EDS PRACTICAL 5

Name – Mayur Kapgate **Roll No. –** 428

Division – D2 **PRN No. –** 202201040065

CSV File –

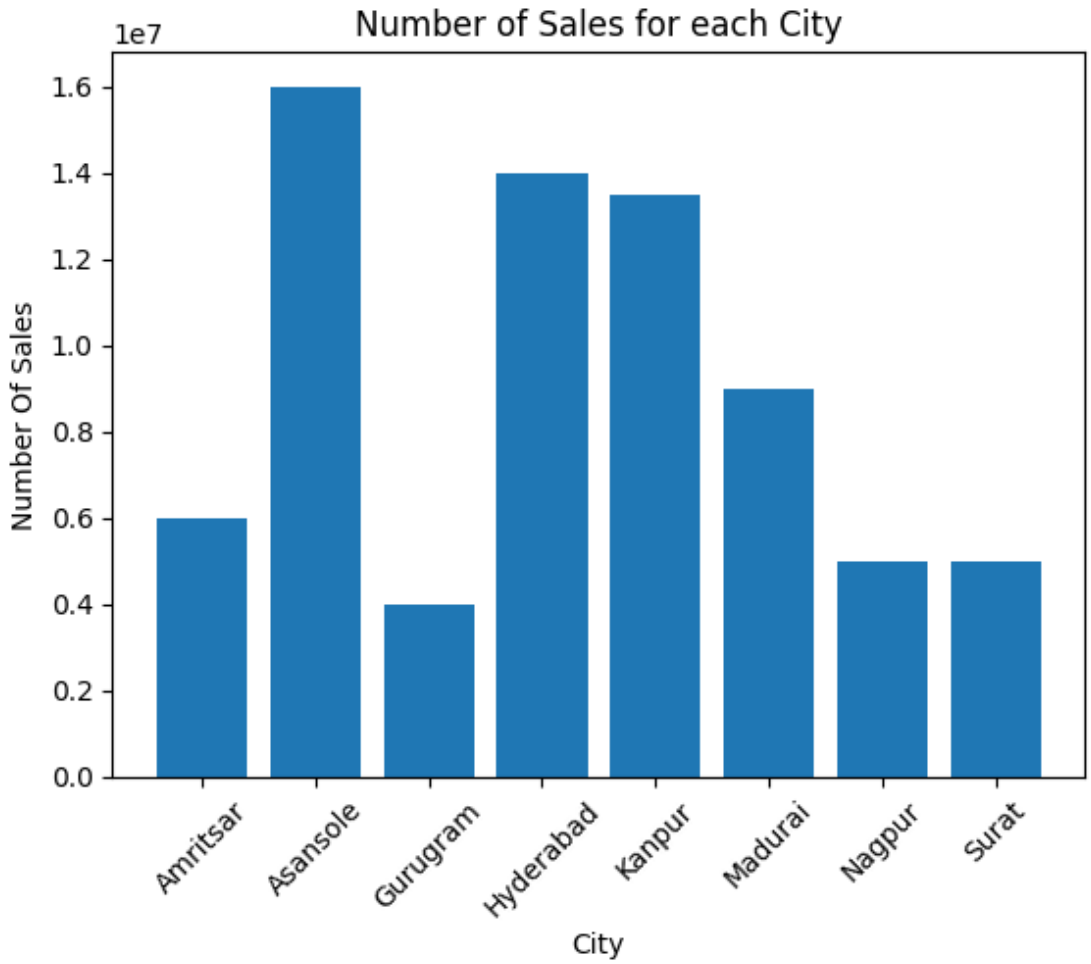
Transaction_ID	GrainName	State	City	Months	Year	Sales
1	Ragi	Maharashtra	Nagpur	JAN	2023	1000000
2	Bajra	Panjab	Amritsar	FEB	2023	1500000
3	Ragi	Maharashtra	Nagpur	JAN	2023	1000000
4	Bajra	Panjab	Amritsar	FEB	2023	1500000
5	Ragi	Maharashtra	Nagpur	JAN	2023	1000000
6	Bajra	Panjab	Amritsar	FEB	2023	1500000
7	Oats	Hariyana	Gurugram	MARCH	2023	2000000
8	Sattu	Gujarat	Surat	APRIL	2023	2500000
9	Sooji	Tamil Nadu	Madurai	MAY	2023	3000000
10	Brown rice	Telangana	Hyderabad	JUNE	2023	3500000
11	Wheat	West Bengol	Asansole	JULY	2023	4000000
12	Corn	UP	Kanpur	AUG	2023	4500000
13	Ragi	Maharashtra	Nagpur	JAN	2023	1000000
14	Bajra	Panjab	Amritsar	FEB	2023	1500000
15	Oats	Hariyana	Gurugram	MARCH	2023	2000000
16	Sattu	Gujarat	Surat	APRIL	2023	2500000
17	Sooji	Tamil Nadu	Madurai	MAY	2023	3000000
18	Brown rice	Telangana	Hyderabad	JUNE	2023	3500000
19	Wheat	West Bengol	Asansole	JULY	2023	4000000
20	Corn	UP	Kanpur	AUG	2023	4500000
21	Sooji	Tamil Nadu	Madurai	MAY	2023	3000000
22	Brown rice	Telangana	Hyderabad	JUNE	2023	3500000
23	Wheat	West Bengol	Asansole	JULY	2023	4000000
24	Corn	UP	Kanpur	AUG	2023	4500000
25	Ragi	Maharashtra	Nagpur	JAN	2023	1000000
26	Brown rice	Telangana	Hyderabad	JUNE	2023	3500000
27	Wheat	West Bengol	Asansole	JULY	2023	4000000

