa coluna. Subtraia esse valor da produção correspondente ($60 - 40 = 20$). Esse resultado do quadro 38 servirá para o cálculo do quadro 39.

Quadro 38 – Cálculo utilizando o método do Custo Mínimo II

Para De	Local 1	Local 2	Local 3	Produção
Filial A	36 -	41 -	25 25	0
Filial B	28 40	25	31 60	20
Filial C	26 -	30	41	55
Demanda	30 40	0	70 45	140

Fonte: Elaboração própria, 2024.

3º passo – No quadro 39, identifique o menor valor de tempo ou custo entre os valores que não sofreram nenhuma alocação de recursos. Nesse caso é o 26, ou seja, o tempo gasto da **Filial C** com destino ao **Local 1** – célula C1 (corresponde ao menor valor não alterado em comparação aos outros: B1 = 28; B3 = 31; C3 = 41). Depois, encontre os valores que dizem respeito à produção e à demanda (nesse caso, a produção é 55 e a demanda é 30). Insira o menor valor entre o da produção e o da demanda na célula C1 (aqui, a demanda, que corresponde a 30). Em seguida, zere a coluna da demanda do **Local 1**, pois esse valor foi alocado

na célula C1. Subtraia esse valor da demanda do valor da produção correspondente ($55 - 30 = 25$). Esse resultado será o valor de produção da linha A (quadro 40).

Quadro 39 – Cálculo utilizando o método do Custo Mínimo III

Para De	Local 1		Local 2		Local 3		Produção
Filial A	-	36	-	41	25	25	0
Filial B	-	28	40	25	31	60	20
Filial C	30	26	-	30	41	55	25
Demanda	30	0	40	0	70	45	140

Fonte: Elaboração própria, 2024.

4º passo – No quadro 40, identifique o menor valor de tempo ou de custo entre os valores que não sofreram nenhuma alocação de recursos (B3 e C3). Nesse caso é o 31, ou seja, o tempo gasto pela Filial B com destino ao Local 3 – célula **B3** (corresponde ao menor valor não alterado em comparação aos outros: $C3 = 41$). A seguir, encontre os valores correspondentes da produção e da demanda (na situação dada a produção é 20 e a demanda é 45). Insira o menor valor entre o da produção e o da demanda na célula B3 (aqui, a produção, que é 20). Em seguida, zere o valor da produção da **Filial B**, pois esse valor foi alocado na célula B3. Subtraia esse valor da produção do valor da demanda correspondente ($45 - 2 = 25$). Esse resultado, que é o valor que sobrou da demanda da localidade do **Local 3**, será o valor de demanda dessa localidade. Como esse é igual ao valor da produção da **Filial C**, aloque-o na célula **C3**, solucionando o problema (Quadro 40). Então, em **C3** aloque o valor de 25, presente na Produção da Filial C, e o 25 existentes na Demanda

do Local 3, alocando para **C3** e zerando as Produção e a demanda desses dois locais relatados anteriormente.

Quadro 40 – Cálculo utilizando o método do Custo Mínimo IV

Para De	Local 1		Local 2		Local 3		Produção	
Filial A		36		41		25	25	0
	-		-		25			
Filial B		28		25		31	60	20
	-		40		20			
Filial C		26		30		41	55	25
	30		-		25			
Demanda	30	0	40	0	70	45	25	140

Fonte: Elaboração própria, 2024.

5º passo – Calcule o tempo mínimo total de distribuição das mercadorias nas três localidades pelas três filiais (A, B e C). Para isso, multiplique o tempo de cada célula (em marrom) pelo recurso alocado na respectiva célula (em rosa), e em seguida realize o somatório de todos esses itens. O traço, que aparece em algumas células no lugar do recurso alocado, indica que o valor ali é zero. Como qualquer valor multiplicado por zero é igual a zero, não são considerados na fórmula de minimização que corresponde à função objetiva da PL:

$\text{Min (T)} = (25 \times 25) + (40 \times 25) + (20 \times 31) + (30 \times 26) + (25 \times 41) = \mathbf{4.050 \text{ minutos por dia}}$

6º passo – Interprete os resultados do problema, com base no Quadro 40.

A Filial A deve entregar:

- 25 unidades para o Local 3. Essa entrega leva 625 minutos, pois $25 \times 25 = 625$.

A Filial B deve entregar:

- 40 unidades para o Local 2. Essa entrega leva 1.000 minutos, pois $40 \times 25 = 1.000$.
- 20 unidades para o Local 3. Essa entrega leva 620 minutos, pois $20 \times 31 = 620$

A Filial C deve entregar:

- 30 unidades para o local 1. Essa entrega leva 780 minutos, pois $30 \times 26 = 780$.
- 25 unidades para o local 3. Essa entrega leva 1.025 minutos, pois $25 \times 41 = 1.025$

O tempo mínimo total gasto para distribuir todas as unidades pelo método do Custo Mínimo, ou seja, o tempo ótimo calculado por esse método, é de 4.050 minutos. Esse valor corresponde ao somatório da entrega das mercadorias pelas Filial A (625 minutos), B ($1.000 + 620 = 1.620$ minutos) e C ($780 + 1.025 = 1.805$ minutos). Para a empresa familiar, não importa entre as filiais quem demorará mais ou menos tempo entregando a mercadoria, e sim o tempo total de entrega.

Já o método de Vogel, também conhecido como método das Penalidades, consiste em fazer o transporte (alocar os recursos) na célula em que estiver o menor tempo/custo unitário da linha ou coluna que apresentar a maior penalidade. Em PO, o termo “penalidade” significa a diferença positiva entre os dois menores valores de tempo/custo unitários da mesma linha ou coluna (isto é, o resultado da subtração entre esses dois valores). Se o resultado for negativo, consideramos o valor absoluto (desconsideramos o sinal). Essa subtração se denomina “técnica da penalidade”.

Para mostrar como funciona a Técnica de Vogel, vamos usar o mesmo exemplo da empresa familiar passo a passo (Pereira, 2010).

1º passo – Subtraia os dois menores valores de tempo presentes em cada linha e em cada coluna (Quadro 41). Depois, identifique o maior valor entre os resultados dessas subtrações (penalidades), que pode estar na linha ou na coluna. Se a penalidade estiver na linha, identifique o menor valor referente ao tempo unitário dessa. Caso a maior penalidade estiver na coluna, identifique o menor valor de tempo unitário nela. Perceba que, no Quadro 41, a maior penalidade é 11, valor originado da subtração entre os dois menores valores da linha da Filial A ($A1 - A3 = 36 - 25 = 11$). Nessa linha, o menor valor de tempo unitário é 25 (célula A3). Comece balanceando essa célula, alocando o menor valor entre a produção e a demanda referentes a ela (nesse caso, o menor valor é a produção é 25). Em seguida, zere o valor da produção dessa linha (pois esse valor foi realocado na célula A3). A demanda da região do Local 3 passa a ser 45, que é a diferença entre o valor da demanda dessa localidade e a produção (alocada) da Filial A ($70 - 25 = 45$).

Quadro 41 – Cálculo utilizando o método de Vogel I

Para De	Local 1	Local 2	Local 3	Produção	Subtração entre os menores valores de tempo/custo presentes em cada linha	Menor valor de tempo unitário presente na linha com a maior penalidade.
Filial A	36	41	25	25	$36 - 25 = 11$ (Maior Valor)	← Maior Penalidade
Filial B	28	25	31	60	$28 - 25 = 3$	
Filial C	26	30	41	55	$30 - 26 = 4$	
Demanda	30	40	70 45	140		
Subtração entre os menores valores de tempo/custo presentes em cada coluna	$28 - 26 = 2$	$30 - 25 = 5$	$31 - 25 = 6$			

Fonte: Elaboração própria, 2024.

2º passo – Considere apenas as células que ainda não tenham sido alocadas (nesse caso, todas presentes nas linhas das filiais A e B). Refaça, em todas as células dessas linhas, a subtração entre os dois menores valores de tempo unitário, conforme o pri-

meiro passo. Em seguida, identifique o maior resultado entre essas operações (penalidade). Esse valor pode estar em uma linha ou em uma coluna que ainda não tenha sido zerada. Se a penalidade estiver na linha, identifique o menor valor referente ao tempo unitário dessa . Caso esteja na coluna, identifique o menor valor de tempo unitário nela. Se, durante a análise, houver penalidades com valores iguais, escolha a que corresponder ao menor valor de tempo unitário. Se esses valores também forem iguais, a escolha pode recair sobre qualquer uma entre elas. Perceba que a maior penalidade no Quadro 42 é 10, valor originado da subtração entre os dois menores valores da coluna do Local 3 ($B3 - C3 = 41 - 31 = 10$). Nessa coluna, o menor valor de tempo unitário é 31, relativo à célula B3 (em comparação à C3, que é 41). Em seguida, aloque zero na célula C3. Em relação ao menor valor entre a produção e a demanda, tem-se que o valor da Demanda (45) é menor que o da Produção (60). Após isso, zere o valor da demanda ($45 - 45 = 0$) dessa coluna (pois esse valor foi realocado na célula B3). A produção da Filial B passa a ser 15, que é a diferença entre a produção da Filial B e o valor alocado em B2 ($60 - 45 = 15$).

Quadro 42 – Cálculo utilizando o método de Vogel II

Para De	Local 1	Local 2	Local 3	Produção	Subtração entre os menores valores de tempo/custo presentes em cada linha
Filial A	-	36	41	25	x
Filial B	28	25	45	60	28 - 25 = 3
Filial C	26	30	-	55	30 - 26 = 4
Demanda	30	40	70	45	140
Subtração entre os menores valores de tempo/custo presentes em cada coluna	28 - 26 = 2	30 - 25 = 5	41 - 31 = 10		

Menor valor de tempo unitário presente na coluna com a maior penalidade.

↑
Maior Penalidade

Fonte: Elaboração própria, 2024.

3º passo – Considere apenas as células que ainda não tenham sido alocadas (nesse caso, as células B1, B2, C1 e C2). No quadro 43, são dos tempos unitários (em marrom) em que não foram alocados valores (em vermelho). Identifique, entre os resultados das subtrações efetuadas no 1º passo, o maior valor (penalidade – cor rosa – Quadro 43). Se a maior penalidade estiver na linha, identifique o menor valor referente ao tempo unitário dessa. Caso esteja na coluna, identifique o menor valor de tempo unitário dessa coluna.

Perceba que a maior penalidade no quadro 43 é 5, valor originado da subtração entre os dois menores valores da linha da Filial C ($C2 - C1 = 30 - 25 = 5$). Nessa linha, o menor valor de tempo unitário é 25, relativo à célula B2 (em comparação C2, que é 30). Em seguida, aloque nessa célula (B2) o menor valor entre a produção e a demanda referentes a ela (nesse caso, o menor valor é a demanda = 15). Após isso, zere o valor da produção do Filial B (pois esse valor foi realocado na célula B2). A demanda do Local 2 passa a ser 25, que é a diferença entre o valor anterior da produção e o valor alocado em C1 ($40 - 15 = 25$).

Quadro 43 – Cálculo utilizando o método de Vogel III

Para De	Local 1	Local 2	Local 3	Produção	Subtração entre os menores valores de tempo/custo presentes em cada linha
Filial A	36 -	41 -	25 25	25	x
Filial B	28 -	15 15	25 45	60 45	28 - 25 = 3
Filial C	26 -	30 -	41 -	55 25	30 - 26 = 4
Demanda	30	25 25	40 70	45 140	
Subtração entre os menores valores de tempo/custo presentes em cada coluna	28 - 26 = 2	30 - 25 = 5	x		

Menor valor de tempo unitário presente na coluna com a maior penalidade.

Maior Penalidade

Fonte: Elaboração própria, 2024.

4º passo – Quando não for mais possível subtrair os menores valores das linhas e os menores valores das colunas de forma a se fazer comparações para detectar o maior e o menor resultado, aloque os valores nas células que ainda não foram alocadas (nesse caso, C1 e C2), com base em sua observação. Comece com o menor entre os valores das demandas ou das produções (demanda Local 1 = 30; produção Filial C = 55; demanda do Local 2 = 25). Portanto, inicialmente a célula C2, pois é a que possui o menor valor (25 referente à demanda do Local 2). Ao alocar o 25 na célula C2, zera a demanda do Local 2 (25-25=0) e a demanda do Local 2 passa de 25 para 0 (25 – 25 = 0). Observe que o valor da produção da Filial C era 55 e passa para 30, e o valor atual da demanda do local 1 com a operação anterior passou a ser 30. Assim, para que a célula C1 seja zerada, aloque esse valor (30) nessa célula. Então, zera-se a Produção de C e a demanda do Local 1.

Quadro 44 – Cálculo utilizando o método de Vogel IV

Para De	Local 1	Local 2	Local 3	Produção	Subtração entre os menores valores de tempo/custo presentes em cada linha
Filial A	36 -	41 -	25 25	25	x
Filial B	28 -	25 15	31 45	60 45 0	
Filial C	26 30	30 25	41 -	55 25 30 0	
Demanda	30 0	40 25 0	70 45	140	
Subtração entre os menores valores de tempo/custo presentes em cada coluna	x		x		

Fonte: Elaboração própria, 2024.

5º passo – Calcule o tempo mínimo total de distribuição das mercadorias nas três Localidades (1,2 e 3) para as três filiais (A, B e C).

Para isso, multiplique o tempo de cada célula (em marrom) pelo recurso nela alocado (em vermelha Quadro 44) e, em seguida, realize o somatório de todos esses itens. O traço, que aparece em algumas células no lugar do recurso alocado, indica que o valor ali é zero. Como qualquer valor multiplicado por zero é igual a zero, não são considerados para esse cálculo, cuja fórmula de minimização corresponde à função objetiva da PL:

Min (T) = $(25 \times 25) + (15 \times 25) + (45 \times 31) + (30 \times 26) + (25 \times 30) = 3.925$ minutos por dia

6º passo – Interprete os resultados do problema, com base no quadro 44:

A Filial A deve entregar:

- 25 unidades para o Local 3. Essa entrega leva 625 minutos, pois $25 \times 25 = 625$.

A Filial B deve entregar:

- 15 unidades para o Local 2. Essa entrega leva 375 minutos, pois $15 \times 25 = 375$.
- 45 unidades para o Local 3. Essa entrega leva 1.395 minutos, pois $45 \times 31 = 1.395$

A Filial C deve entregar:

- 30 unidades para o Local 1. Essa entrega leva 780 minutos, pois $30 \times 26 = 780$.
- 25 unidades para o Local 2. Essa entrega leva 750 minutos, pois $30 \times 25 = 750$

O tempo mínimo total gasto para distribuir todas as unidades pelo método de Vogel, ou seja, o tempo ótimo calculado por esse método, é de 3.925 minutos. Esse valor corresponde ao

somatório da entrega das mercadorias pelas Filial A (625 minutos), B ($375 + 1.395 = 1.770$ minutos) e C ($780 + 750 = 1.530$ minutos). Para a empresa familiar, não importa entre as filiais quem demorará mais ou menos tempo entregando a mercadoria, e sim o tempo total de entrega.

