import **pandas** as **pd**

import **numpy** as **np**

import **matplotlib**.**pyplot** as **plt**

import **seaborn** as **sns**

df= **pd**.**read\_csv**("dados1.csv")

**print**("")

**print**("---Fazendo a leitura das 5 primeiras linhas do csv.---")

**print**(df.**head**())

**print**("-"\*30)

**print**("---Verificando numero de Linhas/Colunas e tipos de dados---")

linhas\_colunas= df.**info**()

**print**(linhas\_colunas)

**print**("-"\*30)

**print**("---Fazendo a leitura de valores nulos.---")

valores\_nulos= df.**isnull**().**sum**()

**print**(valores\_nulos)

**print**("-"\*30)

**print**("---Calculando a Média, Mediana, Desvio-Padrão e variancia.--- ")

media = df[["Passageiros", "Distância (km)", "Ocupação (%)", "Receita (R$)"]].**mean**().**reset\_index**()

**print**(media)

**print**("-"\*30)

**print**("---Calculando a Mediana---")

mediana = df[["Passageiros", "Distância (km)", "Ocupação (%)", "Receita (R$)"]].**median**().**reset\_index**()

**print**(mediana)

**print**("-"\*30)

**print**("---Calculando o Desvio padrão---")

desvio\_padrao=df[["Passageiros", "Distância (km)", "Ocupação (%)", "Receita (R$)"]].**std**().**reset\_index**()

**print**(desvio\_padrao)

**print**("-"\*30)

**print**("---Calculando a Variança.---")

varianca = df[["Passageiros", "Distância (km)", "Ocupação (%)", "Receita (R$)"]].**var**().**reset\_index**()

**print**(varianca)

**print**("-"\*30)

**print**("Calculando o Percentil 25%, 50% e 75%")

percentis = **np**.**percentile**(df["Receita (R$)"], [25, 50, 75])

**print**(f"Percentil 25: {percentis[0]}")

**print**(f"Percentil 50: {percentis[1]}")

**print**(f"Percentil 75: {percentis[2]}")

**print**("-"\*30)

df\_agrupado = df.**groupby**('Companhia').agg({

    'Passageiros': 'sum',

    'Receita (R$)': 'sum'

}).**reset\_index**()

df\_ordenado\_receita = df\_agrupado.**sort\_values**(by='Receita (R$)', ascending=False)

**print**("---Companhia com maior receita---")

**print**(df\_ordenado\_receita)

**print**("-"\*30)

df\_ordenado\_passageiros = df\_agrupado.**sort\_values**(by='Passageiros', ascending=False)

**print**(df\_ordenado\_passageiros)

**print**("-"\*30)

**print**("---Contagem de voos por companhia---")

voos\_por\_companhia = df['Companhia'].**value\_counts**().**reset\_index**()

voos\_por\_companhia.columns = ['Companhia', 'Quantidade de Voos']

**print**(voos\_por\_companhia)

**print**("-"\*30)

receita\_media = df.**groupby**(['Companhia', 'Aeroporto Origem'])['Receita (R$)'].**mean**().**reset\_index**()

receita\_media.**rename**(columns={'Receita (R$)': 'Receita Média (R$)'}, inplace=True)

**print**("---Receita média por companhia e por aeroporto de origem---")

**print**(receita\_media)

**print**("-"\*30)

voos\_por\_aeroporto = df['Aeroporto Origem'].**value\_counts**().**reset\_index**()

voos\_por\_aeroporto.columns = ['Aeroporto Origem', 'Quantidade de Voos']

**print**(voos\_por\_aeroporto.**head**(10))

**plt**.**figure**()

**sns**.**histplot**(df["Passageiros"], kde=True, bins=10, color='green')

**plt**.**title**("Passageiros")

**plt**.**show**()

Gráfico, Histograma

O conteúdo gerado por IA pode estar incorreto.

**plt**.**figure**()

**sns**.**boxplot**(x="Ocupação (%)", y="Companhia", data=df, color="orange")

**plt**.**title**("Ocupação x Companhia Aerea")

**plt**.**show**()

Gráfico, Gráfico de caixa estreita

O conteúdo gerado por IA pode estar incorreto.

