In [2]:
```
pip install numpy
```

```
Collecting numpy
  Downloading numpy-1.24.3-cp311-cp311-win_amd64.whl (14.8 MB)
     ---------------------------------------- 14.8/14.8 MB 3.0 MB/s eta 0:00:00
Installing collected packages: numpy
Successfully installed numpy-1.24.3
Note: you may need to restart the kernel to use updated packages.
```

[notice] A new release of pip available: 22.3.1 -> 23.1.2
[notice] To update, run: python.exe -m pip install --upgrade pip

In [5]:
```
pip install pandas
```

```
Note: you may need to restart the kernel to use updated packages.Collecting pandas
  Downloading pandas-2.0.1-cp311-cp311-win_amd64.whl (10.6 MB)
     ---------------------------------------- 10.6/10.6 MB 3.6 MB/s eta 0:00:00
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\shanm\appdata\local\programs\python\python311\lib\site-packages (from pandas) (2.8.2)
Collecting pytz>2020.1
  Downloading pytz-2023.3-py2.py3-none-any.whl (502 kB)
     ---------------------------------------- 502.3/502.3 kB 3.2 MB/s eta 0:00:00
Collecting tzdata>=2022.1
  Downloading tzdata-2023.3-py2.py3-none-any.whl (341 kB)
     ---------------------------------------- 341.8/341.8 kB 2.7 MB/s eta 0:00:00
Requirement already satisfied: numpy>=1.21.0 in c:\users\shanm\appdata\local\programs\python\python311\lib\site-packages (from pandas) (1.24.3)
Requirement already satisfied: six>=1.5 in c:\users\shanm\appdata\local\programs\python\python311\lib\site-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
Installing collected packages: pytz, tzdata, pandas
Successfully installed pandas-2.0.1 pytz-2023.3 tzdata-2023.3
```

[notice] A new release of pip available: 22.3.1 -> 23.1.2
[notice] To update, run: python.exe -m pip install --upgrade pip

In [8]:
```
pip install matplotlib
```

```
Collecting matplotlib
  Downloading matplotlib-3.7.1-cp311-cp311-win_amd64.whl (7.6 MB)
     ---------------------------------------- 7.6/7.6 MB 4.0 MB/s eta 0:00:00
Collecting contourpy>=1.0.1
  Downloading contourpy-1.0.7-cp311-cp311-win_amd64.whl (162 kB)
     ---------------------------------------- 163.0/163.0 kB 3.2 MB/s eta 0:00:00
Collecting cycler>=0.10
  Downloading cycler-0.11.0-py3-none-any.whl (6.4 kB)
Collecting fonttools>=4.22.0
  Downloading fonttools-4.39.4-py3-none-any.whl (1.0 MB)
     ---------------------------------------- 1.0/1.0 MB 3.1 MB/s eta 0:00:00
Collecting kiwisolver>=1.0.1
  Downloading kiwisolver-1.4.4-cp311-cp311-win_amd64.whl (55 kB)
     ---------------------------------------- 55.4/55.4 kB 2.8 MB/s eta 0:00:00
Requirement already satisfied: numpy>=1.20 in c:\users\shanm\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (1.24.3)
Requirement already satisfied: packaging>=20.0 in c:\users\shanm\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (23.1)
Collecting pillow>=6.2.0
  Downloading Pillow-9.5.0-cp311-cp311-win_amd64.whl (2.5 MB)
     ---------------------------------------- 2.5/2.5 MB 3.3 MB/s eta 0:00:00
Collecting pyparsing>=2.3.1
  Downloading pyparsing-3.0.9-py3-none-any.whl (98 kB)
     ---------------------------------------- 98.3/98.3 kB 5.5 MB/s eta 0:00:00
Requirement already satisfied: python-dateutil>=2.7 in c:\users\shanm\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (2.8.2)
Requirement already satisfied: six>=1.5 in c:\users\shanm\appdata\local\programs\python\python311\lib\site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)
Installing collected packages: pyparsing, pillow, kiwisolver, fonttools, cycler, contourpy, matplotlib
Successfully installed contourpy-1.0.7 cycler-0.11.0 fonttools-4.39.4 kiwisolver-1.4.4 matplotlib-3.7.1 pillow-9.5.0 pyparsing-3.0.9
Note: you may need to restart the kernel to use updated packages.
```

