LAB PROGRAM OUTPUTS

Gokul Shanthosh. M

192124203

1.Write a Pandas program to select distinct department id from employees file.

```
    Dept_id         dept_name  id  locationid
0        10    Administration   1        1500
1        20         Marketing   2        1600
2        30        Purchasing   3        1700
3        40   Human Resources   4        1800
4        50          Shipping   5        1900
0    10
1    20
2    30
3    40
4    50
Name: Dept_id, dtype: int64
```

----

2. Write a Pandas program to display the ID for those employees who did two or more jobs in the past.

```
    EMPLOYEE_ID  START_DATE    END_DATE     JOB_ID  DEPARTMENT_ID
0           102  2001-01-13  2006-07-24    IT_PROG             60
1           101  1997-09-21  2001-10-27  AC_ACCOUNT            110
2           101  2001-10-28  2005-03-15     AC_MGR            110
3           201  2004-02-17  2007-12-19     MK_REP             20
4           114  2006-03-24  2007-12-31   ST_CLERK             50
5           122  2007-01-01  2007-12-31   ST_CLERK             50
6           200  1995-09-17  2001-06-17    AD_ASST             90
7           176  2006-03-24  2006-12-31     SA_REP             80
8           176  2007-01-01  2007-12-31     SA_MAN             80
9           200  2002-07-01  2006-12-31  AC_ACCOUNT            90
EMPLOYEE_ID
101    2
102    1
114    1
122    1
176    2
200    2
201    1
Name: JOB_ID, dtype: int64
Int64Index([101, 176, 200], dtype='int64', name='EMPLOYEE_ID')
```

----

3. Write a Pandas program to display the details of jobs in descending sequence on job title.

```
 sorted by job title
          JOB_ID                     JOB_TITLE  MIN_SALARY  MAX_SALARY
11        ST_MAN                   Stock Manager        5500        8500
12      ST_CLERK                     Stock Clerk        2008        5000
13      SH_CLERK                  Shipping Clerk        2500        5500
8         SA_REP             Sales Representative        6000       12008
7         SA_MAN                   Sales Manager       10000       20080
9         PU_MAN              Purchasing Manager        8000       15000
10      PU_CLERK                Purchasing Clerk        2500        5500
18        PR_REP  Public Relations Representative       4500       10500
6     AC_ACCOUNT                Public Accountant        4200        9000
14       IT_PROG                      Programmer        4000       10000
0        AD_PRES                       President       20080       40000
16        MK_REP         Marketing Representative        4000        9000
15        MK_MAN               Marketing Manager        9000       15000
17        HR_REP   Human Resources Representative       4000        9000
3         FI_MGR                 Finance Manager        8200       16000
1         AD_VP    Administration Vice President       15000       30000
2        AD_ASST        Administration Assistant        3000        6000
5         AC_MGR              Accounting Manager        8200       16000
4     FI_ACCOUNT                      Accountant        4200        9000
```
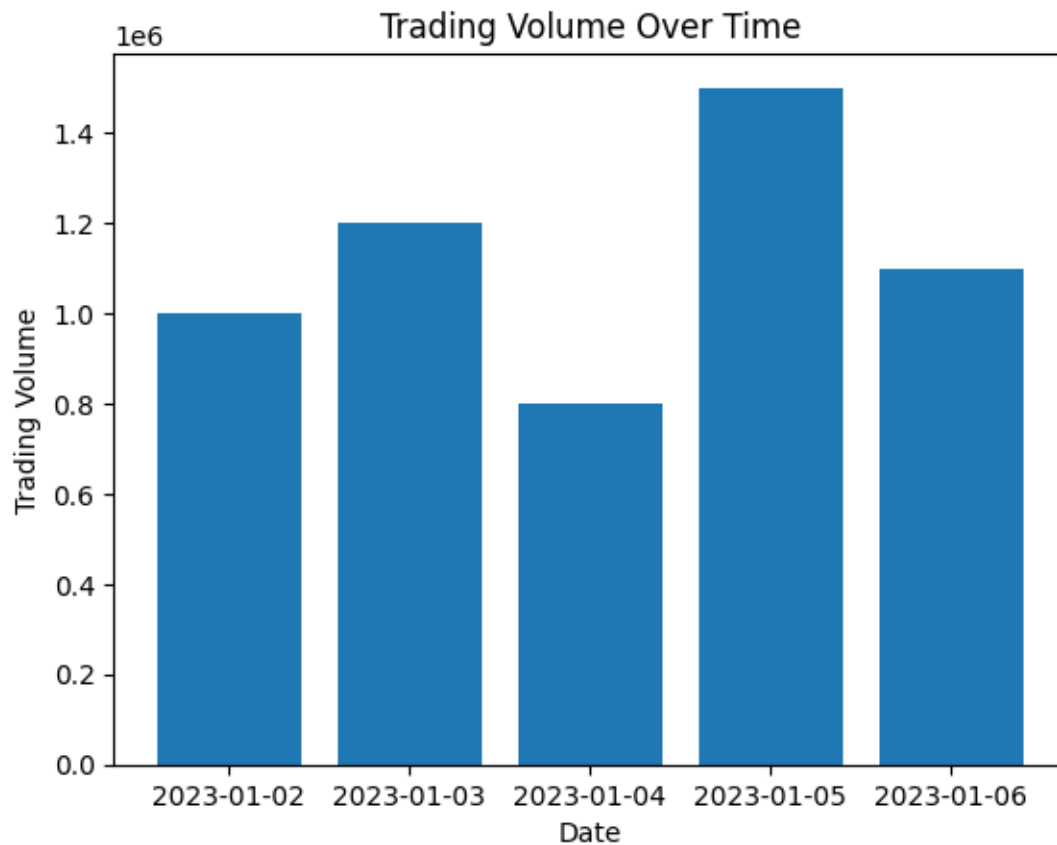
----

4. Write a Pandas program to create a line plot of the historical stock prices of Alphabet Inc. between two specific dates.

