Experiment No :- 05

PAGE No.		
DATE	/	/

Aim :- To study data visualisation methods and tools.

Theory :-

Data visualization is very important in data mining result using python. The libraries in python comes with lot of different features that enable users to make highly customized elegant & interactive plot.

Useful packages for visualization in python. The

Matplotlib :-

MatPlotlib is a visualization library in python for 2D plots. of arrays. Matplot is written in python & makes use of the numpy library. It can be used in Python & Ipython shells Jupyter notebook, & web application servers. MatPlotlib comes with a wide variety of plots like line, bar, scatter, histogram.

Seaborn :-

Seaborn is a dataset-oriented library for making representations in python. It is developed a top matPlotlib & to create different visualizations. It is integrated with pandas data structures the library internally performs the required mapping & aggregation to create informative visuals. It is recommended to use Jupyter/IPython interface in matplotlib mode.

PAGE No.

DATE / /

Bokech :-

Bokech in an interactive visualization library for modern web browsers. It is suitable for streaming data assets and can be used to develop interactive plots and dashboards.

gg plot :-

gg plot is is a python implementation of the grammer of graphics. The Grammer of Graphics refers to the mapping of data to aesthic attributes.

1) Barchort :-

A bar chart is used when we want to compare metric values across different subgroup of data.

2) Scatter plot :-

Scatter plot can be leveraged to identify relationship between two variables.

3) Making Venn diagram :-

4) Displot :

Displot is used basically for univarient set of observations and visualizes it through a histogram i.e. only one observation & hence we choose one particular column of dataset.

5) Box plot :- A box plot is used to show the shape of the distribution, its central value & its variability.

6) Pie-chart :- A pie chart is a pictorial representation of data

7) Histogram :- A histogram is a visualization tool that plots frequency distribution for a variable.