**plt**.**figure**()

**sns**.**barplot**(x="Receita (R$)", y="Companhia", data=df, color="brown", estimator="mean")

**plt**.**title**("Receita Média x Companhia Aerea")

**plt**.**show**()

Gráfico, Gráfico de barras

O conteúdo gerado por IA pode estar incorreto.

**plt**.**figure**()

**sns**.**scatterplot**(x="Distância (km)", y="Receita (R$)", hue="Companhia",  data=df)

**plt**.**title**("Distância x Receita")

**plt**.**show**()

Gráfico, Gráfico de dispersão

O conteúdo gerado por IA pode estar incorreto.

**plt**.**figure**()

**sns**.**heatmap**(df[["Passageiros", "Distância (km)", "Ocupação (%)", "Receita (R$)"]].**corr**(), annot=True, cmap="coolwarm")

**plt**.**title**("Mapa entre variaveis")

**plt**.**show**()

Gráfico, Gráfico de mapa de árvore

O conteúdo gerado por IA pode estar incorreto.

* Qual companhia tem maior participação em número de voos?

R: A companhia Azul é a companhia com maior participação em numero de voos.

* A distância influencia a receita?

R: Sim a distância influencia a receita

* Os voos com maior ocupação são necessariamente os de maior receita?

R: não

* Quais aeroportos de origem concentram mais voos?

Aeroporto Origem Quantidade de Voos

0 CGH 32

1 GIG 28

2 SDU 22

3 GRU 20

4 BSB 18

O CGH é aeroporto de origem com mais voos.

**Extensão do trabalho**

df['Data'] = **pd**.**to\_datetime**(df['Data'])

df['Mês'] = df['Data'].dt.**to\_period**('M').**astype**(**str**)

receita\_media\_mes = df.**groupby**('Mês')['Receita (R$)'].**mean**().**reset\_index**()

receita\_media\_mes.**rename**(columns={'Receita (R$)': 'Receita Média (R$)'}, inplace=True)

**print**("---Receita média por mês---")

**print**(receita\_media\_mes)

**print**("-" \* 30)

Texto

O conteúdo gerado por IA pode estar incorreto.

Gráfico, Gráfico de barras

O conteúdo gerado por IA pode estar incorreto.

ocupacao\_media = df.**groupby**('Companhia')['Ocupação (%)'].**mean**().**reset\_index**()

ocupacao\_media.**rename**(columns={'Ocupação (%)': 'Ocupação Média (%)'}, inplace=True)

**print**("---Ocupação média por companhia---")

**print**(ocupacao\_media)

**print**("-" \* 30)

Texto

O conteúdo gerado por IA pode estar incorreto.

Gráfico, Gráfico de barras

O conteúdo gerado por IA pode estar incorreto.

df['Rota'] = df['Aeroporto Origem'] **+** " → " **+** df['Aeroporto Destino']

ocupacao\_rota = df.**groupby**(['Companhia', 'Rota'])['Ocupação (%)'].**mean**().**reset\_index**()

ocupacao\_rota\_ordenado = ocupacao\_rota.**sort\_values**(by='Ocupação (%)', ascending=False)

**print**("---Top 5 rotas mais eficientes por Ocupação Média---")

**print**(ocupacao\_rota\_ordenado.**head**(5))

**print**("-" \* 30)

Texto

O conteúdo gerado por IA pode estar incorreto.

df['Data'] = **pd**.**to\_datetime**(df['Data'])

df['Mês'] = df['Data'].dt.**to\_period**('M').**astype**(**str**)

passageiros\_mensal = df.**groupby**(['Companhia', 'Mês'])['Passageiros'].**sum**().**reset\_index**()

**print**("---Evolução mensal do total de passageiros por companhia---")

**print**(passageiros\_mensal.**head**())

**print**("-" \* 30)

Tabela

O conteúdo gerado por IA pode estar incorreto.