

NOME:

TURNO:

PROFESSOR: CLEBER PINHEIRO

DISCIPLINA: ANÁLISE ESTATÍSTICA DE DADOS

SEMESTRE:

DATA:

EXERCÍCIOS DE APRENDIZAGEM

2ª Lista

QUESTÃO 1

O teste de hipótese é uma ferramenta essencial para verificar as afirmações e suposições em análises estatísticas. Como exemplo de aplicação, auditorias utilizam tal instrumento estatístico para avaliar, por exemplo, o valor médio das contas a receber de empresas. A partir dos resultados do teste, a gestão financeira poderá tomar decisões mais embasadas sobre a saúde financeira empresarial, garantindo maior acurácia dos dados e uma análise sólida para o gerenciamento das finanças e a tomada de decisões estratégicas.

Estudo de caso:

Uma empresa está preocupada com a precisão do valor médio de suas contas a receber. Para ter uma estimativa confiável, o auditor foi contratado para realizar um teste estatístico de hipótese e fornecer uma avaliação objetiva sobre a média de tais contas. A decisão de aceitar ou rejeitar o valor hipotético é importante para garantir uma gestão financeira adequada e tomar decisões informadas com base nos dados coletados. Neste caso, um auditor deseja testar a hipótese de que o valor médio de todas as contas a receber em uma dada firma é \$260,00, tomando para tanto uma amostra de 18 contas (veja a tabela abaixo). Ele deseja rejeitar o valor hipotético de \$260,00 somente se tal valor for claramente contraditado pela média da amostra.

Conta	Valor (\$)
1	258
2	255
3	245
4	245
5	258
6	258
7	256
8	256
9	252
10	249
11	252
12	258
13	259
14	252
15	259
16	246
17	260
18	258

Crie e execute um programa em **python** que, a partir dos dados acima armazenados num datagrama, determine:

- o valor da média amostral referente à variável “contas a receber”;
- o desvio padrão amostral referente à variável “contas a receber”;
- o tipo de teste estatístico apropriado para averiguar a afirmação dada;
- o valor da estatística de teste;
- o resultado do teste de hipótese onde deve aparecer, via terminal de saída, as seguintes possibilidades de mensagens:

“Aceita-se o fato de que o valor médio de todas as contas a receber em uma dada firma é \$260,00!”;

ou

“Rejeita-se o fato de que o valor médio de todas as contas a receber em uma dada firma é \$260,00!”.

QUESTÃO 2

O **APRENDIZ** é um programa de radiodifusão/reality show sobre gestão e empreendedorismo onde há um grupo seletivo de candidatos que concorrem a um contrato em uma empresa, por pelo menos um ano, pago pelo patrocinador do programa. Divididos em duas equipes, os candidatos devem realizar uma série de tarefas envolvendo vendas, marketing, promoções, negócios imobiliários, finanças, investimentos, ações beneficentes, publicidade e gerenciamento de negócios, bem como arrecadação de fundos, administração de bens e realização de eventos de grandes proporções.



Episódio do programa O APRENDIZ – fonte: Google Imagens

Em um dos episódios do programa (link do vídeo: <https://www.youtube.com/watch?v=Zc0V9OhNMAw> - veja o vídeo entre os instantes **6min53s** e **9min08s**), uma tarefa é atribuída às equipes, onde ambas terão que operar no mercado financeiro usando um crédito fictício de R\$ 1.000.000,00 para aquisição de ações na bolsa, baseando-se em dados REAIS.

A equipe que conseguir gerar maior lucratividade/rentabilidade em investimentos vence a tarefa. Neste sentido, considere que ambas as equipes escolheram um mesmo ativo financeiro para investimento e que adotaram uma metodologia estatística para a tomada de decisão. A série histórica de 10 dias deste ativo é mostrada na figura ao lado. A ideia inicial é verificar o valor mais provável para a cotação de um ativo.