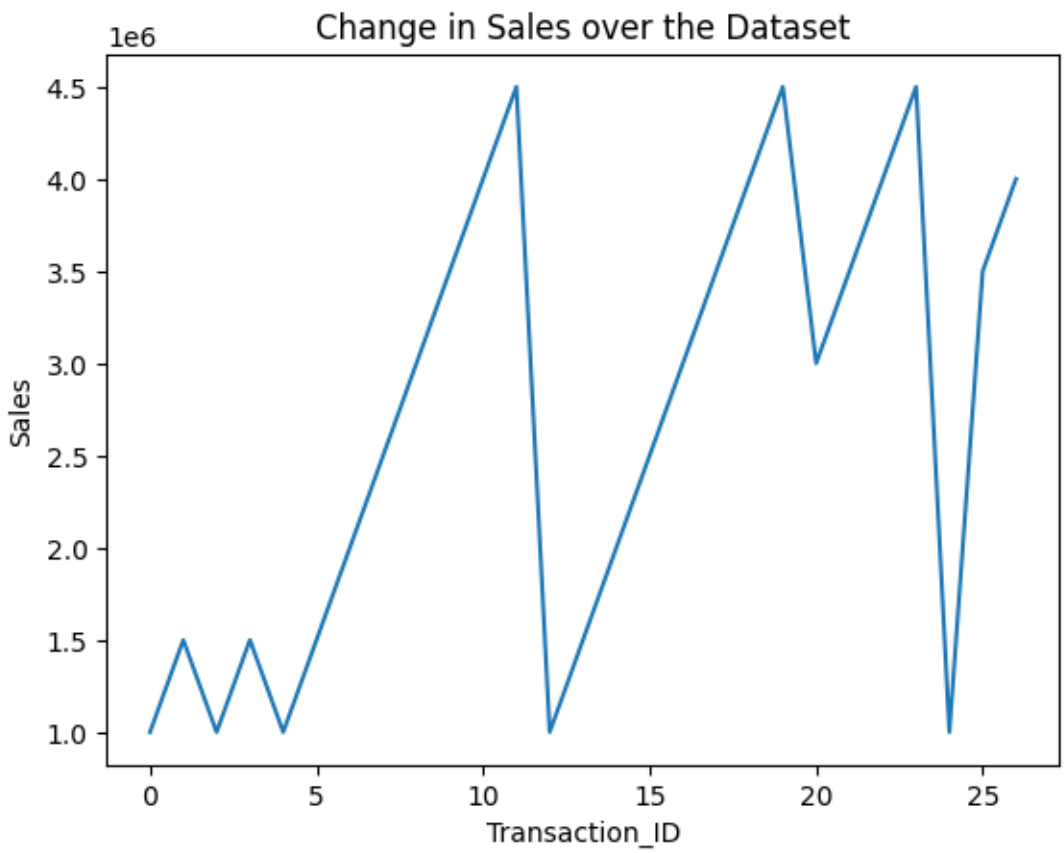
Code –

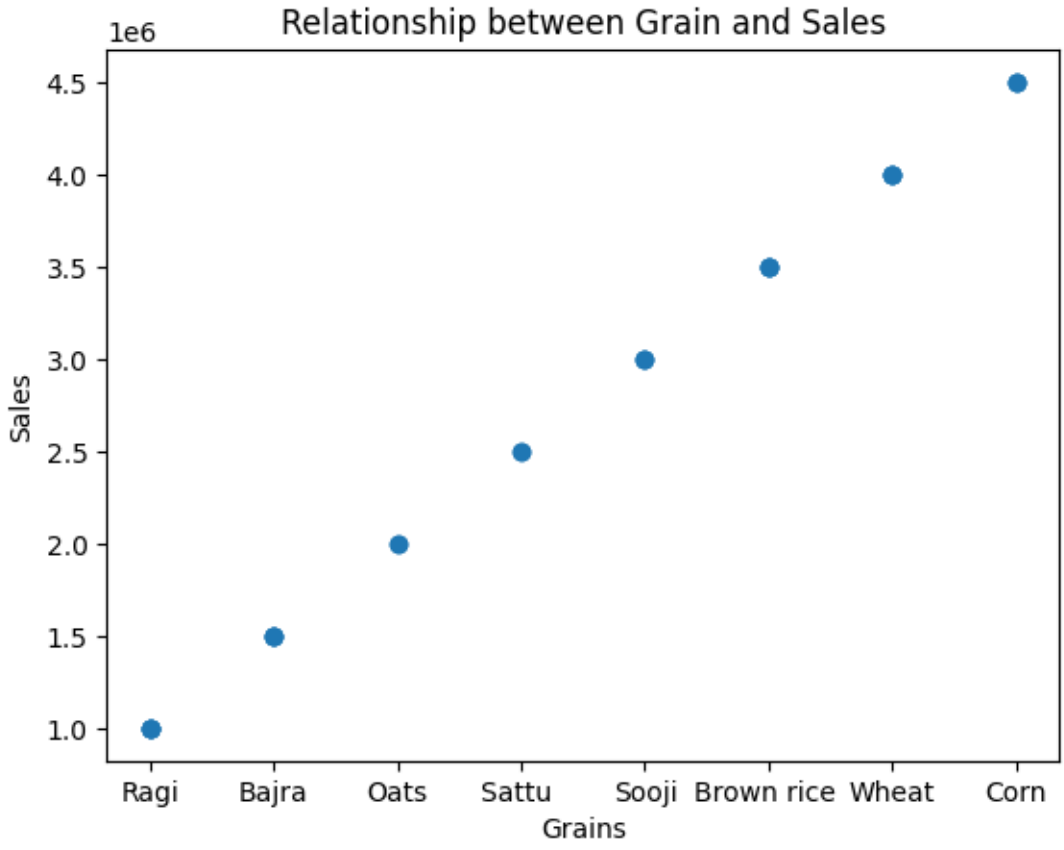
```
import pandas as pd
import matplotlib.pyplot as plt
```