5.Write a Pandas program to create a bar plot of the trading volume of Alphabet Inc. stock between two specific dates.

----

6. Write a Pandas program to create a scatter plot of the trading volume/stock prices of Alphabet Inc. stock between two specific dates.

----
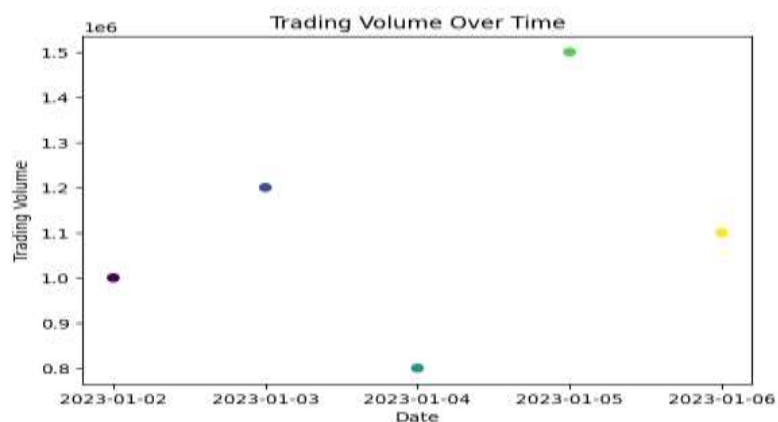
7. Write a Pandas program to create a Pivot table and find the maximum and minimum sale value of the items.(refer sales_data table)

```
    name  sales  profit
0     A    100      12
1     B    200      14
2     C    300      23
3     D    350       7
          max    min
        sales  sales
profit
7        350    350
12       100    100
14       200    200
23       300    300
profit
7     350
12    100
14    200
23    300
Name: (max, sales), dtype: int64
profit
7     350
12    100
14    200
23    300
Name: (min, sales), dtype: int64
```

----

8. Write a Pandas program to create a Pivot table and find the item wise unit sold. .(refer sales_data table).

```
Original DataFrame:
   name  sales  profit
0     A    100      12
1     B    200      14
2     C    300      23
3     D    350       7

Pivot table for item-wise unit sold:
profit    7    12   14    23
name
A         0   100    0     0
B         0     0  200     0
C         0     0    0   300
D       350     0    0     0
```

----

9. Write a Pandas program to create a Pivot table and find the total sale amount region wise, manager wise, salesman wise. . (refer sales data table)

```
                        Sale_amt
Region   Manager
Central  Hermann     298018.0
         Martha      131780.0
         Timothy       6075.0
East     Douglas      40500.0
         Martha      203646.0
West     Douglas      38336.0
         Timothy      67088.0
```

----

10. Create a data frame of ten rows, four columns with random values. Write a Pandas program to highlight the negative numbers red and positive numbers black.

	A	B	C	D
0	0.496714	-0.138264	0.647689	1.523030
1	-0.234153	-0.234137	1.579213	0.767435
2	-0.469474	0.542560	-0.463418	-0.465730
3	0.241962	-1.913280	-1.724918	-0.562288
4	-1.012831	0.314247	-0.908024	-1.412304
5	1.465649	-0.225776	0.067528	-1.424748
6	-0.544383	0.110923	-1.150994	0.375698
7	-0.600639	-0.291694	-0.601707	1.852278
8	-0.013497	-1.057711	0.822545	-1.220844
9	0.208864	-1.959670	-1.328186	0.196861

11.. Create a data frame of ten rows, four columns with random values. Convert some values to nan values. Write a Pandas program which will highlight the nan values.

	A	B	C	D
0	0.496714	-0.138264	0.647689	1.523030
1	-0.234153	-0.234137	1.579213	0.767435
2	-0.469474	nan	-0.463418	-0.465730
3	0.241962	-1.913280	-1.724918	-0.562288
4	-1.012831	0.314247	-0.908024	-1.412304
5	1.465649	-0.225776	0.067528	nan
6	-0.544383	0.110923	-1.150994	0.375698
7	-0.600639	-0.291694	-0.601707	1.852278
8	-0.013497	-1.057711	nan	-1.220844
9	0.208864	-1.959670	-1.328186	0.196861

----

12. Create a data frame of ten rows, four columns with random values. Write a Pandas program to set dataframe background Color black and font color yellow.

	Column1	Column2	Column3	Column4
0	0.573429	0.441949	0.796874	0.065566
1	0.820037	0.259022	0.179464	0.912874
2	0.560891	0.596891	0.784673	0.088654
3	0.350762	0.655286	0.970128	0.199388
4	0.543500	0.275695	0.362812	0.471953
5	0.879589	0.857972	0.087886	0.440997
6	0.114097	0.888724	0.343833	0.765551
7	0.031439	0.285061	0.571110	0.012744
8	0.952810	0.659560	0.165513	0.680663
9	0.288743	0.972120	0.631616	0.275603

----

13. Write a Pandas program to detect missing values of a given DataFrame. Display True or False.

```
        A       B       C
0   False   False   False
1   False    True   False
2    True   False   False
3   False   False   False
```

