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

+---------------+----------------------+------------+-------------+

| DEPARTMENT\_ID | DEPARTMENT\_NAME | MANAGER\_ID | LOCATION\_ID |

+---------------+----------------------+------------+-------------+

| 10 | Administration | 200 | 1700 |

| 20 | Marketing | 201 | 1800 |

| 30 | Purchasing | 114 | 1700 |

| 40 | Human Resources | 203 | 2400 |

| 50 | Shipping | 121 | 1500 |

| 60 | IT | 103 | 1400 |

| 70 | Public Relations | 204 | 2700 |

| 80 | Sales | 145 | 2500 |

| 90 | Executive | 100 | 1700 |

| 100 | Finance | 108 | 1700 |

| 110 | Accounting | 205 | 1700 |

| 120 | Treasury | 0 | 1700 |

| 130 | Corporate Tax | 0 | 1700 |

| 140 | Control And Credit | 0 | 1700 |

| 150 | Shareholder Services | 0 | 1700 |

| 160 | Benefits | 0 | 1700 |

| 170 | Manufacturing | 0 | 1700 |

| 180 | Construction | 0 | 1700 |

| 190 | Contracting | 0 | 1700 |

| 200 | Operations | 0 | 1700 |

| 210 | IT Support | 0 | 1700 |

| 220 | NOC | 0 | 1700 |

| 230 | IT Helpdesk | 0 | 1700 |

| 240 | Government Sales | 0 | 1700 |

| 250 | Retail Sales | 0 | 1700 |

| 260 | Recruiting | 0 | 1700 |

| 270 | Payroll | 0 | 1700 |

+---------------+----------------------+------------+-------------

1. 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 |

+-------------+------------+------------+------------+---------------+

| 102 | 2001-01-13 | 2006-07-24 | IT\_PROG | 60 |

| 101 | 1997-09-21 | 2001-10-27 | AC\_ACCOUNT | 110 |

| 101 | 2001-10-28 | 2005-03-15 | AC\_MGR | 110 |

| 201 | 2004-02-17 | 2007-12-19 | MK\_REP | 20 |

| 114 | 2006-03-24 | 2007-12-31 | ST\_CLERK | 50 |

| 122 | 2007-01-01 | 2007-12-31 | ST\_CLERK | 50 |

| 200 | 1995-09-17 | 2001-06-17 | AD\_ASST | 90 |

| 176 | 2006-03-24 | 2006-12-31 | SA\_REP | 80 |

| 176 | 2007-01-01 | 2007-12-31 | SA\_MAN | 80 |

| 200 | 2002-07-01 | 2006-12-31 | AC\_ACCOUNT | 90 |

+-------------+------------+------------+------------+---------------+

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

+------------+---------------------------------+------------+------------+

| JOB\_ID | JOB\_TITLE | MIN\_SALARY | MAX\_SALARY |

+------------+---------------------------------+------------+------------+

| AD\_PRES | President | 20080 | 40000 |

| AD\_VP | Administration Vice President | 15000 | 30000 |

| AD\_ASST | Administration Assistant | 3000 | 6000 |

| FI\_MGR | Finance Manager | 8200 | 16000 |

| FI\_ACCOUNT | Accountant | 4200 | 9000 |

| AC\_MGR | Accounting Manager | 8200 | 16000 |

| AC\_ACCOUNT | Public Accountant | 4200 | 9000 |

| SA\_MAN | Sales Manager | 10000 | 20080 |

| SA\_REP | Sales Representative | 6000 | 12008 |

| PU\_MAN | Purchasing Manager | 8000 | 15000 |

| PU\_CLERK | Purchasing Clerk | 2500 | 5500 |

| ST\_MAN | Stock Manager | 5500 | 8500 |

| ST\_CLERK | Stock Clerk | 2008 | 5000 |

| SH\_CLERK | Shipping Clerk | 2500 | 5500 |

| IT\_PROG | Programmer | 4000 | 10000 |

| MK\_MAN | Marketing Manager | 9000 | 15000 |

| MK\_REP | Marketing Representative | 4000 | 9000 |

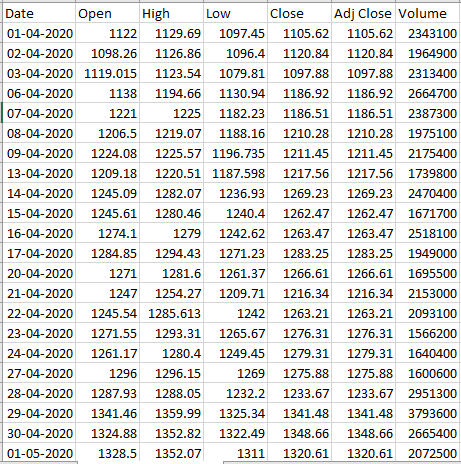
| HR\_REP | Human Resources Representative | 4000 | 9000 |