Considere que foi feita uma afirmação de que o preço médio deste ativo **não seja superior** a R\$ 19,00. Desse modo, construa um código em **python** que verifique a veracidade desta afirmação, usando um nível de significância de 1%. Para tanto, deve-se seguir com determinação de variáveis e parâmetros de acordo com o roteiro abaixo:

Dia	Preço
1	R\$ 15,01
2	R\$ 18,92
3	R\$ 20,00
4	R\$ 15,78
5	R\$ 17,17
6	R\$ 16,23
7	R\$ 18,23
8	R\$ 18,00
9	R\$ 17,88
10	R\$ 19,02

- o valor da média amostral referente à variável “preço do ativo”;
- o desvio padrão amostral referente à variável “preço do ativo”;
- o tipo de teste estatístico apropriado para averiguar a afirmação dada;

Faculdade de Tecnologia e Ciências Sociais Aplicadas – FATECS
Curso de Ciência de Dados

- d) o valor da estatística de teste;
e) o resultado do teste de hipótese onde deve aparecer, via terminal de saída, as seguintes possibilidades de mensagens:
“Aceita-se o fato de que o preço do ativo não é superior a R\$ 19,00!”;
ou
“Rejeita-se o fato de que o preço do ativo não é superior a R\$ 19,00!”.

QUESTÃO 3

A tabela abaixo apresenta as quantias das indenizações pagas por estragos em automóveis, para uma amostra aleatória de 30 motoristas segurados, envolvidos em acidentes de pequena gravidade em uma determinada área geográfica.

Indenizações pagas por estragos em automóveis					
R\$ 984,00	R\$ 1.175,00	R\$ 1.302,00	R\$ 1.189,00	R\$ 1.432,00	R\$ 1.031,00
R\$ 1.181,00	R\$ 1.215,00	R\$ 934,00	R\$ 1.151,00	R\$ 1.344,00	R\$ 1.359,00
R\$ 1.414,00	R\$ 975,00	R\$ 1.106,00	R\$ 1.118,00	R\$ 1.332,00	R\$ 1.414,00
R\$ 1.159,00	R\$ 1.487,00	R\$ 1.364,00	R\$ 1.435,00	R\$ 1.301,00	R\$ 1.079,00
R\$ 1.067,00	R\$ 1.241,00	R\$ 1.295,00	R\$ 1.417,00	R\$ 1.018,00	R\$ 1.368,00

Considere que foi feita uma afirmação de que o valor médio das indenizações na população amostrada é de \$1.000,00. Desse modo, construa um código em **python** que verifique a veracidade desta afirmação, usando um nível de significância de 10%. Para tanto, deve-se seguir com determinação de variáveis e parâmetros de acordo com o roteiro abaixo:

- a) o valor da média amostral referente à variável “Indenização por estragos em automóveis”;
b) o desvio padrão amostral referente à variável “Indenização por estragos em automóveis”;
c) o tipo de teste estatístico apropriado para averiguar a afirmação dada;
d) o valor da estatística de teste;
e) o resultado do teste de hipótese onde deve aparecer, via terminal de saída, as seguintes possibilidades de mensagens:
“Aceita-se o fato de que o valor médio das indenizações na população amostrada é de \$1.000,00!”;
ou
“Rejeita-se o fato de que o valor médio das indenizações na população amostrada é de \$1.000,00!”.

QUESTÃO 4

O teste de hipótese pode ser utilizado para verificar a viabilidade comercial de um empreendimento dentro de certa localidade (exemplos de empreendimentos compreendem redes varejistas e atacadistas, supermercados, franquias, clínicas médicas, lojas de vestuário etc). Para isso, deve-se analisar algumas características importantes para que se possa avaliar tal viabilidade. Uma delas é o nível de renda familiar numa região-alvo de interesse: a partir de dados de renda, pode-se estabelecer critérios de viabilidade para a instauração de um tipo de negócio. A aplicação do referido teste estatístico ajudará o empreendedor a tomar decisões estratégicas e financeiras relacionadas ao empreendimento, de forma válida e confiável.

Estudo de caso:

O representante de um grupo comunitário informa, a um empreendedor que está interessado em estabelecer um centro comercial na região, que a renda média familiar na área é de R\$ 15.000,00. Suponha que a renda média familiar tem distribuição aproximadamente normal, onde o desvio padrão populacional como sendo $\sigma = R\$ 2.000,00$, com base em um estudo anterior. Para uma amostra aleatória de 30 famílias, a renda média familiar foi R\$ 14.000,00.

