



ELDORADO

# *Trilha Python*

## *Aula 09*

***Chamada!***



***Faísca***

# Atividades de hoje

- ◆ Matplotlib;
- ◆ Storytelling;
- ◆ Atividade em grupos.



***Um poquinho  
mais de Matplotlib***



# Matplotlib

O matplotlib, assim como a própria linguagem Python, possui uma estrutura orientada a objetos com alguns elementos chaves para entendermos como uma imagem, no nosso caso um gráfico, é construído, são eles:

“**Figure**” (Figura), “**Axes**” (Eixos), “**Axis**” (Eixo) e “**Artist**” (Artista).

# Matplotlib

**Figure:** Representa todo o espaço disponível para o gráfico e informações adicionais (títulos, legendas). Pense nela como uma tela em branco.

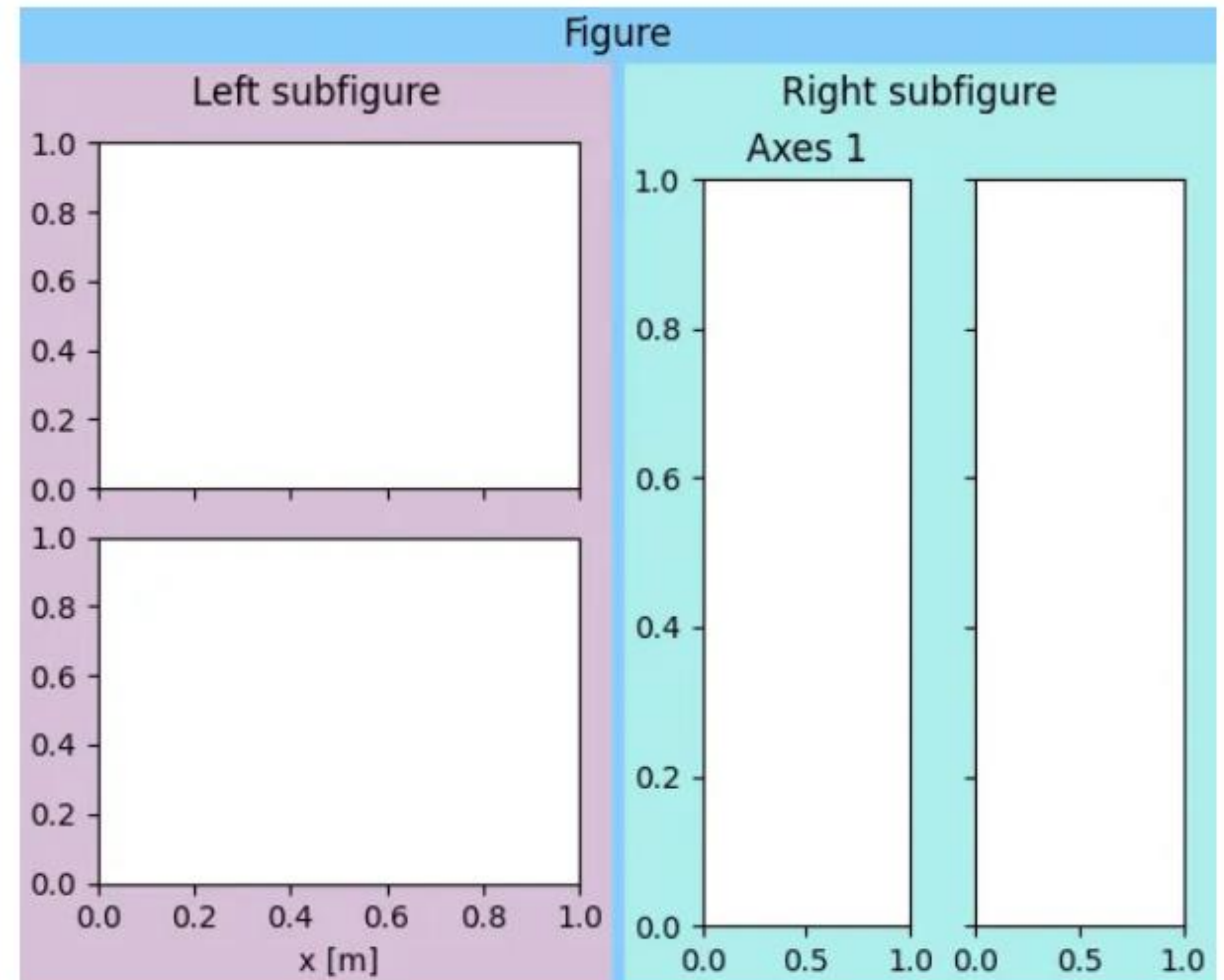
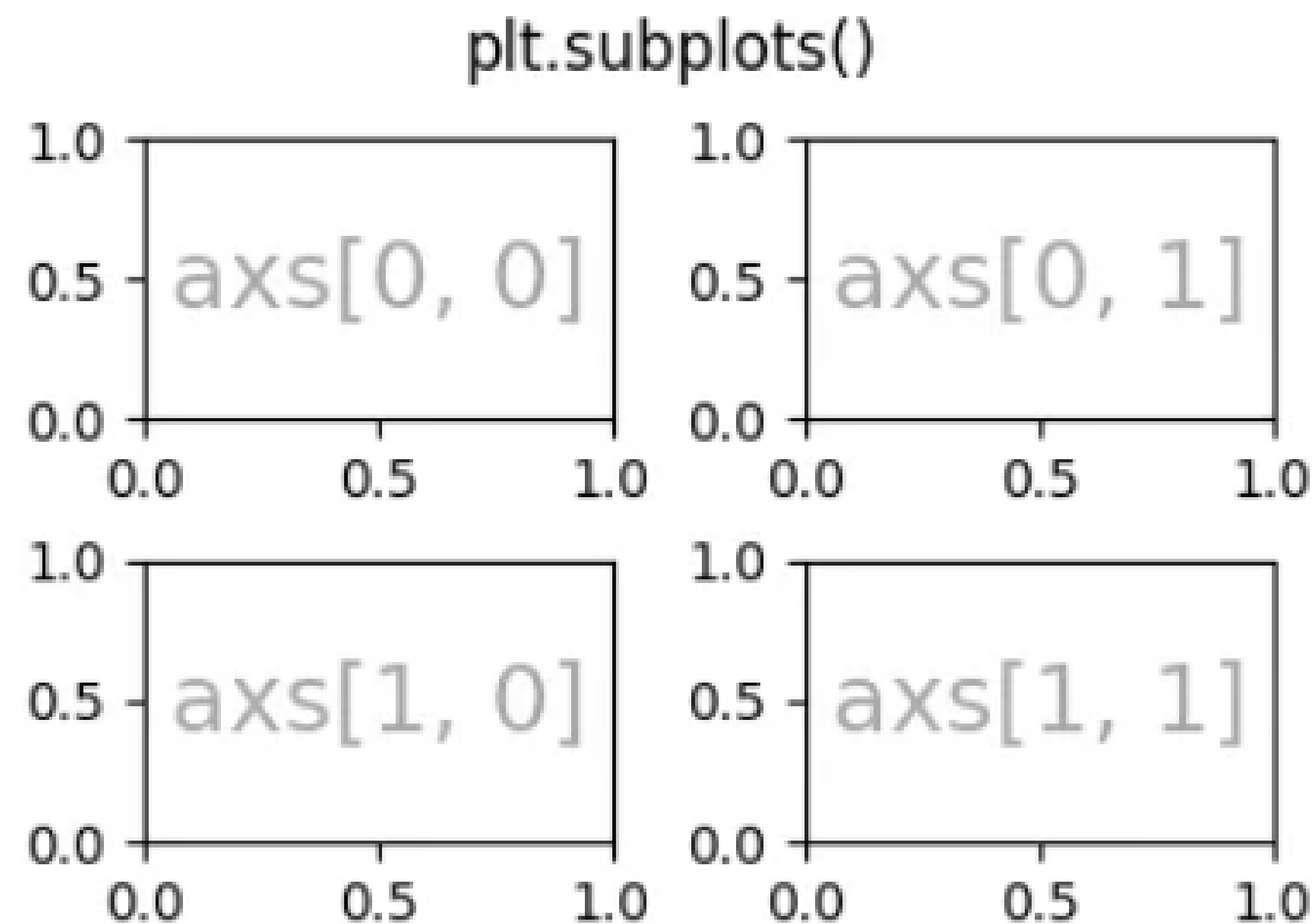


