import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# Load the dataset

data = pd.read\_csv('employees.csv')

# 1. Count of Males/Females in the organization

gender\_count = data['Gender'].value\_counts()

# 2. Count of Males/Females in each department and location

gender\_department = data.groupby(['Department', 'Gender']).size()

gender\_location = data.groupby(['Location', 'Gender']).size()

# 3. Department with the highest average salary

highest\_avg\_salary\_department = data.groupby('Department')['Salary'].mean().idxmax()

# 4. Location with the highest average salary

highest\_avg\_salary\_location = data.groupby('Location')['Salary'].mean().idxmax()

# 5. Ratings analysis

rating\_counts = data['Ratings'].value\_counts(normalize=True) \* 100

good\_rating\_percentage = rating\_counts.get('Good', 0) + rating\_counts.get('Very Good', 0)

poor\_rating\_percentage = rating\_counts.get('Poor', 0) + rating\_counts.get('Very Poor', 0)

average\_rating = data['Ratings'].mean()

# 6. Gender pay gap for each department

gender\_pay\_gap\_department = data.groupby('Department').apply(lambda x: x[x['Gender'] == 'Male']['Salary'].mean() - x[x['Gender'] == 'Female']['Salary'].mean())

# 7. Gender pay gap for each location

gender\_pay\_gap\_location = data.groupby('Location').apply(lambda x: x[x['Gender'] == 'Male']['Salary'].mean() - x[x['Gender'] == 'Female']['Salary'].mean())

# Visualization

plt.figure(figsize=(12, 6))

sns.countplot(data=data, x='Gender', hue='Department')

plt.title('Gender Distribution by Department')

plt.show()

# Statistical methods to explore relationships

sns.boxplot(x='Gender', y='Salary', data=data)

plt.title('Salary Distribution by Gender')

plt.show()

sns.boxplot(x='Location', y='Salary', data=data)

plt.title('Salary Distribution by Location')

plt.xticks(rotation=45)

plt.show()

sns.boxplot(x='Department', y='Salary', data=data)

plt.title('Salary Distribution by Department')

plt.xticks(rotation=45)

plt.show()