Desse modo, construa um código em **python** que verifique a veracidade desta afirmação, usando um nível de significância de 5%. Para tanto, deve-se seguir com determinação de variáveis e parâmetros de acordo com o roteiro abaixo:

- a) o valor da média amostral referente à variável “renda média familiar”;
- b) o desvio padrão amostral referente à variável “renda média familiar”;
- c) o tipo de teste estatístico apropriado para averiguar a afirmação dada;
- d) o valor da estatística de teste;
- e) o resultado do teste de hipótese onde deve aparecer, via terminal de saída, as seguintes possibilidades de mensagens:

“Aceita-se o fato de que o valor da renda média familiar é igual a R\$15.000,00!”;

ou

“Rejeita-se o fato de que o valor da renda média familiar é igual a R\$15.000,00!”.

QUESTÃO 5

Considerando a questão anterior, considere que o desvio padrão amostral seja igual a R\$ 2.000,00, sendo σ desconhecido. Além disso, o tamanho da amostra será igual a 15 famílias. Neste caso, elabore e execute um código em **python** que verifique a veracidade desta afirmação, usando um nível de significância de 5%. Para tanto, deve-se seguir com determinação de variáveis e parâmetros de acordo com o roteiro abaixo:

- a) o valor da média amostral referente à variável “renda média familiar”;
- b) o desvio padrão amostral referente à variável “renda média familiar”;
- c) o tipo de teste estatístico apropriado para averiguar a afirmação dada;
- d) o valor da estatística de teste;
- e) o resultado do teste de hipótese onde deve aparecer, via terminal de saída, as seguintes possibilidades de mensagens:

“Aceita-se o fato de que o valor da renda média familiar é igual a R\$15.000,00!”;

ou

“Rejeita-se o fato de que o valor da renda média familiar é igual a R\$15.000,00!”.

QUESTÃO 6

Para classificar os estilos de liderança, Fiedler (1965) desenvolveu uma medida chamada escala **LPC** (*Least preferred co-worker* – o colaborador menos preferido). Esta medida é obtida em formato de questionário, onde é solicitado aos respondentes que descrevam o colega de trabalho menos preferido com quem já trabalharam, levando em consideração uma lista de 16 adjetivos bipolares numa escala de 1 a 8. As respostas obtidas são somadas, dando o resultado do parâmetro **LPC** para cada respondente (pode-se obter um valor médio para os respondentes).

Se o **LPC** for elevado, isto é, se o colaborador menos preferido for descrito em termos de adjetivos relativamente “positivos/favoráveis”, isto significa (de acordo com Fiedler) que o estilo de liderança do respondente é orientado para as relações humanas, ou seja, a motivação primária deste é ter uma relação mais próxima ao grupo. Ao contrário, se o **LPC** for baixo, ou seja, o colega de trabalho menos preferido é descrito por adjetivos relativamente “negativos/desfavoráveis”, o estilo de liderança do respondente é orientado para as tarefas. A classificação dos valores da escala **LPC** encontra-se na tabela abaixo:

ESCALA LPC – CLASSIFICAÇÃO	
73 ou acima	LPC elevado
64 a 72	LPC intermediário
63 ou inferior	LPC baixo

Considere uma situação em que foi feita uma hipótese de que a média da escala **LPC** de um determinado setor de uma empresa tenha um valor **baixo**. Para averiguar tal afirmação, foi realizada uma entrevista com um grupo amostral de 15 respondentes aleatoriamente selecionados. A média amostral dos valores **LPC** obtidos é igual a 70. O desvio padrão amostral para o valor **LPC** é igual a 6,23. Supõe-se que os valores **LPC** tenham distribuição normal na população. Com base neste resultado amostral, faça o teste de hipóteses usando um nível de significância de 5%.