Imagem de um 'Figure' do Matplotlib. Fonte: [Matplotlib — Introduction to Figures](#)

# Matplotlib

**Axes:** Representa uma **área de plotagem** dentro de uma 'Figure'. Cada 'Axes' tem seus próprios eixos, rótulos e título. Logo, uma 'Figure' **não é o gráfico** em si, já que pode conter um ou mais gráficos.

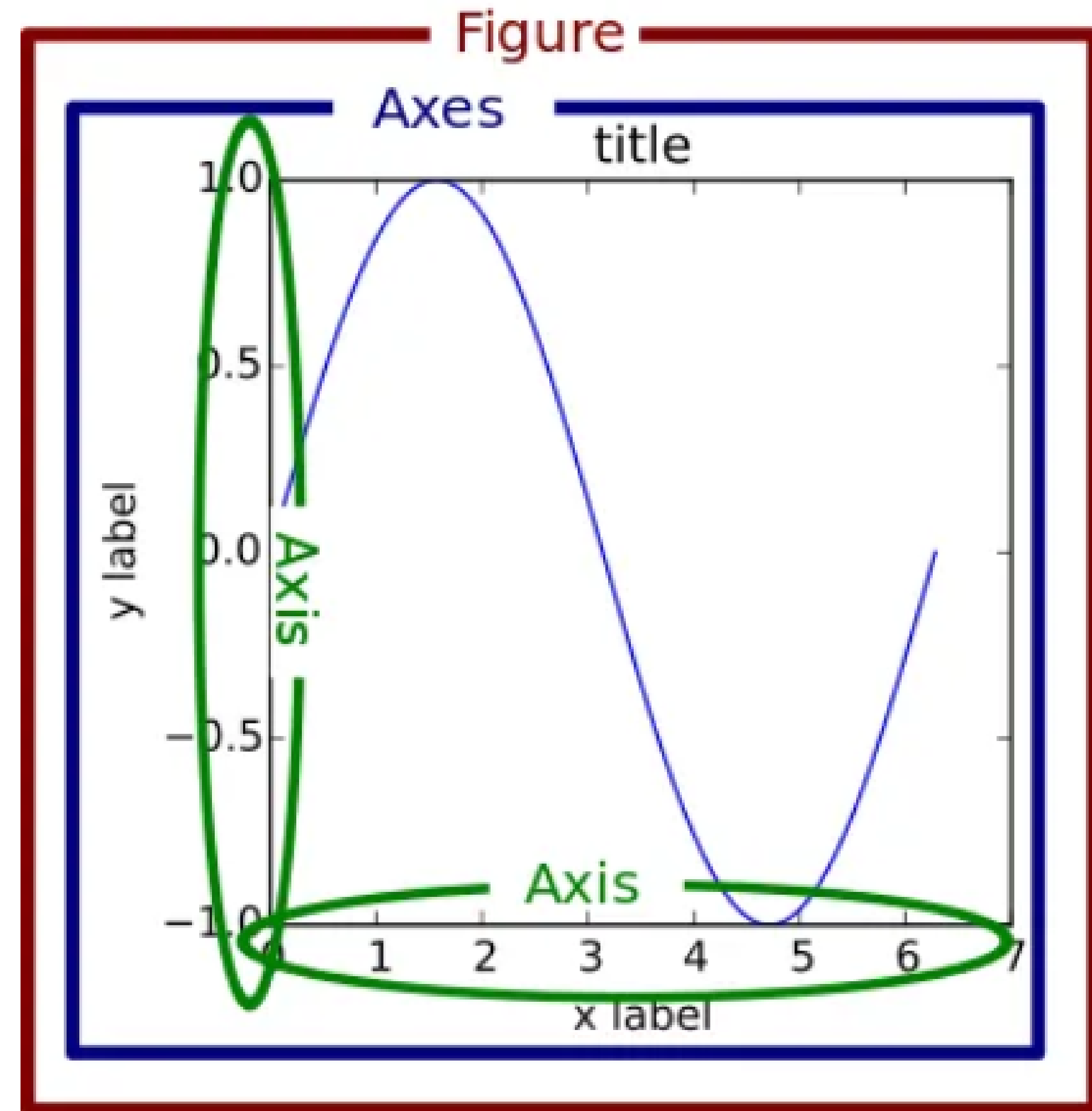


Axes dentro de uma 'Figure'. Fonte: [Axes and subplots](#)



# Matplotlib

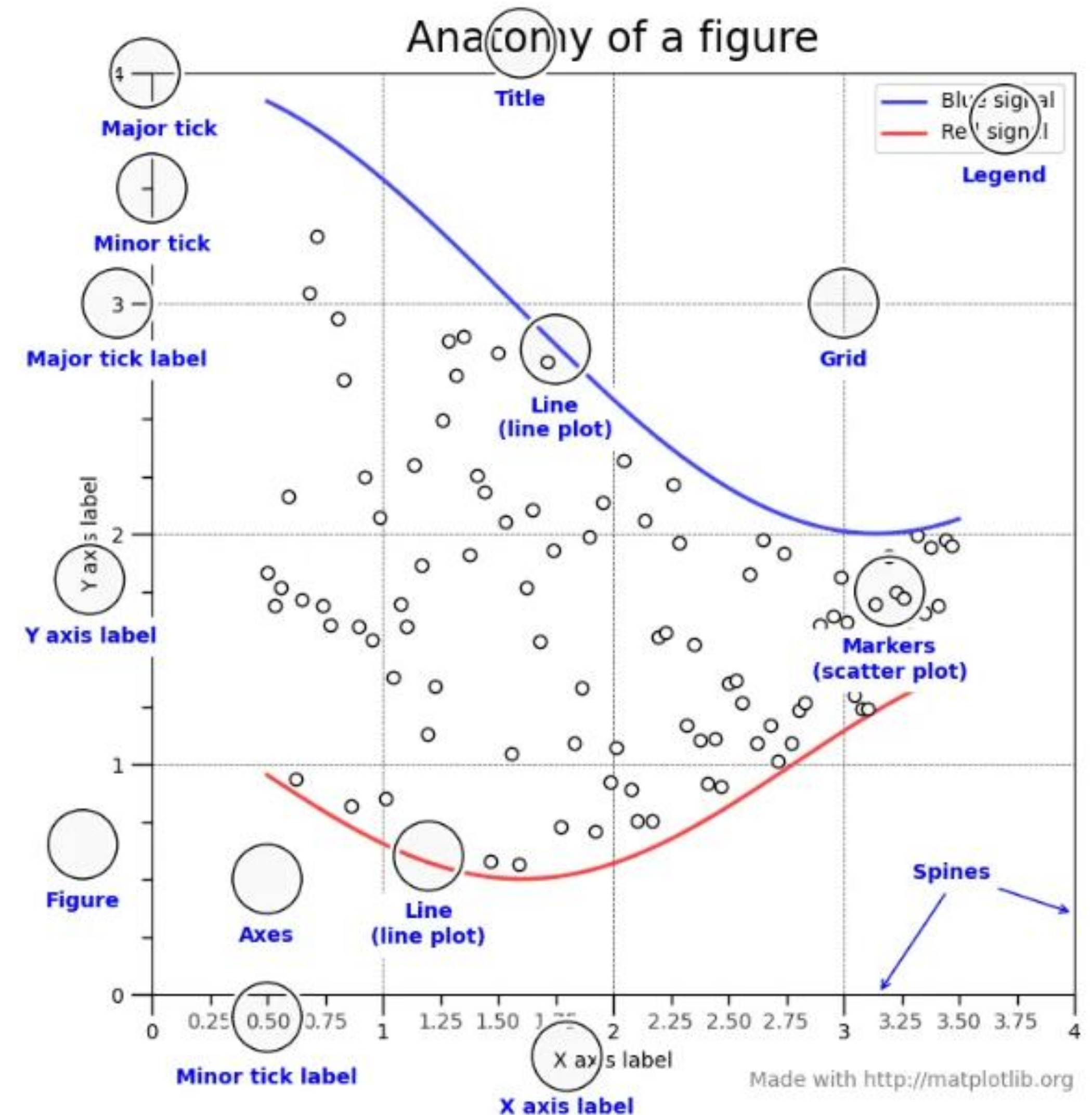
**Axis:** Refere-se a cada eixo individualmente (eixo x ou y), acessado via XAxis ou YAxis.