Então, a técnica de Vogel ofereceu um menor de distribuição de mercadorias (3.925 min) do que a do Custo Mínimo (**4.050**), com uma diferença de 125 minutos na entrega total das mercadorias.

4.2.5 Uso do software LINDO para Programação Linear

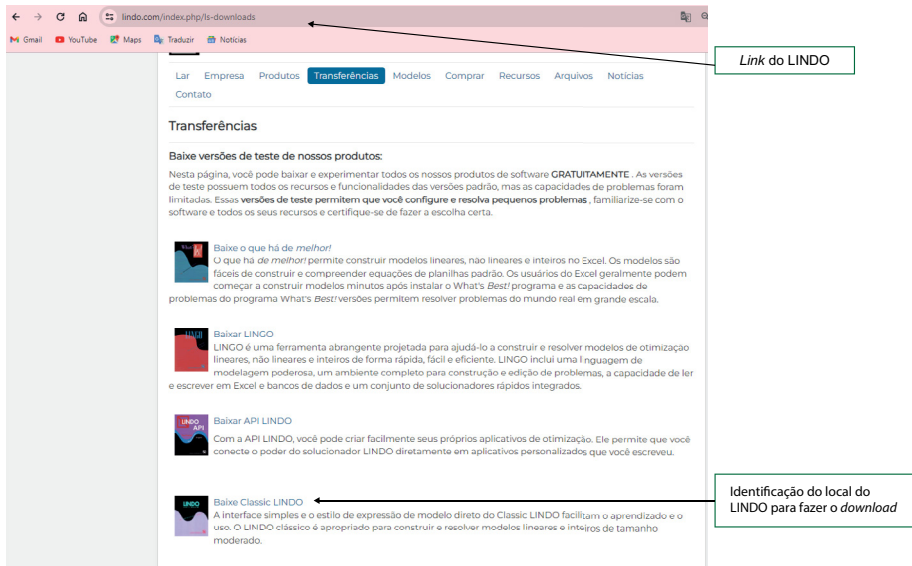
Curiosidade

“É um software livre que cria aplicações de otimização para a resolução de problemas de programação linear, inteira e não linear (quadrática). Foi desenvolvido pela LINDO Systems Inc., em Chicago, nos Estados Unidos no ano de 1979. As empresas utilizam esse software para maximizar o lucro e diminuir os custos, servindo de ferramenta para decisões que envolvam planejamento de produção, transporte, finanças, orçamento de capital, fusão, planejamento, estoque, alocação de recursos, entre outros.” (Pereira, 2010, p. 50).

O *Link* para o *download* do LINDO é <https://www.lindo.com/index.php/ls-downloads>. Chegando na página (Figura 34), clique na palavra “Baixe Classic LINDO”.



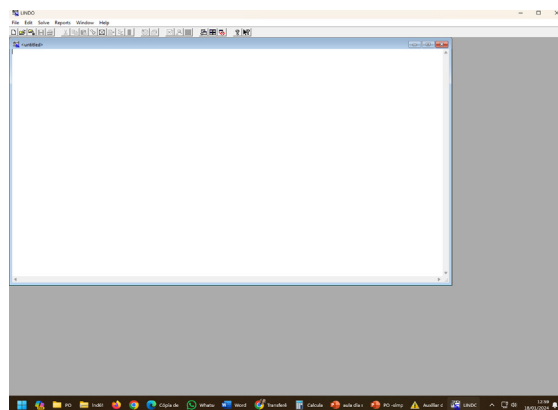
Figura 34 – Link e local onde realizar download do LINDO



Fonte: Elaboração própria, 2024.

Ao abrir o LINDO, a aparência é da Figura 35. Basta apenas digitar o modelo de PL desejador.

Figura 35 – Aparência do LINDO em formato de documento em branco



Fonte: Elaboração própria, 2024.

Vamos recapitular os modelos de PL anteriores de maximização e minimização para colocar no LINDO para cálculo com o uso de *software* e interpretações – Quadro 45.

Quadro 45 – Modelo de Maximização e Minimização do Exemplo do Capítulo 4

• 2 Variáveis de decisões, que são: • X_1 = Calça Comum • X_2 = Calça do Bolso Invisível • FO: $\text{Max}(L) = 120X_1 + 150X_2$ • Restrição Técnica: ◦ RT_1(quantidade e funcionários em unidade): $2X_1 + 4X_2 \leq 100$ ◦ RT_2 (espaço em m^2): $3X_1 + 2X_2 \leq 90$ ◦ RT_3 (quantidade de equipamentos em unidade): $5X_1 + 3X_2 \leq 120$ • Restrição de não negatividade: $X_1 \geq 0$ e $X_2 \geq 0$.	• 9 Variáveis de decisões, que são: • A1: Filial A para Local 1 • A2: Filial A para Local 2 • A3: Filial A para local 3 • B1: Filial B para Local 1 • B2: Filial B para Local 2 • B3: Filial B para Local 3 • C1: Filial C para Local 1 • C2: Filial C para Local 2 • C3: Filial C para Local 3 Min(T) = $36X_{A1} + 41X_{A2} + 25X_{A3} + 28X_{B1} + 25X_{B2} + 31X_{B3} + 26X_{C1} + 30X_{C2} + 41X_{C3}$ • Restrição técnica: • Produção da Filial A: $X_{A1} + X_{A2} + X_{A3} = 25$ • Produção da Filial B: $X_{B1} + X_{B2} + X_{B3} = 60$ • Produção da Filial C: $X_{C1} + X_{C2} + X_{C3} = 55$ • Demanda do Local 1: $X_{A1} + X_{B1} + X_{C1} = 30$ • Demanda do Local 2: $X_{A2} + X_{B2} + X_{C2} = 40$ • Demanda do Local 3: $X_{A3} + X_{B3} + X_{C3} = 70$ • Restrição de não negatividade: $X_{A1}; X_{A2}; X_{A3}; X_{B1}; X_{B2}; X_{B3}; X_{C1}; X_{C2}; X_{C3} \geq 0$
---	--

Fonte: Elaboração própria, 2024.

No LINDO, o modelo deve estar conforme quadro 46. A FO deve vir com Maximização e Minimização respectivamente, representando pela palavra Max e Min, entretanto não colocar o sinal de igual para separar a palavra “máx” (ou “mín”) dos números relativos às variáveis de decisão.

O sistema do LINDO não faz distinção entre letra maiúscula ou minúscula, bem como de espaçamento entre as variáveis. Outro aspecto é não colocar ponto nos números, a não ser que ele substitua uma vírgula. Qualquer palavra, frase ou texto colocado no LINDO, e que não tenha qualquer relação com a modelagem de PL, é necessário que se coloque o ponto de exclamação “!” ou o parêntese “()”.

O sistema deve reconhecer quando termina a função objetiva e quando começa a restrição técnica: digite, em uma linha, apenas a sigla “ST”, que significa *subject to* (sujeito a) em inglês. E nas restrições técnicas o sinal de $\leq$ é substituído por $\leq$, e o de $\geq$ é substituído por $\geq$.

Não é necessário colocar as restrições de não negatividade, pois o programa do LINDO é próprio para calcular a PL.

Para finalizar a modelagem, digite “END”, que significa “fim” em inglês.

Quadro 46 – Modelo de Maximização e Minimização do Exemplo do Capítulo 4 no LINDO

Modelo de Maximização da Empresa Familiar	Modelo de Minimização da Empresa Familiar
! Nome do aluno	! Nome do aluno
Max120X1 + 150X2	Min 36XA1 + 41XA2 + 25XA3 + 28XB1 + 25XB2 + 31XB3 + 26XC1 + 30XC2 + 41XC3
ST	ST
RT1) 2X1 + 4X2 <= 100!Quantidade e funcionários em unidade	Filial A) XA1 + XA2 + XA3 = 25
RT2) 3X1 + 2X2 <= 90!Espaço em m2	Filial B) XB1 + XB2 + XB3 = 60
RT3) 5X1 + 3X2 <= 120!Quantidade de Equipamentos em unidade	Filial C) XC1 + XC2 + XC3 = 55
END	Local 1) XA1 + XB1 + XC1 = 30
	Local 2) XA2 + XB2 + XC2 = 40
	Local 3) XA3 + XB3 + XC3 = 70
	END

Fonte: Elaboração própria, 2024.

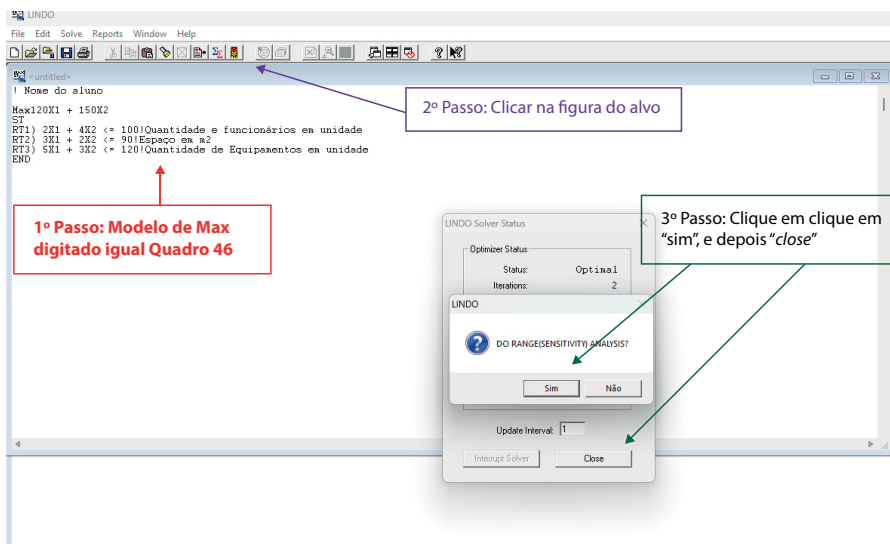
Vamos mostrar passo a passo (Pereira, 2010) como colocar a modelagem de Maximização no LINDO para realizarmos seus cálculos e interpretações. A fim de facilitar a visualização, cada etapa foi identificada na sua respectiva Figura com cores distintas. Então, iremos começar:

1º passo – No espaço em branco em que o cursor estiver piscando no LINDO, copie exatamente o Modelo de Maximização da Empresa Familiar do Quadro 46 (Lado esquerdo).

2º passo – Clique em “solve”, no menu principal, ou no botão com imagem de um alvo na barra de ferramentas (Figura 36).

3º passo -Ainda nessa tela (Figura 36), clique em “sim”, para que o sistema calcule automaticamente a análise de sensibilidade. Em seguida, clique no botão “close”.

Figura 36 – Cálculo do LINDO do modelo de Maximização - Exemplo da Empresa Familiar



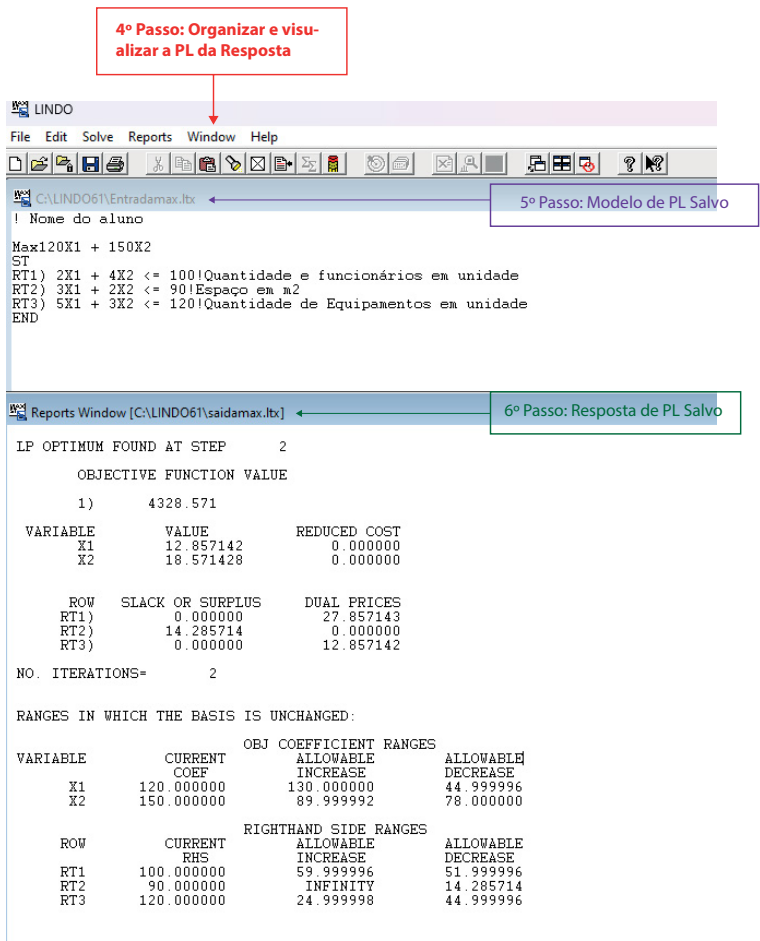
Fonte: Elaboração própria, 2024.

4º passo – Aparecerão duas janelas sobrepostas, para organizá-las, clique em “*window*”, em seguida em “*tile*” e, depois, em “*vertical*” (Figura 37). Ao clicar na palavra “*vertical*”, uma janela ficará ao lado da outra. Se optar por clicar em “*horizontal*”, as tabelas ficarão uma sobre a outra. No meu caso, optei pela horizontal. Na parte superior da tela do LINDO, tem a PL; e na inferior a resposta do problema (Figura 37).

5º passo – Para salvar as informações inseridas até aqui referentes apenas ao modelo do PL (parte superior da Figura 37), selecione a opção “*file*” no menu principal e, em seguida, “*save as*”. Selecione o diretório onde deseja salvar no seu computador, lembrando que o LINDO é um *software* usado para computador não tem como usá-lo em celular. Digite o nome do arquivo com até 10 letras e a extensão (.*ltx*) e, em seguida, clique no botão “OK”. Não se esqueça de apagar o asterisco (*). Lembre-se de que essa é a entrada do seu sistema, então salvei com o nome “Entradamax.ltx” nesse exemplo.

6º passo – Salve o resultado com o nome “saídamax”: no menu principal, clique em “*file*” e, em seguida, em “*save as*” (Figura 37). Afinal, o programa tem que ter dois salvamentos: a entrada que corresponde à modelagem e, depois, a saída, que é o resultado.