| PR\_REP | Public Relations Representative | 4500 | 10500 |

+------------+---------------------------------+------------+------------+

1. Write a Pandas program to create a line plot of the historical stock prices of Alphabet Inc. between two specific dates.
2. Write a Pandas program to create a bar plot of the trading volume of Alphabet Inc. stock between two specific dates.
3. Write a Pandas program to create a scatter plot of the trading volume/stock prices of Alphabet Inc. stock between two specific dates.

**alphabet\_stock\_data:**

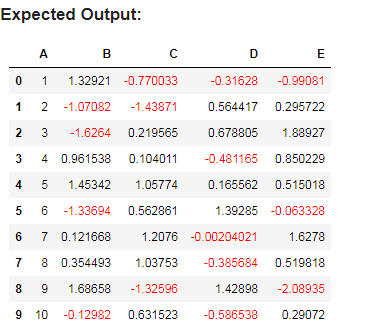


1. Write a Pandas program to create a Pivot table and find the maximum and minimum sale value of the items.(refer sales\_data table)
2. Write a Pandas program to create a Pivot table and find the item wise unit sold. .(refer sales\_data table)
3. Write a Pandas program to create a Pivot table and find the total sale amount region wise, manager wise, sales man wise. .(refer sales\_data table)

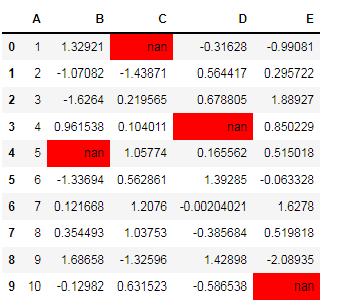
**Sales\_data:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **OrderDate** | **Region** | **Manager** | **SalesMan** | **Item** | **Units** | **Unit\_price** | **Sale\_amt** |
| 1-6-18 | East | Martha | Alexander | Television | 95 | 1,198.00 | 1,13,810.00 |
| 1-23-18 | Central | Hermann | Shelli | Home Theater | 50 | 500.00 | 25,000.00 |
| 2-9-18 | Central | Hermann | Luis | Television | 36 | 1,198.00 | 43,128.00 |
| 2-26-18 | Central | Timothy | David | Cell Phone | 27 | 225.00 | 6,075.00 |
| 3-15-18 | West | Timothy | Stephen | Television | 56 | 1,198.00 | 67,088.00 |
| 4-1-18 | East | Martha | Alexander | Home Theater | 60 | 500.00 | 30,000.00 |
| 4-18-18 | Central | Martha | Steven | Television | 75 | 1,198.00 | 89,850.00 |
| 5-5-18 | Central | Hermann | Luis | Television | 90 | 1,198.00 | 1,07,820.00 |
| 5-22-18 | West | Douglas | Michael | Television | 32 | 1,198.00 | 38,336.00 |
| 6-8-18 | East | Martha | Alexander | Home Theater | 60 | 500.00 | 30,000.00 |
| 6-25-18 | Central | Hermann | Sigal | Television | 90 | 1,198.00 | 1,07,820.00 |
| 7-12-18 | East | Martha | Diana | Home Theater | 29 | 500.00 | 14,500.00 |
| 7-29-18 | East | Douglas | Karen | Home Theater | 81 | 500.00 | 40,500.00 |
| 8-15-18 | East | Martha | Alexander | Television | 35 | 1,198.00 | 41,930.00 |
| 9-1-18 | Central | Douglas | John | Desk | 2 | 125.00 | 250.00 |
| 9-18-18 | East | Martha | Alexander | Video Games | 16 | 58.50 | 936.00 |
| 10-5-18 | Central | Hermann | Sigal | Home Theater | 28 | 500.00 | 14,000.00 |
| 10-22-18 | East | Martha | Alexander | Cell Phone | 64 | 225.00 | 14,400.00 |