Fonte: [Real Python](#)

# Matplotlib

**Artist:** É a base de todos os elementos gráficos, incluindo figuras, eixos, títulos e legendas, além de itens específicos como patch, collections e line2D. Basicamente, **tudo o que é visível em um gráfico no Matplotlib é considerado um “Artist”**



Anatomia de uma 'Figure' do Matplotlib. Fonte: [Quick start guide](#) — Matplotlib

# Principais Métodos

**pyplot.subplots():** Cria a figura e os eixos. Pode definir o número de linhas e colunas, como (2,2) para criar 4 eixos. Sem parâmetros, cria 1 eixo.

**Axes.plot():** Adiciona dados ao eixo, precisa de um conjunto de dados e pode incluir parâmetros como cor e legenda.

**Axes.set\_title():** Adiciona um título ao eixo (recebe uma string).

# Principais Métodos

**`Axes.set_xlabel()` / `Axes.set_ylabel()`:** Adiciona títulos aos eixos x e y (recebem strings).

**`Axes.legend()`:** Adiciona legenda aos dados plotados, desde que os labels tenham sido definidos.

**`pyplot.show()`:** Renderiza o gráfico na tela. “Artist”

# Exemplos

[https://matplotlib.org/stable/users/explain/quick\\_start.html](https://matplotlib.org/stable/users/explain/quick_start.html)



# ***Como usar cada tipo de gráfico***

## Gráfico de Linhas

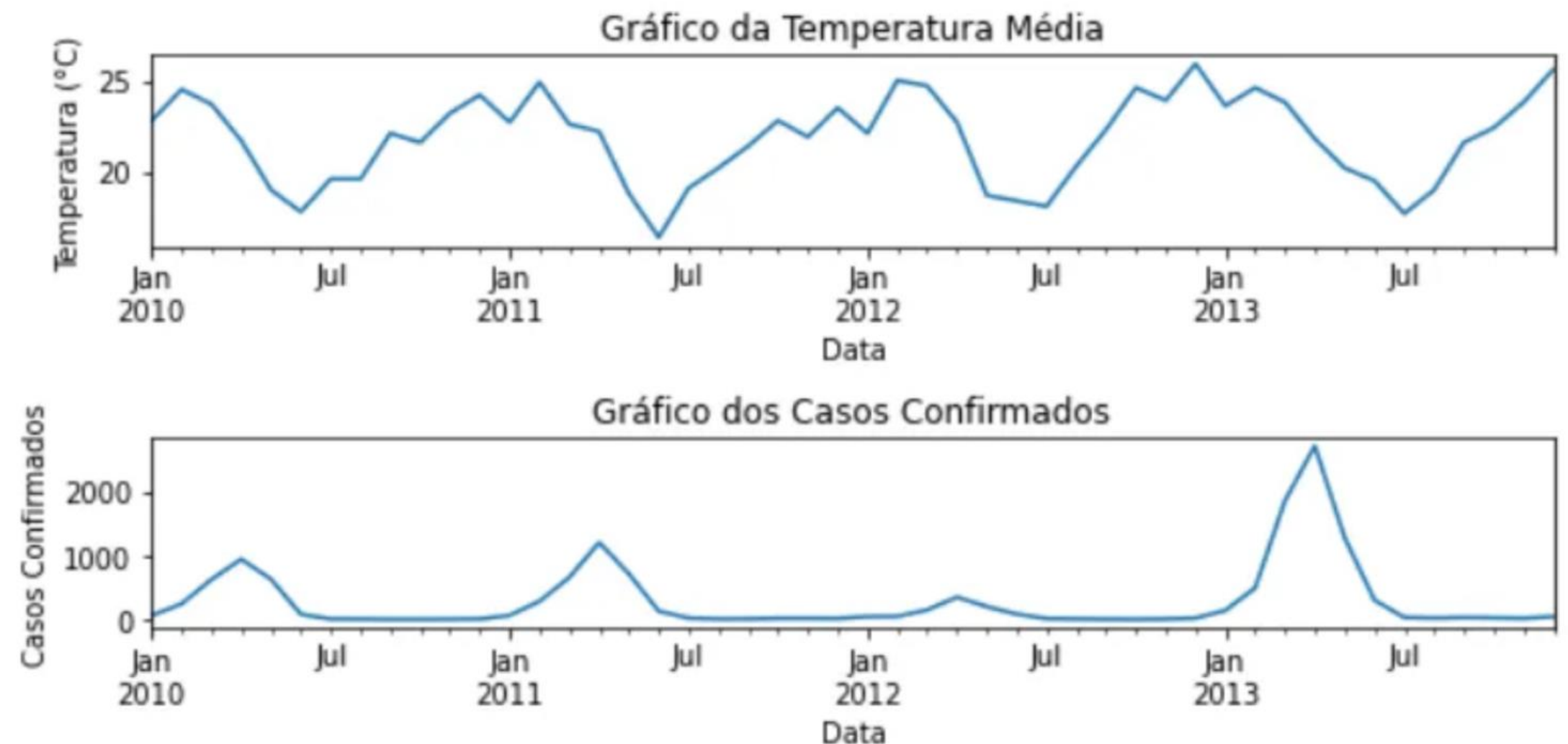
**Visualizar padrões ou tendências ao longo do tempo.**

Usados frequentemente em dados financeiros, climáticos e demográficos..