Conclusion :
        Hence, we studied various visualization methods and plotted above visualization tools in python.
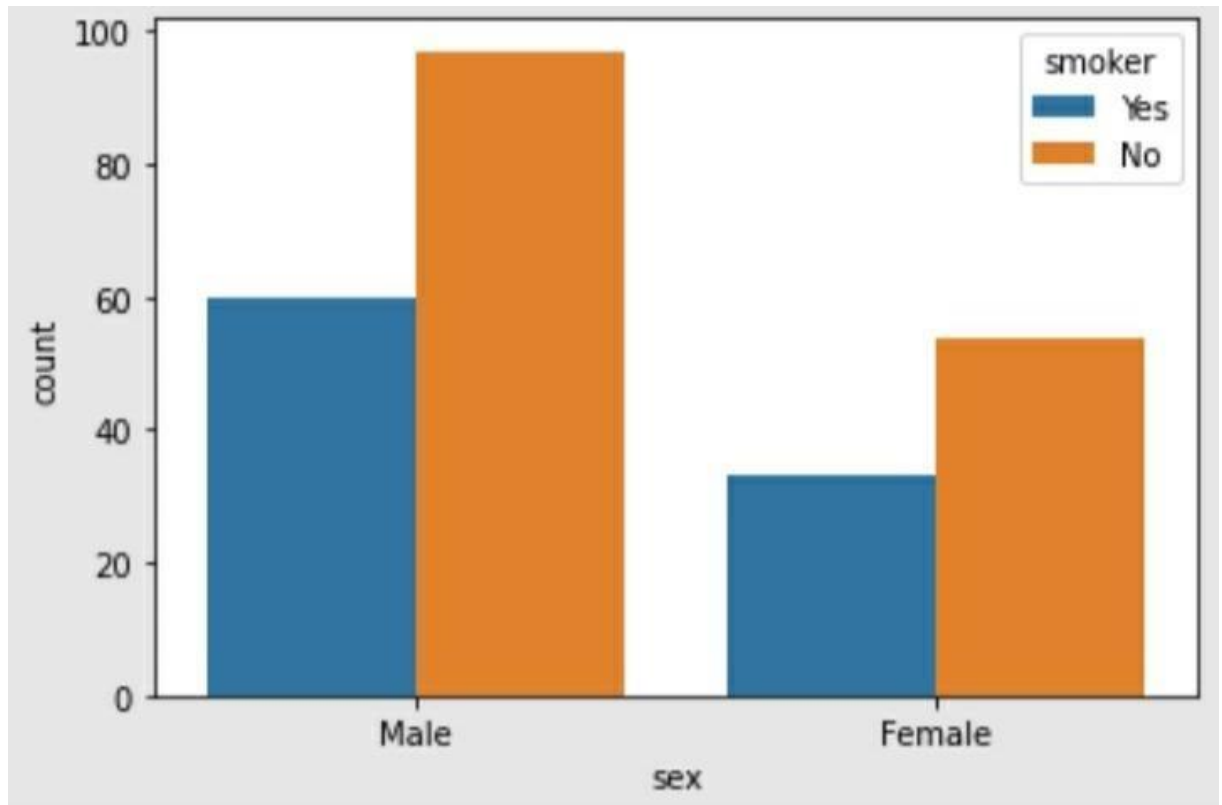
1) Bar chart

```
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
#Creating the dataset
df = sns.load_dataset('titanic')
df_pivot = pd.pivot_table(df,
values="fare",index="who",columns="class",
aggfunc=np.mean)
#Creating a grouped bar chart
ax = df_pivot.plot(kind="bar",alpha=0.5)
#Adding the aesthetics
