



*Fundamentos de Linguagem Python para Análise de Dados e Data Science*

# Fundamentos de Linguagem Python Para Análise de Dados e Data Science

## Conceitos Fundamentais em Estatística



A Estatística é um ramo da Matemática que lida com a coleta, análise, interpretação, apresentação e organização de dados. Ela fornece técnicas e métodos para lidar com informações quantitativas e qualitativas, auxiliando na tomada de decisões e no estudo de fenômenos complexos. A Estatística pode ser dividida em duas áreas principais: Estatística Descritiva e Estatística Inferencial.

A Estatística Descritiva tem como objetivo resumir, descrever e analisar dados coletados. Ela utiliza medidas de tendência central (média, mediana, moda), medidas de dispersão (amplitude, variância, desvio padrão) e gráficos (histograma, boxplot, gráfico de dispersão) para apresentar os dados de forma clara e concisa.

A Estatística Inferencial utiliza os dados coletados de uma amostra para fazer generalizações e tirar conclusões sobre uma população maior. Isso envolve o uso de técnicas como testes de hipóteses, intervalos de confiança e análise de regressão. A Estatística Inferencial ajuda a tomar decisões informadas com base nos dados disponíveis e a estimar parâmetros populacionais.

Os dados em estatística podem ser classificados em dois tipos de variáveis: quantitativas e qualitativas.

**a) Variáveis Quantitativas:** São variáveis que expressam quantidades e podem ser medidas numericamente. Elas podem ser classificadas em dois subtipos:

- **Variáveis Discretas:** Representam contagens e só podem assumir valores inteiros. Exemplos incluem número de filhos, quantidade de produtos vendidos e número de erros em um teste.
- **Variáveis Contínuas:** Representam medidas e podem assumir qualquer valor em um intervalo contínuo. Exemplos incluem altura, peso e temperatura.

**b) Variáveis Qualitativas:** São variáveis que expressam características, atributos ou categorias e não possuem uma escala numérica. Elas podem ser classificadas em dois subtipos:

- **Variáveis Nominais:** Representam categorias sem uma ordem ou hierarquia específica. Exemplos incluem gênero, cor dos olhos e tipo de veículo.
- **Variáveis Ordinais:** Representam categorias com uma ordem ou hierarquia específica, mas sem uma escala numérica fixa. Exemplos incluem classificações em uma competição, níveis de escolaridade e faixas de renda.

Entender os conceitos fundamentais da Estatística, incluindo os tipos de variáveis, é crucial para aplicar adequadamente os métodos estatísticos na análise de dados e na tomada de decisões baseadas em evidências.