

## ✓ Ficha de Exercícios - Visualização com Matplotlib

### Contextualização do Dataset: Iris

O dataset **Iris** é um dos conjuntos de dados mais conhecidos em estatística e ciência de dados. Foi publicado pelo estatístico britânico **Ronald A. Fisher** em 1936, e contém informações sobre 150 amostras de flores da espécie *Iris*, divididas em três subespécies:

- *Iris setosa*
- *Iris versicolor*
- *Iris virginica*

Para cada flor, foram recolhidas quatro medidas:

- **Comprimento da sépala** (*sepal\_length*)
- **Largura da sépala** (*sepal\_width*)
- **Comprimento da pétala** (*petal\_length*)
- **Largura da pétala** (*petal\_width*)

Estas medidas estão expressas em centímetros (cm) e permitem distinguir facilmente entre as espécies com recurso a visualizações.

Este dataset é ideal para aprender a utilizar bibliotecas de visualização como o **Matplotlib**, explorando relações entre variáveis, distribuições e comparações entre grupos.

[+ Code](#)
[+ Text](#)

```
# 📦 Importação de bibliotecas
import pandas as pd
import matplotlib.pyplot as plt

# 📄 Carregamento do dataset
url = "https://raw.githubusercontent.com/uiuc-cse/data-fa14/gh-pages/data/iris.csv"
df = pd.read_csv(url)

# 📊 Visualização inicial dos dados
df.head()
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

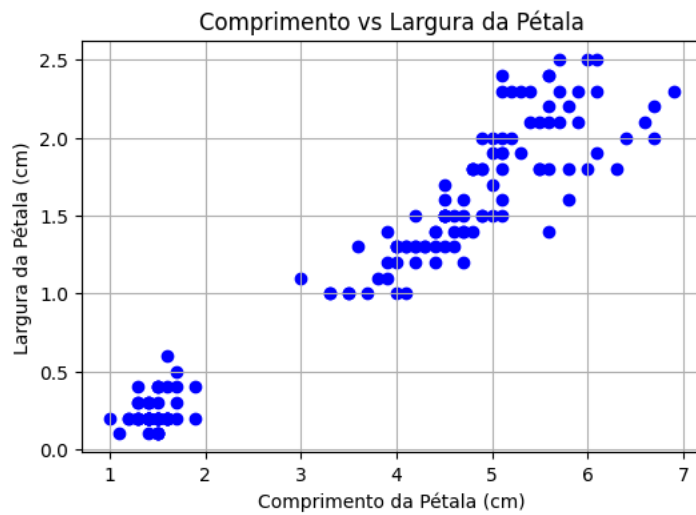
Next steps:

[Generate code with df](#)
[View recommended plots](#)
[New interactive sheet](#)

## ✓ Exemplo 1: Gráfico de Dispersão

Visualização da relação entre duas variáveis numéricas com um gráfico de dispersão (scatter plot).

```
# Relação entre comprimento e largura da pétala
plt.figure(figsize=(6, 4))
plt.scatter(df['petal_length'], df['petal_width'], color='blue')
plt.title("Comprimento vs Largura da Pétala")
plt.xlabel("Comprimento da Pétala (cm)")
plt.ylabel("Largura da Pétala (cm)")
plt.grid(True)
plt.show()
```



### Exercício 1:

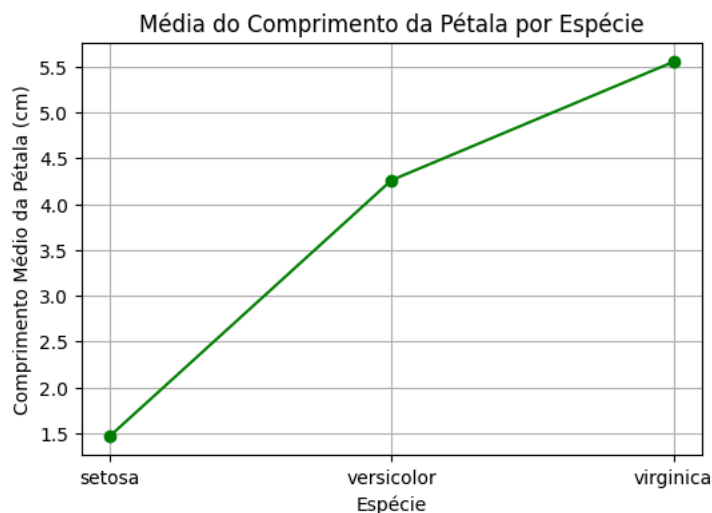
Cria um gráfico de dispersão para analisar a relação entre o **comprimento e a largura da sépala**. Inclui os rótulos dos eixos e o título do gráfico.

### Exemplo 2: Gráfico de Linhas

Visualização de valores agregados (como médias) por categoria usando linhas.