----

14. . Write a Pandas program to find and replace the missing values in a given DataFrame which do not have any valuable information.

```
Original DataFrame:
     A    B   C
0   1.0  5.0   9
1   2.0  NaN  10
2   NaN  7.0  11
3   4.0  8.0  12

DataFrame after replacing missing values:
          A         B   C
0  1.000000  5.000000   9
1  2.000000  6.666667  10
2  2.333333  7.000000  11
3  4.000000  8.000000  12
```

15.. Write a Pandas program to keep the rows with at least 2 NaN values in a given DataFrame.

```
Original DataFrame:
     A    B     C
0  1.0  5.0   9.0
1  NaN  NaN  10.0
2  3.0  7.0   NaN
3  NaN  NaN   NaN

DataFrame with rows having at least 2 NaN values:
     A    B    C
0  1.0  5.0  9.0
2  3.0  7.0  NaN
```

----

16. Write a Pandas program to split the following data frame into groups based on school code. Also check the type of Group By object.

```
school
s001      [Alberto Franco, Eesha Hinton]
s002        [Gino Mcneill, Gino Mcneill]
s003                        [Ryan Parkes]
s004                       [David Parkes]
Name: name, dtype: object
```

17.Write a Pandas program to split the following dataframe by school code and get mean, min, and max value of age for each school.

	age			height			weight		
	min	max	mean	min	max	mean	min	max	mean
school									
s001	12	13	12.5	167	173	170.0	30	35	32.5
s002	12	14	13.0	151	192	171.5	31	32	31.5
s003	13	13	13.0	186	186	186.0	33	33	33.0
s004	12	12	12.0	159	159	159.0	32	32	32.0

----

18. Write a Pandas program to split the following given dataframe into groups based on school code and class.

```
(('s001', 'V'),    class school           name date_Of_Birth  age  height  weight  address
0      V   s001  Alberto Franco    15/05/2002   12     173      35  street1)
(('s001', 'VI'),   class school           name date_Of_Birth  age  height  weight  address
3     VI   s001   Eesha Hinton    25/09/1998   13     167      30  street1)
(('s002', 'V'),    class school           name date_Of_Birth  age  height  weight  address
1      V   s002   Gino Mcneill    17/05/2002   12     192      32  street2
4      V   s002   Gino Mcneill    11/05/2002   14     151      31  street2)
(('s003', 'VI'),   class school           name date_Of_Birth  age  height  weight  address
2     VI   s003    Ryan Parkes    16/02/1999   13     186      33  street3)
(('s004', 'VI'),   class school           name date_Of_Birth  age  height  weight  address
5     VI   s004   David Parkes    15/09/1997   12     159      32  street4)
```

----

19. Write a Pandas program to display the dimensions or shape of the World alcohol consumption dataset. Also extract the column names from the dataset.

```
(5, 5)
```


20. Write a Pandas program to find the index of a given substring of a DataFrame column.

```
        Column1
0         apple
1        banana
2        cherry
3          date
4    elderberry
0    -1
1    -1
2     2
3    -1
4     6
Name: Column1, dtype: int64
```

----

21. Write a Pandas program to swap the cases of a specified character column in a given DataFrame.

```
        Name               City
0       jOHN           New York
1      aLICE        Los Angeles
2        bOB            Chicago
3        eVE      San Francisco
```
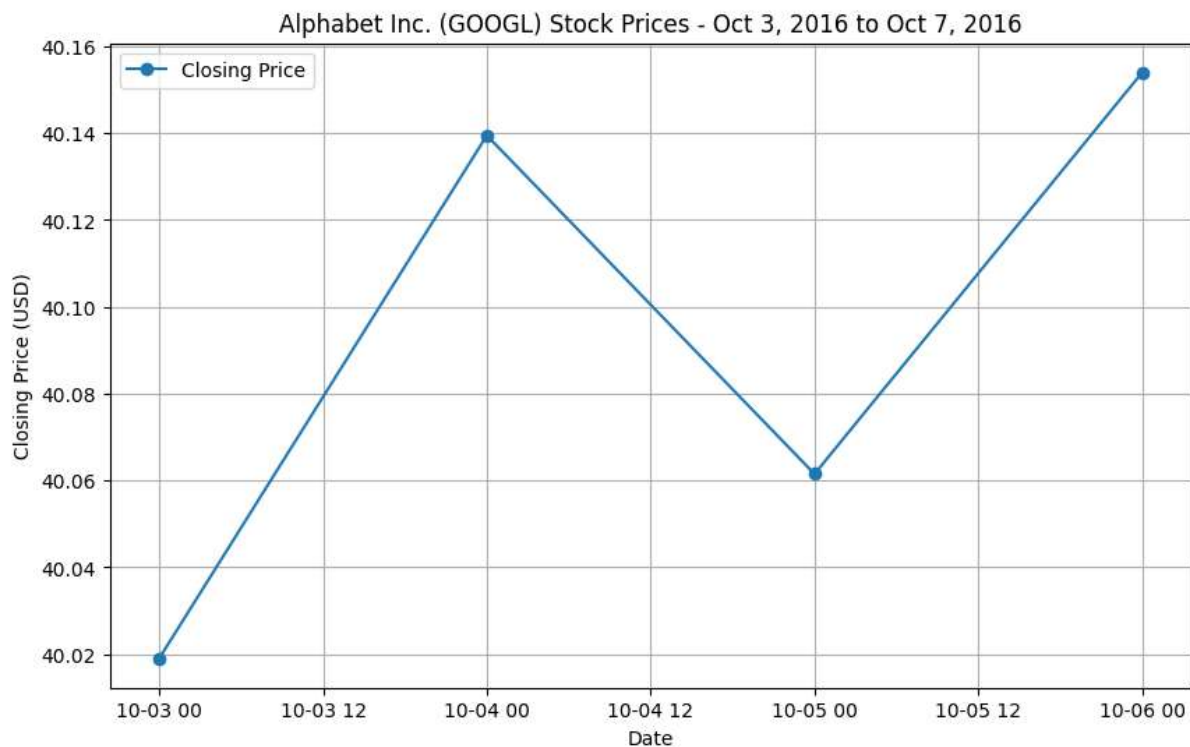
----