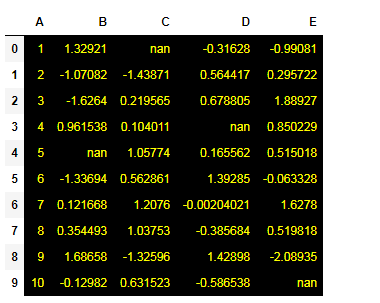
10.Create a dataframe of ten rows, four columns with random values. Write a Pandas program to highlight the negative numbers red and positive numbers black.



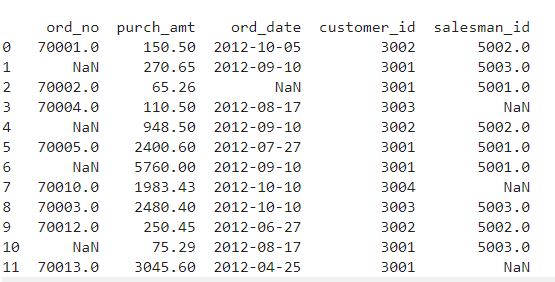
11.Create a dataframe of ten rows, four columns with random values. Convert some values to nan values. Write a Pandas program which will highlight the nan values.



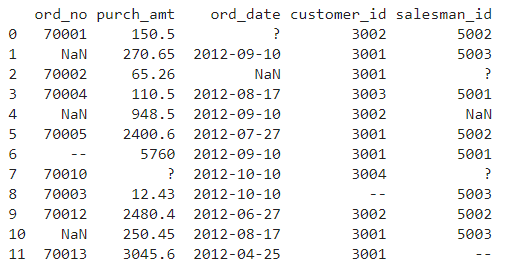
12.Create a dataframe of ten rows, four columns with random values. Write a Pandas program to set dataframe background Color black and font color yellow.



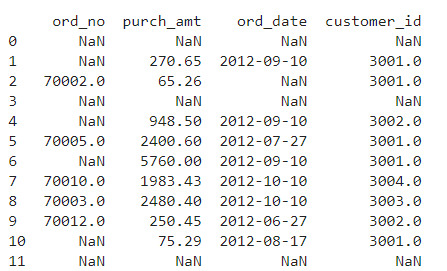
13.Write a Pandas program to detect missing values of a given DataFrame. Display True or False.



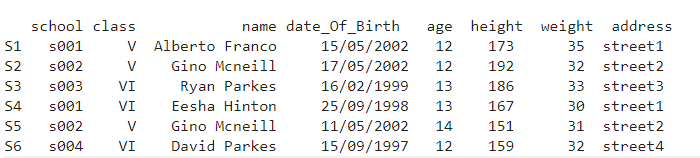
14. Write a Pandas program to find and replace the missing values in a given DataFrame which do not have any valuable information.



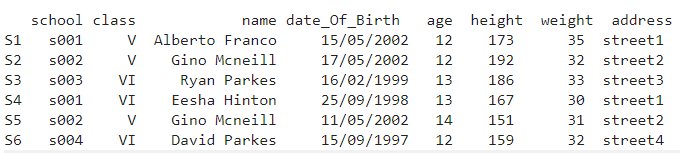
15.Write a Pandas program to keep the rows with at least 2 NaN values in a given DataFrame.



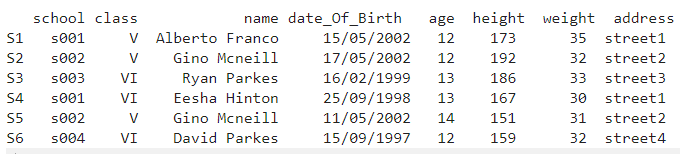
16.Write a Pandas program to split the following dataframe into groups based on school code. Also check the type of GroupBy object.



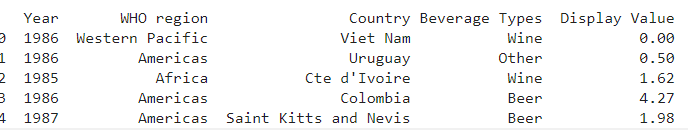
17.Write a Pandas program to split the following dataframe by school code and get mean, min, and max value of age for each school.