ESCALAS DE RESPOSTAS

1	Agradável	8	7	6	5	4	3	2	1	Desagradável
2	Amigoso	8	7	6	5	4	3	2	1	Inamigoso
3	Rejeitador	1	2	3	4	5	6	7	8	Acolhedor
4	Prestativo	8	7	6	5	4	3	2	1	Frustrante
5	Sem entusiasmo	1	2	3	4	5	6	7	8	Entusiasta
6	Tenso	1	2	3	4	5	6	7	8	Calmo
7	Distante	1	2	3	4	5	6	7	8	Íntimo
8	Frio	1	2	3	4	5	6	7	8	Caloroso
9	Cooperativo	8	7	6	5	4	3	2	1	Não-cooperativo
10	Apoiador	8	7	6	5	4	3	2	1	Hostil
11	Maçante	1	2	3	4	5	6	7	8	Interessante
12	Desequilibrado	1	2	3	4	5	6	7	8	Equilibrado
13	Seguro de si	8	7	6	5	4	3	2	1	Hesitante
14	Eficiente	8	7	6	5	4	3	2	1	Ineficiente
15	Sombrio	1	2	3	4	5	6	7	8	Alegre
16	Expansivo	8	7	6	5	4	3	2	1	Reservado

Resposta: $t_{\text{crítico}} = 1,761$; $t_{\text{teste}} \approx 4,352$; **rejeita-se a hipótese de que o valor médio da escala LPC seja considerado baixo num determinado setor de uma empresa**

QUESTÃO 7

A decisão a ser tomada acerca de como disponibilizar essas marcas nas prateleiras de comércio varejistas é uma empreitada de impacto substancial. Qualquer resolução adotada pode repercutir diretamente nas escolhas dos consumidores. A probabilidade de eleição entre as opções disponíveis será intrinsecamente associada a uma miríade de variáveis, abrangendo o perfil demográfico de seus clientes, suas predileções pessoais e, primordialmente, o atributo que valorizam com maior ênfase.

A determinação de uma estratégia prudente e acertada no tocante à disposição e promoção dessas marcas requer, portanto, a condução de pesquisas de mercado aprofundadas, visando a uma compreensão mais ampla das preferências dos consumidores em sua região de atuação. Além disso, a análise de tendências de vendas passadas, aliada a estratégias de marketing direcionadas, se mostra imprescindível para a tomada de decisão embasada.

Na posição de gestor de aquisições, a escolha de marcas e a forma de precificação são fundamentais para a eficiência na lucratividade no cenário varejista. Outro ponto é a diversificação na comercialização de bens e serviços, visando atratividade comercial para o seu público-alvo de interesse.

Como exemplo ilustrativo, considere que no mercado três produtos de marcas concorrentes de molho de tomate tipo “ketchup”: **Produto 1**, **Produto 2** e **Produto 3**. Neste contexto, observa-se a manifestação de preferências diversas por parte dos clientes. Nesse caso, o ponto a ser observado é a alocação de tais produtos nas prateleiras e promovê-las, a fim de fomentar as vendas (veja a imagem ao lado).



Para tal fim, deve-se realizar uma pesquisa prévia de preferência junto a presumíveis clientes: a proporção de escolhas irá dirimir a organização de tais produtos nas prateleiras, bem como destacar as características de tais produtos futuramente. Neste caso, deseja-se averiguar se há ou não preferência uniforme entre os produtos supracitados

Deseja-se realizar um teste de hipóteses para averiguar se há distribuição uniforme quanto à preferência entre os produtos supracitados

Com base nos resultados desta pesquisa (disponível no endereço via link [Acesso ao questionário - escolha de produtos](#)), construa um código em python que determine:

- o valor da estatística de teste apropriada;
- o resultado do teste de hipóteses usando uma confiança de 99%;

No mesmo código, [registre](#) os textos abaixo e [relacione-os](#) com os produtos supracitados:

Marca A: Este molho de tomate é consagrado pela sua riqueza de sabor, evocando a autenticidade caseira e fazendo ostentação de ingredientes inteiramente naturais. A marca ressalta a qualidade de seus ingredientes, isentos de conservantes ou corantes artificiais. Notavelmente, seu preço é posicionado em um patamar intermediário, tornando-o uma alternativa de acesso viável.

Marca B: Este molho de tomate conquistou notoriedade graças ao seu design de rótulo moderno e apelativo, agregando apelo visual à sua embalagem. A marca enfatiza o compromisso com ingredientes orgânicos e sustentáveis. Todavia, é importante ressaltar que o seu preço é ligeiramente superior em comparação às demais marcas.