22. Write a Python program to draw a line with suitable label in the x axis, y axis and a title.
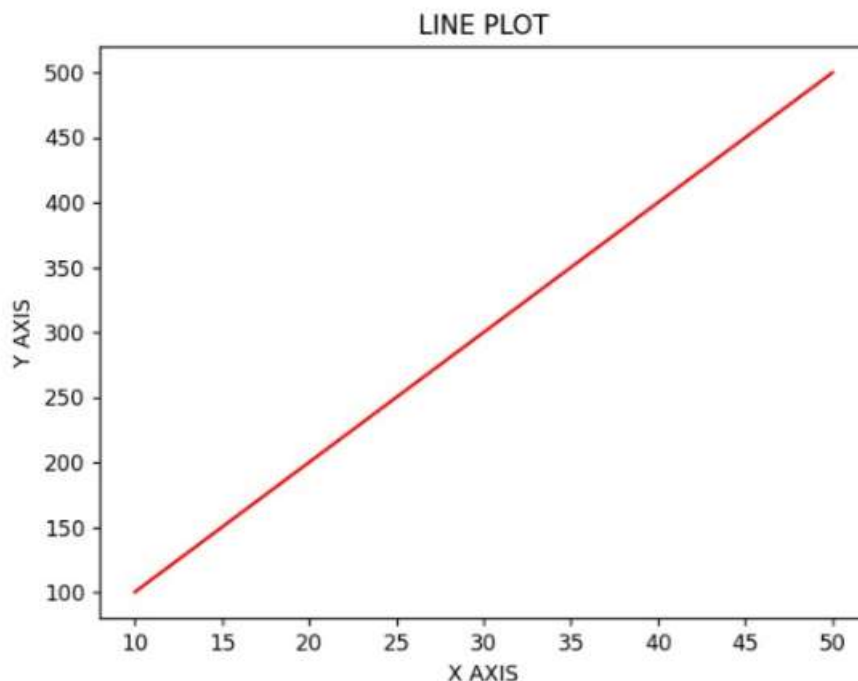
23.Write a Python program to draw a line using given axis values taken from a text file, with suitable label in the x axis, y axis and a title.

----
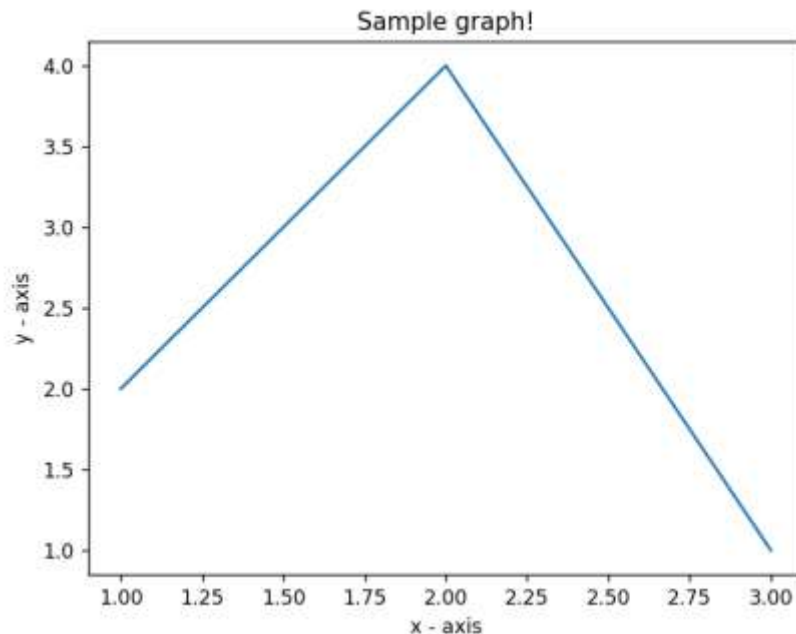
24. Write a Python program to draw line charts of the financial data of Alphabet Inc. between October 3, 2016 to October 7, 2016.
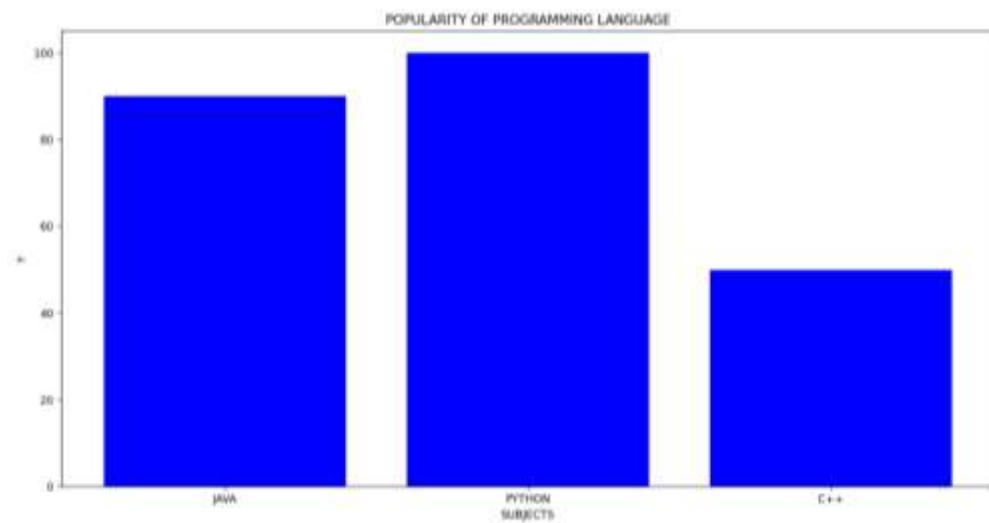

