**MODULE 6 PACKAGES**

**Name:- Avishek Kymar Verma**

**Batch:- 05012021\_10A.M**

**Implement in Python**

1. For the dataset “Indian\_cities”,
2. Find out top 10 states in female-male sex ratio
3. Find out top 10 cities in total number of graduates
4. Find out top 10 cities and their locations in respect of total effective\_literacy\_rate.

**Ans:-**

import pandas as pd

cities = pd.read\_csv("C:\\Users\\admin\\Desktop\\D.S-360\\1.Python\\Python Assignment\\Indian\_cities.csv")

type(cities)

cities

#a. #sorting sex\_ratio column

sex\_ratio\_values = cities.sort\_values(by="sex\_ratio",ascending = False).sex\_ratio

sex\_ratio\_values[0:10]

sex\_ratio = cities.sort\_values(by="sex\_ratio",ascending = False).name\_of\_city

sex\_ratio[0:10]

#b.

graduates\_value = cities.sort\_values(by="total\_graduates",ascending = False).total\_graduates

graduates\_value[0:10]

graduates\_top\_cities = cities.sort\_values(by = "total\_graduates",ascending = False).name\_of\_city

graduates\_top\_cities[0:10]

literracy\_value = cities.sort\_values(by="effective\_literacy\_rate\_total",ascending = False).effective\_literacy\_rate\_total

literracy\_value[0:10]

literracy\_cities = cities.sort\_values(by="effective\_literacy\_rate\_total",ascending = False).name\_of\_city

literracy\_cities[0:10]

literracy\_location = cities.sort\_values(by="effective\_literacy\_rate\_total",ascending = False).location

literracy\_location[0:10]

1. For the data set “Indian\_cities”
2. Construct histogram on literates\_total and comment about the inferences
3. Construct scatter plot between male graduates and female graduates

Ans:-

import matplotlib.pyplot as plt

#a. Histogram

plt.hist(cities['literates\_total'])

#histogram is positively scwed i.e mean>median>mode

#there is also persence of Outliers

#b. Scatter plot

plt.scatter(cities['male\_graduates'],cities['female\_graduates'])

1. For the data set “Indian\_cities”
2. Construct Boxplot on total effective literacy rate and draw inferences
3. Find out the number of null values in each column of the dataset and delete them.

Ans:-

#a. Boxplot

plt.boxplot(cities['effective\_literacy\_rate\_total'])

#distribution is little negetively scwed, and there is also persence of outliers.

#b. null values

cities.isna()

cities.dropna()