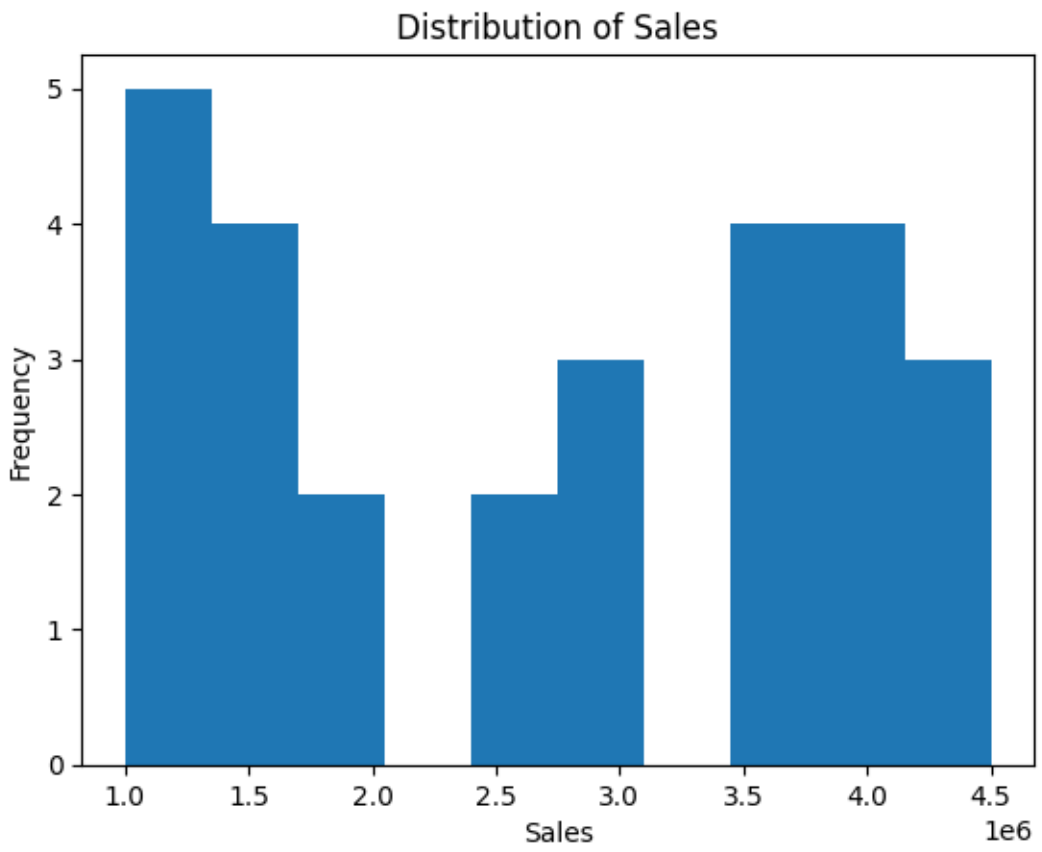
```
# Read the CSV file into a pandas DataFrame

data = pd.read_csv('/content/grainsales.csv')

# Bar Chart - Number of sales for each city

City_Sales = data.groupby('City')['Sales'].sum()
plt.bar(City_Sales.index, City_Sales.values)
plt.xlabel('City')
plt.ylabel('Number Of Sales')
plt.title('Number of Sales for each City')
plt.xticks(rotation=45)
plt.show()

# Line Chart - Change in Sales over the dataset

plt.plot(data['Sales'])
plt.xlabel('Transaction_ID')
plt.ylabel('Sales')
plt.title('Change in Sales over the Dataset')
plt.show()

# Scatter Plot - Relationship between Grain and Sales

plt.scatter(data['GrainName'], data['Sales'])
plt.xlabel('Grains')
plt.ylabel('Sales')
plt.title('Relationship between Grain and Sales')
plt.show()

# Histogram - Distribution of Sales

plt.hist(data['Sales'], bins=10)
plt.xlabel('Sales')
plt.ylabel('Frequency')
plt.title('Distribution of Sales')
plt.show()

# Pie Chart - Distribution Of Grain

plt.figure(figsize=(8, 8))
top_10_grains =
data['GrainName'].value_counts().head(10).index.tolist()
subset_data = data[data['GrainName'].isin(top_10_grains)]
subset_data['GrainName'].value_counts().plot.pie(autopct='%1.1f%%')
plt.title('Grain Distribution')
plt.show()
```