Figura 37 – Salvar no LINDO do modelo de Maximização (Entrada) e a Resposta (Saída)-Exemplo da Empresa Familiar



Fonte: Elaboração própria, 2024.

7º passo – Interprete os dados, buscando informações para a área de administração no caso do exemplo. Cada informação na janela do resultado da parte inferior do Quadro 38 será apresentada a seguir com a respectiva explicação na mesma cor, para facilitar a visualização.

LP OPTIMUM FOUND
AT STEP 2

LP OPTIMUM FOUND AT STEP – Número de vezes que foi feito o cálculo do SIMPLEX para chegarmos ao resultado positivo. No caso da Empresa Familiar, chegamos ao resultado na segunda tentativa – Ver item 4.2.2 (Explicação) – Quadros 25 e 33.

OBJECTIVE FUNCTION VALUE

1) 4328.571

OBJECTIVE FUNCTION VALUE – Valor encontrado na função objetiva. Corresponde ao lucro máximo, que no caso é de R\$ 4.328,57. Pode ser confirmado pelo quadro 33 do SIMPLEX e pela Figura 32 através do método gráfico.

VARIABLE – Identificação da variável de decisão X_1 (Calça Comum) e X_2 (Calça de Bolso Invisível).

VARIABLE

VALUE

REDUCED COST

X1

12.857142

0.000000

X2

18.571428

0.000000

REDUCED COST – Custo reduzido de uma variável. Corresponde ao valor em que o coeficiente da função objetiva da variável deve ser aumentado para que sua solução seja diferente de zero. Nesse caso, não há sugestão para aumento, pois ambas deram zero

VALUE – Valor das variáveis de decisão. $X_1 = 12,85$ significa que devem ser vendidos 13 Calças Comuns, e $X_2 = 18,57$ significa que devem ser vendidos 19 Calças de Bolso Invisível.

ROW

SLACK OR
SURPLUS

DUAL PRICES

RT1)

0.000000

27.857143

RT2)

14.285714

0.000000

RT3)

0.000000

12.857142

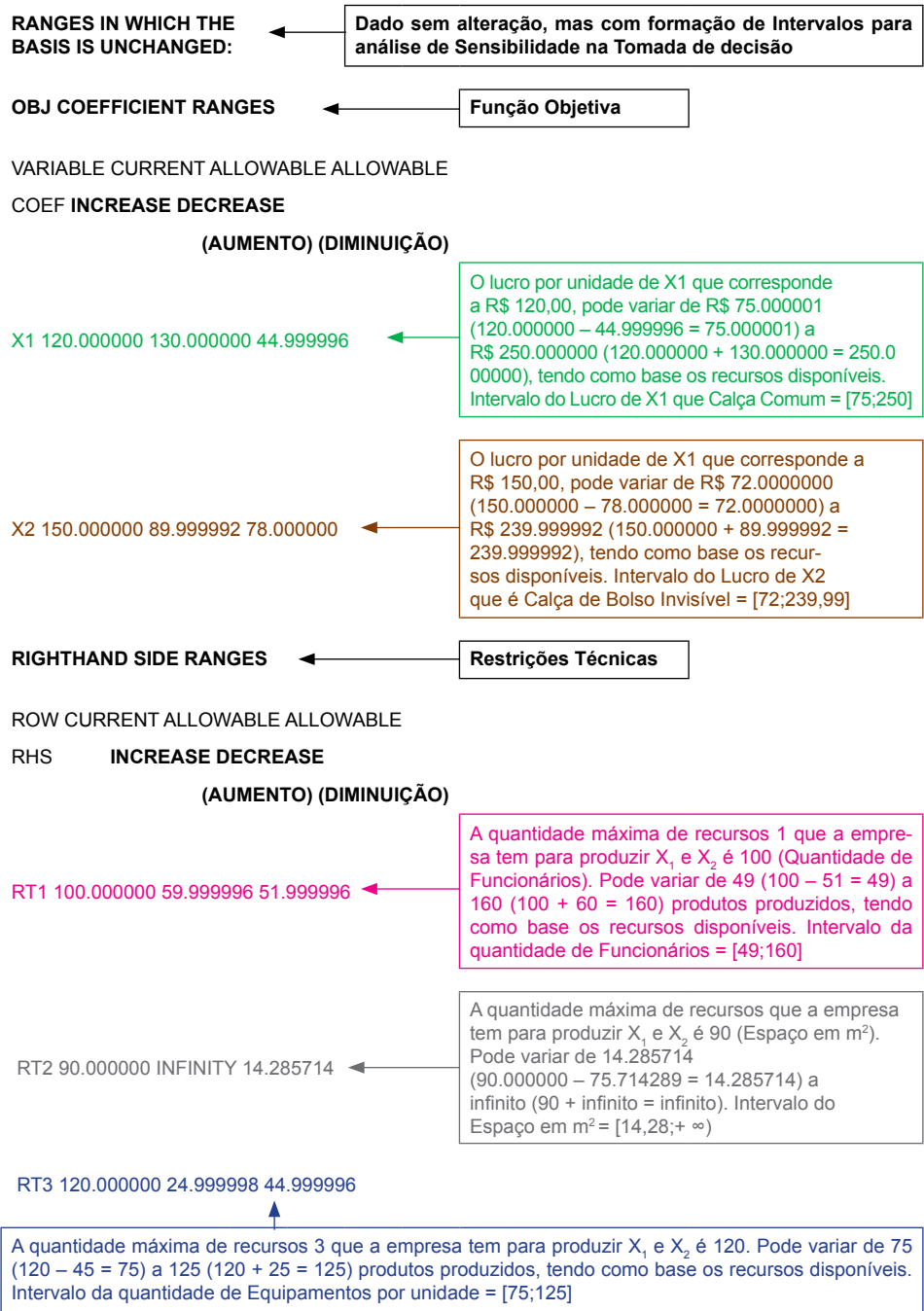
ROW – Resultado das linhas referentes às restrições técnicas

DUAL PRICES – Representa o aumento na função objetiva a cada unidade aumentada no recurso empresarial. Por exemplo, a primeira restrição técnica é $3X_1 + 2X_2 \leq 100$ (R1) e está na segunda linha da modelagem. Se o recurso aumentar de 100 para 101, a função objetiva aumentará mais R\$ 27.857143. O lucro passaria de R\$ 4328.571 para R\$ 4356.3281 (R\$ 4328.571 + R\$ 27.857143). Se o gestor preferir aumentar a demanda de R3 de 120 para 121, o lucro aumentará mais R\$ 12.85714, passando de R\$ 4328.571 para R\$ 4341.4281 (R\$ 4328.571 + R\$ 12.857142).

SLACK OR SURPLUS – Variável de folga. Corresponde ao valor que sobra para chegarmos aos recursos impostos na restrição referente a cada linha especificada. Assim, no caso da empresa familiar, há apenas a linha referente a $RT2(3X_1 + 2X_2 \leq 90)$, que corresponde ao espaço em m^2 , que sobrou 14,29 m^2 . Então, $3(12,8571429) + 2(18,571429) = 75,7142859$, ou seja $90 - 75,71 = 14,29 m^2$. Já RT1 (Quantidade e funcionários em unidade) e RT3(Quantidade de Equipamentos em unidade). Ambas deram zero, significando que usou todos os recursos, que são respectivamente 100 funcionários e 120 equipamentos.

LP OPTIMUM FOUND
AT STEP 2

LP OPTIMUM FOUND AT STEP – Número de vezes que foi feito o cálculo do SIMPLEX para chegarmos ao resultado positivo. No caso da Empresa Familiar, chegamos ao resultado na segunda tentativa – Ver item 4.2.2(Explicação) – Quadros 25 e 33



Observação:

- 1) Não há recursos e nem lucros infinitos, é necessário adequar o valor da análise da sensibilidade o mais próximo possível da realidade na hora da análise e tomada de decisão.
- 2) Caso deseje ver na PL dados estatísticos vistos nos capítulos 2 e 3, bem como o valor do SIMPLEX calculado de forma Tabular, basta ir na barra principal, clicar “*Report*” e depois “*Solution*”.
- 3) Se optar em fazer gráfico de linha das Restrições técnicas, basta ir na barra principal, clicar “*Report*” e depois “*Parametrics*”.
- 4) Se optar em fazer gráfico da Função objetiva, basta ir na barra principal, clicar “*Report*” e depois “*Peruse*”.

Fonte: Elaboração própria, 2024.

.....
Vídeo de Maximização do Exemplo da Empresa Familiar com as Interpretações (LINDO)



Link: <https://youtu.be/ocYFVEwcfUQ>

.....

Ao realizar o modelo de PL de minimização usando o LINDO do Quadro 46, tem-se a resposta abaixo. Basta seguir os mesmos passos do exemplo anterior de maximização.

LP OPTIMUM FOUND AT STEP 5

LP OPTIMUM FOUND AT STEP – Número de vezes que foi feito o cálculo do SIMPLEX para chegarmos ao resultado positivo. No caso da Empresa Familiar que possui três filiais para entregar em três locais distintos, chegamos ao resultado na quinta tentativa. Neste livro utilizamos O Custo Mínimo (Quadro 40) e o Vogel (Quadro 44)

OBJECTIVE FUNCTION VALUE

1) 3925.000

OBJECTIVE FUNCTION VALUE – Valor encontrado na função objetiva. Corresponde ao mínimo tempo, que, no caso, é de 3.925 min. Pode ser confirmado pelo quadro 44 do método de Vogel.

VARIABLE – Identificação da variável de decisão. Por ser um exemplo de alocação de recurso, a primeira letra após o X corresponde à Origem e a segunda, que no caso é um número, é o destino. Então, XA1 quer dizer que A Filial A leva para o local 1 e assim sucessivamente.

VARIABLE	VALUE	REDUCED COST
XA1	0.000000	21.000000
XA2	0.000000	22.000000
XA3	25.000000	0.000000
XB1	0.000000	7.000000
XB2	15.000000	0.000000
XB3	45.000000	0.000000
XC1	30.000000	0.000000
XC2	25.000000	0.000000
XC3	0.000000	5.000000

Reduced cost – Custo reduzido de uma variável. Corresponde ao valor a ser aumentado no coeficiente da função objetiva da variável para que sua solução seja diferente de zero.

Corresponde ao Valor das variáveis. Assim, somente identificarei as que não são zero. XA3 = 25; XB2 = 15;XB3 = 45;XC2 = 30 e XC2 = 25. Ao substituir os números acima nos seus respectivos lugares no Quadro 36, tem-se o Quadro 47, pois se refere a quantidade que deve ser levada. Ou seja, $Min(T) = (25 \times 25) + (15 \times 25) + (45 \times 31) + (30 \times 26) + (25 \times 30) = 3.925$ minutos por dia (mesma resposta e interpretação referente ao Quadro 44, feito manualmente).

Quadro 47 – Substituição dos números encontrados no LINDO, para ver alocação dos recursos

Para De	Local 1	Local 2	Local 3	Produção
Filial A	36 -	41 -	25 25	25
Filial B	28 -	25 15	31 45	60
Filial C	26 30	30 25	41 -	55
Demanda	30	40	70	140

Fonte: Elaboração própria, 2024.

Slack or surplus – Variável de folga. Corresponde à quanto sobra para chegar aos recursos impostos na restrição referente a cada linha especificada

ROW	SLACK OR SURPLUS	DUAL PRICES
FILIAL A)	0.000000	11.000000
FILIAL B)	0.000000	5.000000
FILIAL D)	0.000000	0.000000
LOCAL 1)	0.000000	-26.000000
LOCAL 2)	0.000000	-30.000000
LOCAL 3)	0.000000	-36.000000

Dual prices – Representa o aumento na função objetiva a cada unidade aumentada no recurso empresarial

NO. ITERATIONS = 5

RANGES IN WHICH THE BASIS IS UNCHANGED:

OBJ COEFFICIENT RANGES

Obj coefficient ranges – Valor máximo e mínimo da função objetiva.

VARIABLE CURRENT ALLOWABLE ALLOWABLE
COEF INCREASE DECREASE

XA1 36.000000 INFINITY 21.000000
XA2 41.000000 INFINITY 22.000000
XA3 25.000000 21.000000 INFINITY
XB1 28.000000 INFINITY 7.000000
XB2 25.000000 7.000000 5.000000
XB3 31.000000 5.000000 21.000000

XC1 26.000000 7.000000 INFINITY
XC2 30.000000 5.000000 7.000000
XC3 41.000000 INFINITY 5.000000

RIGHTHAND SIDE RANGES

Righthand side ranges – Valor máximo e mínimo da demanda e da produção.

ROW CURRENT ALLOWABLE ALLOWABLE
RHS INCREASE DECREASE
FILIAL A 25.000000 0.000000 0.000000
FILIAL B 60.000000 0.000000 0.000000
FILIAL D 55.000000 0.000000 0.000000
LOCAL 1 30.000000 0.000000 0.000000
LOCAL 2 40.000000 0.000000 0.000000
LOCAL 3 70.000000 0.000000 0.000000

Vídeo de Minimização do Exemplo da Empresa Familiar com as Interpretações (LINDO)



Link: https://youtu.be/w-2hx68f_I0

4.3 TRANSPORTE AVANÇADO E SUA RELAÇÃO COM PERT E CPM

Observamos que, no problema de Transporte anterior, há uma trajetória que se inicia em uma origem e existe um único caminho para o destino. Já nos problemas mais avançados de Transporte, há diversas opções de caminho entre uma origem e um destino.

Imagine que a mesma Empresa Familiar necessita levar as Calças do Ponto 1 até o ponto 5. Assim, os números (Figura 38) correspondem aos cinco locais (denominados de Nó) e as letras ao percurso da trajetória do caminho, conforme Figura 38:

- A é representado pelo conjunto de nós 1,2;
- B é representado pelo conjunto de nós 2,4;
- C é representado pelo conjunto de nós 4,5;
- D é representado pelo conjunto de nós 3,4;
- E é representado pelo conjunto de nós 1,3; e
- F é representado pelo conjunto de nós 3,5.

Assim, como no exemplo da questão anterior, o objetivo é atribuir uma duração ou um custo a cada tarefa para se determinar em quanto tempo ou custo é possível completar a atividade.