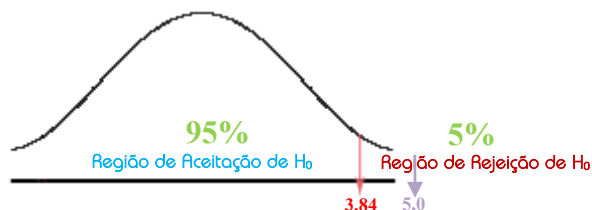
Marca C: Este molho de tomate é favorecido por uma clientela que busca efetuar economias substanciais. Configura-se como a alternativa mais acessível em termos de preço, angariando adeptos que valorizam a relação custo-benefício. A marca, por sua vez, não promove ingredientes de cariz orgânico ou distinções especiais em seus produtos, mas se destaca por sua competitividade no que tange ao preço.

QUESTÃO 8

Afirma-se que igual número de homens e mulheres são clientes da uma loja especializada em venda de calças “jeans”. Observa-se uma amostra aleatória de 80 clientes, sendo 50 homens e 30 mulheres. Teste a hipótese nula de que o número total de clientes homens e de clientes mulheres é igual, utilizando um nível de significância de 5%.

	Clientes		Total
	Homens	Mulheres	
Quantidade consultada	50	30	80

Resposta: $\chi^2_{\text{crítico}} = 3,84$; $\chi^2_{\text{teste}} = 5,00$; **rejeita-se a hipótese nula quanto à igualdade do número total de clientes homens e de clientes mulheres**



QUESTÃO 9

A gestão de marketing desempenha um papel vital na promoção bem-sucedida de produtos e serviços: apreende a compreensão do mercado, verifica as preferências de consumidores e a concorrência, permitindo a criação de estratégias eficazes. O marketing posiciona a marca, desenvolve estratégias de promoção, otimiza preços e estimula a demanda. Além disso, a gestão de marketing mantém as empresas competitivas, adaptando-se às tendências de mercado e coletando feedback para a melhoria contínua. Em um ambiente de negócios dinâmico, o marketing é essencial para alcançar e manter o sucesso a longo prazo.

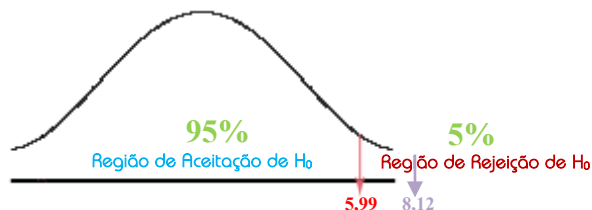
Estudo de caso:

Como exemplo prático, considere um cenário em que o desafio enfrentado pela gestão de marketing de um fabricante de refrigeradores torna-se essencial na tomada de decisões estratégicas. A empresa, que apresenta três segmentos de produtos distintos, categorizados como "baixo", "intermediário" e "alto" em relação aos preços, planeja lançar uma campanha promocional destinada a realçar as características dos refrigeradores de alto preço.

Antes da realização de uma promoção de vendas destinada a destacar as qualidades de refrigeradores de preço "alto", as quantidades de vendas das três categorias eram, respectivamente, 23, 15 e 12. De uma amostra aleatória de 50 refrigeradores vendidos depois da promoção, as quantidades vendidas nas categorias de preço "baixo", "intermediário" e "alto" foram, respectivamente, 15, 15 e 20. Teste a hipótese de que o atual padrão de vendas não difere do padrão histórico, utilizando um nível de significância de 5%.

	Categoria de preço			Total
	Baixo	Intermediário	Alto	
Quantidade vendida	15	15	20	50
Quantidade esperada de vendas	23	15	12	50

Resposta: $\chi^2_{\text{crítico}} = 5,99$; $\chi^2_{\text{teste}} = 8,12$; **rejeita-se a hipótese nula de que o atual padrão de vendas não difere do padrão histórico**



QUESTÃO 10

Os dados fictícios abaixo representam os pesos dos indivíduos que se comprometeram com três tipos de dieta:

Pesos de indivíduos após uso de dietas		
Dieta 1	Dieta 2	Dieta 3
198	214	174
211	200	176
240	259	213
189	194	201
178	188	158

Usando o teste ANOVA, verifique a afirmação de que a média populacional de pesos são iguais para diferentes tipos de dietas, usando um nível de significância de 5%.