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



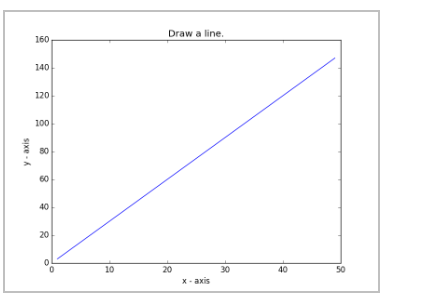
19.Write a Pandas program to display the dimensions or shape of the World alcohol consumption dataset. Also extract the column names from the dataset.



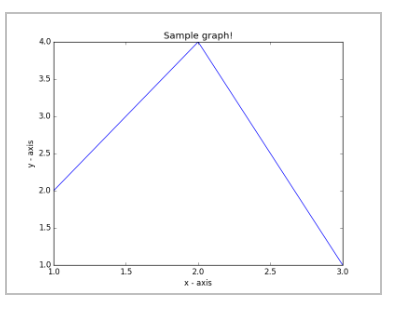
20.Write a Pandas program to find the index of a given substring of a DataFrame column.

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

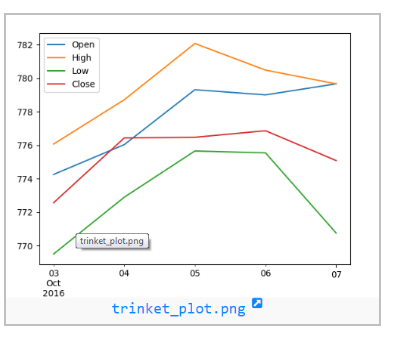
22.Write a Python program to draw a line with suitable label in the x axis, y axis and a title.



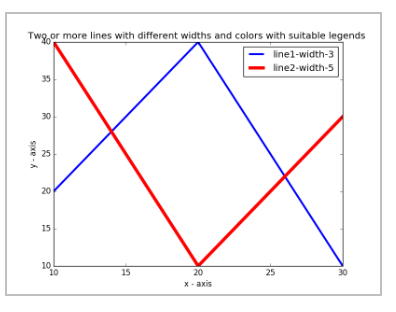
23.Write a Python program to draw a line using given axis values taken from a text file, with suitable label in the x axis, y axis and a title.  
Test Data:  
test.txt  
1 2  
2 4  
3 1



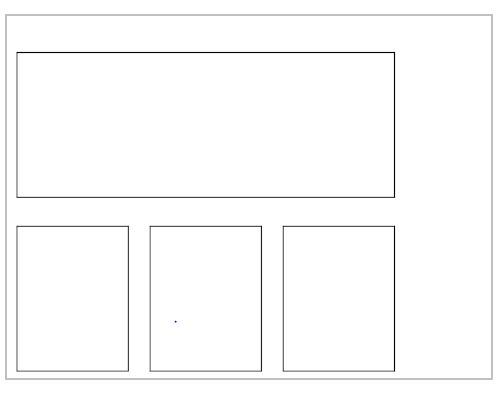
24.Write a Python program to draw line charts of the financial data of Alphabet Inc. between October 3, 2016 to October 7, 2016.  
Sample Financial data (fdata.csv):  
Date,Open,High,Low,Close  
10-03-16,774.25,776.065002,769.5,772.559998  
10-04-16,776.030029,778.710022,772.890015,776.429993  
10-05-16,779.309998,782.070007,775.650024,776.469971  
10-06-16,779,780.47998,775.539978,776.859985  
10-07-16,779.659973,779.659973,770.75,775.080017



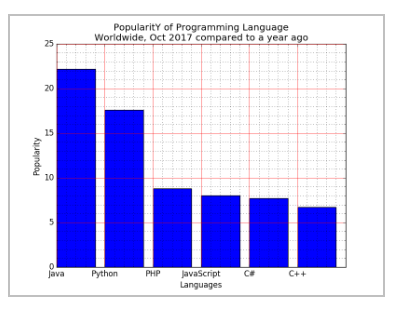
25.Write a Python program to plot two or more lines with legends, different widths and colors.