Para montar o PL desse sistema da Figura 38, devemos considerar que:

- FO deve representar o somatório do custo ou do tempo de transporte de todas as rotas possíveis.
- RT deve respeitar:
 - A saída de cada Nó;
 - O bi nessa questão todos serão zero, pois não teve produção ou demanda a serem levados como na questão anterior. Caso tivesse limitação de recursos no bi, como ocorreu na questão anterior para produção e demanda, seria necessário usar os sinais de $\geq$ (maior ou igual), $\leq$ (menor ou igual), $>$ (maior), $<$ (menor) ou $=$ (igual), conforme limitação de recurso existente;
 - Na RT, as setas para frente devem ter sinal (+), e caso tenham setas que retornem, o sinal da variável em RT é negativo. No exemplo da Figura 38 todas as setas são

para frente, não havendo de retorno, assim todas as variáveis de RT serão positivas. Entretanto, conforme a Figura 38, quando se recebe a seta o sinal é positivo, se transmitindo a seta, o sinal é negativo. A primeira RT de origem deve ter sinal (-), ou seja, ser multiplicada por (-1) e igual a -1. Por outro lado, a chegada, no caso o nó 6 na Figura 38, deve ser multiplicada por (1). Então, no exemplo da Figura 38, ficará (+) e igual a 1. Isso somente ocorrerá quando não tiver bi no primeiro nó de origem e no último nó de destino, conforme Quadro 47.

Para facilitar e compreender a montagem do LINDO da Figura 38 em relação à RT, foi feito o Quadro 47

Quadro 47 – Nó do exemplo Empresa familiar com seus respectivos números para colocar no LINDO

NÓ	X12	X24	X45	X56	X13	X35	X34
1	-1				-1		
2	1	-1					
3					1	-1	-1
4		1	-1				
5				-1		1	
6				1			

Fonte: Elaboração própria, 2024.

Diante da explicação da Figura 38, teremos o seguinte modelo a ser jogado no LINDO:

Min $3x_{12} + 5x_{24} + 4x_{45} + 3x_{56} + 7x_{13} + 2x_{35} + 3x_{34}$

ST

Nó 1) $-x_{12} - x_{13} = -1$

Nó 2) $x_{12} - x_{24} = 0$

Nó 3) $x_{13} - x_{34} - x_{35} = 0$

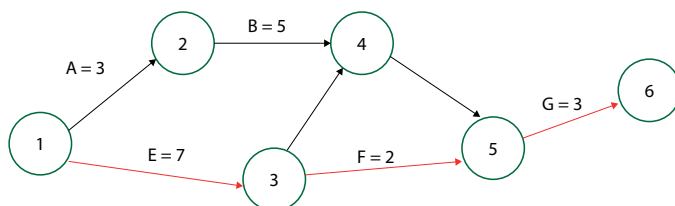
Nó 4) $x_{24} - x_{45} = 0$

Nó 5) $x_{35} - x_{56} = 0$

Nó 6) $x_{56} = 1$

END

Figura 38 – Exemplo de Transporte em Rede da Empresa Familiar



Fonte: Adaptado de Martins e Laugeni (2006, p. 419)

A resposta do LINDO foi:

LP OPTIMUM FOUND AT STEP 2

LP OPTIMUM FOUND AT STEP – Número de vezes que foi feito o cálculo do SIMPLEX para chegarmos ao resultado positivo (2)

OBJECTIVE FUNCTION VALUE

1) 12.00000

OBJECTIVE FUNCTION VALUE – Valor encontrado na função objetiva. Corresponde ao mínimo tempo da Trajetória da Figura 38, que, no caso, é de 12 minutos

VARIABLE VALUE REDUCED COST

X12 0.000000 0.000000

X24 0.000000 8.000000

X45 0.000000 4.000000

X56 1.000000 0.000000

X13 1.000000 0.000000

X35 1.000000 0.000000

X34 0.000000 10.000000

O melhor caminho do percurso é:

1 - 3:E = 7

3 - 5:F = 2

5 - 6:G = 3

O somatório desse valor é 12, achado na cor azul anteriormente. É identificado pelo número 1 na coluna Value. Assim, a Figura 38 mostra o caminho ótimo na cor vermelha; e este tem o seguinte somatório: $7 + 2 + 3 = 12$ min

Se fossemos fazer manualmente, veríamos que é possível três caminhos (Quadro 48):

- A-B-C-G, que teria uma duração de 15 minutos para a entrega;
- E-D-C-G, que teria uma duração de 17 minutos para a entrega (maior caminho);
- E-F-G, que teria uma duração de 12 minutos para a entrega (menor caminho identificado no cálculo do LINDO, cor azul anteriormente).

Quadro 48 – Cálculo do Caminho Crítico

Caminho	Cálculo da duração a partir valores trajetória Figura 49	Duração	
A-B-C-G	3+5+4+3	15	
E-D-C-G	7+3+4+3	17	Maior Caminho
E-F-G	7+2+3	12	Menor Caminho- Identificado no LINDO

Fonte: Elaboração própria, 2024.

Para calcularmos a variável de folga, pegaremos o valor do maior caminho e subtrairemos dos outros valores dos três caminhos do quadro 48, ou seja, é um tipo de desvio denominado amplitude (Quadro 49).

Quadro 49 – Cálculo da Folga dos 3 caminhos Quadro 48

Folga de cada caminho	Cálculo da Folga	Folga
A-B-C-G	17 - 15	2
E-D-C-G	17 - 17	0
E-F-G	17 - 12	5

Fonte: Elaboração própria, 2024.

Curiosidade

Esse tipo de técnica de transporte também pode ser utilizada para o planejamento e o controle de empreendimentos ou projetos, como com relações existentes entre as pessoas em rede social, em que cada nó é uma pessoa, sendo um tipo de relacionograma.

Durante a operação, é possível calcular isso através do PERT (*Program Evaluation and Review Technique*) ou CPM (*Critical Path Method*). No primeiro, a cada atividade se atribuem três durações distintas (Otimista, Provável e Pessimista). No segundo, há uma única duração para cada atividade e aplicar o algoritmo do caminho crítico.

Os dados mostrados anteriormente representam o cálculo do COM. Para realizarmos o PERT, montamos a partir da Figura 38 o Quadro 50 hipotético, em que colocamos que cada atividade possui três durações distintas: a Otimista, a Provável e a Pessimista. Os dados adquiridos das frequências referentes as três durações podem ser oriundas, por exemplo, de fontes secundárias, observações reais, uso de *softwares* (como Google Maps/Waze) ou prazos empresariais de entrega de projetos. Os dados das frequências das três durações devem levar em consideração a influência de fatores que podem interferir no cálculo, como por exemplo, congestionamento (que interfere na entrega), falta de algum recurso durante o processo, dentre outros

Quadro 50 – Três Durações das atividades

Atividades	Predecessores imediatos	Duração Otimista	Duração Provável	Duração Pessimista
A	não tem	2	3	5
B	A	3	5	7
C	B,D	2	4	6
D	E	2	3	5
E	não tem	5	7	9
F	E	2	2	3
G	C,F	1	3	4

Fonte: Elaboração própria, 2024.

A partir das informações do Quadro 50, foi montado o Quadro 51, em que a Duração esperada do evento corresponde à fórmula:

$$1/6(\text{Duração otimista} + 4(\text{Duração provável}) + \text{Duração Pessimista})$$

Já o desvio padrão de cada percurso é calculado pela fórmula:

$$\text{Desvio Padrão de cada Percurso} = (\text{Duração Pessimista} - \text{Duração Otimista}) / 6$$

Então, tem-se o Quadro 51 para cada trajetória, em que:

- **A** tem duração média de 3,17 minutos e o desvio padrão de 0,5. Somando e subtraindo a média aritmética pelo desvio padrão, conforme visto no capítulo 3, teremos um intervalo em que é possível realizar o trajeto entre [2,67 min;3,67].
- **B** tem duração média de 5 minutos e o desvio padrão de 0,7. Assim, seu intervalo para realizar o trajeto é entre [2,67 min;3,67].
- **C** tem duração média de 4 minutos e o desvio padrão de 0,7. Assim, seu intervalo para realizar o trajeto é entre [3,33 min;4,67].
- **D** tem duração média de 3,17 minutos e o desvio padrão de 0,5. Assim, seu intervalo para realizar o trajeto é entre [2,67min;3,67].
- **E** tem duração média de 7 minutos e o desvio padrão de 0,7. Assim, seu intervalo para realizar o trajeto é entre [6,33 min;7,67].
- **F** tem duração média de 2,17 minutos e o desvio padrão de 0,2. Assim, seu intervalo para realizar o trajeto é entre [2 min;2,33].
- **G** tem duração média de 2,83 minutos e o desvio padrão de 0,5. Assim, seu intervalo para realizar o trajeto é entre [2,33 min;3,33].

Quadro 51 – Tempo de Duração esperada da atividade e seu respectivo desvio padrão

Atividade	Duração esperada(min)	Desvio padrão(min)	Menor Intervalo Esperado (Min)	Maior Intervalo Esperado (Min)
A	3,17	0,5	2,67	3,67
B	5,00	0,7	4,33	5,67
C	4,00	0,7	3,33	4,67
D	3,17	0,5	2,67	3,67
E	7,00	0,7	6,33	7,67
F	2,17	0,2	2,00	2,33
G	2,83	0,5	2,33	3,33

Fonte: Elaboração própria, 2024.

Depois, podemos calcular Data mais cedo do início (DCI), correspondente ao valor da duração esperada anteriormente ao percurso em estudo. Assim, A e E não tem, é zero, pois não possuem nenhum trajeto antes dele conforme Figura 38. Entretanto, B possui e é A, em que seu valor da duração Esperada (Figura 51) é 3,17. Temos que fazer isso para todas as Trajetórias presentes no Quadro 52. Já a Data mais Cedo do Término (DCT) (conforme Quadro 52) corresponde:

O valor do percurso que vem antes + duração esperada

Assim, ao fazer o cálculo do DCI, do menor tempo admitido no percurso e do DCT maior tempo admitido no percurso, podemos encontrar o que sobra entre os dois, calculando a folga (Quadro 52).

Quadro 52 – Cálculo da Data mais cedo do início (DCI), Data mais cedo do término e da Folga

Atividade	Data mais cedo do início (DCI)	Data mais cedo do término (DCT)	Folga (Amplitude DCT-DCI)
A	0,00	3,17	$3,17 - 0 = 3,17$
B	3,17	$3,17 + 5 = 8,17$	$8,17 - 3,17 = 5,00$
C	8,17	$8,17 + 4 = 12,17$	$12,17 - 8,17 = 4,00$
D	7,00	$7 + 3,17 = 10,17$	$10,17 - 7 = 3,17$
E	0,00	$0 + 7 = 7,00$	$7 - 0 = 7,00$
F	7,00	$7 + 2,17 = 9,17$	$9,17 - 7 = 2,17$
G	12,17	$12,17 + 2,83 = 15,00$	$15 - 12,17 = 2,83$

Fonte: Elaboração própria, 2024.

Outra forma de chegar ao cálculo dos três caminhos encontrados no Quadro 49 é fazendo o cálculo do Quadro 53, só que dessa vez é realizado o somatório a partir da duração esperada no Quadro 51

Quadro 53 – Outra forma de calcular as três durações das atividades

Caminho	Cálculo da duração Esperada Quadro 51	Duração	
A-B-C-G	$3,17 + 5 + 4 + 2,83$	15	
E-D-C-G	$7 + 3,17 + 4 + 2,83$	17	Maior Caminho
E-F-G	$7 + 2,17 + 2,83$	12	Menor Caminho – Identificado no LINDO

Fonte: Elaboração própria, 2024.

O valor do desvio padrão da amostra é de 1,5 minuto, com uma variância de 2,1 minutos (Quadro 54). Para chegar a esse resultado, é preciso pegar a 3ª coluna do Quadro 51 e elevar os valores ao quadrado. O somatório disso é a variância, e a raiz da variância é o desvio padrão (conforme capítulo 3)

Quadro 54 – Valor da variância e do desvio da amostra

Atividade	desvio padrão	(desvio padrão) ²
A	0,5	0,3
B	0,7	0,4
C	0,7	0,4
D	0,5	0,3
E	0,7	0,4
F	0,2	0,0
G	0,5	0,3
Total (Variância)		2,1
Raiz do valor da Variância		1,5

Fonte: Elaboração própria, 2024.

Observação: Há vários programas para trabalhar CPM e PERT, como por exemplo Ms Project e o Gephi. Fiz a opção do Gephi por ser um *software* livre, como o LINDO, e melhor para visualização dos Nós e Caminhos, inclusive fazendo cálculos estatísticos mais detalhados para o uso em redes sociais e relacionograma. O Ms Project é melhor para atividades relacionadas a Projetos.

Uma forma de ver os Nós e trajetória é através do *software* Gephi (*link* para *download*: <https://gephi.org/users/download/>). Vamos supor que um dos fornecedores do jeans da calça comum e da calça do bolso invisível faça justamente o trajeto da Figura 38. O Quadro 55 mostra as principais características referente à entrega, mas poderiam ser inseridas outras colunas para avaliar outros atributos importantes. Para colocar no Gephi, é necessário elaborar duas planilhas diferentes. A primeira deve ter o *ID* (Identificador), que corresponde à numeração da quantidade de Nós, seguida da identificação do Nó (Coluna do *Label*). Depois, é preciso colocar em quantas colunas o pesquisador achar necessário em relação

aos atributos do estudo, sempre escrevendo na planilha as palavras com a mesma fonte e com a mesma quantidade de espaço para o programa não fazer a leitura errada.

Quadro 55 – Atributos da empresa de entrega de matéria prima a partir da Figura 38 para uso do Gephi (1ª Planilha de Excel)

ID é a identificação de cada elemento obrigado a ter

ID	Label	Atributos do Processo de Entrega
1	Local 1	Matéria Prima
2	Local 2	Produção Calça Comum
3	Local 3	Produção Calça Bolso Invisível
4	Local 4	Demanda1
5	Local 5	Demanda2
6	Local 6	Demanda3

Label nome dos elementos conectados – obrigado a ter

Fonte: Elaboração própria, 2024.

Já a segunda coluna de Excel deve se identificar na 1ª coluna a origem (*Source*) e na segunda coluna o destino (*Target*), sempre a partir do *ID* da primeira planilha. Na terceira coluna, da palavra *Type*, para dizer que os Nós tem uma direção, usar a palavra “*Directed*”; na quarta coluna, coloca-se o valor do caminho ou o peso ponderado do mesmo (geralmente 1), e na última coluna a letra desse caminho, conforme Quadro 56.