----

25. Write a Python program to plot two or more lines with legends, different widths and colours.
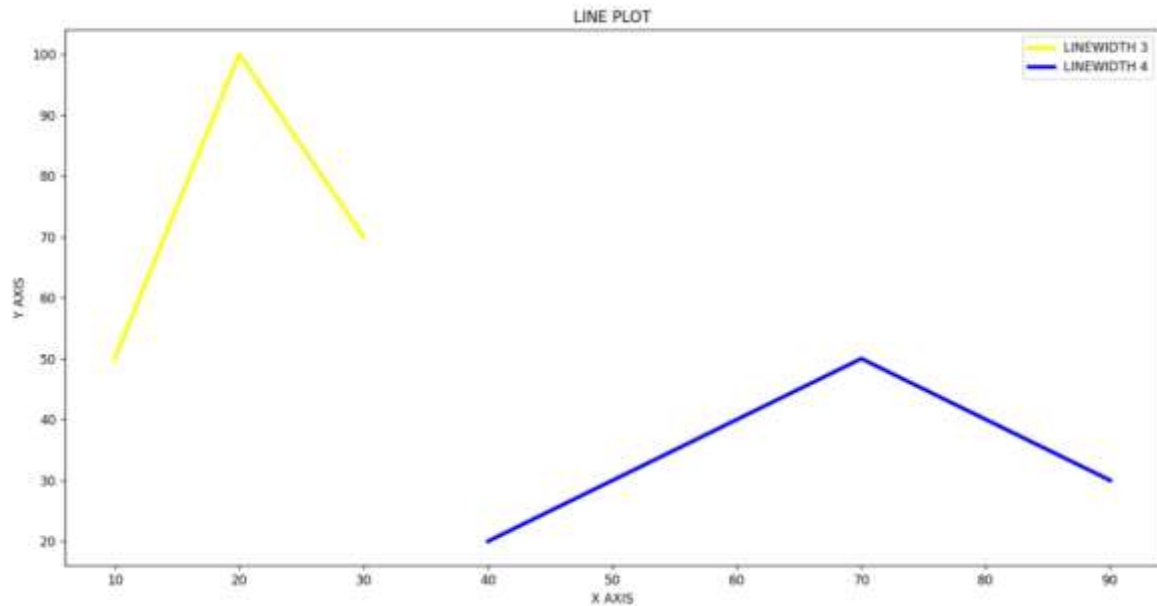
26. Write a Python program to create multiple plots.


27.Write a Python programming to display a bar chart of the popularity of programming Languages.

POPULARITY OF PROGRAMMING LANGUAGE

----

28. Write a Python programming to display a horizontal bar chart of the popularity of programming Languages.

POPULARITY OF PROGRAMMING LANGUAGE

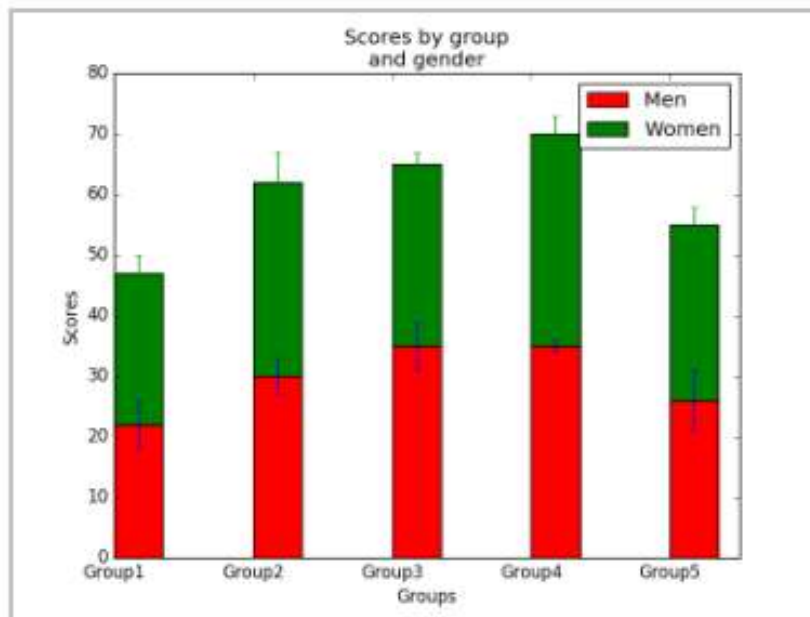--
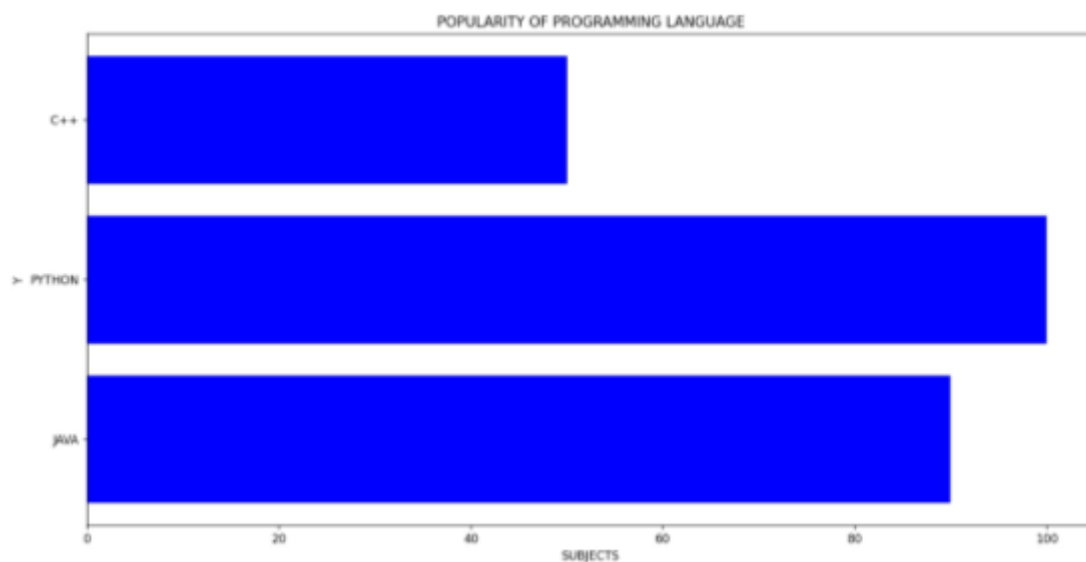
29. Write a Python programming to display a bar chart of the popularity of programming Languages. Use different color for each bar.