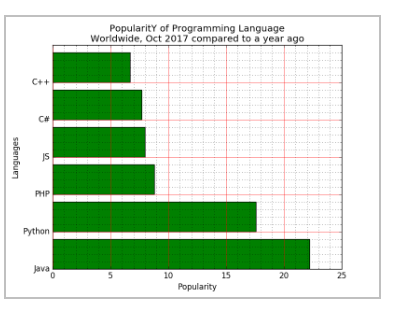
26.Write a Python program to create multiple plots.



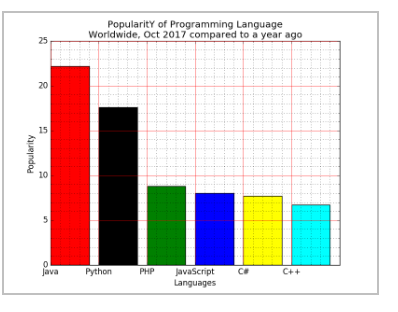
27.Write a Python programming to display a bar chart of the popularity of programming Languages.  
Sample data:  
Programming languages: Java, Python, PHP, JavaScript, C#, C++  
Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7



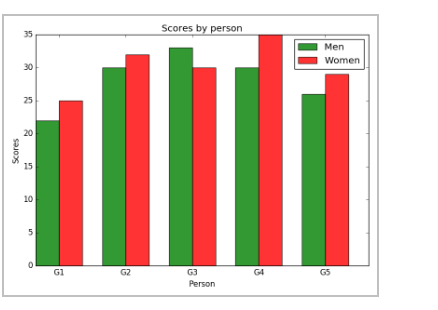
28.Write a Python programming to display a horizontal bar chart of the popularity of programming Languages.  
Sample data:  
Programming languages: Java, Python, PHP, JavaScript, C#, C++  
Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7



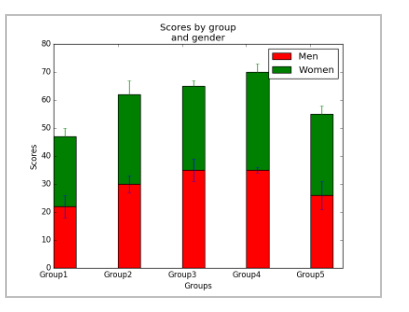
29.Write a Python programming to display a bar chart of the popularity of programming Languages. Use different color for each bar.  
Sample data:  
Programming languages: Java, Python, PHP, JavaScript, C#, C++  
Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7



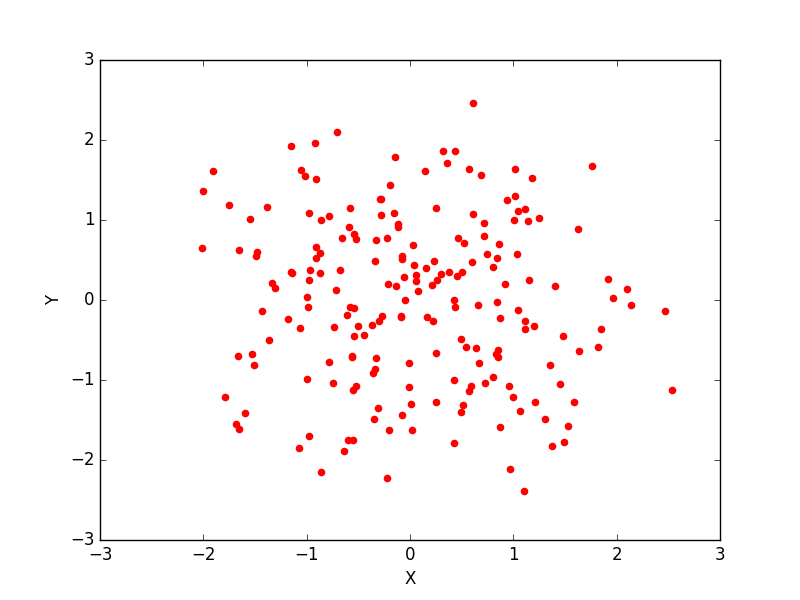
30.Write a Python program to create bar plot of scores by group and gender. Use multiple X values on the same chart for men and women.  
  
