

## Módulo 5 – Técnicas de programação II | python

Turma 815

### Perguntas e respostas

**P:** Notei que aprendemos sobre o plotly express. Lendo a documentação, vi que tem um plotly mais "completo". Vocês já usaram ele? Sabem qual a diferença para o express o quando é melhor usar a versão completa?

**R:** No geral, utilizo mais a versão express, por ser bem completa e simples de utilizar. Mas quando quero gerar muitos gráficos diferentes numa mesma figura, acabo precisando utilizar o graph\_objects. Além disso, alguns tipos de gráficos ainda não são possíveis de serem feitos pelo express (gráficos 3D, por exemplo).

Na documentação do plotly, sempre que você for procurar pela construção de um tipo de gráfico, sempre irá encontrar como fazer esse gráfico utilizando o plotly express, e depois, utilizando o graph\_objects ([exemplo](#)). Vocês verão que, no fundo, o express utiliza o graph\_objects por trás, mas transforma um código de várias linhas em um código super reduzido (vide exemplo abaixo, retirado da própria documentação):

```
import pandas as pd

df = pd.DataFrame({
    "Fruit": ["Apples", "Oranges", "Bananas", "Apples", "Oranges", "Bananas"],
    "Contestant": ["Alex", "Alex", "Alex", "Jordan", "Jordan", "Jordan"],
    "Number Eaten": [2, 1, 3, 1, 3, 2],
})
```

```
# Plotly Express
import plotly.express as px

fig = px.bar(df, x="Fruit", y="Number Eaten", color="Contestant", barmode="group")
fig.show()
```

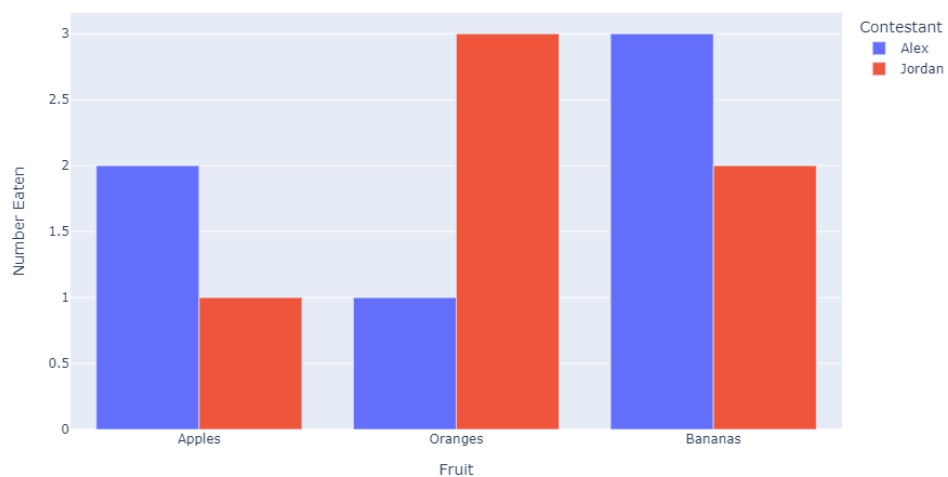


Figura 1 - Gráfico feito utilizando plotly.express

```
# Graph Objects
import plotly.graph_objects as go

fig = go.Figure()

for contestant, group in df.groupby("Contestant"):
    fig.add_trace(go.Bar(x=group["Fruit"], y=group["Number Eaten"], name=contestant,
        hovertemplate="Contestant=%s<br>Fruit=%{x}<br>Number Eaten=%{y}<extra></extra>"% contestant))

fig.update_layout(legend_title_text = "Contestant")
fig.update_xaxes(title_text="Fruit")
fig.update_yaxes(title_text="Number Eaten")
fig.show()
```