plt.title('GIT COMPUTER Data ANalysis')
plt.xlabel('X axis  title')
plt.ylabel('Y axis title')
# Show the plot
plt.show()
```
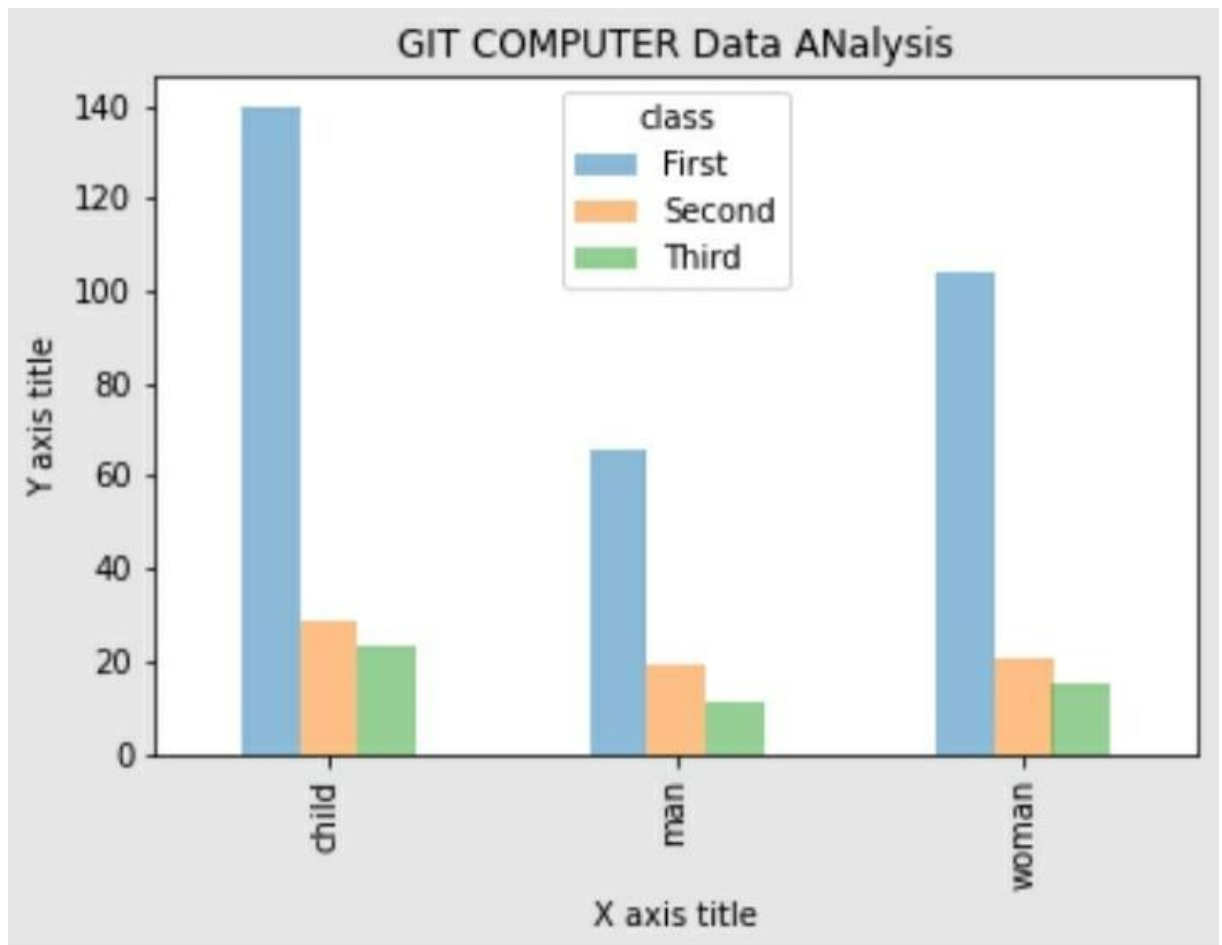
#Bar chart
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
import numpy as np
# importing the required library
# read a tips.csv file from seaborn libraray
df = sns.load_dataset('tips')
# count plot on two categorical variable
sns.countplot(x ='sex', hue = "smoker", data =
df)
# Show the plot
plt.show()