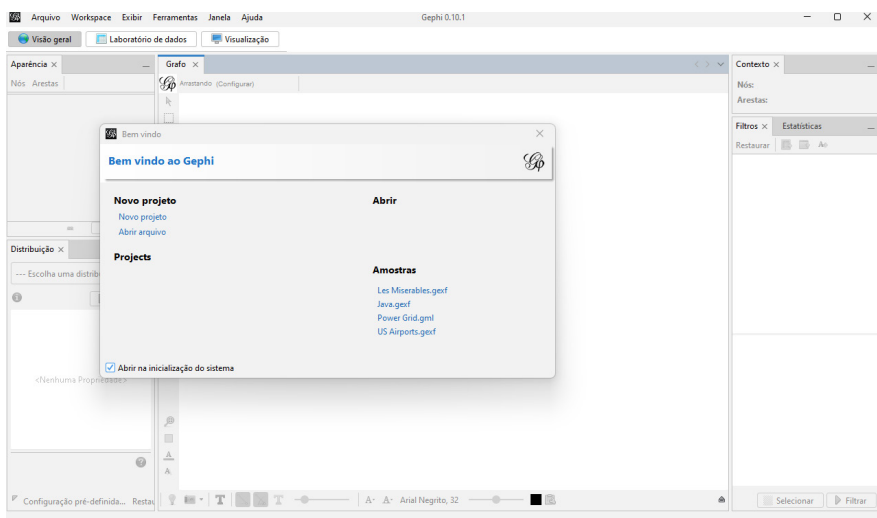
Quadro 56 – Identificação dos Nós e Arestas (caminho) para uso no Gephi (2ª planilha)

Source	Target	Type	weight	Letra
1	2	Directed	3	A
1	3	Directed	7	E
2	4	Directed	5	B
3	5	Directed	2	F
3	4	Directed	3	D
4	5	Directed	4	C
5	6	Directed	3	G

Fonte: Elaboração própria, 2024.

Depois de Instalar o *Gephi*, ele tem essa aparência da Figura 39. Clique em novo projeto. Vamos ver agora o passo a passo para colocar as duas planilhas.

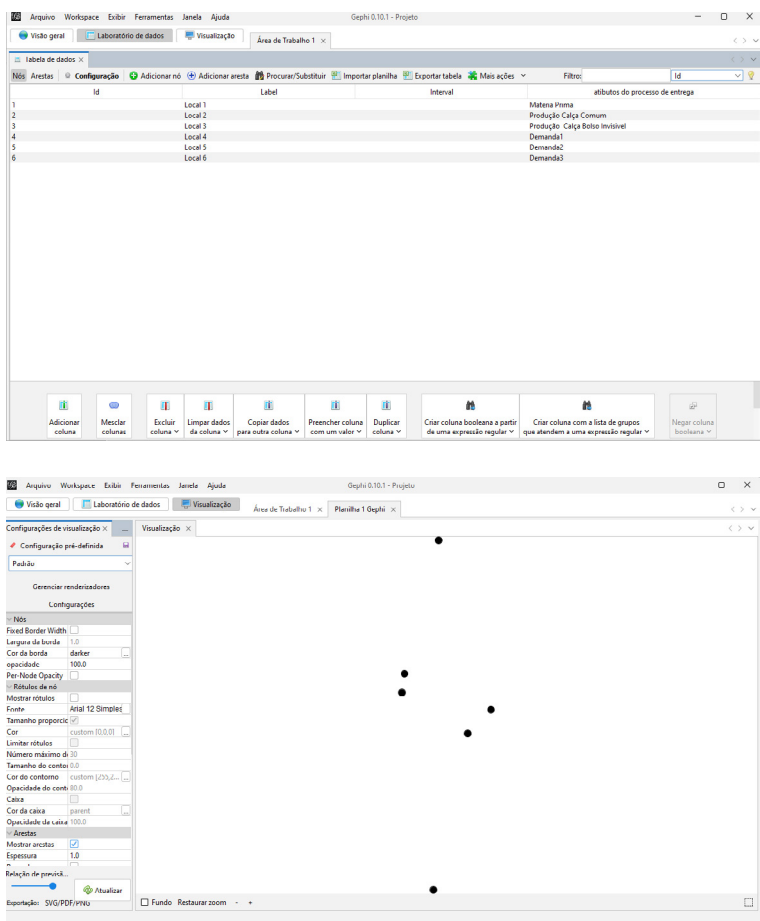
Figura 39 – Aparência do Gephi



Fonte: Elaboração própria, 2024.

1º passo – Clique “Laboratório de Dados”, depois Nós, e depois Importar planilha. Busque no seu computador dos dados do quadro 55 e basta aceitar os outros passos. A Figura 40 (parte superior) mostra os dados importados da 1ª planilha do Excel referente ao Quadro 55. Ao clicar no botão “Visualização” na Barra ferramenta principal, tem-se a Figura 40 (parte inferior) dos seis pontos da trajetória.

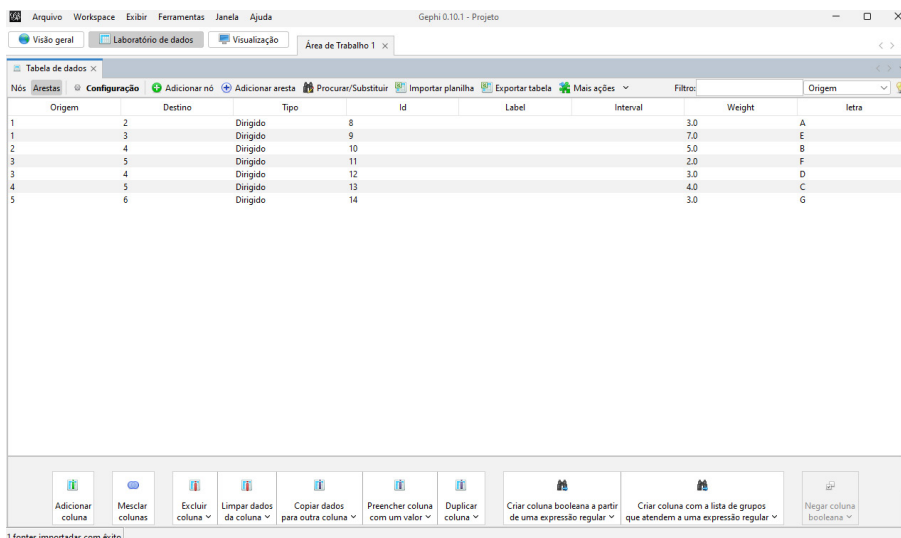
Figura 40 – Importação do Gephi primeira planilha



Fonte: Elaboração própria, 2024.

2º passo- Clique “Laboratório de Dados”, depois Arestas, e depois Importar planilha. Busca no seu computador dos dados do Quadro 56 e basta aceitar os outros passos (Figura 41).

Figura 41 – Importação do Gephi segunda planilha



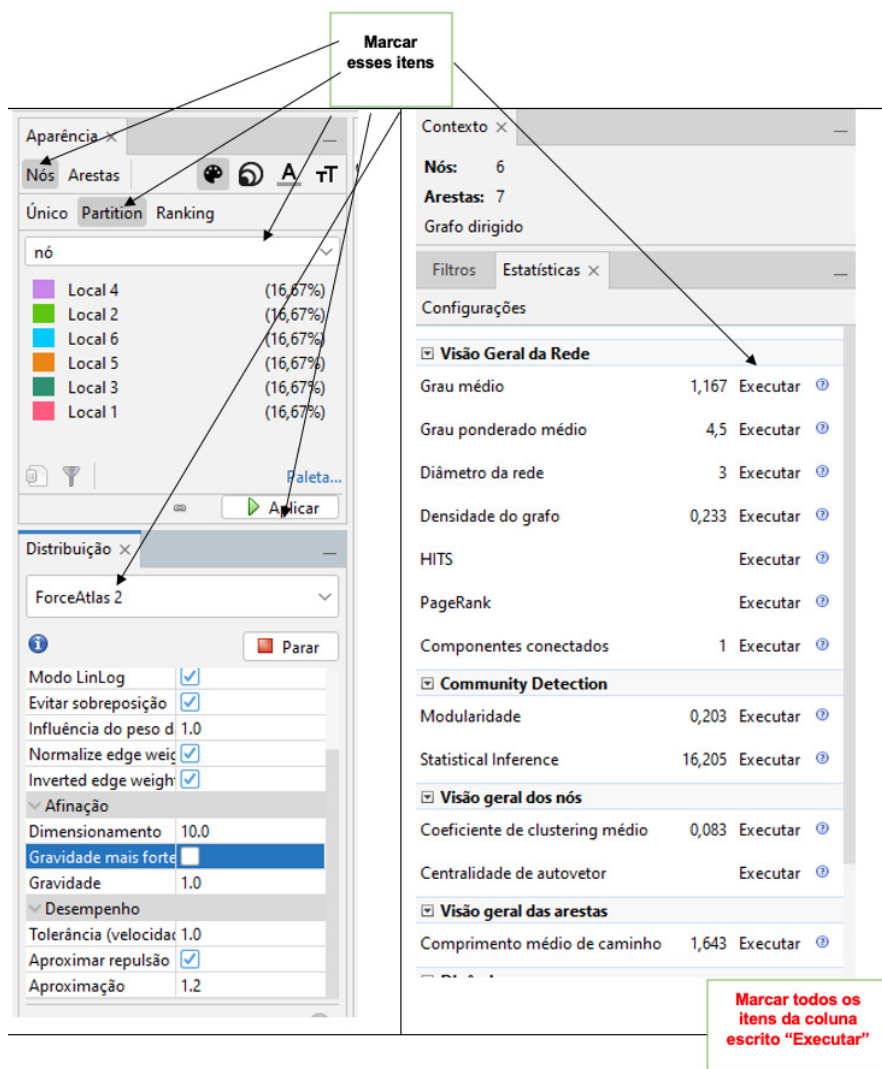
	Origem	Destino	Tipo	Id	Label	Interval	Weight	letra
1	2	Dirigido	8			3.0		A
1	3	Dirigido	9			7.0		E
2	4	Dirigido	10			5.0		B
3	5	Dirigido	11			2.0		F
3	4	Dirigido	12			3.0		D
4	5	Dirigido	13			4.0		C
5	6	Dirigido	14			3.0		G

Fonte: Elaboração própria, 2024.

3º passo- Vá em visualização Geral no menu principal (Figura 42), depois Nós-Partition e, dentro da barra de Partition, Nó e o botão. Aplicar na cor verde e no item Distribuição (ForceAtlas2 – Executar). Quando achar desenho visível, parar. Também marcar itens abaixo de ForceAtlas2, conforme Figura 42.

Do lado direito do Programa, mantendo a tela de Visão Geral do Gephi, basta clicar em Estatística e em executar em cada item estatístico presente no programa.

Figura 42 – Locais para clicar para fazer gráfico e cálculo estatístico



Fonte: Elaboração própria, 2024.

O Quadro 57 mostra as principais explicações estatística existente no programa Gephi para o exemplo da Figura 38.

Quadro 57 – Principais Explicações estatísticas do resultado do Gephi para o caso da empresa familiar

Métrica	Valor	Explicação
Visão Geral da Rede		
Grau Médio	1,167	Representa o número de conexão, que em média são os Nós que uma rede possui. Valor útil para comparar com outras redes, visto que não possuem uma distribuição em torno da média, e sim exponencial.
Grau Ponderado Médio	4,5	Representa o número de conexões ponderadas que os Nós de uma rede possuem em média. Assim, ver a intensidade dessa conexão, visto que as redes não possuem uma distribuição em torno da média e sim exponencial.
Diâmetro da Rede	3	- Centralidade de intermediação: Mede a frequência com que o Nó aparece no caminho mais curto entre os Nós da rede. - Centralidade de Proximidade: Distância média de um determinado Nó inicial para todos os demais da rede. - Excentricidade: Distância de um determinado Nó inicial até o Nó mais distante dele na rede. Representa a maior distância entre dois nós. Distância que corresponde ao número de conexões intermediárias existentes entre eles. Valor útil para comparar redes e avaliar uma certa medida, além da maior distância que seus nós devem percorrer para se conectarem.
Densidade do grafo	0,333	Corresponde à taxa de quantas conexões existem no grafo em relação a todas as possíveis, considerando que todos os nós estiveram ligados entre si diretamente. Medida útil para comparação de redes, em especial pequenas. No entanto, tende a ser baixa para a maioria das redes: quanto maior forem, mas baixa será, não sendo muito útil para análise.
Componentes conectados	1 (Fraco) 6 (Forte)	Determina o número de componentes conectados em uma rede. Esses correspondem a um conjunto de Nós no qual há pelo menos um caminho de conexão entre todos eles, que podem estar fracamente ou fortemente conectados.

continua...

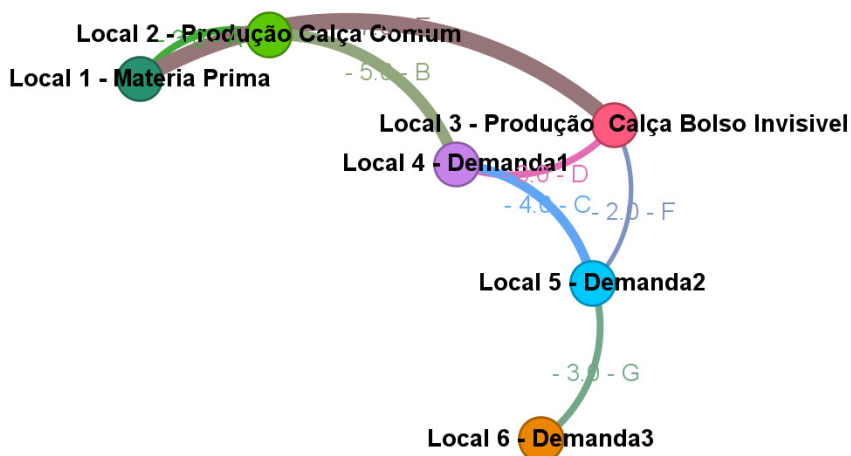
conclusão

Métrica	Valor	Explicação
Atributos comunidade		
Modularidade	0,203. Existem três comunidades conectadas	Mostra o quanto uma rede pode ser decomposta em comunidades modulares. Um alto valor indica uma rede com uma complexa estrutura de comunidade interna. É uma medida interessante para comparar diferentes redes se buscarmos mapear seus modos de organização em comunidades de Nós.
Estatística Inferencial	16,205. Existe uma comunidade conectadas	Calcula as chances de um evento acontecer com base em dados e um conhecimento prévio, através de Redes bayesianas.
Visão Geral dos Nós		
Coeficiente do Clustering Médio	0,083, com 12 interações	Determina o coeficiente médio em que os Nós de uma rede formam grupalidade entre si, ou seja, estão mais ou menos conectados e formando grupos de conexão direta. O Coeficiente Médio de Clustering é o valor médio dos coeficientes individuais.
Centralidade do Autovetor	0.07300610477809102 São 12	É uma medida de centralidade que considera as conexões de um Nó para considerar seu grau de importância. Assim esse item considera os nós que tem maior centralidade em relação a sua conexão com nós de menores valores de centralidade.
Visão Geral da Aresta		
Cumprimento Médio do Caminho	1,643	Representa o comprimento médio do caminho entre dois Nós em uma rede. Cálculo útil para difusão de informação ou recursos a partir de seus caminhos.

