

Nome: Gabriela Braz

Curso: Análise e Desenvolvimento de Sistemas

Link do dataset utilizado: <https://www.kaggle.com/datasets/mohithsairamreddy/salary-data>

Link do meu video de apresentação: <https://www.loom.com/share/52991a790a39426e82de26c7e29032cc>

Plot 1 – Gráfico de Setores

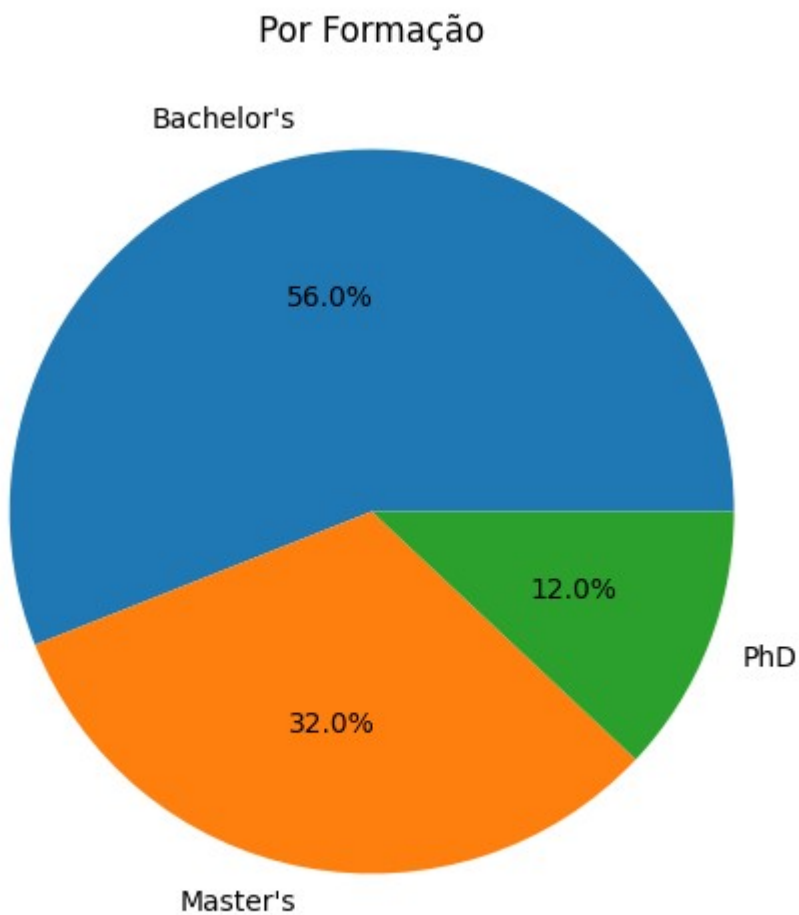


Gráfico de setores mostra como toda a base está dividida em relação à formação acadêmica dos funcionários.

