EXPERIMENT 1B

Matplot Library – Data Visualization

Aim:

Analyze and visualize the distribution of various data science roles (Data Analyst, Data Engineer, Data Scientist, etc.) from a dataset.

Algorithm:

1. Import pandas, matplotlib, and seaborn libraries.
2. Create a dataset containing different data science job titles.
3. Convert the dataset into a pandas DataFrame.
4. Count the frequency of each job title using value\_counts().
5. Plot a pie chart showing the percentage distribution of job roles.
6. Add a title and make the pie chart circular using axis('equal').
7. Plot a bar graph showing the count of each job role.
8. Add labels, title, and adjust layout for better readability.

Program:

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

data = {

'Job\_Title':[

'Data Analyst', 'Data Scientist', 'Data Engineer', 'ML Engineer',

'Business Analyst', 'Data Scientist', 'Data Analyst', 'ML Engineer',

'Data Engineer', 'Data Scientist', 'Business Analyst', 'Data Engineer',

'Data Analyst', 'Data Scientist', 'ML Engineer']

}

df = pd.DataFrame(data)

print(df.head())

role\_counts = df['Job\_Title'].value\_counts()

print(role\_counts)

# Pie Chart

plt.figure(figsize=(8, 8))

plt.pie(role\_counts, labels=role\_counts.index, autopct='%1.1f%%', startangle=140, colors=sns.color\_palette('Set3'))

plt.title('Distribution of Data Science Job Roles')

plt.axis('equal')

plt.show()

# Bar Plot

plt.figure(figsize=(8, 5))

sns.barplot(x=role\_counts.index, y=role\_counts.values)

plt.title('Count of Data Science Job Roles', fontsize=14)

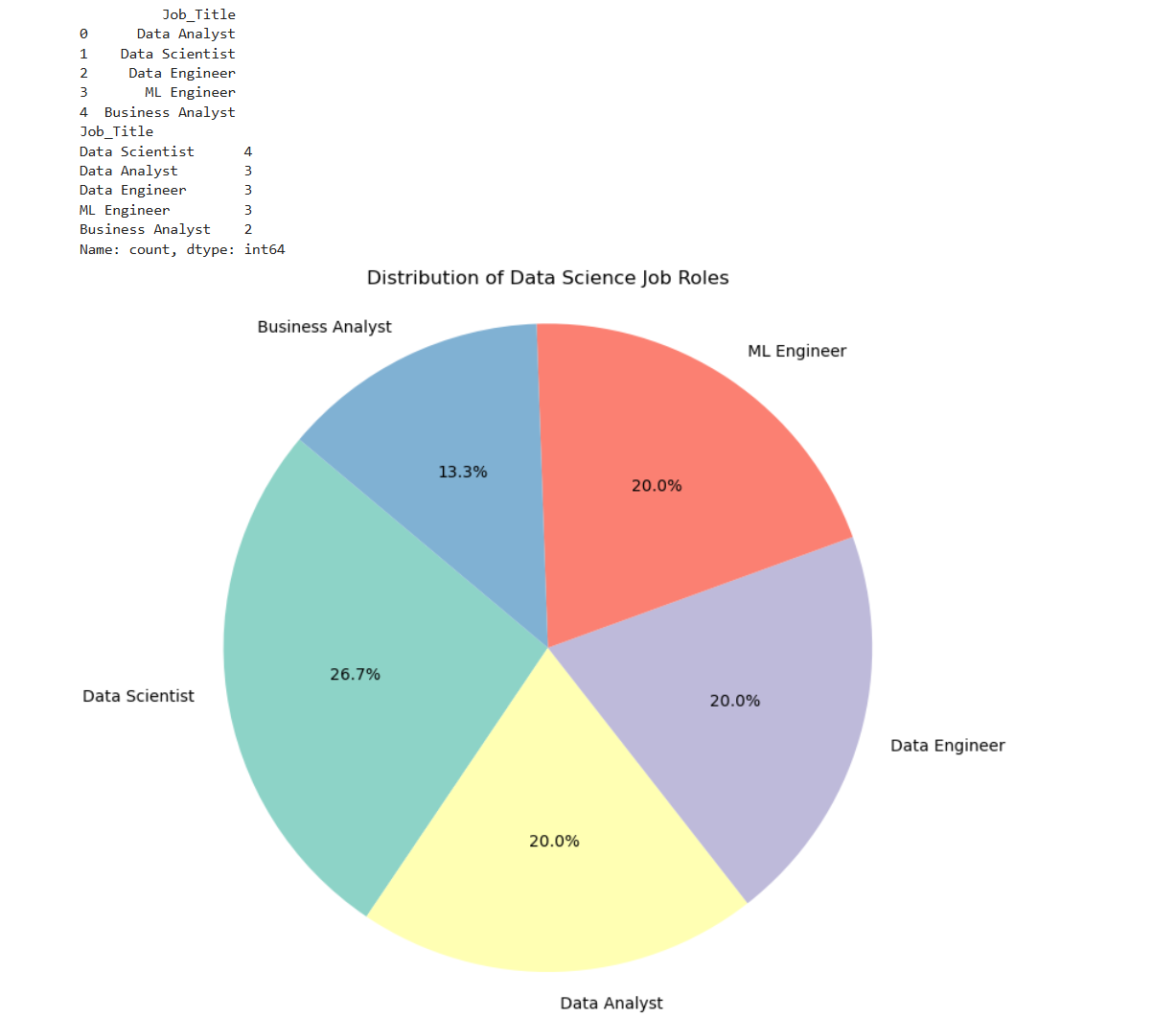
plt.xlabel('Job Role', fontsize=12)