Fonte: Elaboração própria, 2024.

Uma das formas que ficou de resposta da Figura 38 foi através da Figura 43. O pesquisador pode formatar o desenho de várias, e essa é apenas uma das maneiras.

Figura 43 – Desenho da Rede no Gephi da Figura 38



Fonte: Elaboração própria, 2024.

Vídeo para elaborar o Desenho da Figura 39 a partir do *software* Gephy e os cálculos estatísticos



Link: <https://youtu.be/DJ5LtnTU3c>

O próximo capítulo são exercícios para treinar os assuntos referentes aos Capítulos 1, 2, 3 e 4. Mas antes, faremos umas atividades referentes aos conteúdos apenas desse capítulo.

4.4 ATIVIDADES PARA PRATICAR COM RESPECTIVAS RESPOSTAS

Faça um *quiz* com 13 perguntas sobre os conceitos básicos de Pesquisa Operacional e de Programação Linear, tanto de maximização como de minimização, no *link*: <https://www.goconqr.com/pt-BR/quiz/39648956/4-a-pesquisa-operacional-e-os-seus-modelos-estatisticos-capitulo-4-estatistica-passo-a-passo>



CASO 1 (ADAPTADO) – VENDA DE “CAUDAS DE SEREIA” RENDE FATURAMENTO DE ATÉ R\$ 60 MIL POR MÊS

A empresária investiu R\$ 140 mil no desenvolvimento do produto, que inicialmente foi pensado para o público infantil e conquistou adultos.

Primeiro, com a insistência da filha Giulia, Luana Schrappe precisou importar dos Estados Unidos a tal “cauda de sereia”. O produto chegou, Giulia se encantou, mas quando entrou na água descobriu que não dava para fazer todos os movimentos que queria. Luana percebeu uma oportunidade de negócio. Com a ajuda de uma arquiteta, ela começou a fazer protótipos até chegar ao modelo ideal. A missão era melhorar a performance na água e garantir a segurança das crianças.

O preço da fantasia de sereia infantil é de R\$ 329,00; com lucro unitário de R\$ 170,00 por peça. Já a fantasia de sereia adulta é de R\$ 400,00; com lucro unitário de R\$ 200,00. Percebe-se que produto foi inicialmente pensado apenas para o público infantil, mas acabou conquistando também mulheres adultas.

Ao mês, a empresa vende até 350 fantasias de sereia, tanto infantis como adultas. O mercado infantil corresponde a 60% das vendas, enquanto o

adulto é de 40%. A empresária investiu R\$ 140 mil no desenvolvimento do produto, que já está em processo de patente.

A invenção acabou rendendo outro tipo de negócio: para os fotógrafos, que agora fazem *book* da sereia. A empresa de Luana ainda nem completou um ano, mas as vendas já superaram as expectativas.

Pergunta: Qual o modelo de Programação Linear que representa o Caso 1? Após a elaboração do modelo, interprete os dados de forma a otimizar o estudo de caso empresarial.

Resposta: Identificar as variáveis que estão sempre ligadas ao que se quer maximizar ou minimizar.

	X1(Caudas infantis)	X2(Caudas adultas)	Bi
FO – max(L)	170	200	
RT1 (Demanda total)	0.6	0.4	< 350
RT2 (Demanda infantil)			< 210
RT3 (Demanda adulta)			< 140
Não negatividade	$X1 > 0$	$X2 > 0$	

1º PASSO: Identificar a função objetiva (FO)

Max(L) = 170x1 + 200x2

PRIMAL

FO = 170x1 + 200x2

RT1 = 0.6x1 + 0.4x2 ≤ 350

RT2 = 1x1 ≤ 210

RT3 = 1x2 ≤ 140

X1; X2 ≥ 0

DUAL

FO = 350y1 + 210y2 + 140y3

RT1 = 0.6y1 + 1y2 + 0y3 ≥ 170

RT2 = 0.4y1 + 0y2 + 1y3 ≥ 200

2º PASSO: Inverter o sinal das variáveis de decisão

Max(L) = -170x1 - 200x2

3º PASSO: Identificar o maior valor absoluto (independente do sinal). Para identificar a entrada.

Max(L) = -170x1 - 200x2

Entrada em X2

4º PASSO: Identificar a SAÍDA- Dividindo Bi pelo valor da entrada de cada restrição técnica.

RESTRIÇÕES	Bi	X2	Bi/valor de entrada
RT1 = 0.6x1 + 0.4x2 < 350	350	0.4	875
RT2 = x1 ≤ 210	210	0	0
RT3 = x2 ≤ 140	140	1	140 (Saída)

Saída = Menor valor não negativo do resultado da divisão.

5º PASSO: Acrescentar a variável de folga.

Depende do sinal de cada restrição técnica:

$\leq$: + XF $\geq$: - XF = Não vai haver variável de folga, variável auxiliar (A)

Sendo:

$RT1 = 0.6x1 + 0.4x2 \leq 350$
$RT2 = x1 \leq 210$
$RT3 = x2 \leq 140$

Então:

$$RT1 = 0.6x1 + 0.4x2 + XF1 = 350$$

$$RT2 = x1 + XF2 = 210$$

$$RT3 = x2 + XF3 = 140$$

6º PASSO: Organizar o sistema, identificando entrada e saída.

	Z	X1	X2	XF1	XF2	XF3	Bi
F0	1	-170	-200	0	0	0	0
RT1	0	0.6	0.4	1	0	0	350
RT2	0	1	0	0	1	0	210
RT3	0	0	①	0	0	1	140

1 = Elemento pivô

7º PASSO: Encontrar uma nova linha pivô: Dividir cada um dos seus elementos pelo número pivô.

	Z	X1	X2	XF1	XF2	XF3	Bi
Linha pivô	0	0	1	0	0	1	140
Linha/ elemento pivô	0	0	1	0	0	1	140

- Bi da nova linha pivô tem que ser igual ao valor de saída (nesse caso, é 140).

8º PASSO: Calcular as novas linhas com base na nova linha pivô.

1. Multiplicar cada elemento da nova linha pivô pelo valor de entrada da linha que se quer, com sinal invertido.
2. Escrever a linha original.
3. Somar cada resultado da multiplicação por cada elemento original.

NOVA LINHA FO:

Nova PIVÔ	0	0	1	0	0	1	140
PIVÔ X Valor de entrada de FO com sinal invertido (200).	0	0	200	0	0	200	28.000
F0	1	-170	-200	0	0	0	0
Soma	1	-170	0	0	0	200	28.000

NOVA LINHA RT1

Nova PIVÔ	0	0	1	0	0	1	140
PIVÔ X Valor de entrada de RT1 com sinal invertido (-0.4).	0	0	-0.4	0	0	-0.4	-56
RT1	0	0.6	0.4	1	0	0	350
Soma	0	0.6	0	1	0	-0.4	294

NOVA LINHA RT2

Nova PIVÔ	0	0	1	0	0	1	140
PIVÔ X Valor de entrada de RT2 com sinal invertido (0).	0	0	0	0	0	0	0
RT2	0	1	0	0	1	0	210
Soma	0	1	0	0	1	0	210

9º PASSO: Nova tabela com as linhas encontradas.

	Z	X1	X2	XF1	XF2	XF3	BI
F0	1	-170	0	0	0	200	28.000
RT1	0	0.6	0	1	0	-0.4	294
RT2	0	1	0	0	1	0	210
RT3	0	0	1	0	0	1	140

- As variáveis de decisão não podem conter valores negativos. Restrição de não negatividade: $X1 \geq 0$ e $X2 \geq 0$

10º PASSO: Identificar na nova FO o maior valor absoluto.

	Z	X1	X2	XF1	XF2	XF3	BI
F0	1	-170	0	0	0	200	28.000

Entrada em X1.

11º PASSO: Identificar a SAÍDA, Dividindo Bi pelo valor da entrada de cada restrição técnica da nova tabela.

Bi	X1	Bi/valor de entrada
294	0.6	490
210	1	210 (saída)
140	0	0

12º PASSO: Nova Linha Pivô

	Z	X1	X2	XF1	XF2	XF3	BI
F0	1	-170	0	0	0	200	28.000
RT1	0	0.6	0	1	0	-0.4	294
RT2	0	1	0	0	1	0	210
RT3	0	0	1	0	0	1	140

Elemento Pivô = 1

	Z	X1	X2	XF1	XF2	XF3	Bi
Linha pivô	0	1	0	0	1	0	210
Linha/ elemento pivô	0	1	0	0	1	0	210

13º PASSO: Calcular as novas linhas com base na nova tabela e na nova linha pivô .

NOVA LINHA FO:

Nova PIVÔ	0	1	0	0	1	0	210
PIVÔ X Valor de entrada de FO com sinal invertido (170).	0	170	0	0	170	0	35.700
F0	1	-170	0	0	0	200	28.000
Soma	1	0	0	0	170	200	63.700

NOVA LINHA RT1:

Nova PIVÔ	0	1	0	0	1	0	210
PIVÔ X Valor de entrada de RT1 com sinal invertido (-0,6).	0	-0,6	0	0	-0,6	0	-126
RT1	0	0.6	0	1	0	-0.4	294
Soma	0	0	0	1	-0,6	-0,4	168

NOVA LINHA RT3:

Nova PIVÔ	0	1	0	0	1	0	210
PIVÔ X Valor de entrada de RT3 com sinal invertido (0).	0	0	0	0	0	0	0
RT3	0	0	1	0	0	1	140
Soma	0	0	1	0	0	1	140

14º PASSO: Nova Tabela

	Z	X1	X2	XF1	XF2	XF3	Bi
F0	1	0	0	0	170	200	63.700
RT1	0	0	0	1	-0,6	-0,4	168
RT2	0	1	0	0	1	0	210
RT3	0	0	1	0	0	1	140

ANÁLISE

15º PASSO – Interpretar valores com a intenção de se obter informações para as decisões empresariais.

Temos uma função ótima, já que todos os valores das variáveis são positivos. As variáveis assumem o valor de Bi nas linhas que possuem o número 1, sendo assim, teremos:

- Variáveis não básicas: Não há, pois se encontrou o valor 1 em todas as variáveis.
- Variáveis básicas:

Z = 63.700. Esse será o lucro obtido com a venda das caudas por mês; X1 = 210 Venda mensal de caudas infantis;

X2 = 140 Venda mensal de caudas adultas. Variáveis de folga: XF1 = 168, XF2 = 210 e XF3 = 140

O lucro das caudas adultas é maior, porém a demanda máxima é de 40% do máximo de 350 caudas vendidas mensalmente. Portanto, devemos produzir as 140 caudas adultas (os 40% de 350) e o máximo de caudas infantis, que é de 210 (demanda de 60% de 350), obtendo-se o lucro:

$$FO: \text{Max}(L) = (170 \times 210) + (200 \times 140) = 63.700$$

Mesmo que o lucro das caudas adultas seja maior, é importante pensar que o principal público da empresa é o infantil, sendo o maior volume de

vendas. A organização poderá então investir em ações para alavancar a venda das caudas adultas e tornar esse produto mais atraente para as consumidoras realizando, como por exemplo, ensaios fotográficos e vendendo caudas com a proposta “tal mãe, tal filha”. Também é interessante criar ações para datas específicas, usando como exemplo o Carnaval, visto que muitas pessoas compram fantasias. Para ajudar nesse aumento e fazer com que mais pessoas comprem, poderá ser aplicado um desconto de 20% na segunda peça, estimulando o interesse dos clientes, principalmente no que se diz respeito aos que influenciam na compra de outras pessoas.

LINDO

MAX

$$170x_1 + 200x_2$$

ST

$$0.6x_1 + 0.4x_2 \leq 350$$

$$x_1 \leq 210$$

$$x_2 \leq 140$$

END

P OPTIMUM FOUND AT STEP 2

2

Número de vezes que foi feito o cálculo do Simplex para chegar ao resultado positivo.

OBJECTIVE FUNCTION VALUE

1) 63700.00

Lucro máximo

VARIABLE VALUE REDUCED COST

X1 210.000000 0.000000

X2 140.000000 0.000000

Variáveis de decisão

X1 = 210 (caudas infantis)

X2 = 140 (caudas adultas)

Restrições Técnicas

ROW SLACK OR SURPLUS DUAL PRICES

2) 168.000000 0.000000

3) 0.000000 170.000000

4) 0.000000 200.000000

Aumento do lucro a cada unidade aumentada.

Lucro X1 = 170

Lucro X2 = 200

NO. ITERATIONS =

2

SIMPLEX encontrou o resultado na segunda vez.

RANGES IN WHICH THE BASIS IS UNCHANGED:

OBJ COEFFICIENT RANGES

VARIABLE CURRENT ALLOWABLE ALLOWABLE

COEF	INCREASE	DECREASE	
X1	170.000000	INFINITY	170.000000
X2	200.000000	INFINITY	200.000000



Tendo como base os recursos disponíveis, o lucro pode variar de:

X1: Pode variar de 0 (170 - 170) a infinito (170 + infinito = infinito)

X2: O lucro pode variar de 0 a infinito.

RIGHTHAND SIDE RANGES

ROW CURRENT ALLOWABLE ALLOWABLE RHS INCREASE
DECREASE

2	350.000000	INFINITY	168.000000	
3	210.000000	(+) 280.000000	(-) 210.000000	
4	140.000000	420.000000	140.000000	



Tendo como base os recursos disponíveis, a demanda pode variar de: 2) Total: de 182 (350 - 168) a infinito (350 + infinito = infinito)

3) X1: De 0 a 490

4) X2: 0 a 560.

Caso 2 – Uma empresa realiza a compra de materiais que serão utilizadas nos telhados em três fornecedores, que encaminham para suas unidades na Bahia localizadas em A, B e C. Na unidade da A, a demanda é de 210 unidades mensais; em B, 40 unidades; e em C, 152 unidades. Já a capacidade de produção do Local 1 é de 200 unidades, do Local 2 é de 174 unidades e do Local 3 é de 126 unidades. Segue abaixo o tempo de envio dos produtos para cada unidade, e a Produção e Demanda mensais:

Para De	A	B	C	Produção
Local 1	85	66	93	200
Local 2	64	45	89	174
Local 3	92	58	71	126
Demanda	210	138	152	500

Pergunta-se :

- Qual o modelo de programação Linear para o LINDO?
- Qual o caminho que deve ser seguido ?