QUESTÃO 11

Os dados abaixo representam o quantitativo de desempenho médio de quinze funcionários que estão em treinamento para o desenvolvimento de um nível de habilidade dentro de uma determinada empresa. Neste caso, tais funcionários recebem um treinamento prévio com diferentes métodos de instrução. Após o término do treinamento, todos são submetidos a uma avaliação para verificar o desempenho. Neste caso, os tratamentos são os métodos de ensino/instrução. As pontuações obtidas nas provas estão listadas na tabela. Use o procedimento de análise da variância para testar a hipótese nula de que as três amostras foram obtidas da mesma população, usando um nível de significância de 5% para o teste.

Método de Instrução	Desempenho nas provas				
A ₁	86	79	81	70	84
A ₂	90	76	88	82	89
A ₃	82	68	73	71	81

Resposta: a hipótese nula de que as pontuações médias para os três métodos de instrução são mutuamente iguais não pode ser rejeitada com o nível de significância de 5%

QUESTÃO 12

A tabela abaixo apresenta as vendas para um período de 11 anos de uma companhia de programas gráficos (fictícia) incorporada em 1990. Determine:

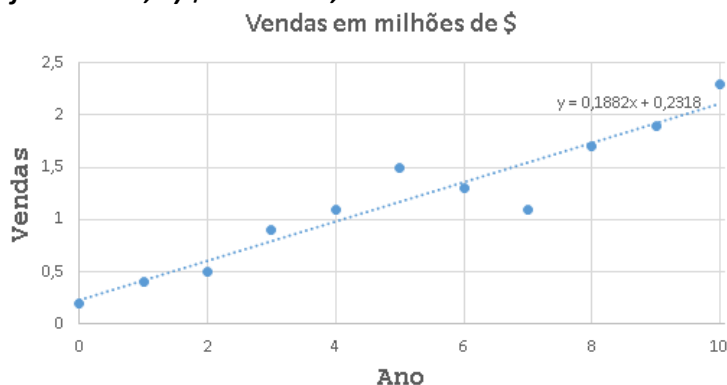
- A)** um gráfico de dispersão mostrando o comportamento das vendas em função do tempo (anos);
- B)** a equação de tendência linear para estes dados através do método dos mínimos quadrados, codificando 1990 como o ano "0";
- C)** Com base na resposta do item anterior, determine a projeção de vendas para o ano de 2001.

Ano	Vendas em milhões de \$
1990	0,2
1991	0,4
1992	0,5
1993	0,9
1994	1,1
1995	1,5
1996	1,3

Faculdade de Tecnologia e Ciências Sociais Aplicadas – FATECS
Curso de Ciência de Dados

1997	1,1
1998	1,7
1999	1,9
2000	2,3

Resposta: A) e B): vide gráfico abaixo; C) \$ 2.302.000,00



QUESTÃO 13

Considerando os dados da questão anterior, determine a componente cíclica para a referida série temporal. Para tanto, deve-se determinar o valor do componente relativo cíclico, dada pela expressão:

$$\text{Relativo Cíclico} = 100\% \times \frac{Y_{\text{observado}}}{Y_{\text{estimado}}}$$

onde $Y_{\text{observado}}$ é o valor real/observado do dado para a variável dependente no problema e Y_{estimado} é o valor da projeção/estimativa da referida variável usando a equação de tendência/regressão linear.

Resposta:

Ano	Ano Codificado	Y	Yestimado	Relativo Cíclico
1990	0	0,2	0,2318	86,28%
1991	1	0,4	0,42	95,24%
1992	2	0,5	0,6082	82,21%
1993	3	0,9	0,7964	113,01%
1994	4	1,1	0,9846	111,72%
1995	5	1,5	1,1728	127,90%
1996	6	1,3	1,361	95,52%
1997	7	1,1	1,5492	71,00%
1998	8	1,7	1,7374	97,85%
1999	9	1,9	1,9256	98,67%
2000	10	2,3	2,1138	108,81%

QUESTÃO 14

A tabela a seguir representa o comportamento das vendas trimestrais de uma empresa no período entre 1.995 e 2.000. Determine:

- os valores da média móvel de vendas;
- o valor da média das médias móveis (média móveis “centralizadas” para quatro trimestres);
- o valor da variação relativa sazonal;
- o valor da “média modificada” por trimestre, caracterizada pela média das médias móveis obtidas entre 1.995 e 2.000, eliminando-se o maior e o menor valor das médias móveis por trimestre;