Sample Data:  
Means (men) = (22, 30, 35, 35, 26)  
Means (women) = (25, 32, 30, 35, 29)



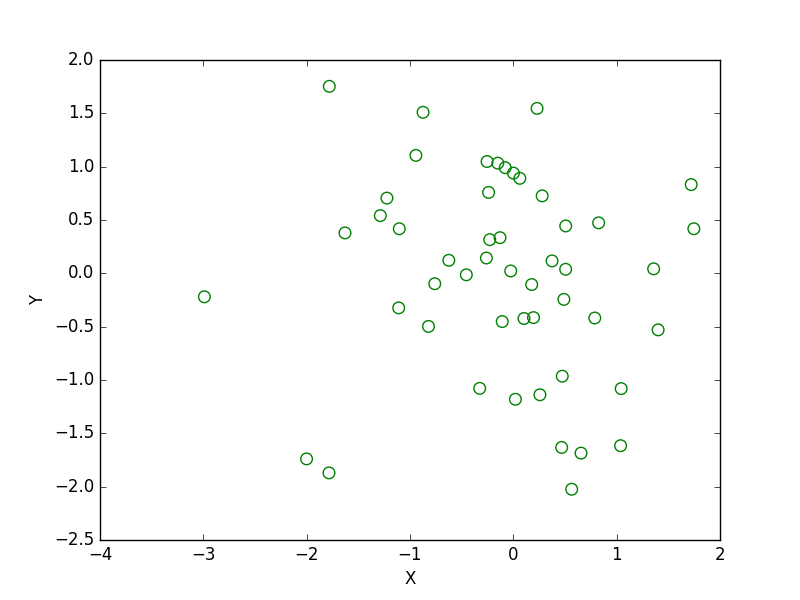
31.Write a Python program to create a stacked bar plot with error bars.  
Note: Use bottom to stack the women?s bars on top of the men?s bars.  
Sample Data:  
Means (men) = (22, 30, 35, 35, 26)  
Means (women) = (25, 32, 30, 35, 29)  
Men Standard deviation = (4, 3, 4, 1, 5)  
Women Standard deviation = (3, 5, 2, 3, 3)



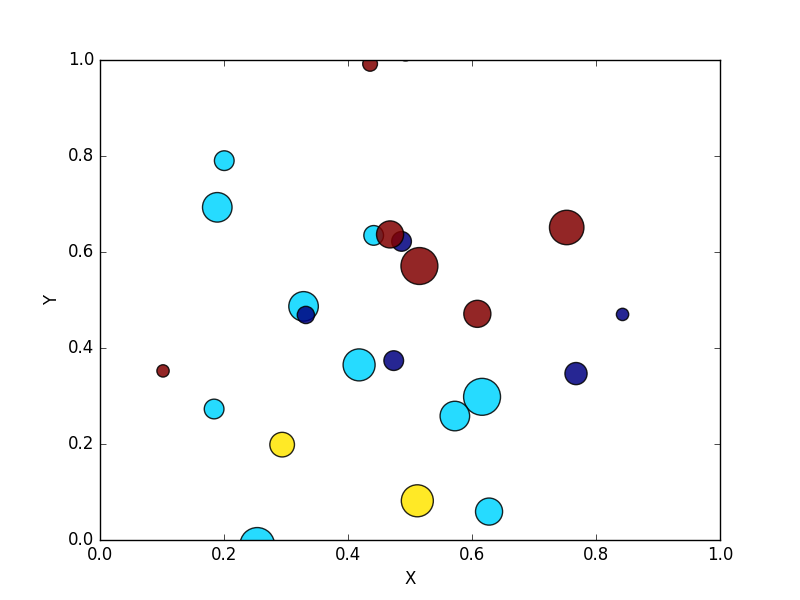
32.Write a Python program to draw a scatter graph taking a random distribution in X and Y and plotted against each other.



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.