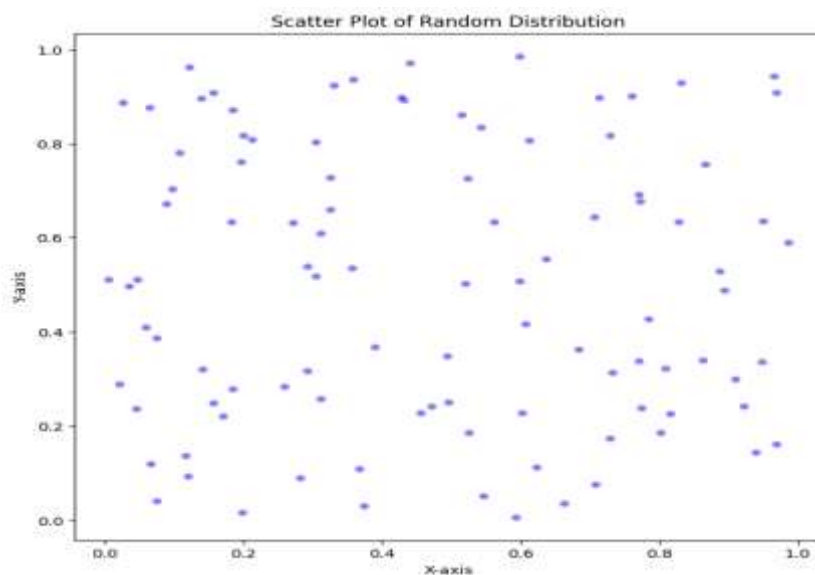--
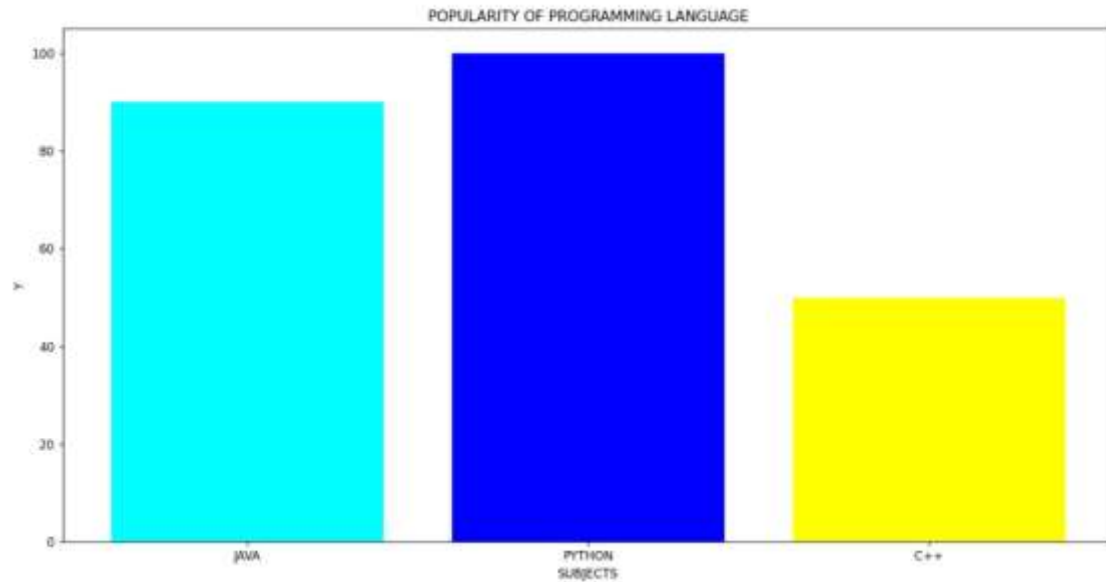
30. Write a Python program to create bar plot of scores by group and gender. Use multiple X values on the same chart for men and women.