```
# Média do comprimento da pétala por espécie
medias = df.groupby("species")["petal_length"].mean()

plt.figure(figsize=(6, 4))
plt.plot(medias.index, medias.values, marker='o', linestyle='-', color='green')
plt.title("Média do Comprimento da Pétala por Espécie")
plt.xlabel("Espécie")
plt.ylabel("Comprimento Médio da Pétala (cm)")
plt.grid(True)
plt.show()
```



### Exercício 2:

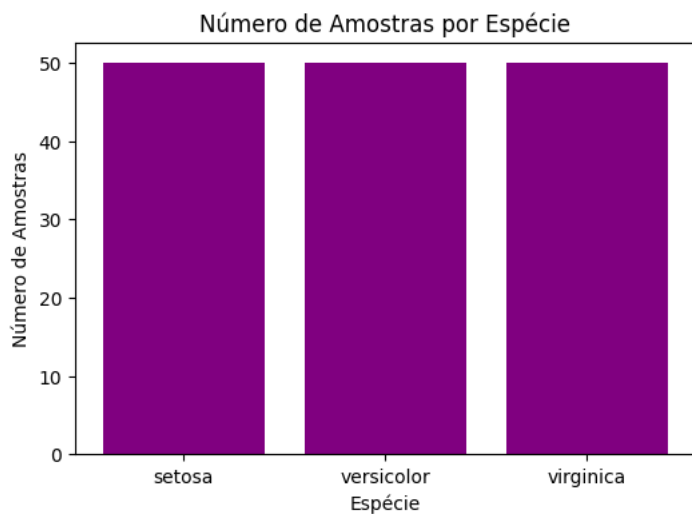
Cria um gráfico de linhas que represente a **média da largura da sépala** por espécie. Inclui marcadores nos pontos e personaliza o gráfico.

### Exemplo 3: Gráfico de Barras

Visualização de quantidades ou valores agregados por categoria com barras verticais.

```
# Número de registos por espécie
contagem = df["species"].value_counts()
```

```
plt.figure(figsize=(6, 4))
plt.bar(contagem.index, contagem.values, color='purple')
plt.title("Número de Amostras por Espécie")
plt.xlabel("Espécie")
plt.ylabel("Número de Amostras")
plt.show()
```



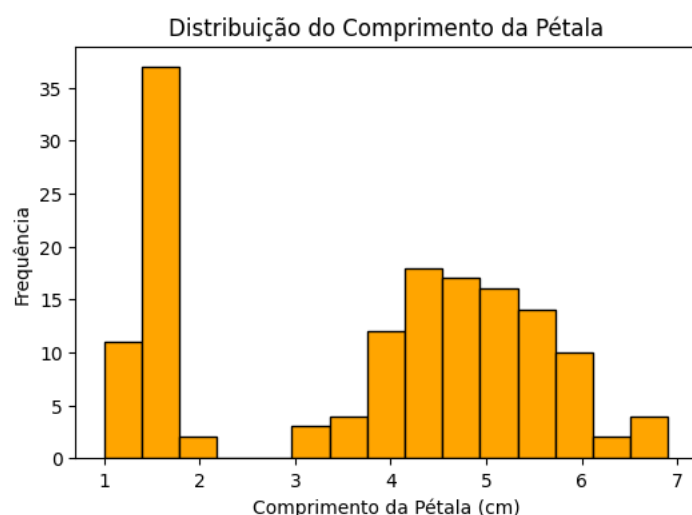
### Exercício 3:

Cria um gráfico de barras com a **média do comprimento da sépala** por espécie. Adiciona título e rótulos nos eixos.

### Exemplo 4: Histograma

Visualização da distribuição de uma variável contínua.

```
# Distribuição do comprimento da pétala
plt.figure(figsize=(6, 4))
plt.hist(df["petal_length"], bins=15, color='orange', edgecolor='black')
plt.title("Distribuição do Comprimento da Pétala")
plt.xlabel("Comprimento da Pétala (cm)")
plt.ylabel("Frequência")
plt.show()
```



### Exercício 4:

Cria um histograma para representar a **distribuição da largura da sépala**. Usa pelo menos 10 intervalos (bins).

### Exemplo 5: Gráfico de Setores (Pizza)

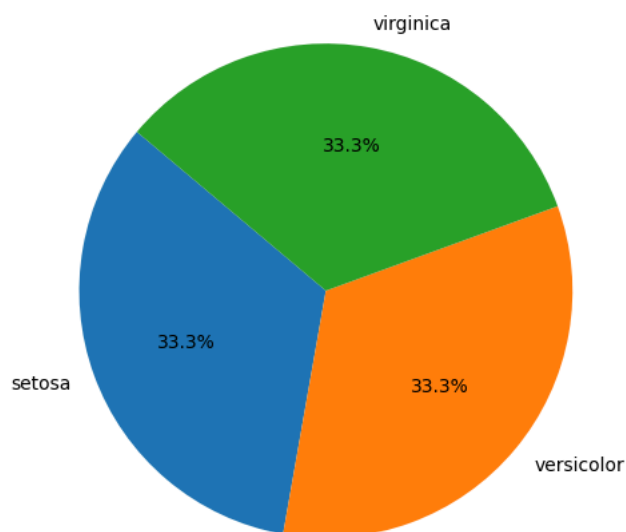
Visualização da proporção de categorias numa variável qualitativa.

```
# Proporção de espécies no dataset  
contagem = df["species"].value_counts()
```

```
plt.figure(figsize=(6, 6))  
plt.pie(contagem.values, labels=contagem.index, autopct='%1.1f%%', startangle=140)  
plt.title("Distribuição de Espécies no Dataset")  
plt.show()
```



Distribuição de Espécies no Dataset



### Exercício 5:

Cria um gráfico de setores com a **frequência das espécies** no dataset. Inclui percentagens e título.



### Desafio Final



### Desafio:

Cria um gráfico que combine múltiplos tipos (por exemplo, barras e linhas) para representar **duas variáveis diferentes por espécie** (ex: comprimento e largura da pétala).

Podes usar `plt.twinx()` ou criar duas figuras lado a lado.