**UNIVERSIDADE POSITIVO**

**DATA SCIENCE**

**Atividade avaliativa prática**

**Parte 1: Análise de Posição e Dispersão**

1. Calcule a média das vendas ao longo dos 4 anos e discuta como ela representa o valor central das vendas.

Resp: A média (14,72) representa o valor médio de um conjunto de dados. No contexto das vendas, o cálculo da média das vendas é uma maneira de obter um valor representativo do comportamento médio das vendas em um determinado período.

2. Calcule a mediana das vendas para cada ano e descreva como ela difere da média.

Resp: mediana 2024 = 15 2025 = 15,3 2026 = 15,7 2027 = 12,3

A principal diferença é que a mediana é menos sensível a valores extremos do que a média. Isso a torna uma tendência central robusta, adequada para dados com valores atípicos.

3. Calcule a variância e o desvio padrão das vendas de todos os meses. Interprete o significado dessas medidas em relação à dispersão dos dados.

Resp: Variância (5,4) Indica o quão os valores das vendas se afastam da média.

Desvio padrão (2,32) Raiz quadrada da variância que representa a dispersão média dos valores em relação à média.

4. Determine o mês com a maior venda e o mês com a menor venda. Compare esses valores com a média para identificar os meses atípicos.

Resp: meses atípicos janeiro (10,5) e dezembro (19)

**Parte 2: Análise de Quartis**

1. Divida os dados de vendas em quatro quartis. Identifique os valores que definem cada quartil.

Resp: q1 = 12,55 q2 = 15,2 q3 = 16,4

2. Calcule os intervalos interquartil (Quadrantes) para as vendas e explique seu significado.

Resp: Intervalo interquartil (3,84) Representa a faixa que cobre a maioria dos dados de vendas. Valores fora dessa faixa, pode considerá-los como possíveis valores atípicos.

**Parte 3: Previsão (Forecasting)**

1. Use um método de análise de tendência, para criar uma previsão das vendas para o próximo ano (2028) com base nos dados anteriores.

Resp: 189.2

2. Determine o erro médio e o erro percentual médio da sua análise de tendência.

Resp: Erro médio = 15,23 Erro percentual médio = 46,11%

3. Discuta as limitações do método utilizado e sugira como ele pode ser adotado em um cenário real.

Resp: O modelo pode ser sensível a outliers e anomalias nos dados de treinamento, o que pode afetar negativamente as previsões.

Em alguns casos, pode ser necessário ajustar manualmente os parâmetros do modelo para obter resultados precisos.

O método assume que a série temporal é estacionária, o que significa que as propriedades estatísticas da série não mudam com o tempo. Se a série for não estacionária, pode ser necessário realizar transformações nos dados, como diferenciação, para torná-la estacionária.

O método assume que a sazonalidade é fixa ao longo do tempo, o que pode não ser verdadeiro em cenários onde os padrões sazonais mudam.

Para adotar o modelo em um cenário real, é crucial entender as limitações do modelo, realizar a análise abrangente dos dados e estar preparado para ajustes contínuos à medida que a série temporal evolui.

4. Baseado na análise de tendência, faça uma previsão das vendas para os próximos cinco anos (2028 a 2032) e explique as suposições subjacentes.

Resp: 2028 = 189.2 2029 = 189.2 2030 = 189.2 2031 = 189.2 2032 = 189.2

Média estimativa: O método se baseia na média das vendas passadas como uma estimativa das vendas futuras. Isso pressupõe que as vendas futuras serão semelhantes às vendas médias no passado, o que pode não ser verdade se houver mudanças significativas nas condições de mercado.

Variabilidade constante: O método assume que a variabilidade das vendas permanecerá constante ao longo do tempo. No entanto, a variabilidade das vendas pode aumentar ou diminuir devido a certos fatores.

Nenhuma consideração de fatores externos: O método não leva em consideração fatores externos que podem afetar as vendas