--
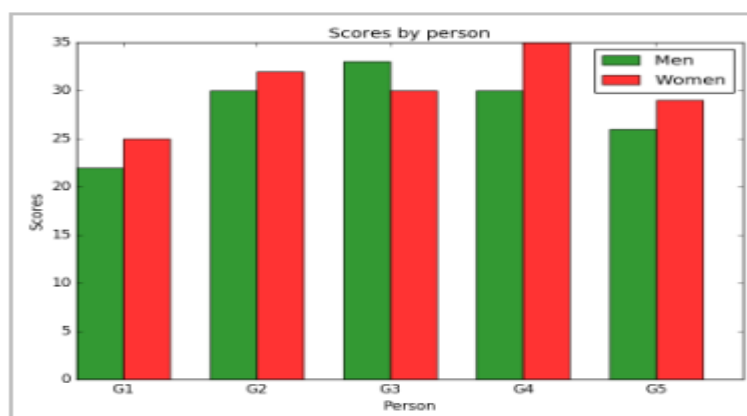
31. Write a Python program to create a stacked bar plot with error bars.

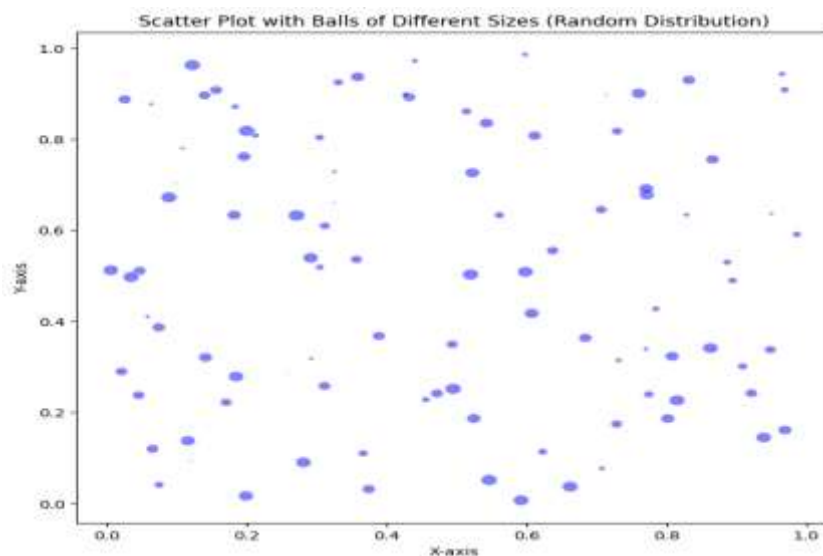**32. Write a Python program to draw a scatter graph taking a random distribution in X and Y and plotted against each other.**

--

33. Write a Python program to draw a scatter plot with empty circles taking a random distribution in X and Y and plotted against each other.

Scatter Plot with Empty Circles (Random Distribution)

---
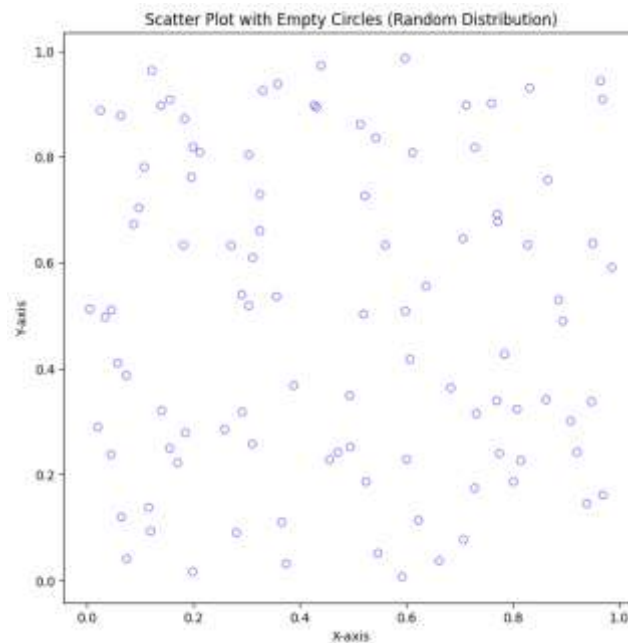
34. Write a Python program to draw a scatter plot using random distributions to generate balls of different sizes.

Scatter Plot with Balls of Different Sizes (Random Distribution)