[notice] A new release of pip available: 22.3.1 -> 23.1.2
[notice] To update, run: python.exe -m pip install --upgrade pip

In [9]:
```
pip install scipy
```

```
Collecting scipy
  Downloading scipy-1.10.1-cp311-cp311-win_amd64.whl (42.2 MB)
     ---------------------------------------- 42.2/42.2 MB 8.3 MB/s eta 0:00:00
Requirement already satisfied: numpy<1.27.0,>=1.19.5 in c:\users\shanm\appdata\local\programs\python\python311\lib\site-packages (from scipy) (1.24.3)
Installing collected packages: scipy
Successfully installed scipy-1.10.1
Note: you may need to restart the kernel to use updated packages.
```

[notice] A new release of pip available: 22.3.1 -> 23.1.2
[notice] To update, run: python.exe -m pip install --upgrade pip

In [7]:
```
import numpy as np
```

In [8]:
```
import pandas as pd
```

In [9]:
```
import matplotlib.pyplot as plt
```

In [10]:
```
import scipy
```

In [11]:
```
pip install openpyxl
```

```
Requirement already satisfied: openpyxl in c:\users\shanm\appdata\local\programs\python\python311\lib\site-packages (3.1.2)
Requirement already satisfied: et-xmlfile in c:\users\shanm\appdata\local\programs\python\python311\lib\site-packages (from openpyxl) (1.1.0)
Note: you may need to restart the kernel to use updated packages.
```

[notice] A new release of pip available: 22.3.1 -> 23.1.2
[notice] To update, run: python.exe -m pip install --upgrade pip

In [12]:
```
data=pd.read_excel('Student_Performance_Data.xlsx')
```

In [27]:
```
data.head(10)
```

Out[27]:

	Student_ID	Semster_Name	Paper_ID	Paper_Name	Marks
0	SID20131143	Sem_1	SEM00012995	Paper 1	44
1	SID20131143	Sem_1	SEM00015183	Paper 2	74
2	SID20131143	Sem_1	SEM00018371	Paper 3	80
3	SID20131143	Sem_1	SEM00015910	Paper 4	44
4	SID20131143	Sem_1	SEM00016208	Paper 5	95
5	SID20131143	Sem_1	SEM00017431	Paper 6	61
6	SID20131143	Sem_1	SEM00014130	Paper 7	90
7	SID20131143	Sem_2	SEM00024747	Paper 1	92
8	SID20131143	Sem_2	SEM00025909	Paper 2	57
9	SID20131143	Sem_2	SEM00022443	Paper 3	91

In [28]:
```
data.tail(5)
```

Out[28]:

	Student_ID	Semster_Name	Paper_ID	Paper_Name	Marks
209606	SID20189989	Sem_8	SEM00082598	Paper 3	94
209607	SID20189989	Sem_8	SEM00088030	Paper 4	49
209608	SID20189989	Sem_8	SEM00081794	Paper 5	47
209609	SID20189989	Sem_8	SEM00086600	Paper 6	87
209610	SID20189989	Sem_8	SEM00083259	Paper 7	73

In [29]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 209611 entries, 0 to 209610
Data columns (total 5 columns):
 #   Column       Non-Null Count   Dtype 
---  ------       --------------   ----- 
 0   Student_ID   209611 non-null  object
 1   Semster_Name  209611 non-null  object
 2   Paper_ID     209611 non-null  object
 3   Paper_Name   209611 non-null  object
 4   Marks        209611 non-null  int64 
dtypes: int64(1), object(4)
memory usage: 8.0+ MB
```

In [34]:
```
#to find the location
data.loc[0]
```

Out[34]:
```
Student_ID      SID20131143
Semster_Name          Sem_1
Paper_ID      SEM00012995
Paper_Name        Paper 1
Marks                  44
Name: 0, dtype: object
```

In [36]:
```
data.loc[[0,1]]
```

Out[36]:

	Student_ID	Semster_Name	Paper_ID	Paper_Name	Marks
0	SID20131143	Sem_1	SEM00012995	Paper 1	44
1	SID20131143	Sem_1	SEM00015183	Paper 2	74

In [37]:
```
print(pd.options.display.max_rows)
```