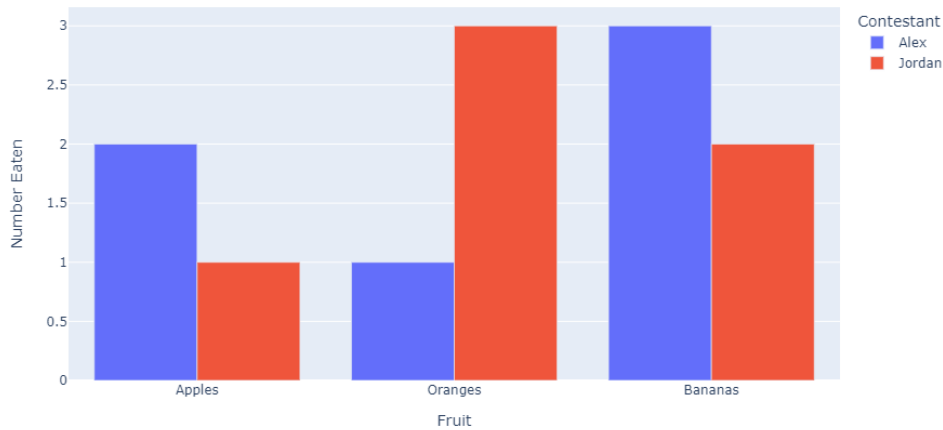


Figura 2 - Gráfico feito utilizando `graph_objects`

**P:** Casos de uso para cada gráfico; Quais dados são melhores para cada gráfico?

**R:** Não existe uma resposta certa para essa pergunta. Dependendo de cada contexto, o tipo de gráfico a exibir melhor a informação pode ser diferente. No geral, podemos seguir com os exemplos mostrados na primeira aula: Gráficos de linhas são ótimos para mostrar tendências ao longo do tempo; Histograma e box-plot servem bem para observar a distribuição de uma única variável contínua em uma amostra de dados; Scatter plot ajuda a ver possíveis correlações entre variáveis; Gráficos de barra e pizza são muito utilizados para analisar proporções de uma variável por grupos. Mas isso não é regra ditada, é possível criar uma mesma visualização com gráficos diferentes, e tudo vai depender de contexto e preferências: por exemplo, muitas vezes o box-plot é uma ótima visualização para analisar a distribuição de uma variável por grupos, porém não é um gráfico muito trivial de se entender – pode não ser uma boa ideia utilizá-lo, dependendo do cliente para quem você for exibir esse gráfico.

---

**P:** Como Montar uma Data-Viz do jeito que as empresas gostam?

**R:** Também não tem resposta certa para essa pergunta. Cada cliente tem suas preferências, muito a depender da área em que se está trabalhando. Mas no geral, gráficos limpos, com legendas e títulos intuitivos são bem-vistos – busque sempre montar um gráfico que alguém possa interpretar sozinho, sem precisar que você explique cada detalhe.

---

**P:** Quais são as melhores aplicações para cada biblioteca? Por exemplo, quando é recomendado usar Matplotlib, qual gráfico funcionaria melhor em plotly? Qual a vantagem de usar o seaborn em determinada apresentação?

**R:** Na maioria das vezes, a melhor biblioteca é a que você tiver mais facilidade para mexer. Todas as bibliotecas de visualização do python têm suas vantagens e desvantagens, mas conseguem gerar gráficos ótimos de todos os tipos. O que muda é a quantidade de código necessário para criar o gráfico, e os parâmetros das funções.

---

**P:** Quais são os tipos de gráficos que vocês mais usam no dia-a-dia? Quando vão fazer um EDA, como fazem? Tem exemplos? Se pudéssemos fazer uma pequena simulação com algum DataFrame, seria ótimo.

**R:** Mais uma vez, não tem resposta certa! Depende muito do tipo de dado que você está utilizando, e seu objetivo no trabalho. Numa análise exploratória, se não for ser exibido para nenhum cliente, não é necessário gastar muito tempo escolhendo as melhores cores e títulos para o gráfico – basta que ele esteja suficientemente bom para você responder à sua pergunta. O projeto é uma ótima forma de treinar a escolha do gráfico.

---

**P:** Vocês tem exemplos de gráficos que já fizeram para apresentar para outras áreas? Eu gostaria de saber mais de boas práticas de apresentação e como vocês gostam de usar os templates.

**R:** Muito difícil conseguir mostrar um case real, pois muitas vezes possuem dados sensíveis. Deixo aqui um exemplo público: <https://repositorio.usp.br/directbitstream/f86d986a-01b5-4745-8a82-4a13387349a7/2926912.pdf> (um dos meus TCCs). Estava em clima de copa do mundo, então utilizei as cores amarelo, azul e verde para fazer os gráficos! 😊 (páginas 32 a 53).