----

35. Write a Python program to draw a scatter plot comparing two subject marks of Mathematics and Science. Use marks of 10 students.
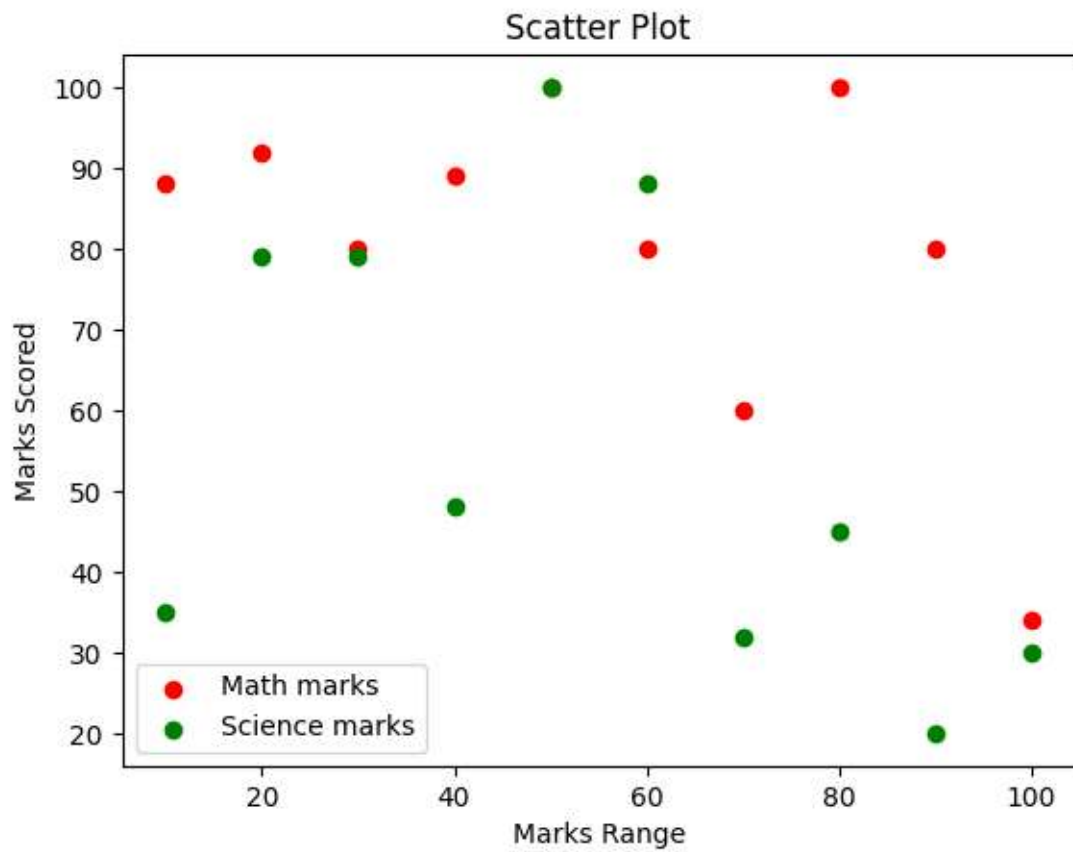
36.Write a Python program to draw a scatter plot for three different groups comparing weights and heights.

Group wise Weight vs Height scatter plot
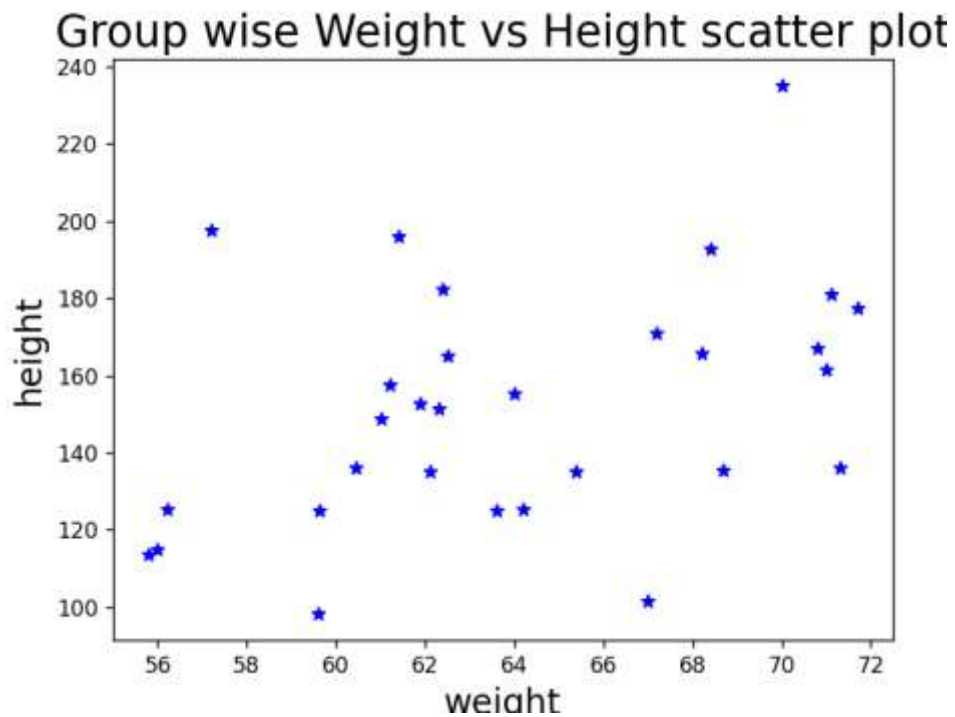
37.Write a Pandas program to create a dataframe from a dictionary and display it.

```
       Name  Age            City
0      John   25        New York
1     Alice   28   San Francisco
2       Bob   22     Los Angeles
3       Eva   24         Chicago
```

----

38. Write a Pandas program to create and display a DataFrame from a specified dictionary data which has the index labels.

```
      Name  Age             City
0     John   25         New York
1    Alice   28    San Francisco
2      Bob   22      Los Angeles
3      Eva   24          Chicago
0     John
1    Alice
2      Bob
3      Eva
Name: Name, dtype: object
```

----

39. Write a Pandas program to get the first 3 rows of a given DataFrame.

```
      Name  Age             City
0     John   25         New York
1    Alice   28    San Francisco
2      Bob   22      Los Angeles
3      Eva   24          Chicago
      Name  Age             City
0     John   25         New York
1    Alice   28    San Francisco
2      Bob   22      Los Angeles
```

----

40. Write a Pandas program to select the 'name' and 'score' columns from the following DataFrame.

```
      Name  Age             City
0     John   25         New York
1    Alice   28    San Francisco
2      Bob   22      Los Angeles
3      Eva   24          Chicago
0     John
1    Alice
2      Bob
3      Eva
Name: Name, dtype: object
0     25
1     28
2     22
3     24
Name: Age, dtype: int64
```