

Estatística

1. Média

A média é o valor obtido pela soma de todos os números em um conjunto de dados, dividida pelo número total de dados. É frequentemente chamada de "média aritmética". A fórmula para calcular a média é:

$$\text{Média} = \frac{\text{Soma de todos os valores}}{\text{Número total de valores}}$$

Visualmente, você pode pensar na média como o ponto de equilíbrio de um conjunto de números, onde a soma de todos os valores é distribuída igualmente.

Suponha que tenhamos os seguintes números: 3, 5, 7, 9 e 11.

- Soma dos valores: $3 + 5 + 7 + 9 + 11 = 35$
- Número total de valores: 5
- Média: $\frac{35}{5} = 7$

2. Mediana

A mediana é o valor que divide o conjunto de dados em duas partes iguais. Para encontrar a mediana, primeiro você precisa ordenar os dados em ordem crescente ou decrescente e, em seguida, encontrar o valor do meio. Se houver um número par de dados, a mediana é a média dos dois valores do meio. Se houver um número ímpar de dados, a mediana é o valor do meio exato. Abaixo está a fórmula básica:

- Se o número de observações (n) for ímpar: A mediana é o valor na posição $(n + 1)/2$.
- Se o número de observações (n) for par: A mediana é a média dos valores nas posições $(n/2)$ e $(n/2 + 1)$.

Visualmente, a mediana é o ponto no meio dos dados quando estão organizados em ordem crescente.

Suponha que temos os seguintes números: 3, 7, 2, 8, 5, 4, 9.

- Primeiro, vamos ordená-los: 2, 3, 4, 5, 7, 8, 9.
- Como o número de dados é ímpar (7), a mediana é o valor na posição $(7 + 1)/2 = 4$.
- Portanto, a mediana é 5.

3. Média e Mediana na Análise de Dados

Tanto a média quanto a mediana são medidas estatísticas fundamentais que fornecem insights valiosos sobre um conjunto de dados. Abaixo estão algumas razões pelas quais essas informações são úteis para análise de dados:

- **Resumo Estatístico:**

A média e a mediana são formas eficientes de resumir grandes conjuntos de dados em um único número representativo. Elas fornecem uma visão geral do "centro" dos dados, auxiliando os analistas a entenderem o que é típico ou representativo dentro do conjunto.

- **Tendência Central:**

A média e a mediana ajudam a identificar a tendência central dos dados, ou seja, onde a maioria dos valores se concentra. Se os valores estiverem uniformemente distribuídos em torno da média ou mediana, isso sugere uma distribuição simétrica dos dados. Por outro lado, se houver uma grande diferença entre a média e a mediana, isso pode indicar que os dados estão distorcidos por valores extremos.

- **Identificação de Outliers:**

Outliers são valores extremos que se afastam significativamente da maioria dos dados. A média é sensível a outliers, pois ela leva em consideração todos os valores. Se houver outliers presentes, a média pode ser distorcida. Nesse caso, a mediana pode ser uma medida mais robusta, pois ela é menos afetada por valores extremos.

- **Decisões Empresariais:**

Na análise de dados empresariais, entender a média e a mediana pode ajudar na tomada de decisões. Por exemplo, ao analisar o desempenho de vendas, a média de vendas por mês pode fornecer uma ideia geral do desempenho médio, enquanto a mediana pode ajudar a entender o desempenho típico, desconsiderando vendas extremamente altas ou baixas que podem distorcer a média.

- **Comparação de Grupos:**

Ao comparar diferentes grupos ou subconjuntos de dados, a média e a mediana podem revelar diferenças significativas em suas distribuições. Por exemplo, ao comparar os salários médios de diferentes departamentos de uma empresa, a média

pode mostrar uma diferença, mas a mediana pode revelar se essa diferença é devido a valores extremos em um dos grupos.

Em resumo, a média e a mediana são ferramentas poderosas na análise de dados, fornecendo insights importantes sobre a distribuição, tendências e características dos dados, e ajudando na tomada de decisões informadas.