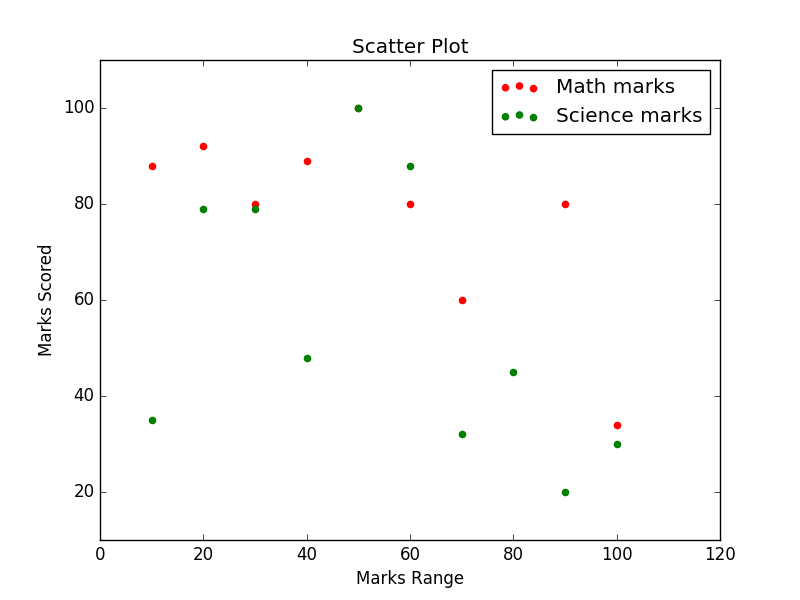


34.Write a Python program to draw a scatter plot using random distributions to generate balls of different sizes.

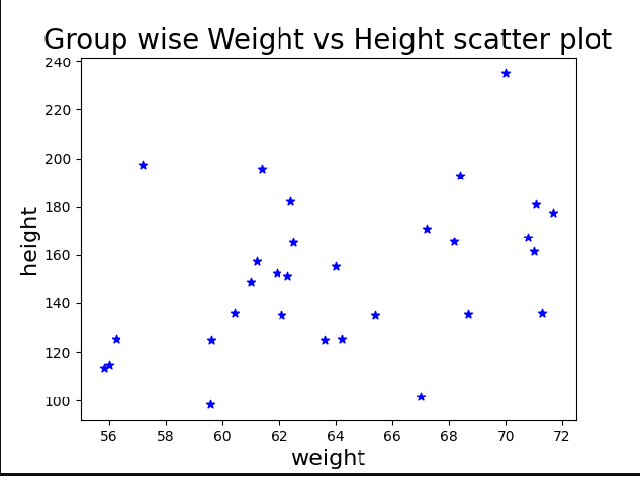


35.Write a Python program to draw a scatter plot comparing two subject marks of Mathematics and Science. Use marks of 10 students.  
Sample data:

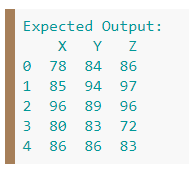
Test Data:  
math\_marks = [88, 92, 80, 89, 100, 80, 60, 100, 80, 34]  
science\_marks = [35, 79, 79, 48, 100, 88, 32, 45, 20, 30]  
marks\_range = [10, 20, 30, 40, 50, 60, 70, 80, 90, 100]



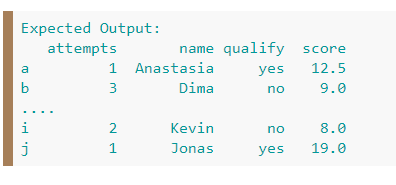
36.Write a Python program to draw a scatter plot for three different groups comparing weights and heights.



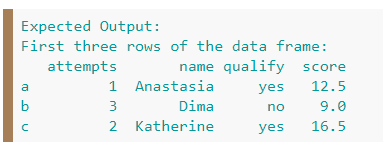
37.Write a Pandas program to create a dataframe from a dictionary and display it.  
Sample data: {'X':[78,85,96,80,86], 'Y':[84,94,89,83,86],'Z':[86,97,96,72,83]}



38.Write a Pandas program to create and display a DataFrame from a specified dictionary data which has the index labels.  
Sample Python dictionary data and list labels:  
exam\_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],  
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],  
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],  
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}  
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']



39.Write a Pandas program to get the first 3 rows of a given DataFrame.  
Sample Python dictionary data and list labels:  
exam\_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],  
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],  
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],  
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}  
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']



1. Write a Pandas program to select the 'name' and 'score' columns from the following DataFrame.  
   Sample Python dictionary data and list labels:  
   exam\_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],  
   'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],  
   'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],  
   'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}  
   labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