- Eixo-x: ordem cronológica.
- Eixo-y: variação de uma variável dependente.

Mais adequados para variáveis numéricas contínuas, não-categóricas.

**Regra geral:** use gráficos de linha para mostrar como os dados se conectam e evoluem ao longo do tempo (princípio da conexão de Gestalt).



# Histograma

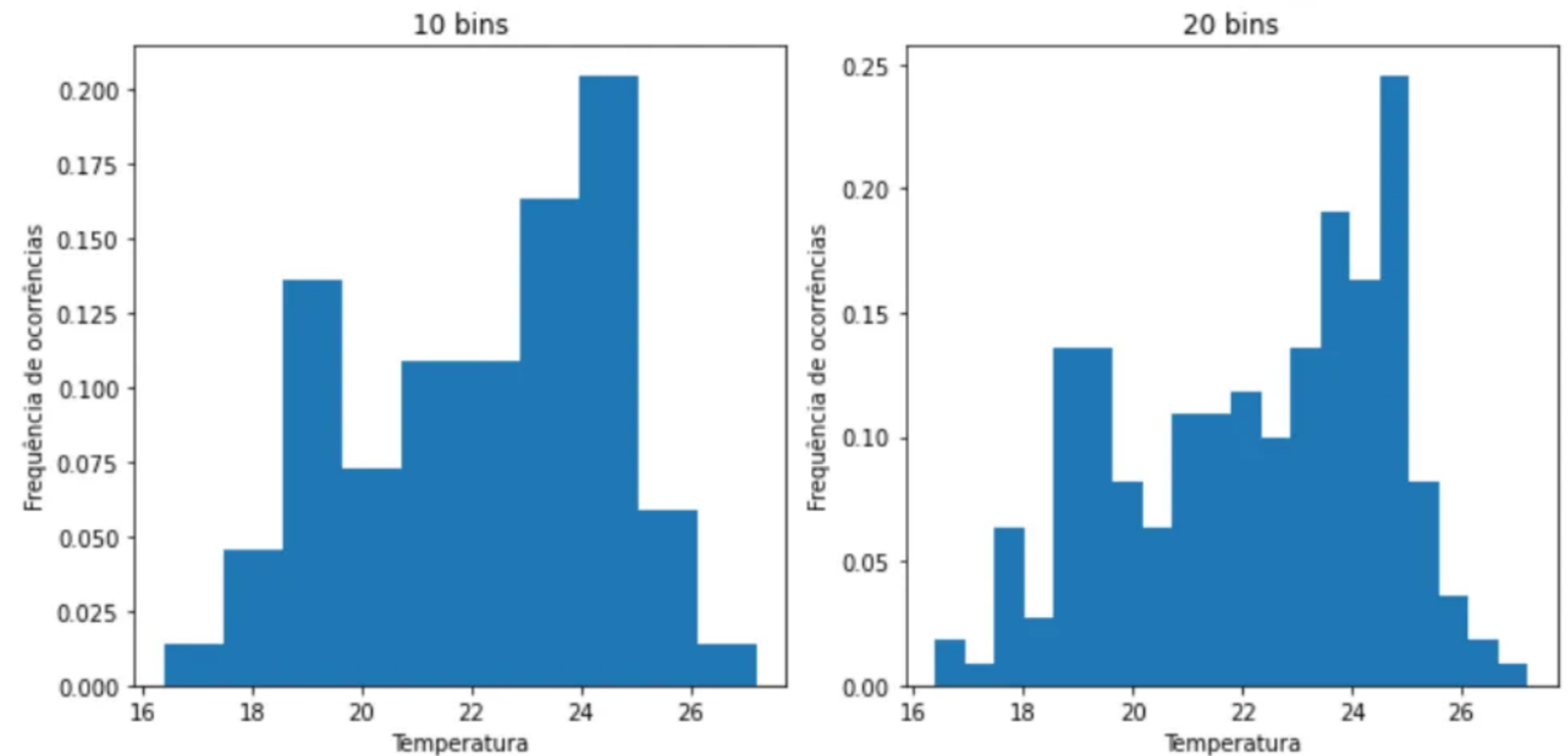
O histograma é usado para visualizar a distribuição de uma variável contínua.

Representa a frequência de uma variável ao dividir seus valores em intervalos (ou bins) e contar as ocorrências em cada intervalo.

## Escolha de bins:

- Poucos bins simplificam demais a distribuição.
- Muitos bins dificultam a visualização dos padrões.

A quantidade ideal de bins depende do tamanho do dataset e do detalhamento desejado.



## Regra comum:

**Freedman-Diaconis** usa o **IQR** (intervalo entre o primeiro e terceiro quartil) para calcular o número de bins. Use histograma para analisar padrões e a distribuição de frequência de uma variável contínua.

$$bins = 2 \frac{IQR}{n^{1/3}}$$

Regra de Freedman-Diaconis para determinar o número de bins de um histograma.