Vendas trimestrais para uma empresa em milhares de \$						
Trimestre	1995	1996	1997	1998	1999	2000
1	500	450	350	550	550	750
2	350	350	200	350	400	500
3	250	200	150	250	350	400
4	400	300	400	550	600	650

Resposta: itens A, B e C:

Trimestre	Vendas	Média Móvel - 4 períodos trimestrais	Média das médias móveis (2 períodos)	Relativo Sazonal (%)
1/1995	500			
2/1995	350			
3/1995	250			
4/1995	400	375		110,34%
1/1996	450	362,5		126,32%
2/1996	350	362,5		103,70%
3/1996	200	350	356,25	64,00%
4/1996	300	325	337,5	106,67%
1/1997	350	300	312,5	136,59%
2/1997	200	262,5	281,25	76,19%
3/1997	150	250	256,25	50,00%
4/1997	400	275	262,5	116,36%
1/1998	550	325	300	146,67%
2/1998	350	362,5	343,75	86,15%
3/1998	250	387,5	375	58,82%
4/1998	550	425	406,25	127,54%
1/1999	550	425	425	122,22%
2/1999	400	437,5	431,25	85,33%
3/1999	350	462,5	450	70,00%
4/1999	600	475	468,75	111,63%
1/2000	750	525	500	134,83%
2/2000	500	550	537,5	87,91%
3/2000	400	562,5	556,25	
4/2000	650	575	568,75	

Trimestre	1995	1996	1997	1998	1999	2000	Média Modificada
1		126,32%	136,59%	146,67%	122,22%	134,83%	132,58%
2		103,70%	76,19%	86,15%	85,33%	87,91%	86,46%
3	67,80%	64,00%	50,00%	58,82%	70,00%		63,54%
4	110,34%	106,67%	116,36%	127,54%	111,63%		112,78%


QUESTÃO 15



Uma característica proeminente das commodities agrícolas é a sua suscetibilidade a variações sazonais. A análise de séries temporais permite a identificação de padrões recorrentes associados a estações específicas, possibilitando que agricultores e formuladores de políticas antecipem e planejem para as flutuações sazonais na produtividade. Além disso, o reconhecimento de tendências cíclicas, influenciadas por fatores como rotação de culturas e ciclos econômicos, fornece informações valiosas para a tomada de decisões estratégicas no setor agrícola.

A aplicação de técnicas estatísticas avançadas na análise de séries temporais viabiliza a modelagem e previsão de comportamentos futuros das commodities agrícolas. Modelos como projeção com uso de médias móveis, projeções cíclicas, ajustes sazonais e análise de tendência possibilitam a captura de padrões complexos e a projeção de tendências, contribuindo para a elaboração de estratégias de plantio, colheita e comercialização mais eficientes.

Estudo de caso:

Considere os dados reais referentes à área colhida em lavouras das “commodities” agrícolas Cana-de-Açúcar, Milho e Soja (em hectares) entre os anos de 1.990 e 2022 (download dos dados das séries históricas – clique neste link do [IBGE](http://ibge.gov.br) ou, se preferir, acesse o arquivo anexo neste documento clicando aqui: ). A partir de tais dados, determine para cada “commodity” citada anteriormente:

- A)** os coeficientes “a” e “b” da linha de tendência para o período **1.990 a 2.020** usando as equações abaixo:

$$a = \frac{n \sum x_i \times y_i - \left(\sum x_i \right) \times \left(\sum y_i \right)}{n \sum x_i^2 - \left(\sum x_i \right)^2} \qquad b = \frac{\sum y_i - a \times \left(\sum x_i \right)}{n}$$

Para tanto, preencha a tabela abaixo para futura determinação dos coeficientes “a” e “b”:

x	y	x × y	x ²
TOTAL			

Faculdade de Tecnologia e Ciências Sociais Aplicadas – FATECS
Curso de Ciência de Dados

- B)** Uma previsão para área colhida para o ano de 2.023 referente à cada commodities mencionada sob as condições do item **A)**;
- C)** Um gráfico contendo os valores do “relativo cíclico” para a commodity agrícola “milho” para o período **1.990 a 2.020**.