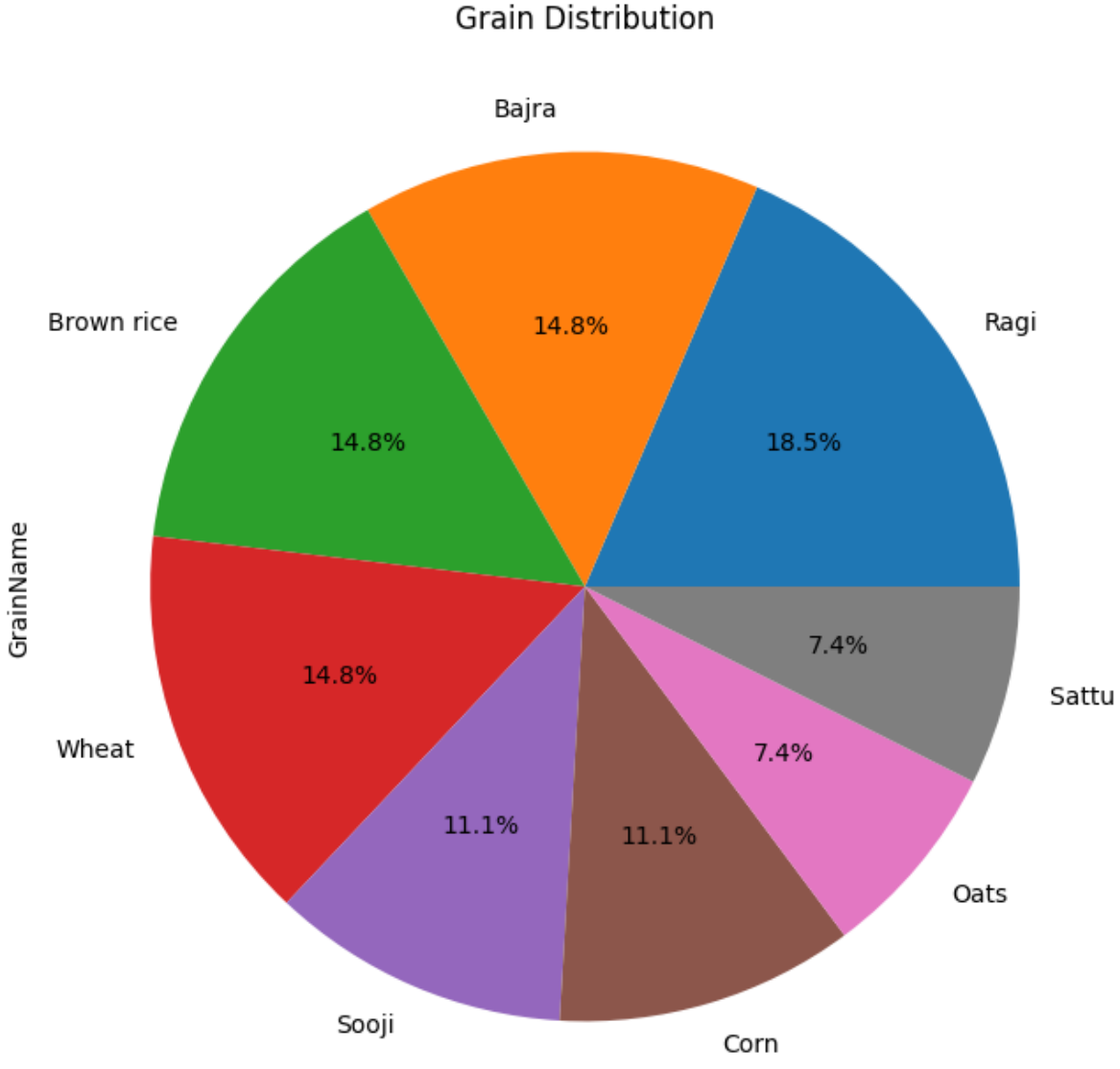
Output –


Number of Sales for each City
1e7
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Number Of Sales
Amritsar
Asansole
Gurugram
Hyderabad
Kanpur
Madurai
Nagpur
Surat
City


Change in Sales over the Dataset
1e6
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
Sales
0
5
10
15
20
25
Transaction_ID


Relationship between Grain and Sales
1e6
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
Sales
Ragi
Bajra
Oats
Sattu
Sooji
Brown rice
Wheat
Corn
Grains


Distribution of Sales
Frequency
Sales
1e6
0
1
2
3
4
5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5


Grain Distribution
GrainName
Bajra
Ragi
Sattu
Oats
Corn
Sooji
Wheat
Brown rice
14.8%
18.5%
7.4%
7.4%
11.1%
11.1%
14.8%
14.8%