RESPOSTA

a) MODELO LINEAR

$$FO = 85X_{11} + 66X_{12} + 93X_{13} + 64X_{21} + 45X_{22} + 89X_{23} + 92X_{31} + 58X_{32} + 71X_{33}$$

$$RT1: X_{11} + X_{12} + X_{13} = 200$$

$$RT2: X_{21} + X_{22} + X_{23} = 174$$

$$RT3: X_{31} + X_{32} + X_{33} = 126$$

$$RT4: X_{11} + X_{21} + X_{31} = 210$$

$$RT5: X_{12} + X_{22} + X_{32} = 138 \quad RT6: X_{13} + X_{23} + X_{33} = 152$$

$$RNN: X_{11}; X_{12}; X_{13}; X_{21}; X_{22}; X_{23}; X_{31}; X_{32}; X_{33} \geq 0$$

b) Caminho a ser realizado – Inequação para colocar no LINDO

```

MIN 85X11 + 66X12 +93X13 +64X21 +45X22 +89X23 + 92X31 +58X32 +71X33
ST
X11 + X12 + X13 = 200! PRODUÇÃO MADEIREIRA 1
X21 + X22 + X23 = 174! PRODUÇÃO MADEIREIRA 2
X31 + X32 + X33 = 126! PRODUÇÃO MADEIREIRA 3
X11 + X21 + X31 = 210! DEMANDA IMBASSAI
X12 + X22 + X32 = 138! DEMANDA PRAIA DO FORTE
X13 + X23 + X33 = 152! DEMANDA ITACIMIRIM
END
    
```

```

LP OPTIMUM FOUND AT STEP      5

      OBJECTIVE FUNCTION VALUE
    1)      34668.00

      VARIABLE            VALUE            REDUCED COST
      X11      174.000000            0.000000
      X12       0.000000            0.000000
      X13      26.000000            0.000000
      X21      36.000000            0.000000
      X22      138.000000            0.000000
      X23       0.000000            17.000000
      X31       0.000000            29.000000
      X32       0.000000            14.000000
      X33      126.000000            0.000000

      ROW    SLACK OR SURPLUS    DUAL PRICES
    2)         0.000000           0.000000
    3)         0.000000          21.000000
    4)         0.000000          22.000000
    5)         0.000000         -85.000000
    6)         0.000000         -66.000000
    7)         0.000000         -93.000000

NO. ITERATIONS=         5
    
```

A Figura abaixo relata, na cor azul e a partir da interpretação do LINDO, o caminho que deve ser seguido de forma a se otimizar o recurso de tempo ou custo

Para De	A	B	C	Produção
Local 1	85 174	66	93 26	200
Local 2	64 36	45 138	89	174
Local 3	92	58	71 126	126
Demanda	210	138	152	500

A partir do caminho acima, é possível jogar no Gephy para ver as redes mais fortes e mais fracas.

5. EXERCÍCIOS DE ESTATÍSTICA COM SUAS PRINCIPAIS RESPOSTAS

OBJETIVOS:

- Praticar os conteúdos das atividades dos capítulos 1, 2, 3 e 4.
- Resolver e interpretar as questões para melhor compreensão dos assuntos dos capítulos 1, 2, 3 e 4.
- Internalizar os assuntos dos capítulos 1, 2, 3 e 4.

CONTEÚDOS PRÁTICOS COM AS DEVIDAS RESPOSTAS:

- a Estatística e suas etapas aplicada a Pesquisa.
- as Frequências e suas representações Estatísticas.
- a Medida de Tendência Central, Separatriz e Principais Desvios Estatísticos.
- a Pesquisa Operacional e os seus modelos Estatísticos.

Explicação. Objetivos e Conteúdos do Capítulo 5

Link: <https://youtu.be/nkjjwODRxow>



Este capítulo possui atividades para treinar os conteúdos dos capítulos 1, 2, 3 e 4. Cada questão tem sua respectiva resposta, pois a autora acredita que o conteúdo de Estatística não pode ser segmentado ou dividido. Até na construção dos capítulos de 1 a 4 há uma continuidade entre eles. Então, os exercícios seguirão o mesmo raciocínio.

QUESTÃO 1 – Um Pesquisador coletou dados em *site* governamental através de Pesquisa exploratória, conforme dados abaixo por estado em ROL decrescente. Como interpretar os dados?

- Distrito Federal — R\$ 2.475;
- São Paulo — R\$ 1.814;
- Rio Grande do Sul — R\$ 1.759;
- Rio de Janeiro — R\$ 1.723;
- Santa Catarina — R\$ 1.632;
- Paraná — R\$ 1.508;
- Mato Grosso do Sul — R\$ 1.488;
- Mato Grosso — R\$ 1.401;
- Espírito Santo — R\$ 1.347;
- Minas Gerais — R\$ 1.314;
- Goiás — R\$ 1.258;
- Rondônia — R\$ 1.169;
- Rio Grande do Norte — R\$ 1.077;
- Tocantins — R\$ 1.060;
- Ceará — R\$ 1.028;
- Sergipe — R\$ 1.028;
- Roraima — R\$ 983;

- Bahia — R\$ 965;
- Acre — R\$ 917;
- Pernambuco — R\$ 897;
- Amapá — R\$ 893;
- Paraíba — R\$ 892;
- Pará — R\$ 883;
- Piauí — R\$ 859;
- Amazonas — R\$ 852;
- Alagoas — R\$ 796; e
- Maranhão — R\$ 676.

Fonte dos dados: <https://economia.ig.com.br/2021-02-26/renda-media-no-brasil-foi-de-r-1380-em-2020-confira-os-maiores-rendimentos.html>



RESPOSTA – QUESTÃO 1

Pode-se calcular de forma agrupada ou não agrupada as frequências do estado. Também é possível organizar por Regiões do país. Assim, há várias formas de organizar os dados. Optei em colocar as respostas de forma agrupada por Estado.

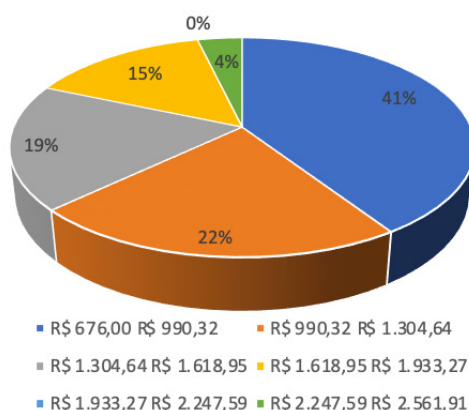
- Fazer Tabela fórmula de Sturges

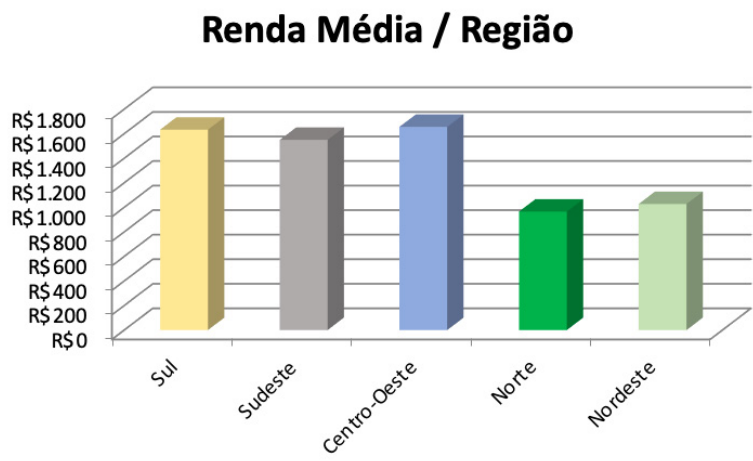
logn	At	k		Intervalo da Tabela
1,43	R\$ 1.799	5,7235	6	R\$ 314

- Elaborar Tabela com frequências

Intervalos		Frequência	Frq Rel (%)	FACA (+)	FACR (%)	FADA (-)	FADR (%)	XI
R\$ 676,00	R\$ 990,32	11	41%	11	41%	27	100%	833,159069
R\$ 990,32	R\$ 1.304,64	6	22%	17	63%	16	59%	1147,47721
R\$ 1.304,64	R\$ 1.618,95	5	19%	22	81%	10	37%	1461,79535
R\$ 1.618,95	R\$ 1.933,27	4	15%	26	97%	5	19%	1776,11349
R\$ 1.933,27	R\$ 2.247,59	0	0%	26	97%	5	19%	2090,43162
R\$ 2.247,59	R\$ 2.561,91	1	4%	27	100%	1	4%	2404,74976

- Elaboração de alguns gráficos para as interpretações





• Valores das Medidas de Tendência Central e dos Desvios

Intervalos		Frequência	XI	XI*FI	XI-X	DESVIO^2	(DESVIO)^2*FI
R\$ 676,00	R\$ 990,32	833,159069	9164,75	-384,167	147583,9874	1623423,861	
R\$ 990,32	R\$ 1.304,64	1147,47721	6884,863	-69,8485	4878,8095	29272,857	
R\$ 1.304,64	R\$ 1.618,95	1461,79535	7308,977	244,4697	59765,41638	298827,0819	
R\$ 1.618,95	R\$ 1.933,27	1776,11349	7104,454	558,7878	312243,808	1248975,232	
R\$ 1.933,27	R\$ 2.247,59	2090,43162	0	873,1059	762313,9844	0	
R\$ 2.247,59	R\$ 2.561,91	2404,74976	2404,75	1187,424	1409975,946	1409975,946	
Total		27		32867,79			4610474,978

MÉDIA ARITMÉTICA	1217,326	
MEDIANA		
EMD	13,5	
	R\$ 1.090,33	
MODA	R\$ 868	
ASSIMETRIA POSITIVA	0,82935559	
EQ1	6,75	
	R\$ 869	
EQ3	20,25	
	1508,94307	
DESVIO QUARTIL	320,033014	Esse desvio é o melhor para interpretação do desvio padrão
MD+/-DQ	R\$ 770,30	
	R\$ 1.410,36	
VARIÂNCIA	177325,9607	AMOSTRAL
Desvio padrão da amostra	421,1008913	RAIZ DA VARIÂNCIA
Erro padrão	81,04090432	DESVIO PADRÃO/RAIZ DE N
COEFICIENTE DE VARIÂNCIA	0,066572903	DESVIO PADRÃO/MÉDIA
MÉDIA+/- DESVIO PADRÃO	796,2247922	
Assimetria Positiva	1638,426575	

QUESTÃO 2 (Adaptada) – Empresa cria banheiro inteligente para *pets* e planeja crescer 40%. Desde 2016, foram vendidas mais de 40 mil unidades, que custam R\$ 474.

06/02/2022 – 07h22

Veja vídeo: <https://globoplay.globo.com/v/10273514/>



Desde 2016, foram vendidas mais de 40 mil unidades, que custam R\$ 474. A empresa que faz o banheiro inteligente para *pet* foi montada pelo casal Fábio Carvalho e Lara Lorga. Ele é engenheiro, e ela fez propaganda e *marketing*. Os dois tinham bons empregos, mas sonhavam com o negócio próprio. O primeiro modelo eles mesmos fizeram, em 2016. O investimento foi de 5 mil reais. Foi tanta procura que seis meses depois montaram a fábrica no interior de São Paulo. Eles fazem versões do produto com a mangueira para jogar a água fora pelo ralo ou com uma bolsa para coletar o líquido. Já venderam mais de 40 mil banheiros. As vendas são feitas pelo *site* e em lojas para *pets*. A estratégia é conquistar com as três vantagens do produto: a praticidade, a economia e a sustentabilidade. Como dura anos, o produto tem baixa recorrência de compra, então os empresários precisam sempre lançar novidades. Hoje a empresa tem linhas para cão, gato e roedores. O preço médio do banheiro é de 474 reais. Ano passado, a empresa teve retração nas vendas. Esse ano, os planos são investir quase um milhão de reais em divulgação nas redes sociais, crescer 40% e exportar.

O preço unitário de venda por hora é de R\$ 6,00 dos banheiros para gato e de R\$ 10,00 para cachorros. No galpão da empresa são utilizados três funcionários para trabalhar com banheiros

de gato e cinco funcionários para os de cachorro, não podendo ultrapassar quinze integrantes por hora na montagem de ambos os banheiros. Outro dado é que o material de estoque para a montagem é de 5 m² para os banheiros de gato e 2 m² para os de cachorro, não podendo ultrapassar 10m².

Mais informações sobre o caso Weasy – Telefone: (17) 99615-7235; E-mail: contato@weasy.com.br Site: www.weasy.com.br; Instagram: <https://www.instagram.com/weasy.br/>

Pergunta-se:

- a) Qual o modelo da PL para promover lucro?
- b) Elabore o SIMPLEX.
- c) Faça a representação gráfica.
- d) Interprete o resultado.
- e) A empresa anterior distribuiu os produtos em 3 origens para 4 locais, de forma semanal. Abaixo estão os custos de distribuição ao mês. Qual deve ser o programa de distribuição mensal de mercadorias?

