Assignment-II

1. **Pandas basics**

   Let df be a pandas Dataframe constructed with the following code:

```
data = np.array([0, 7, 3, 6, 2, 8, 5, 9, 4]).reshape(3, -1)
df = pd.DataFrame(data, index=['One', 'Two', 'Three'], columns=['a', 'b', 'c'])
```

   What is the output of the following code? (Try to write the output without using python.)

   a. print(df)
   b. df[‘a’]
   c. df[‘One’]
   d. df.loc[‘Two’]
   e. df[:2]
   f. df.iloc[:,:2]
   g. list(df.columns)
   h. list(df.index)
   i. df[‘b’][‘Two’]
   j. list(df.iloc[2, :])
   k. df.drop('a', axis=1)
   l. df[df.a !=5]
   m. list(df.sum(axis=0))
   n. df.iloc[:, list(df.sum(axis=0) < 17)]
   o. df.sort_values(by='c')
   p. df.sort_values(by='Two', axis=1)
   q. df.T
   r. (df<=2).any(axis=0)
   s. df.applymap(lambda x: x*2-1)
   t. df.apply(lambda x: max(x), axis=1)

2. **Use pandas to load sample.csv file into a Dataframe called df2 and then do the following.**
   a. Show a boxplot of the data
   b. Use pandas function describe () to print out the summary statistics of the data
   c. Use pandas function hist to show the histogram of each column of the data frame. (Use option normed = True so it plots probability instead of counts.) Decide an appropriate number of bins and whether to apply log transformation on the data.

3. A Data Frame pandas:

	1990	2000	2010
1.	54	345	895
2.	64	485	562
3.	79	690	1100
4.	96	770	890

Write Code to Create:-

(a) A scatter chart from 1990 and 2010 of data framepd.

(b) A line chart from the 1990 and 2000 of data framepd.

(c) A bar chart to plotting the three columns of data framepd.

4. Write a Python programming to display a barchart of the popularity of programming Languages.

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7,6.7