1. Importar as bibliotecas e ler os dados: Importe as bibliotecas necessárias e verifique a cara dos dados e suas informações básicas dos conteúdos das colunas. ok

2. Limpeza e transformação da variável alvo Salário Mensal: verifique o tipo de dado da variával Salário Mensal ok ( Fiz uma alteração no formato de "," para "."), faça as devidas mudanças, plote a distribuição e dispersão dos dados e aplique a transformação matemática para a correção da assimetria.

3. Verificação dos dados da variável Idade: verifique os dados da variável Idade, plote uma distribuição e uma dispersão, calcule a correlação com a variável salário mensal e interprete os resultados.

4. Verificação dos dados da variável Grau de Instrução: verifique os dados da variável Grau de instrução, plote uma distribuição e uma dispersão, calcule a correlação com a variável salário mensal e interprete os resultados.

5. Verificação dos dados da variável Sexo: verifique os dados da variável Sexo, plote uma distribuição e uma dispersão, calcule a correlação com a variável salário mensal e interprete os resultados. Depois disso, faça uma inferência estatística pra verificar se os homens ganham mais que as mulheres.

df['Grau Instrução'].value\_counts().head(1)

df['Grau Instrução'].value\_counts()

df['Grau Instrução'].value\_counts(3)

df2 = df[['Grau Instrução', 'Salário Mensal']]

df2[0:10]

df.sample(200000).head()

%matplotlib inline

df['Grau Instrução'].plot(kind="hist")

df[['Grau Instrução', 'Salário Mensal']].corr(method='kendall')

6. Verificação dos dados da variável Raça Cor: verifique os dados da variável Raça Cor, plote uma distribuição e uma dispersão, calcule a correlação com a variável salário mensal e interprete os resultados. Depois disso, faça uma inferência estatística pra verificar se os brancos ganham mais que os outros grupos étnicos.

7. Elaborar uma hipótese para ser rejeitada ou não-rejeitada por voce de acordo com o resultado da sua análise

'line' : line plot (default)

- 'bar' : vertical bar plot

- 'barh' : horizontal bar plot

- 'hist' : histogram

- 'box' : boxplot

- 'kde' : Kernel Density Estimation plot

- 'density' : same as 'kde'

- 'area' : area plot

- 'pie' : pie plot

- 'scatter' : scatter plot

- 'hexbin' : hexbin plot