DATA VISUALIZATION

EXPERIMENT NO 2:

AIM: WRITE A PROGRAM FOR HISTOGRAM

PROGRAM:

import matplotlib.pyplot as plt

import numpy as np

# Use numpy to generate a bunch of random data in a bell curve around 5.

n = 5 + np.random.randn(1000)

m = [m for m in range(len(n))]

plt.bar(m, n)

plt.title("Raw Data")

plt.show()

plt.hist(n, bins=20)

plt.title("Histogram")

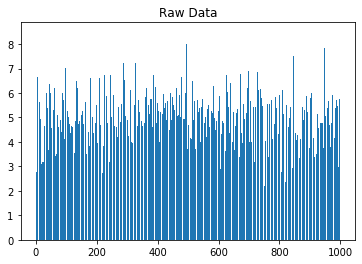
plt.show()

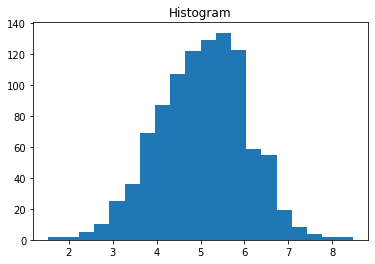
plt.hist(n, cumulative=True, bins=20)

plt.title("Cumulative Histogram")

plt.show()

OUTPUT:





Scatter Plots

import matplotlib.pyplot as plt

x1 = [2, 3, 4]

y1 = [5, 5, 5]

x2 = [1, 2, 3, 4, 5]

y2 = [2, 3, 2, 3, 4]

y3 = [6, 8, 7, 8, 7]

# Markers: https://matplotlib.org/api/markers\_api.html

plt.scatter(x1, y1)

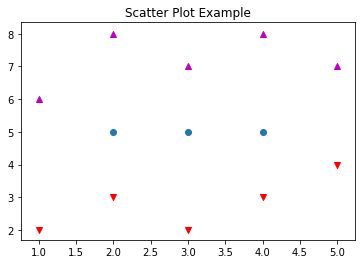
plt.scatter(x2, y2, marker='v', color='r')

plt.scatter(x2, y3, marker='^', color='m')

plt.title('Scatter Plot Example')

plt.show()

OUTPUT:



BUBBLE PLOT

import plotly.graph\_objects as go

fig = go.Figure(data=[go.Scatter(

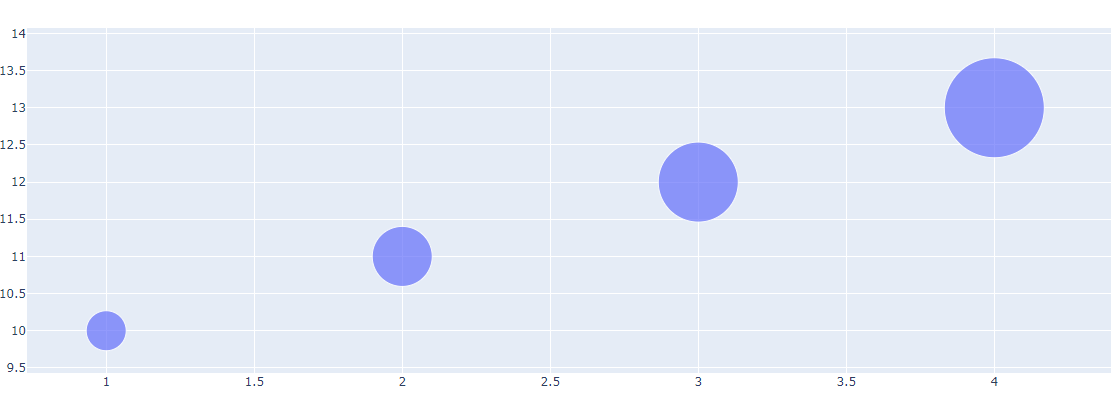
    x=[1, 2, 3, 4], y=[10, 11, 12, 13],

    mode='markers',

    marker\_size=[40, 60, 80, 100])

])

fig.show()



DISTRIBUTION PLOT

import pandas as pd

#create DataFrame

df = pd.DataFrame({'team': ['A', 'A', 'A', 'A', 'A', 'A', 'A', 'A', 'A', 'A',

                            'B', 'B', 'B', 'B', 'B', 'B', 'B', 'B', 'B', 'B'],

                   'points': [3, 3, 4, 5, 4, 7, 7, 7, 10, 11,

                              8, 7, 8, 9, 12, 12, 12, 14, 15, 17]})

#view DataFrame

print(df)

OUTPUT:

eam points

0 A 3

1 A 3

2 A 4

3 A 5

4 A 4

5 A 7

6 A 7

7 A 7

8 A 10

9 A 11

10 B 8

11 B 7

12 B 8

13 B 9

14 B 12

15 B 12

16 B 12

17 B 14

18 B 15

19 B 17