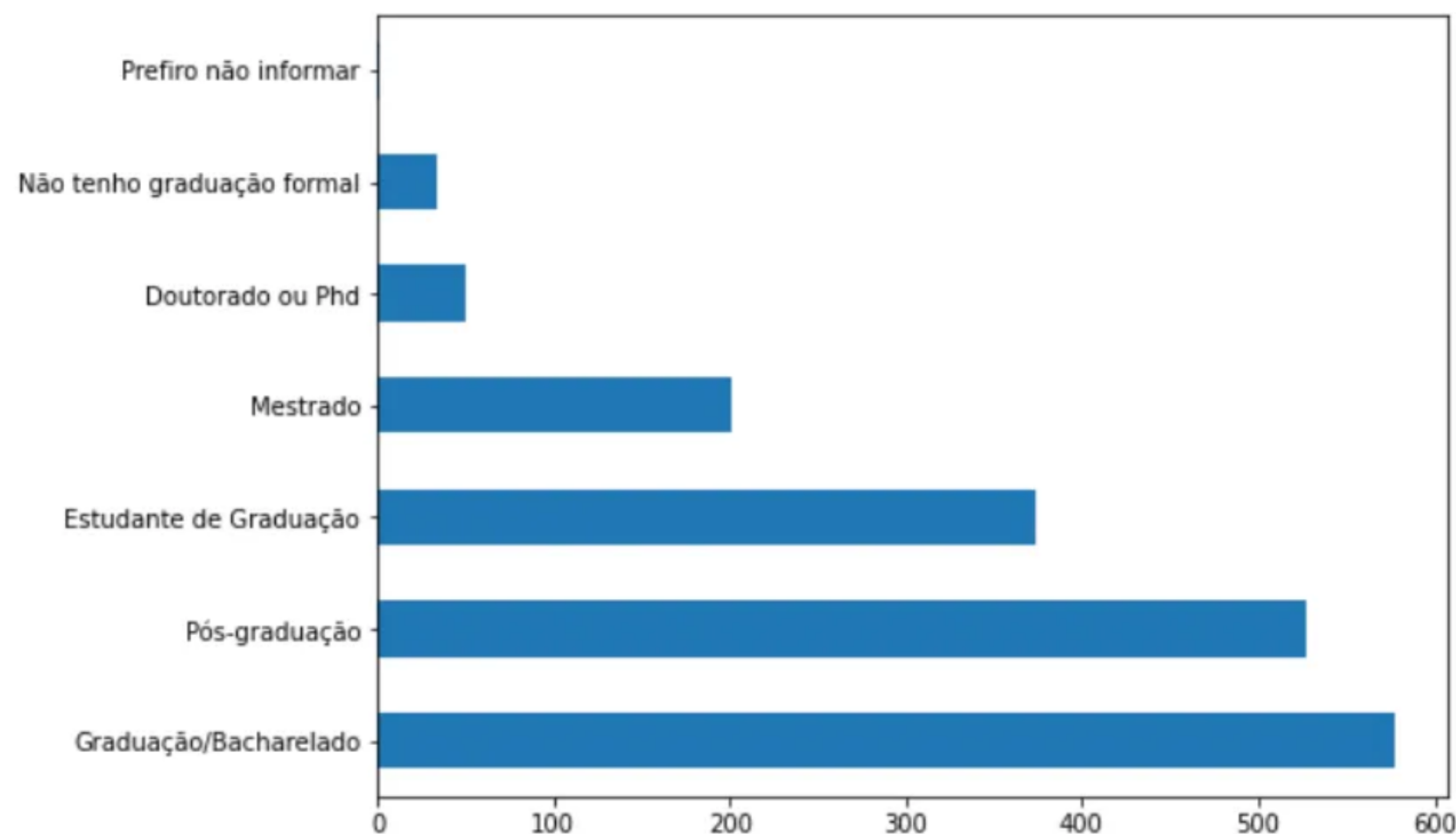
## Gráfico de Barras

O gráfico de barras usa barras proporcionais aos valores representados.

Indicado para dados discretos,  
ao contrário dos histogramas.

Altura das barras representa o  
valor ou frequência de cada  
categoria.

Usado em pesquisas de mercado,  
vendas, finanças, educação e saúde.



Exemplo de gráfico de barras que expõe uma variável categórica de escolaridade.



***Gráfico de Barras***



***Histograma***

Diferenças entre gráfico de barras e histograma:

- **Histograma:** Mostra a distribuição de frequência de uma variável contínua, com barras adjacentes (sem gaps) para indicar continuidade.
- Gráfico de barras: Compara categorias discretas, com barras separadas por gaps para destacar a natureza não-contínua dos dados.