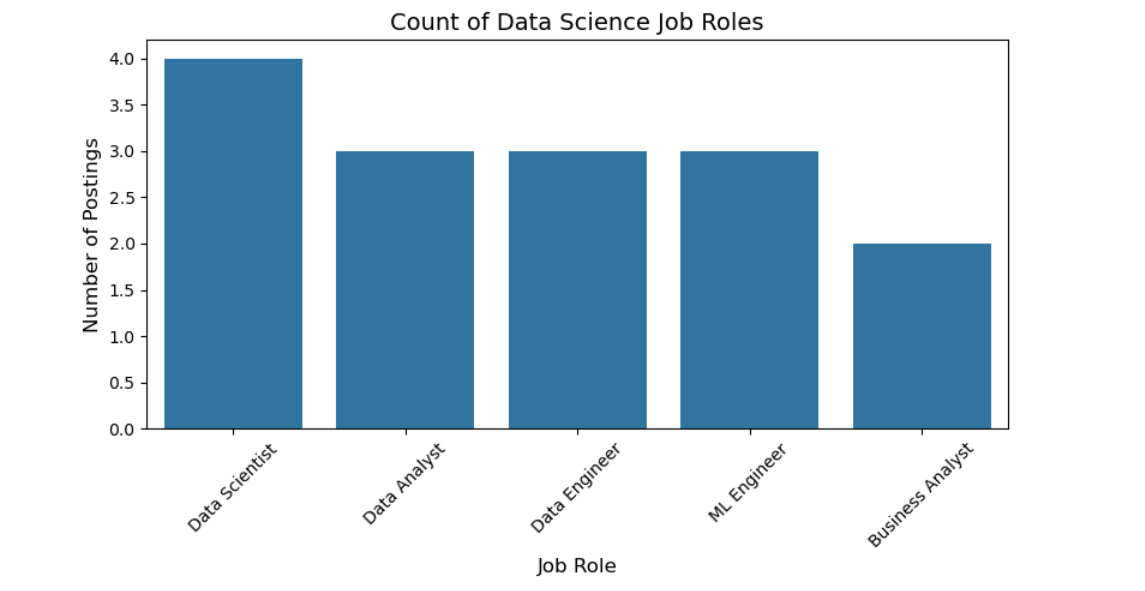
plt.ylabel('Number of Postings', fontsize=12)

plt.xticks(rotation=45)

plt.tight\_layout()

plt.show()

Output:



Result:

Hence a program to analyze and visualize the distribution of various data science roles is written and executed successfully and a bar graph is plotted.