Destino Origem	1	2	3	4	Produção
1	10	7	6	5	9000
2	2	8	9	1	10000
3	11	12	8	4	8000
Demanda	7000	6000	10000	4000	

RESPOSTA – QUESTÃO 2

A) $MAX(L) = 6X1 + 10X2$

ST

$3X1 + 5X2 \leq 15$

$5X1 + 2X2 \leq 10$

$X1;X2 \geq 0$

LINDO – modelo para colocar no programa

$MAX6X1 + 10X2$

ST

$3X1 + 5X2 \leq 15$

$5X1 + 2X2 \leq 10$

End

B)

- Cálculo pelo SIMPLEX

SIMPLEX – cálculo:

PL		B. Gatos “G”	B. Cães “C”	BI
FO	Max	6	10	
RT funcionários		3	5	≤ 15
RT estoque (m²)		5	2	≤ 10
RN		≥ 0	≥ 0	

$Max = 6G + 10C$

$Max - 6G - 10C$

Variável de entrada: C (maior valor absoluto)

Variável de saída:

PL	B. Gatos “G”	B. Cães “C”	BI	V. saída
RT funcionários	3	5	≤ 15	$15/5 = 3$
RT estoque (m ²)	5	2	≤ 10	$10/2 = 5$

Variável de Folga

$$3G + 5C + XF1 \leq 15$$

$$5G + 2C + XF2 \leq 10$$

PL	Z	G	C entrada	XF1	XF2	BI
FO	1	-6	-10	0	0	0
RT func	0	3	5 pivô	1	0	15 saída
RT estoque (m ²)	0	5	2	0	1	10

Linha pivô	0	3	5 pivô	1	0	15 saída
Nova linha pivô	0	0,6	1	0,2	0	3

Nova Tabela:

	Z	G	C	XF1	XF2	BI
1º linha	1	0	0	2	0	30
2º linha	0	0,6	1	0,2	0	3
3º linha	0	3,8	0	-0,4	1	4

Z = 30 valor – máximo da Função Objetiva

- Cálculo pelo LINDO

LP OPTIMUM FOUND AT STEP 1

OBJECTIVE FUNCTION VALUE

(L) 30.00000

VARIABLE VALUE REDUCED COST

X1 0.000000 0.000000

X2 3.000000 0.000000

ROW SLACK OR SURPLUS DUAL PRICES

2) 0.000000 2.000000

3) 4.000000 0.000000

NO. ITERATIONS = 1

RANGES IN WHICH THE BASIS IS UNCHANGED:

OBJ COEFFICIENT RANGES

VARIABLE CURRENT ALLOWABLE ALLOWABLE

COEF INCREASE DECREASE

X1 6.000000 0.000000 INFINITY

X2 10.000000 INFINITY 0.000000

RIGHTHAND SIDE RANGES

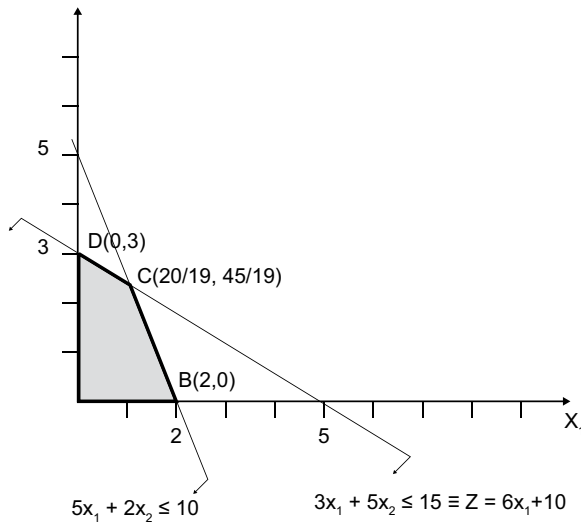
ROW CURRENT ALLOWABLE ALLOWABLE

RHS INCREASE DECREASE

2 15.000000 10.000000 15.000000

3 10.000000 INFINITY 4.000000

c) resposta via gráfico



d) INTERPRETAÇÃO

Conforme ponto ótimo-gráfico e o SIMPLEX, o maior lucro é de R\$ 30 e se deve trabalhar apenas com banheiros de cachorro na quantidade de 3. Trabalhará com os 15 funcionários, não existindo folga nessa variável. Entretanto, sobrá 4 m² dos 10 m² do estoque, que ficará ocioso. Observando o aspecto qualitativo, é necessário trabalhar com o banheiro de gato e o de cachorro.

e)

	D1	D2	D3	D4	PRODUÇÃO
Origem- 1	0	6000	3000	0	9.000
Origem – 2	7000	0	0	3000	10.000
Origem – 3	0	0	7000	1000	8.000
DEMANDAS	7.000	6.000	10.000	4.000	27.000
	D1	D2	D3	D4	PRODUÇÃO
Origem- 1	0	42000	18000	0	60.000
Origem – 2	14000	0	0	3000	17.000
Origem – 3	0	0	56000	4000	60.000
DEMANDAS	14.000	42.000	74.000	7.000	137.000 valor do custo total de distribuição

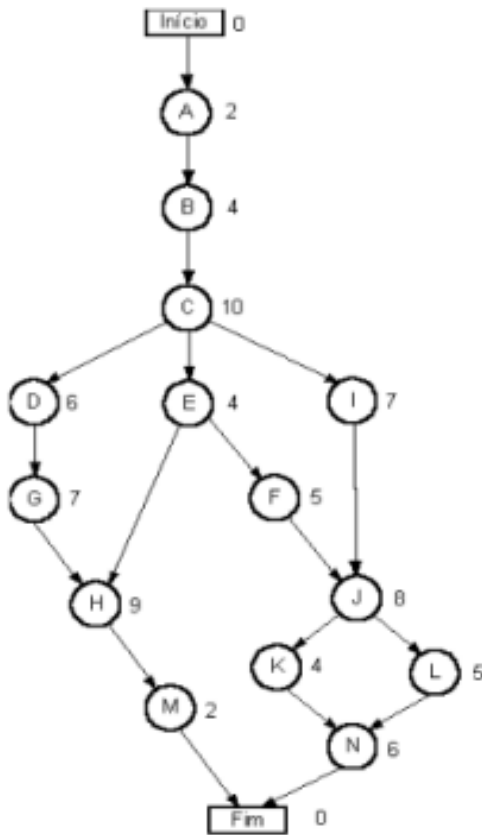
QUESTÃO 3 (Adaptada do Enem) – A empresa familiar do Banheiro Inteligente para *Pet* optou por abrir uma loja que tem o seguinte projeto abaixo. Descreva qual o caminho crítico.

Atividade	Descrição	Atividades Precedentes	Duração Estimada (semanas)
A	Escavação	-	2
B	Fundação	A	4
C	Paredes	B	10
D	Telhado	C	6
E	Encanamento Exterior	C	4
F	Encanamento Interior	E	5
G	Muros	D	7
H	Pintura Exterior	E,G	9
I	Instalação Elétrica	C	7
J	Divisórias	F,I	8
K	Piso	J	4
L	Pintura Interior	J	5
M	Acabamento Exterior	H	2
N	Acabamento Interior	K,L	6

RESPOSTA – QUESTÃO 3

As técnicas denominadas PERT e CPM foram independentemente desenvolvidas para o Planejamento e Controle de Projetos em torno de 1950, porém a grande semelhança entre elas fez com que PERT/CPM seja utilizado corriqueiramente como apenas uma técnica. Os termos são acrônimos de Program Evaluation and Review Technique (PERT) e Critical Path Method (CPM). PERT e CPM utilizam principalmente os conceitos de Redes (grafos) para planejar e visualizar a coordenação das atividades do projeto.

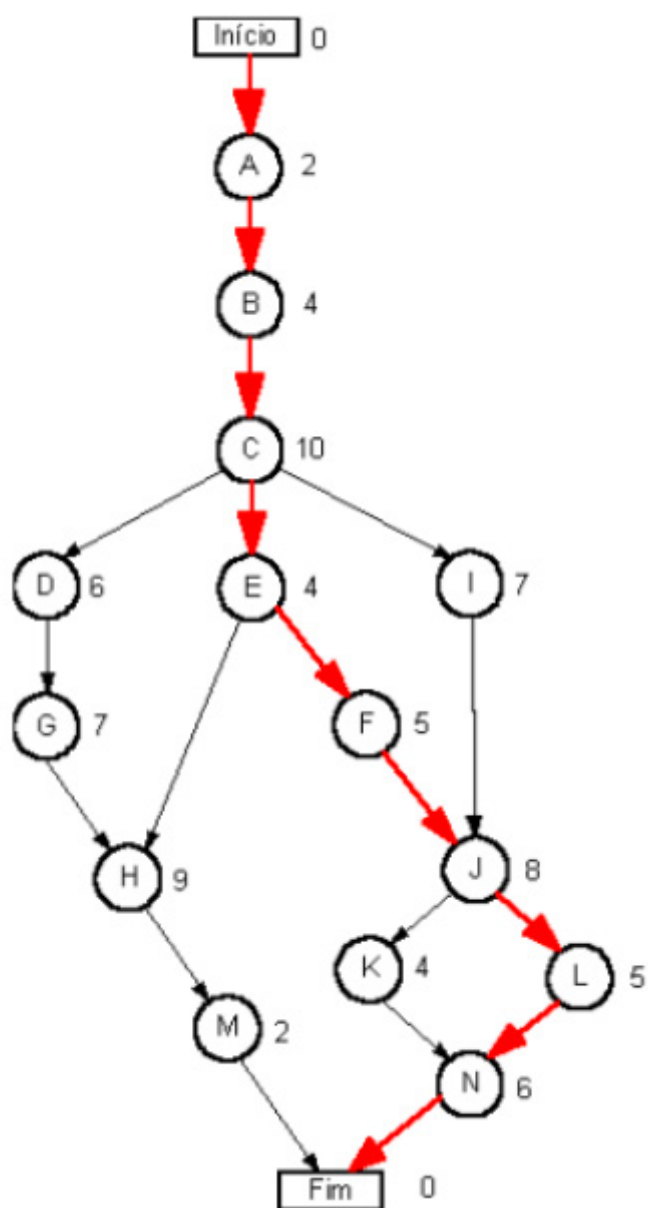
Lembrando que na construção, é necessário colocar os Nós, que correspondem às letras das Atividades, associados ao caminho, que são as semanas do projeto.



Na rede, foram detectados os seguintes Caminhos:

Tabela 2 - Caminhos e seus respectivos Comprimentos	
Caminho	Comprimento (semanas)
Início-A-B-C-D-G-H-M-Fim	$2 + 4 + 10 + 6 + 7 + 9 + 2 = 40$
Início-A-B-C-E-H-M-Fim	$2 + 4 + 10 + 4 + 9 + 2 = 31$
Início-A-B-C-E-F-J-K-N-Fim	$2 + 4 + 10 + 4 + 5 + 8 + 4 + 6 = 43$
Início-A-B-C-E-F-J-L-N-Fim	$2 + 4 + 10 + 4 + 5 + 8 + 5 + 6 = 44$
Início-A-B-C-I-J-K-N-Fim	$2 + 4 + 10 + 7 + 8 + 4 + 6 = 41$
Início-A-B-C-I-J-L-N-Fim	$2 + 4 + 10 + 7 + 8 + 5 + 6 = 42$

O caminho de maior comprimento é o caminho crítico, uma vez que todos os demais devem alcançar o Nó final antes dele. Assim, o tempo final requerido é de 44 semanas para finalizar o projeto. As atividades sobre esse caminho são as atividades críticas, e qualquer atraso em uma delas com certeza afetará a duração para um tempo maior do projeto. Já as outras atividades poderão ou não atrasar o projeto. A figura abaixo mostra na cor vermelha a trajetória do caminho crítico, que devem ser bem monitorados.



REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE ESTATÍSTICA (ABE). São Paulo, 2024. Disponível em: <https://www.redeabe.org.br/>. Acesso: 3 jan. 2024.

AKSENEN, Elisângela Zarpelon; MARTINELLI, Paulo. Estatística aplicada à gestão. Curitiba: Aymará, 2009.

ARENALES, Marcos et al. Pesquisa Operacional. 2a. ed. São Paulo: Elsevier, 2007.

BARBETTA, Pedro Alberto. Estatística aplicada às Ciências Sociais. 8a. ed. Florianópolis: Editora da UFSC, 2012.

BRUNI, Adriano Leal. Estatística para concurso. São Paulo: Atlas, 2008.

CRESPO, Antônio Arnot. Estatística fácil. 19a. ed. São Paulo: Saraiva, 2009.

BRASIL. Ministério da Educação. Fundo de Financiamento ao Estudante do Ensino Superior (FIES). Brasília, DF, 2010. Disponível em: <https://sis-fiesportal.mec.gov.br/>. Acesso em: 18 jan. 2024.

KUBY, Patricia; JOHNSON, Robert. ESTAT. São Paulo: Cengage Learning, 2013.

BRASIL. Instituto Brasileiro de Geografia e Estatística (IBGE). Brasília, DF, 2018. Disponível em: <https://www.ibge.gov.br/>. Acesso em: 18 jan. 2024.

KREJCIE, Robert V.; MORGAN, Daryle W. Determining Sample Size for Research Activities. Educational and Psychological Measurement, Texas, v. 30, p. 607-610, 1970. Disponível em: <https://bit.ly/3YE3gGm>. Acesso em: 14 ago. 2024.

LACHTERMACHER, Gerson. Pesquisa operacional na tomada de decisões. 5a. ed. Rio de Janeiro: LTC, 2016.

LAPONI, Juan Carlos. Estatística usando o excel. São Paulo: Elsevier, 2005

MARTINS, Petronio Garcia; LAUGENI, Fernando. Administração da Produção. São Paulo: Saraiva, 2006

MATTAR, Fauze Najib. Pesquisa de Marketing: Metodologia, Planejamento. São Paulo: Atlas, 2013.

MOREIRA, Daniel Augusto. Pesquisa operacional: curso introdutório. São Paulo: Thomson, 2007.

PASSOS, Eduardo José Pedreira Franco dos. Programação linear como instrumento da pesquisa operacional. São Paulo: Atlas, 2008.

PEREIRA, Aliger dos Santos. Clusters de veículo em Salvador: geoprocessamento e gestão de negócio para micro, pequenas e médias empresas (MPMEs). Salvador: EDUFBA; EDUNEB, 2019.

PEREIRA, Aliger dos Santos. Pesquisa Operacional. Curitiba: Aymará, 2010.

PEREIRA, Aliger dos Santos (Org.). Hospital Santa Izabel: Relatório de Gestão Hoteleira com o Uso de Modelos Matemáticos. Salvador: Universidade do Estado da Bahia, 2018.

PEREIRA, Aliger dos Santos. Inteligência competitiva: micro, pequenas e médias empresas Salvador (Brasil) e Bragança (Portugal). Salvador, EDUNEB, 2024. Disponível em: <https://saberaberto.uneb.br/items/3e77d-c2f-e9cf-48b8-bd44-9b786e52d514>. Acesso em: 24. fev. 2025

PEREIRA, Aliger dos Santos (Org.). Organização, sistemas e métodos: estudo de caso micro e pequenas empresas de Salvador. Salvador: UNEB, 2020.

PRADO, Darci. Programação linear. Belo Horizonte: Instituto de Desenvolvimento Gerencial, 2008.

RODRIGUES, Tatiane Daby de Fatima Faria; OLIVEIRA, Guilherme Saramago de; SANTOS, Josely Alves dos. As pesquisas qualitativas e quantitativas na educação. *Revista Prisma*, Rio de Janeiro, v. 2, n. 1, p. 154-174, 2021. Disponível em: <https://bit.ly/3SNYz9h>. Acesso em: 18 jan. 2024.

SAMARA, Beatriz Santos; BARROS, José Carlos De. Pesquisa de Marketing: Conceitos e Metodologia. São Paulo: Pearson Universidade, 2006.

SHARPE, Norean; DE VEAUX, Richard D.; VELLEMAN, Paul F. Estatística aplicada: administração, economia e negócio. Porto Alegre: Bookman, 2011

SILVA, Ermes Medeiros da et al. Pesquisa operacional para os cursos de economia, administração e ciências contábeis. São Paulo: Atlas, 2007.

TAVARES, Marcelo. Estatística aplicada à administração. Florianópolis: EduCAPES, 2014.

TRIBUNAL DE CONTAS DO ESTADO DO ESPÍRITO SANTO. Escola de Contas Públicas. Capítulo 3 – Gráficos. Vitória, 2024. Disponível em: <https://ava.tcees.tc.br/mod/book/view.php?id=8705&chapterid=1187>. Acesso em: 18 jan. 2024.