#Horizontal BAR chart

import matplotlib.pyplot as plt

import seaborn as sns

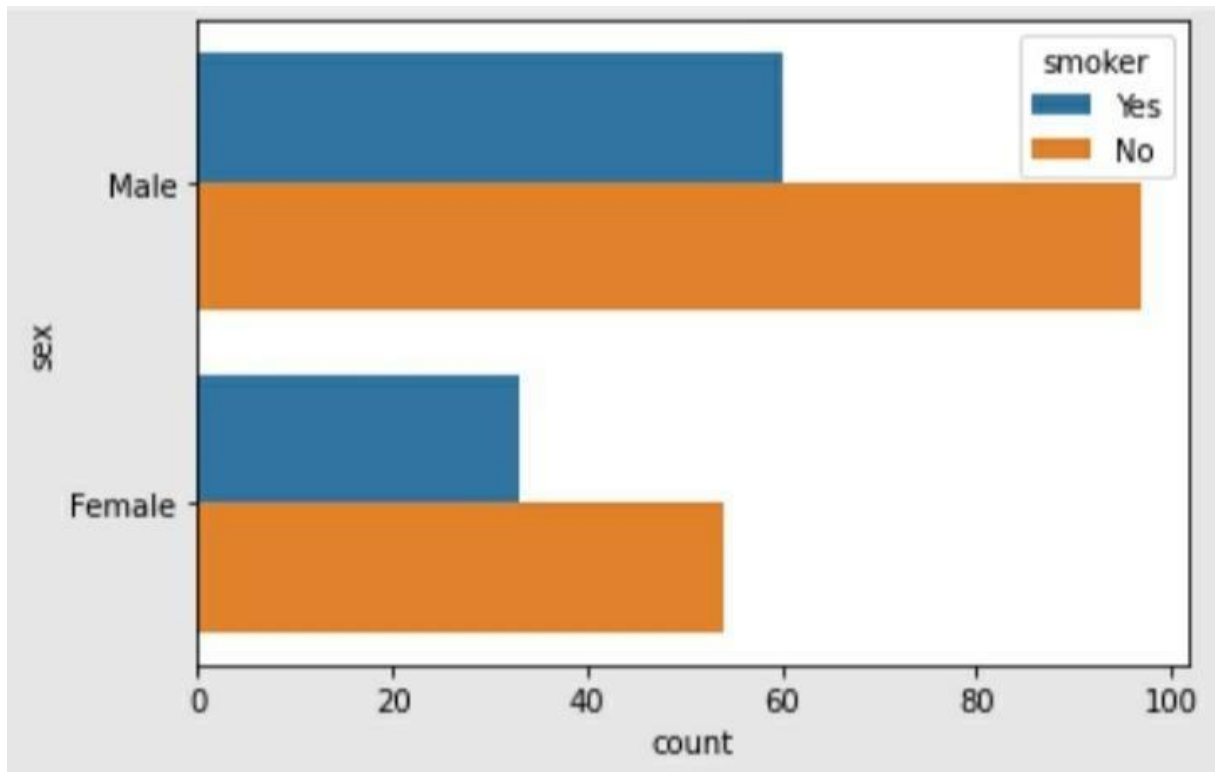# read a tips.csv file from seaborn libraray

df = sns.load_dataset('tips')

