import pandas as pd

import numpy as np

import seaborn as sns

df = pd.read\_csv('Employee.csv')

df.head()

df.shape

#Average Age

df['Age'].mean()

#Most Common City

df['City'].value\_counts()

#Most Common Education

df['Education'].value\_counts()

df['JoiningYear'].value\_counts()

df.sample(10)

df.isnull().any()

df.duplicated().sum()

np.random.seed(42) # Ensures reproducibility

df['Income'] = np.random.randint(30000, 100000, df.shape[0])

df.sample(10)

print("Mean Income:", df['Income'].mean())

print("Minimum Income:", df['Income'].min())

print("Maximum Income:", df['Income'].max())

print("Median Income:", df['Income'].median())

print("Standard Deviation of Income:", df['Income'].std())

print("Variance of Income:", df['Income'].var())

print("Mode of Age:", df['Income'].mode()[0])

print("Mean Age:", df['Age'].mean())

print("Minimum Age:", df['Age'].min())

print("Maximum Age:", df['Age'].max())

print("Median Age:", df['Age'].median())

print("Standard Deviation of Age:", df['Age'].std())

print("Variance of Age:", df['Age'].var())

print("Mode of Age:", df['Age'].mode()[0])

df['Income'].value\_counts()

def classify\_income(income):

if income < 40000:

return 'Low Income'

elif income > 80000:

return 'High Income'

else:

return 'Moderate Income'

def categorize\_age(age):

if age < 25:

return "Experienced"

elif age > 18 and age <= 60:

return "Fresher"

else:

return "Retired"

df['Income Category'] = df['Income'].apply(classify\_income)

df['Age Category'] = df['Age'].apply(categorize\_age)

df.sample(10)

df.groupby(['Age Category']).Income.mean()

df.groupby(['Income Category']).Income.mean()

df.groupby(['Education']).Income.mean()