Plot 2 – Gráfico de Barras

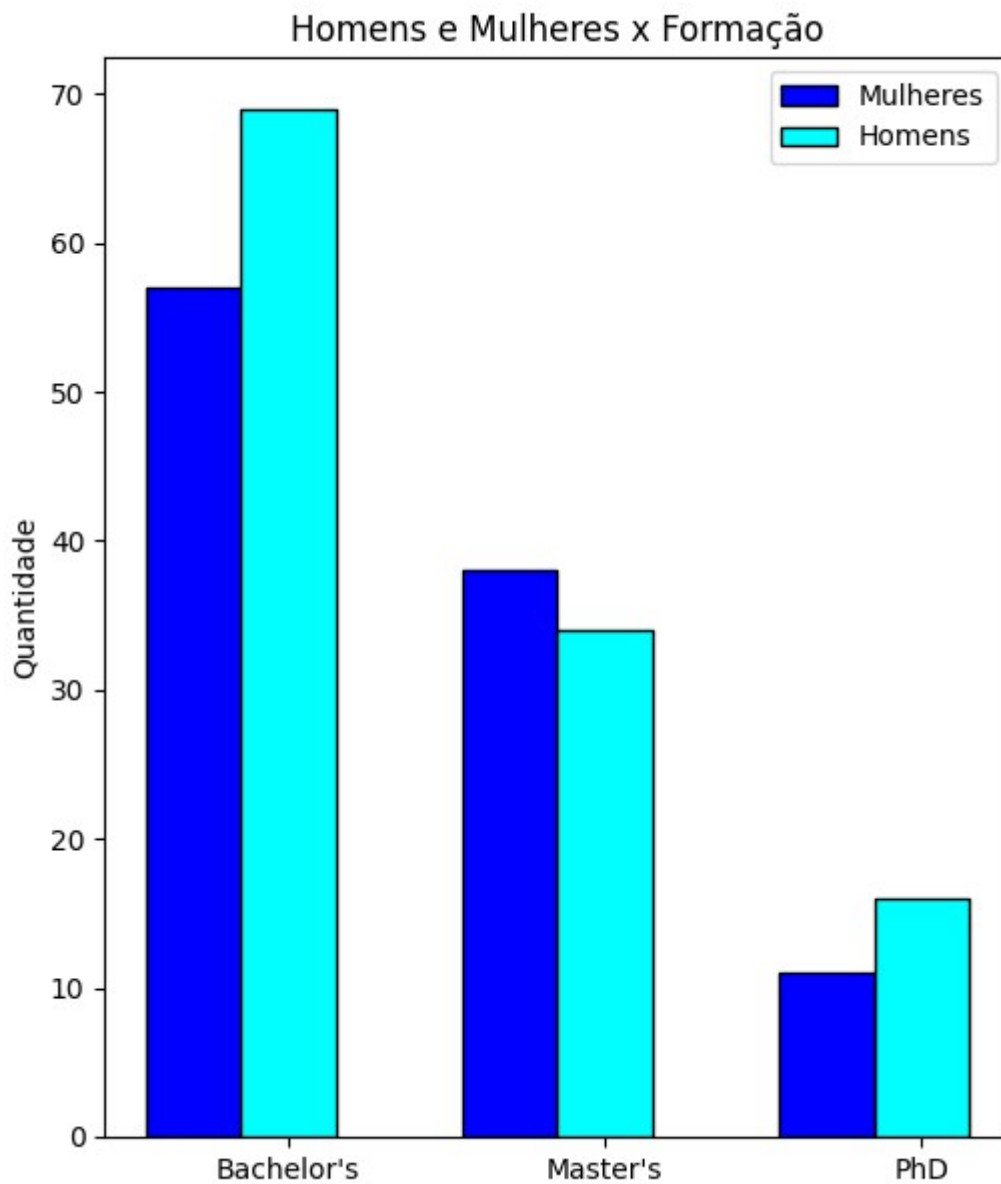


Gráfico de barras mostra como toda a base está dividida em relação ao Gênero dos funcionários dentro das suas formações acadêmicas.