 # count plot along y axis

sns.countplot(y ='sex', hue = "smoker", data = df)

 # Show the plot

plt.show()

2) Scatter Plot

```
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
#Creating the dataset
df = sns.load_dataset("tips")
#Creating the scatter plot
plt.scatter(df['total_bill'],df['tip'],alpha=0.5 )
#Adding the aesthetics
plt.title('Chart title')
plt.xlabel('X axis title')
plt.ylabel('Y axis title')
#Show the plot
plt.show()
```
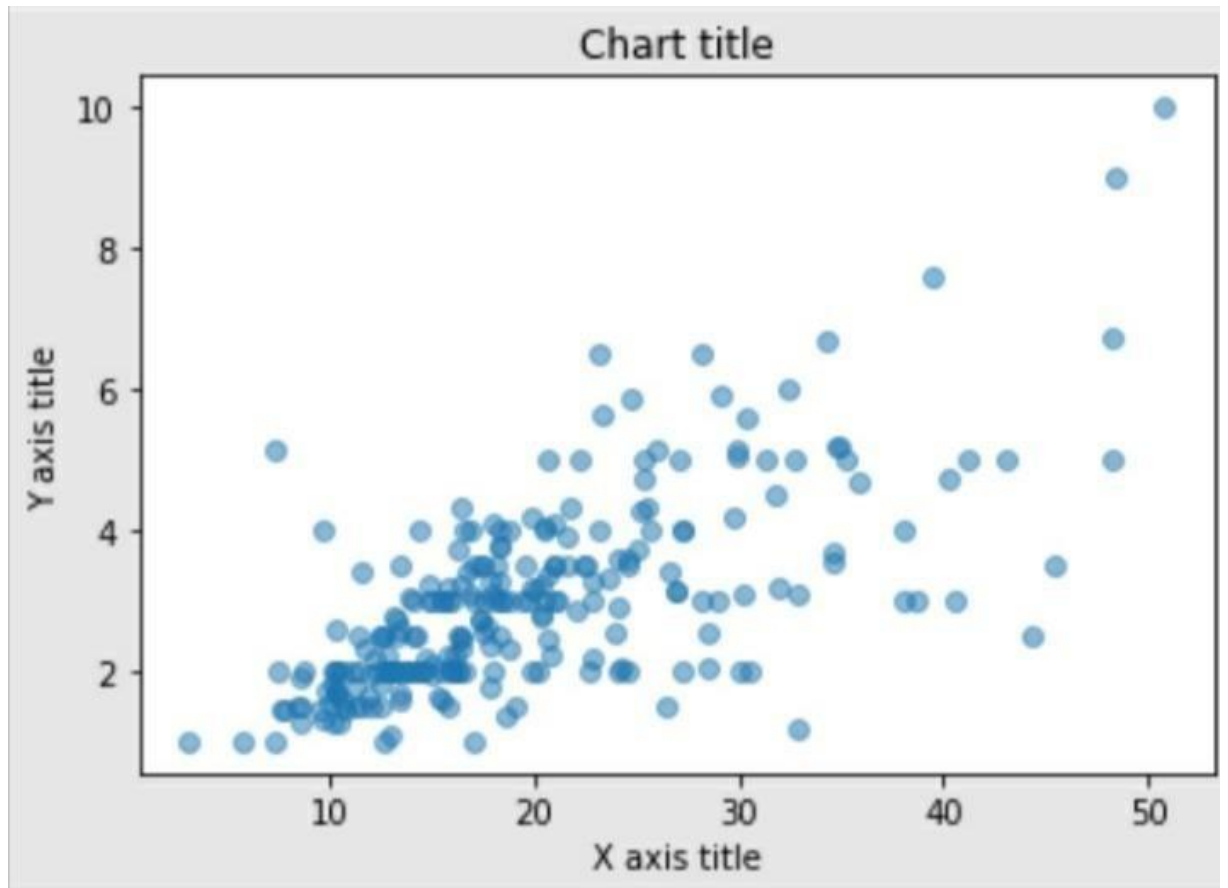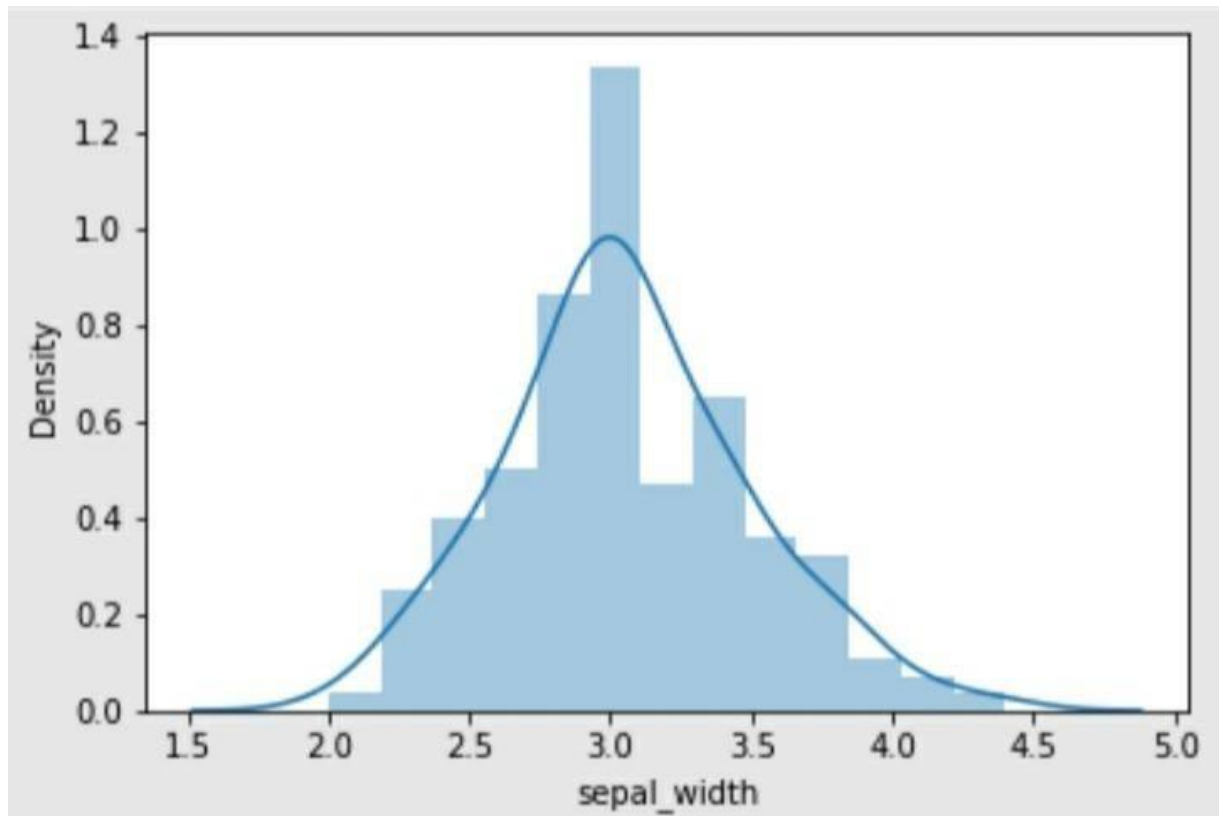
1) **3)** Making venn diagram

```
from matplotlib_venn import venn3
import matplotlib.pyplot as plt
venn3(subsets = (10, 8, 22, 6,9,4,2))
plt.show()
```

4) Displot

```
# importing packages
import seaborn as sns
import matplotlib.pyplot as plt

# loading dataset
data = sns.load_dataset("iris")
 sns.distplot(data['sepal_width'])
       plt.show()
```

