

# Profissão: Cientista de Dados



# GLOSSÁRIO



# Descritiva I



Dica: para encontrar rapidamente a palavra que procura aperte o comando CTRL+F e digite o termo que deseja achar.

- **Conheça os tipos de variáveis**
- **Utilize a biblioteca Matplotlib**
- **Utilize boxplot e histograma**
- **Utilize Seborn e Bivariadas Discreta**



# Conheça os tipos de variáveis



# Conheça os tipos de variáveis

## ● Gráfico de Caixa

Também conhecido como box plot, é um método para representar graficamente grupos de dados numéricos através de seus quartis.

## ● Gráfico de Dispersão

É um gráfico que usa coordenadas cartesianas para exibir valores de um conjunto de dados.

## ● Histograma

É um gráfico de barras que mostra a distribuição de frequência de um conjunto de dados contínuos.



# Conheça os tipos de variáveis

## • Variáveis Contínuas

São variáveis que podem assumir um número infinito de valores dentro de um intervalo específico.

## • Variáveis Qualitativas

São variáveis que podem ser divididas em categorias, mas não têm uma ordem ou medida.

## • Variáveis Discretas

São variáveis que podem assumir um número finito de valores.

## • Variáveis Quantitativas

São variáveis que podem ser medidas e ordenadas. Elas podem ser contínuas ou discretas.



# Utilize a biblioteca Matplotlib



# Utilize a biblioteca Matplotlib

## • Adição de subplots

Criação de múltiplos gráficos em uma única figura, cada um com seu próprio sistema de eixos.

## • Biblioteca Matplotlib

Biblioteca de software para a linguagem de programação Python que fornece funcionalidades para a criação de gráficos estáticos, animados e interativos.





# Utilize boxplot e histograma



# Utilize boxplot e histograma

## • Boxplot

Tipo de gráfico que é usado para visualizar a distribuição de uma variável, mostrando a mediana, quartis, máximo e mínimo dos dados.

## • Função 'distplot'

Ferramenta em Python, especificamente na biblioteca Seaborn, usada para visualizar a distribuição de uma variável.

## • Função 'describe'

Ferramenta em Python que fornece uma visão geral rápida dos dados, incluindo contagem, média, desvio padrão, mínimo, primeiro quartil, mediana, terceiro quartil e máximo.

## • Histograma

Gráfico que mostra a distribuição de uma variável contínua dividindo-a em bins e contando o número de observações que caem em cada bin.



# Utilize Seborn e Bivariadas Discreta



# Utilize Seborn e Bivariadas Discreta

## ● **Análise Bivariada**

É a análise de duas variáveis para determinar a relação entre elas.

## ● **Frequência Cruzada**

É uma tabela que mostra a frequência de ocorrência de duas variáveis categóricas, permitindo analisar a relação entre elas.

## ● **Catplot**

É uma função do Seaborn que permite fazer gráficos de barra onde a altura da barra pode representar algo além da frequência, como a média ou a soma.

## ● **Seaborn**

É uma biblioteca de visualização de dados Python baseada em Matplotlib. Fornece uma interface de alto nível para desenhar gráficos estatísticos atraentes e informativos.



# Bons estudos!