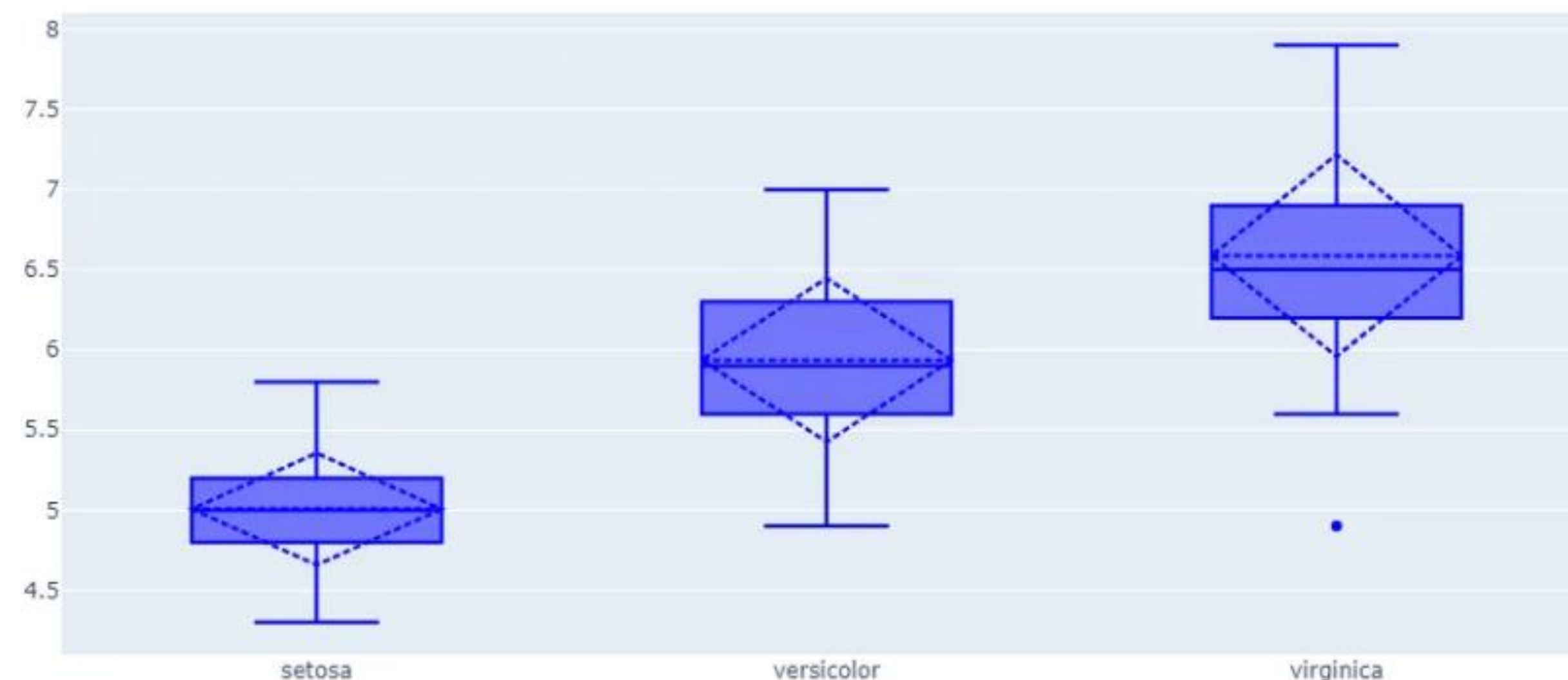
## ***Boxplot***

O boxplot é ótimo para extrair informações estatísticas dos dados.

Use para: Comparar múltiplos datasets e detectar diferenças nas distribuições.

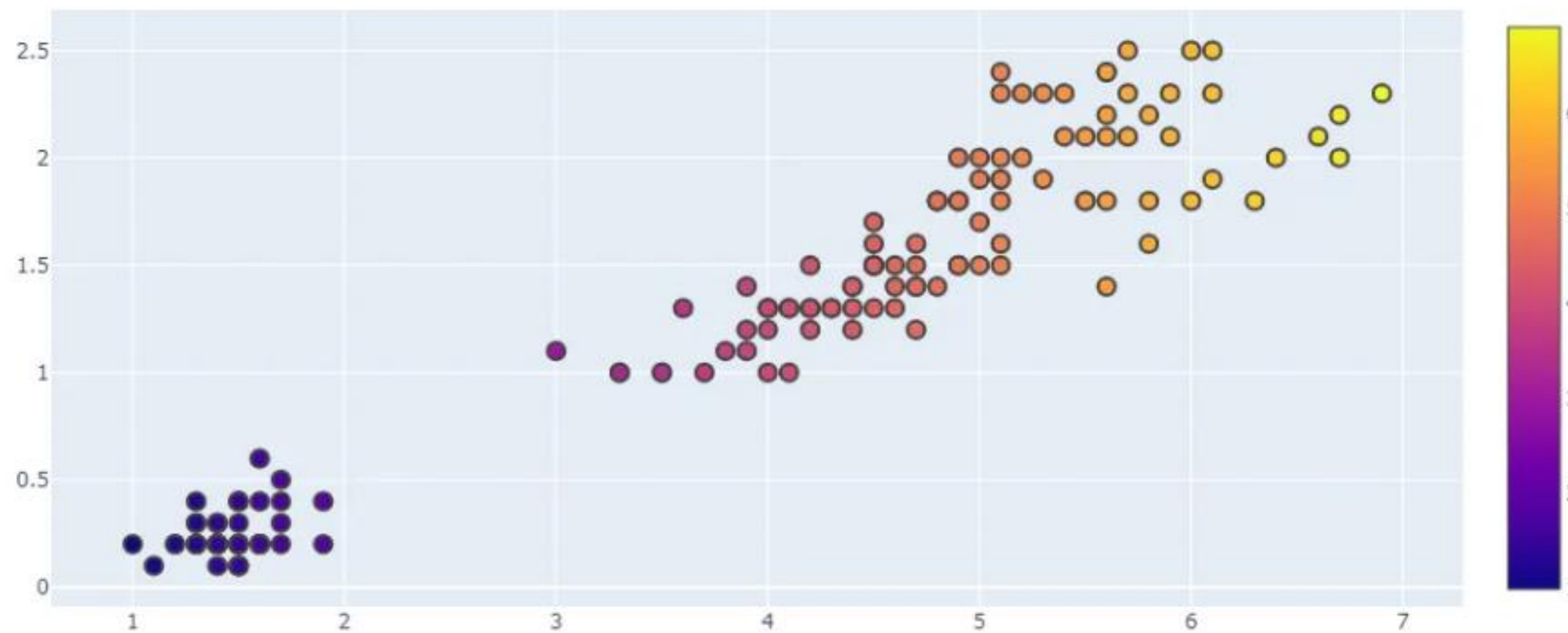
Visualizar a distribuição de datasets extensos.