4) Box plot

```
import seaborn as sns
import matplotlib.pyplot as plt
#Reading the dataset
bill_dataframe = sns.load_dataset("tips")
#Creating boxplots
ax = sns.boxplot(x="day", y="total_bill", hue="smoker", data=bill_dataframe, palette="Set3")
#Adding the aesthetics
plt.title('Chart title')
plt.xlabel('X axis title')
plt.ylabel('Y axis title')
# Show the plot
plt.show()
```
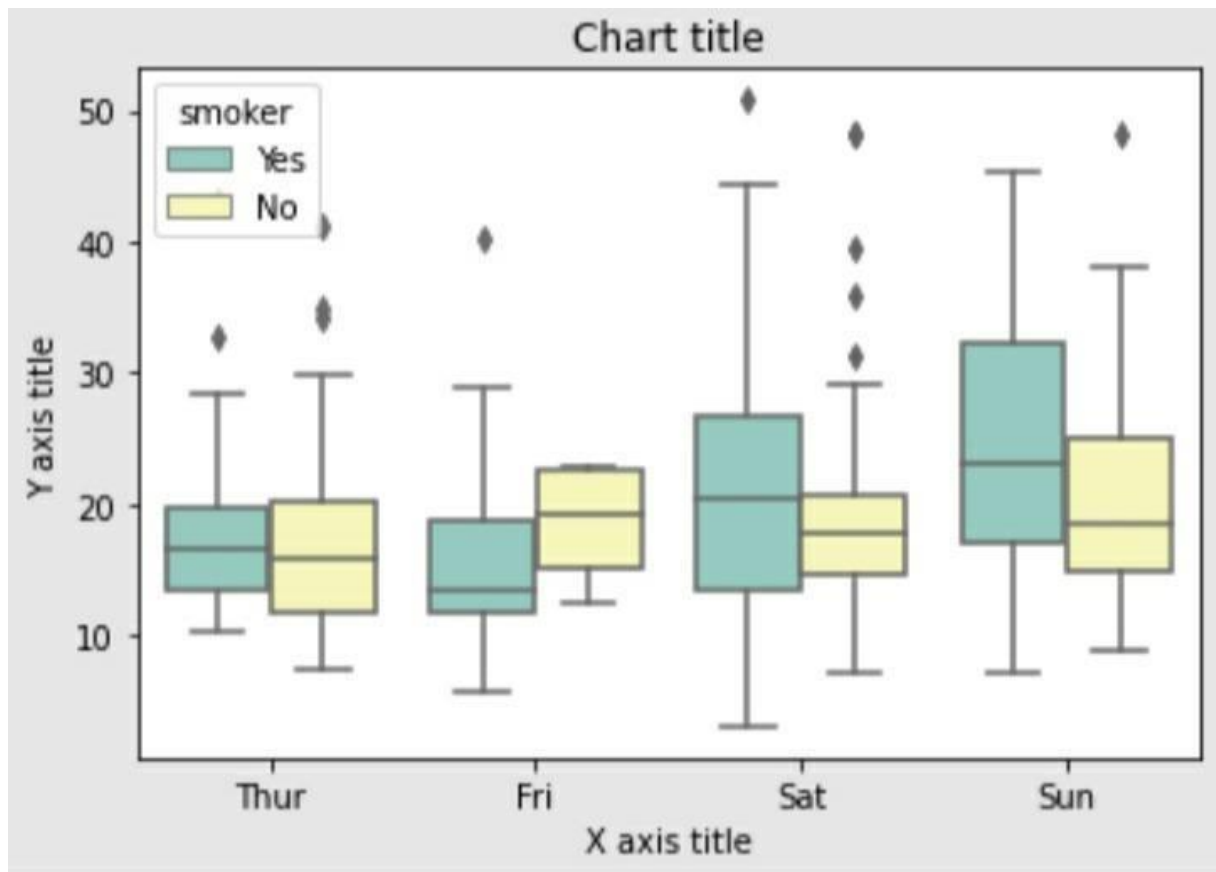
5) Pie-chart

```
# importing packages
import seaborn as sns
import matplotlib.pyplot as plt
#Creating the dataset
cars = ['AUDI', 'BMW', 'NISSAN',
     'TESLA', 'HYUNDAI', 'HONDA']
data = [10, 15, 15, 14, 16, 25]