Plot 3 – Gráfico Polar - Hierárquico

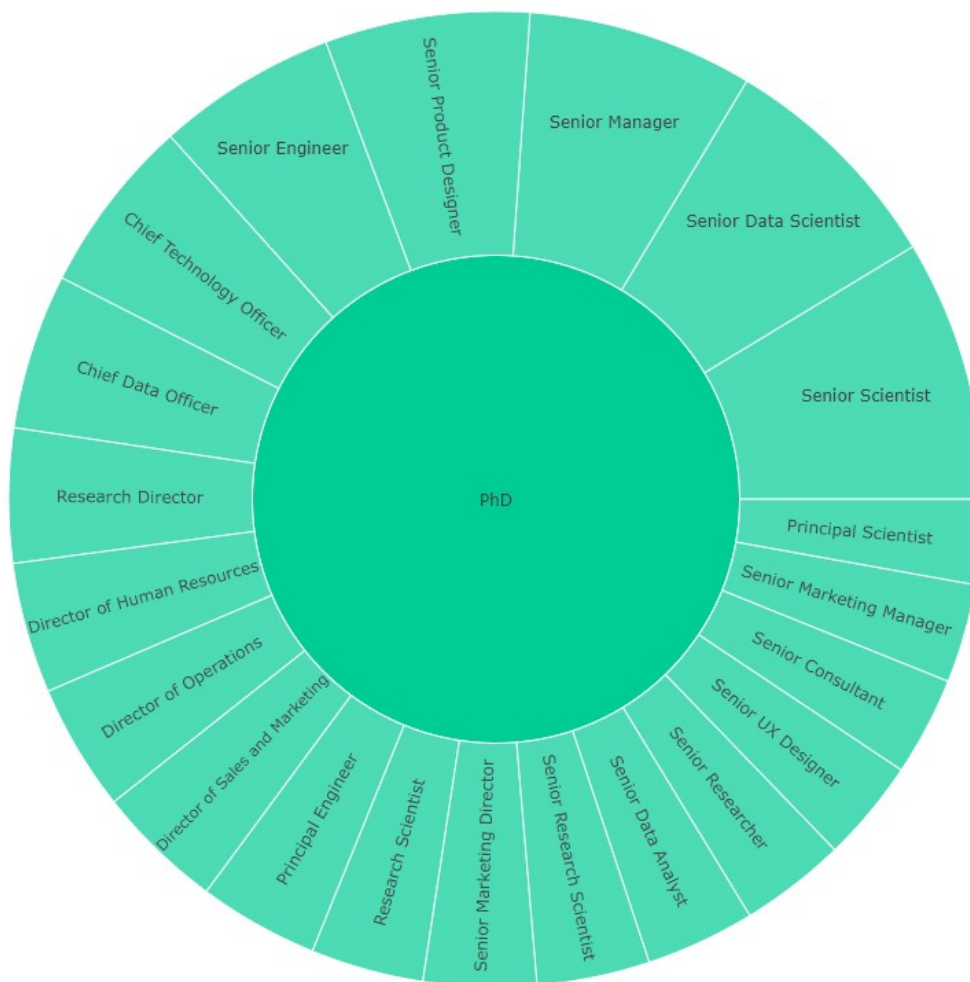


Gráfico Hierárquico mostra como toda a base está dividida em relação aos Cargos dentro das suas formações acadêmicas.

Código fonte Python

```
import plotly.express as px
import pandas as pd
import csv
import matplotlib.pyplot as plt
import numpy as np
from pandas import read_csv

formacao = []
cargo = []
salario = []

series = read_csv(r"salary_data.csv", header=0,
index_col=0, parse_dates=True, sep = ',')
series.head()
```

```

cargo1F = 0
cargo2F = 0
cargo3F = 0

cargo1M = 0
cargo2M = 0
cargo3M = 0

with open('salary_data.csv','r') as csvfile:
    plots = csv.reader(csvfile, delimiter = ',')

    for row in plots:
        if plots.line_num > 1 :
            formacao.append(row[2])
            cargo.append(row[3])
            salario.append(row[5])

            if row[2] == "Bachelor's": #Somente Formação = Bachelor's
                if row[1] == "Female":
                    cargo1F = cargo1F + 1
                if row[1] == "Male":
                    cargo1M = cargo1M + 1

            if row[2] == "Master's": #Somente Formação = Master's
                if row[1] == "Female":
                    cargo2F = cargo2F + 1
                if row[1] == "Male":
                    cargo2M = cargo2M + 1

            if row[2] == "PhD": #Somente Formação = PhD
                if row[1] == "Female":
                    cargo3F = cargo3F + 1
                if row[1] == "Male":
                    cargo3M = cargo3M + 1

plt.rc('figure', figsize = (15, 7))
# Definindo área do gráfico
area = plt.figure()
g1 = area.add_subplot(1,2,1)
g2 = area.add_subplot(1,2,2)

#Dados Agrupados por Formação ***** SETORES / PIZZA
grupo1 = series.groupby('Education Level').Salary

label = grupo1.count().index
valores = grupo1.count().values

g1.pie(valores, labels = label, autopct = '%1.1f%%')
g1.set_title('Por Formação')

#Dados Agrupados por Gênero ***** BARRAS

```

```

# width of the bars
barWidth = 0.3

# Choose the height of the blue bars
bars1 = [cargo1F, cargo2F, cargo3F]

# Choose the height of the cyan bars
bars2 = [cargo1M, cargo2M, cargo3M]

# The x position of bars
r1 = np.arange(len(bars1))
r2 = [x + barWidth for x in r1]

# Create blue bars
plt.bar(r1, bars1, width = barWidth, color = 'blue', edgecolor = 'black', capsize=7,
label='Mulheres')

# Create cyan bars
plt.bar(r2, bars2, width = barWidth, color = 'cyan', edgecolor = 'black', capsize=7,
label='Homens')

# general layout
plt.xticks([r + barWidth for r in range(len(bars1))], ["Bachelor's", "Master's", "PhD"])
plt.ylabel('Quantidade')
plt.title('Homens e Mulheres x Formação')
plt.legend()
# Show graphic
plt.show()

#Dados por Hierarquia (Formação >> Cargos) ***** Sunburst
df = pd.DataFrame(dict(cargo=cargo,formacao=formacao,salario=salario))
df["all"] = "all" #garante um único nó raiz
fig = px.sunburst(df,path=[formacao,cargo],values=salario)
fig.show()

```

Referencias Bibliográficas

Exemplos de códigos fontes para implementação de gráficos de Barras

<https://www.python-graph-gallery.com/8-add-confidence-interval-on-barplot>

acessado em 28/05/2023

Datasets para uso nos exemplos

<https://www.kaggle.com/datasets/mohithsairamreddy/salary-data>

acessado em 28/05/2023

Filtrando dataset no python

<https://www.youtube.com/watch?v=6M0PUNw7faE&t=380s>

acessado em 28/05/2023