**Aryan Mohan**

**500092142**

**Batch- 2**

**Machine Learning Lab**

**Experiment- 3**

**Objective: Preparing synthetic data set to plot scatter plot using M.S.Excel and Python**

**Syntax:**

import pandas as pd

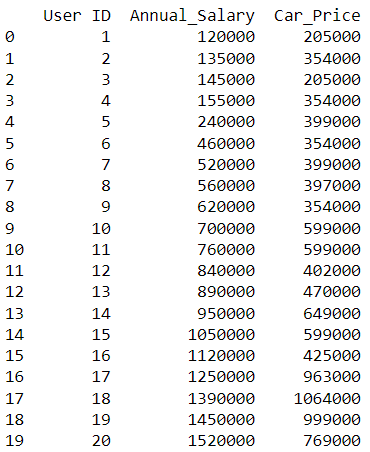
import matplotlib.pyplot as plt

# Experiment 3: Prepairing synthetic dataset to plot a graph to find out the pattern/relationship with scatterplot in ms excel/python. Study fatures of scatter plot.

uidsalcar\_df=pd.read\_csv('C:/Users/aryan/OneDrive/Documents/SalaryCarPrice.csv')

print(uidsalcar\_df)

**Output:**

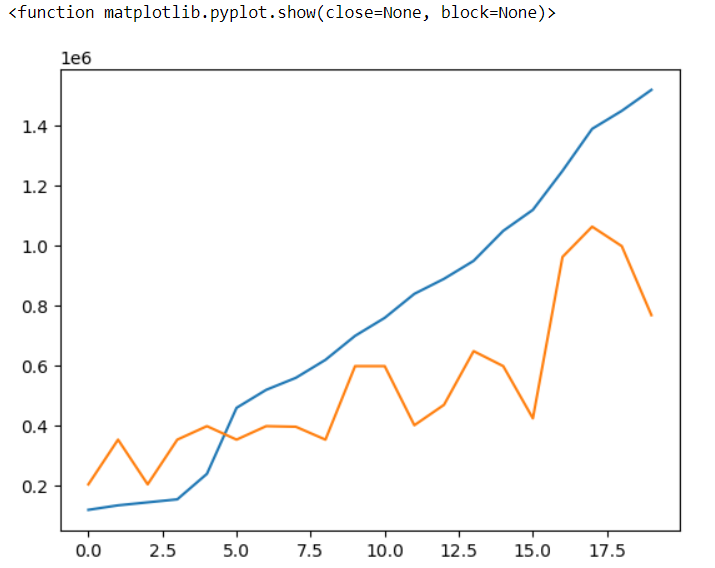


**Syntax:**

plt.plot(uidsalcar\_df)

plt.show

Output:



**Syntax:**

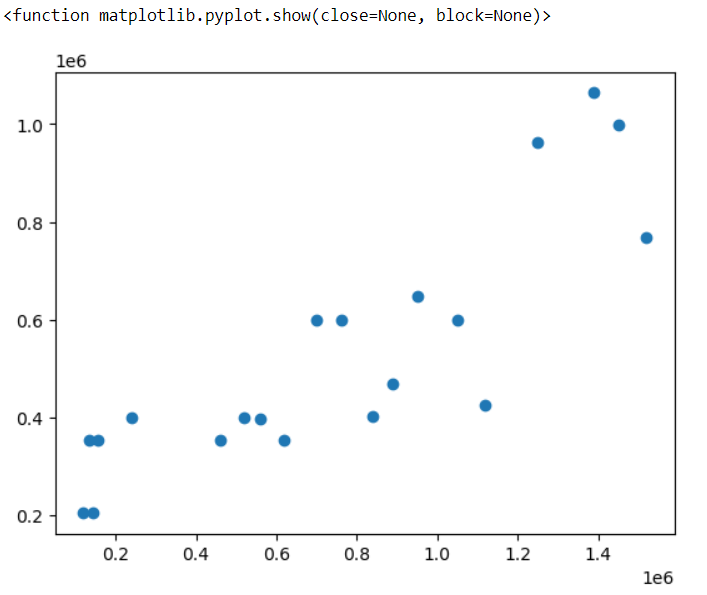
columns=["Annual\_Salary", "Car\_Price"]

uidsalcar\_df=pd.read\_csv('C:/Users/aryan/OneDrive/Documents/SalaryCarPrice.csv', usecols=columns)

plt.scatter(uidsalcar\_df.Annual\_Salary, uidsalcar\_df.Car\_Price)

plt.show

**Output:**



**Features of scatter plot:**

A scatter plot is a graphical representation of two variables that uses dots to represent individual data points. The position of each dot on the x- and y-axes represents values for an individual data point. The following are some of the features of a scatter plot:

* X- and Y-Axes: The x-axis and y-axis represent the two variables being compared. The x-axis is typically used to represent the independent variable and the y-axis is used to represent the dependent variable.
* Data Points: Each dot on the scatter plot represents a single data point and its position on the x- and y-axes represents the values of the two variables for that data point.
* Trends and Patterns: Scatter plots can be used to identify trends and patterns in the data. For example, a positive correlation between two variables will result in a upward slope in the scatter plot, while a negative correlation will result in a downward slope.
* Outliers: Scatter plots can also be used to identify outliers, which are data points that are significantly different from the rest of the data.
* Density: The density of data points on a scatter plot can provide additional information about the relationship between the two variables. For example, a high density of data points in a certain area of the plot may indicate a strong relationship between the two variables in that region.
* Customization: Scatter plots can be customized with different colors, shapes, and sizes to highlight specific data points or groups of data points.