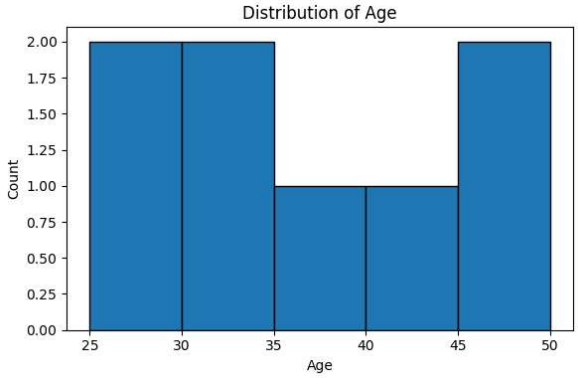
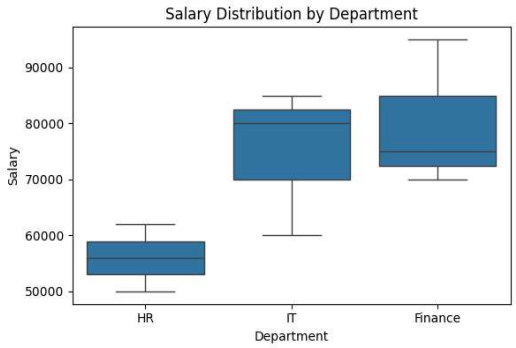
Week 4: Data Visualization with Matplotlib and Seaborn

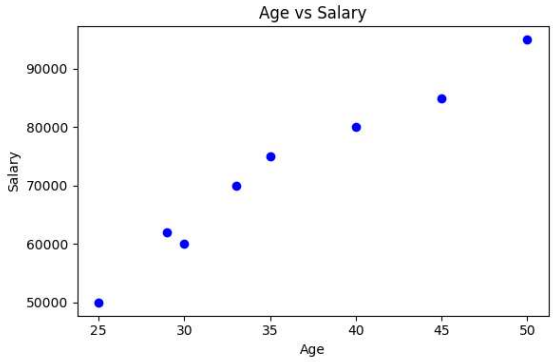
Hands-On: Create visualizations for dataset analysis.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Sample dataset
data = {
    "Age": [25, 30, 35, 40, 29, 50, 45, 33],
    "Salary": [50000, 60000, 75000, 80000, 62000, 95000, 85000, 70000],
    "Department": ["HR", "IT", "Finance", "IT", "HR", "Finance", "IT", "Finance"]
}
df = pd.DataFrame(data)

# 1. Histogram of Age
plt.figure(figsize=(6,4))
plt.hist(df["Age"], bins=5, edgecolor="black")
plt.title("Distribution of Age")
plt.xlabel("Age")
plt.ylabel("Count")
plt.show()

# 2. Scatter Plot (Age vs Salary)
plt.figure(figsize=(6,4))
```

```
plt.scatter(df["Age"], df["Salary"], c="blue")
plt.title("Age vs Salary")
plt.xlabel("Age")
plt.ylabel("Salary")
plt.show()

# 3. Boxplot of Salary by Department
plt.figure(figsize=(6,4))
sns.boxplot(x="Department", y="Salary", data=df)
plt.title("Salary Distribution by Department")
plt.show()
```







Client Project: Create a dashboard for visualizing relationships between features in a dataset (e.g., scatter plots, histograms).

```
import pandas as pd

import plotly.express as px

import plotly.io as pio


# Sample dataset

data = {

   "Age": [25, 30, 35, 40, 29, 50, 45, 33],

   "Salary": [50000, 60000, 75000, 80000, 62000, 95000, 85000, 70000],

   "Department": ["HR", "IT", "Finance", "IT", "HR", "Finance", "IT", "Finance"]

}

df = pd.DataFrame(data)


# Scatter plot (Age vs Salary with Department color)

fig1 = px.scatter(df, x="Age", y="Salary", color="Department", size="Salary",

            title="Age vs Salary by Department")

fig1.show()


# Histogram of Salary

fig2 = px.histogram(df, x="Salary", nbins=6, color="Department",

             title="Salary Distribution")

fig2.show()
```

Boxplot Salary by Department

```
fig3 = px.box(df, x="Department", y="Salary", color="Department",
              title="Salary Distribution by Department")
fig3.show()
```