Não recomendado para amostras pequenas, pois podem não representar bem a distribuição geral.



# Gráfico de Dispersão

São representações bidimensionais onde cada ponto é um par de valores (x, y).



Scatterplot da distribuição das pétalas com o `plotly.graph_objects`. Fonte: Autor.

A variável x é independente e a y é dependente.

Útil para visualizar a correlação entre variáveis e observar tendências e padrões na distribuição dos pontos.

Muito usado em pesquisas científicas para mostrar relações e simplificar problemas.

Excelente para detectar **outliers**, pois valores distantes do aglomerado ou da tendência geral ficam evidentes.

## ***Pesquise, quando usar***

***Spider plot***

***Pie Chart***

***Mapa de Densidade***

***Violin Plots***

***Area Chart***

# ***Storytelling***



## ***Entender o Contexto e o Público***

Identificar o problema que está sendo resolvido, os dados disponíveis e, especialmente, quem vai consumir essa história.

**Definir o objetivo:** O que a narrativa pretende resolver ou esclarecer?

**Identificar o público-alvo:** São executivos, clientes, colegas de equipe?  
Cada grupo precisa de uma abordagem diferente.

**Perguntas-chave:** Quais são as principais perguntas que você precisa responder com seus dados?

## ***Entender o Contexto e o Público***

### **Exemplo:**

Em uma análise de transporte urbano, os gestores da empresa precisam entender quais áreas da cidade são mais problemáticas em termos de atrasos.

Então, sua narrativa deve focar em fornecer respostas práticas sobre como otimizar as operações.

## ***Definir a Pergunta Central (ou Hipótese)***

Escolher uma questão ou hipótese principal que guiará a análise dos dados.  
A narrativa precisa de um foco claro.

**Formule uma pergunta clara:** Qual é o problema que a análise pretende resolver?

**Verifique a relevância:** A pergunta é importante para o público? Vai gerar valor para a tomada de decisão?

**Hipótese inicial:** Faça suposições ou hipóteses que possam ser confirmadas ou refutadas pela análise de dados.

## ***Definir a Pergunta Central (ou Hipótese)***

### **Exemplo:**

"Quais linhas de ônibus têm maior taxa de superlotação durante os horários de pico?"

ou

"Como podemos reduzir os atrasos nas áreas centrais da cidade?"

## ***Explorar os Dados (Ou Criar uma Lógica Baseada em Dados Hipotéticos)***

Realizar uma exploração inicial dos dados ou construir uma lógica que conecte os pontos principais de forma estruturada.

**Identificar padrões:** Quais tendências ou padrões podem ser observados?

**Selecionar insights relevantes:** Escolha apenas os dados que ajudam a responder a pergunta central.

**Filtrar o ruído:** Evite dados que não contribuem para a história principal.



## ***Explorar os Dados (Ou Criar uma Lógica Baseada em Dados Hipotéticos)***

### **Exemplo:**

"Os dados mostram que a linha 45 está constantemente atrasada nos horários entre 7h e 9h, enquanto outras linhas não apresentam problemas graves."

## ***Construir a História: Estrutura Clássica do Storytelling***

A narrativa de dados pode ser estruturada como uma história com início, meio e fim. Seguir uma estrutura facilita a compreensão dos insights e das conclusões.

**Introdução:** Apresente o problema que a análise vai resolver.

**Desenvolvimento:** Mostre como os dados fornecem insights sobre o problema. Aqui você apresenta gráficos, tendências e análises.

**Conclusão:** Baseado nos dados, proponha recomendações ou soluções práticas.

## ***Construir a História: Estrutura Clássica do Storytelling***

### **Exemplo:**

**Introdução:** "A CityMoves está enfrentando atrasos constantes em certas linhas durante o horário de pico, causando insatisfação entre os passageiros."

**Desenvolvimento:** "A análise revela que 60% dos atrasos ocorrem nas linhas que passam pela área central entre 7h e 9h."

**Conclusão:** "Recomendamos o aumento da frota em 30% para as linhas 45 e 92 nesse período para reduzir os atrasos."

## ***Selecionar Visualizações Apropriadas***

As visualizações são ferramentas essenciais para contar uma história com dados. Elas ajudam a tornar os insights mais claros e acessíveis.

**Escolha visualizações que facilitem a interpretação:** Use gráficos de barras, linhas, ou mapas, dependendo do que está sendo explicado.

**Simplifique:** Não sobrecarregue o público com gráficos complexos; foque na clareza.

**Conecte a visualização com a história:** Cada gráfico deve responder a uma parte da pergunta central ou reforçar um insight importante.

## ***Selecionar Visualizações Apropriadas***

### **Exemplo:**

Um gráfico de barras mostrando os atrasos por linha de ônibus, segmentado por horário do dia, ajudaria a visualizar claramente o problema.



# ***Trilha de Análise de dados***

**Módulo 1: Terminar até dia 15/10**

**Módulo 2: Terminar até dia 22/10**

**Módulo 3: Terminar até dia 29/10**

# ***Trilha de Intro Inteligência Artificial***

**Módulo 1: Terminar até dia 25/10**

**Módulo 2: Terminar até dia 01/11**

**Módulo 3: Terminar até dia 08/11**

**Módulo 4: Terminar até dia 15/11**

**Módulo 5: Terminar até dia 22/11**

# ***Trilha de Intro Machine Learning***

**Módulo 1: Terminar até dia**

**Módulo 2: Terminar até dia**

**Módulo 3: Terminar até dia**

**Módulo 4: terminar até dia**

***Até a próxima aula...***

***16/10/24***