#Creating the pie chart
plt.pie(data, labels = cars,colors = ['GREEN','BLUE','#B0E0E6','#7B68EE','#483D8B','red'])
#Adding the aesthetics
plt.title('CAR Analysis Report')
#Show the plot
plt.show()
```
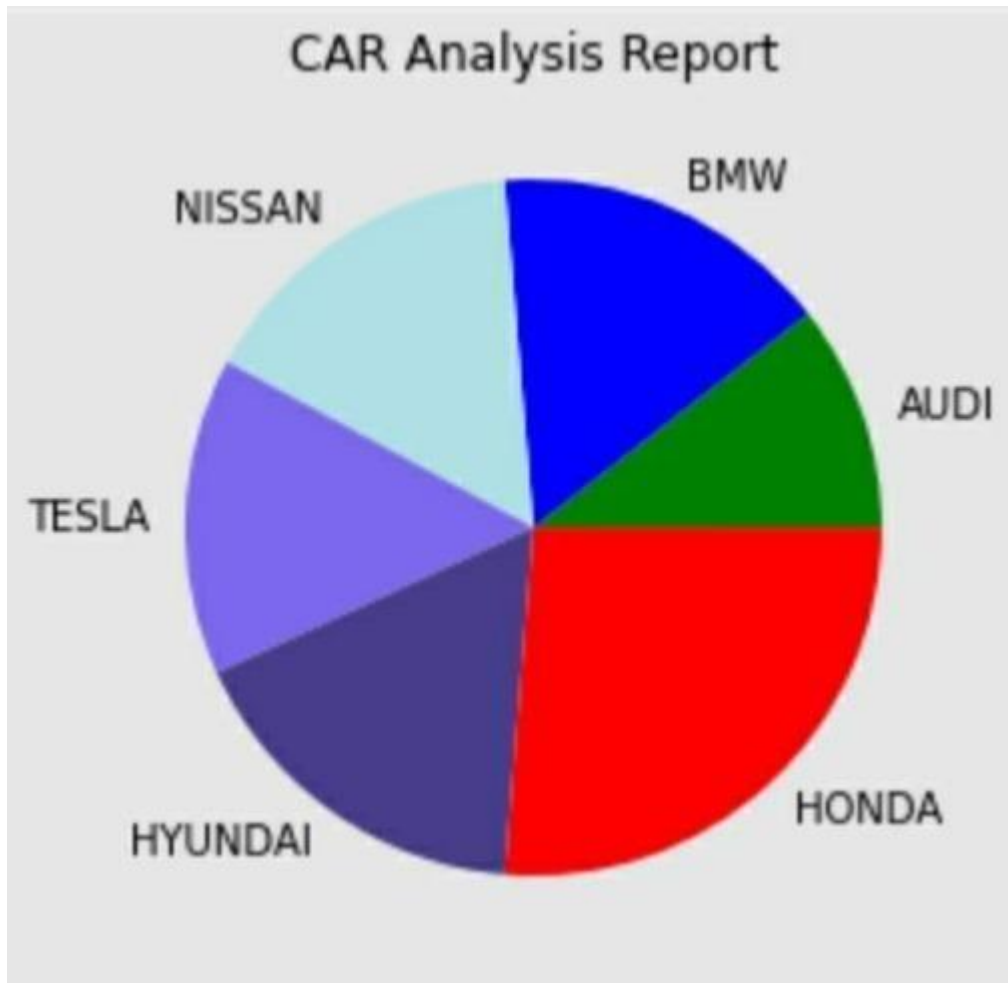
5)Histogram:-

```
import matplotlib.pyplot as plt

data = [87, 53, 66, 61, 67, 68, 62, 110,
     104, 61, 111, 123, 117, 119, 116,
     104, 92, 111, 90, 103, 81, 80, 101,
     51, 79, 107, 110, 129, 145, 128,
     132, 135, 131, 126, 139, 110]

binwidth = 6
plt.hist(data, bins=range(min(data), max(data) + binwidth, binwidth),      edgecolor="yellow",
color="brown")

plt.show()
```