```
60
```

In [4]:
```
import pandas as pd
print(pd.options.display.max_columns)
```

```
20
```

In [39]:
```
pd.options.display.max_rows=10
print(data)
```

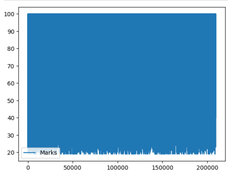
```
       Student_ID Semster_Name     Paper_ID Paper_Name  Marks
0      SID20131143        Sem_1  SEM00012995    Paper 1     44
1      SID20131143        Sem_1  SEM00015183    Paper 2     74
2      SID20131143        Sem_1  SEM00018371    Paper 3     80
3      SID20131143        Sem_1  SEM00015910    Paper 4     44
4      SID20131143        Sem_1  SEM00016208    Paper 5     95
...            ...          ...          ...        ...    ...
209606 SID20189989        Sem_8  SEM00082598    Paper 3     94
209607 SID20189989        Sem_8  SEM00088030    Paper 4     49
209608 SID20189989        Sem_8  SEM00081794    Paper 5     47
209609 SID20189989        Sem_8  SEM00086600    Paper 6     87
209610 SID20189989        Sem_8  SEM00083259    Paper 7     73

[209611 rows x 5 columns]
```

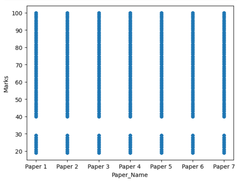
In [41]:
```
print(data.duplicated())
```

```
0         False
1         False
2         False
3         False
4         False
          ...  
209606    False
209607    False
209608    False
209609    False
209610    False
Length: 209611, dtype: bool
```

In [43]:
```
data.plot()
plt.show()
```

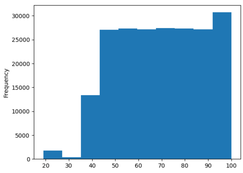


In [55]:
```
data.plot(kind='scatter',x='Paper_Name',y='Marks')
plt.show()
```

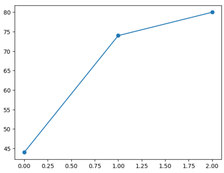


In [47]:
```
data['Marks'].plot(kind='hist')
```

Out[47]:
```
<Axes: ylabel='Frequency'>
```



In [51]:
```
xpoints=np.array([44,74,80])
plt.plot(xpoints,marker = 'o')
plt.show()
```



In [7]:
```
print(data)
```

```
       Student_ID Semster_Name     Paper_ID Paper_Name  Marks
0      SID20131143        Sem_1  SEM00012995    Paper 1     44
1      SID20131143        Sem_1  SEM00015183    Paper 2     74
2      SID20131143        Sem_1  SEM00018371    Paper 3     80
3      SID20131143        Sem_1  SEM00015910    Paper 4     44
4      SID20131143        Sem_1  SEM00016208    Paper 5     95
...            ...          ...          ...        ...    ...
209606 SID20189989        Sem_8  SEM00082598    Paper 3     94
209607 SID20189989        Sem_8  SEM00088030    Paper 4     49
209608 SID20189989        Sem_8  SEM00081794    Paper 5     47
209609 SID20189989        Sem_8  SEM00086600    Paper 6     87
209610 SID20189989        Sem_8  SEM00083259    Paper 7     73

[209611 rows x 5 columns]
```

In [8]:
```
pd.options.display.max_columns=2
print(data)
```

```
       Student_ID  ... Marks
0      SID20131143  ...    44
1      SID20131143  ...    74
2      SID20131143  ...    80
3      SID20131143  ...    44
4      SID20131143  ...    95
...            ...  ...   ...
209606 SID20189989  ...    94
209607 SID20189989  ...    49
209608 SID20189989  ...    47
209609 SID20189989  ...    87
209610 SID20189989  ...    73

[209611 rows x 5 columns]
```

In [9]:
```
pd.options.display.max_rows=2
print(data)
```

```
       Student_ID  ... Marks
0      SID20131143  ...    44
...            ...  ...   ...

[209611 rows x 5 columns]
```

```
...          ...  ...  ...
209610  SID20189089  ...   73

[209611 rows x 5 columns]
```