# Pandas Lab

## 1. Create a data series with marks of students : 75, 80, 79, 60

In [1]:
```
import pandas as pd
marks = pd.Series([75,80,79,60])
print(marks)
print(type(marks))
```

```
0    75
1    80
2    79
3    60
dtype: int64
<class 'pandas.core.series.Series'>
```

## 2. Create a data frame with name of students, id and marks

In [2]:
```
frame = {
    "students": ["Hassan", "Asad", "Zahid", "Saba", "Karima"],
    "id" : [101,201,301,401,501],
    "marks": [75,80,79,60,50]
}
df_std = pd.DataFrame(frame)
display(df_std.head())
type(df_std)
```

| | students | id | marks |
|---|---|---|---|
| **0** | Hassan | 101 | 75 |
| **1** | Asad | 201 | 80 |
| **2** | Zahid | 301 | 79 |
| **3** | Saba | 401 | 60 |
| **4** | Karima | 501 | 50 |

Out[2]:
```
pandas.core.frame.DataFrame
```

## 3. Now read the file 'data.csv' in panda

In [3]:
```
df = pd.read_csv("data for panda exercise.csv")
df.head()
```

Out[3]:

| | Student Code | Degree | Student Name | Mid | Quiz 1 | Quiz 2 | Best of Quizzes | Assignment 1 | Assignment 2 |
|---|---|---|---|---|---|---|---|---|---|
| **0** | 022-14-19987 | BS(CS) | Abdul Basit | 28 | 8.0 | 3.0 | 8 | 7.0 | 9.0 |
| **1** | 022-14-110233 | BS(CS) | Adeel Ahmed | 17 | NaN | 5.0 | 5 | 8.0 | 10.0 |
| **2** | 022-14-110585 | BS(CS) | Afrah Zareen | 18 | 5.0 | 2.0 | 5 | 8.0 | 10.0 |
| **3** | 022-14-19718 | BS(CS) | Ahmed Ali Raza | 14 | 7.0 | 2.0 | 7 | NaN | 2.0 |
| **4** | 022-14-110648 | BS(CS) | Ahsan Ali Vohra | 27 | 7.0 | 6.0 | 7 | 7.0 | 9.0 |

## 4. What are the columns in the dataframe?

In [4]:
```
df.columns
```

Out[4]:
```
Index(['Student Code', 'Degree', 'Student Name', 'Mid', 'Quiz 1',
'Quiz 2',
       'Best of Quizzes', 'Assignment 1', 'Assignment 2',
       'Best of Assignments', 'Total Sessional (50)', 'Final (5
0)',
       'Total (100)', 'Grade'],
      dtype='object')
```

In [5]:
```
df.isnull().sum()
```

Out[5]:
```
Student Code             0
Degree                   0
Student Name             0
Mid                      0
Quiz 1                   3
Quiz 2                   3
Best of Quizzes          0
Assignment 1             3
Assignment 2            12
Best of Assignments      0
Total Sessional (50)     0
Final (50)               1
Total (100)              0
Grade                    0
dtype: int64
```

## 5. Sort the data based on Marks obtained. Fill all the 'na' cells with 0

In [23]:
```
# Sort the DataFrame by Marks obtained in descending order
df.sort_values(by='Total Sessional (50)', ascending=False, inplace=T

# Fill all NaN values with 0
df.fillna(0, inplace=True)
df.head()
```

Out[23]:

| | Student Code | Degree | Student Name | Mid | Quiz 1 | Quiz 2 | Best of Quizzes | Assignment 1 | Assignme |
|---|---|---|---|---|---|---|---|---|---|
| **19** | 022-14-110222 | BS(CS) | Mohammad Hunain | 27 | 9.0 | 6.0 | 9 | 8.0 | 10 |
| **10** | 022-14-110214 | BS(CS) | Fatima Haider Warsi | 30 | 8.0 | 7.0 | 8 | 8.0 | 0 |
| **40** | 022-14-110584 | BS(CS) | Sharif Taqi | 27 | 9.0 | 6.0 | 9 | 8.0 | 10 |
| **0** | 022-14-19987 | BS(CS) | Abdul Basit | 28 | 8.0 | 3.0 | 8 | 7.0 | 9 |
| **45** | 022-14-110400 | BS(CS) | Syeda Sabika Raza | 27 | 9.0 | 6.0 | 9 | 9.0 | 0 |

## 6. Display the top 10 rows

In [24]:
```
df.head(10)
```

Out[24]:

| | Student Code | Degree | Student Name | Mid | Quiz 1 | Quiz 2 | Best of Quizzes | Assignment 1 | Assignme |
|---|---|---|---|---|---|---|---|---|---|
| **19** | 022-14-110222 | BS(CS) | Mohammad Hunain | 27 | 9.0 | 6.0 | 9 | 8.0 | 10 |
| **10** | 022-14-110214 | BS(CS) | Fatima Haider Warsi | 30 | 8.0 | 7.0 | 8 | 8.0 | 0 |
| **40** | 022-14-110584 | BS(CS) | Sharif Taqi | 27 | 9.0 | 6.0 | 9 | 8.0 | 10 |
| **0** | 022-14-19987 | BS(CS) | Abdul Basit | 28 | 8.0 | 3.0 | 8 | 7.0 | 9 |
| **45** | 022-14-110400 | BS(CS) | Syeda Sabika Raza | 27 | 9.0 | 6.0 | 9 | 9.0 | 0 |
| **43** | 022-14-110451 | BS(CS) | Syed Faizan Uddin | 28 | 9.0 | 4.0 | 9 | 6.0 | 8 |
| **37** | 022-14-110230 | BS(CS) | Sadaf Nosheen | 27 | 9.0 | 5.0 | 9 | 9.0 | 0 |
| **7** | 022-14-110388 | BS(CS) | Aneebullah Niazi | 26 | 9.0 | 6.0 | 9 | 8.0 | 10 |
| **27** | 022-14-110452 | BS(CS) | Muhammad Osama Khan | 27 | 8.0 | 6.0 | 8 | 8.0 | 10 |
| **25** | 022-14-110105 | BS(CS) | Muhammad Faraz | 26 | 9.0 | 6.0 | 9 | 8.0 | 10 |

## 7. Display the last 10 rows

In [25]:
```
df.tail(10)
```

Out[25]:

| | Student Code | Degree | Student Name | Mid | Quiz 1 | Quiz 2 | Best of Quizzes | Assignment 1 | Assignme |
|---|---|---|---|---|---|---|---|---|---|
| **22** | 022-14-110398 | BS(CS) | Muhammad Abdullah | 20 | 7.0 | 6.0 | 7 | 7.0 | 0 |
| **2** | 022-14-110585 | BS(CS) | Afrah Zareen | 18 | 5.0 | 2.0 | 5 | 8.0 | 10 |
| **1** | 022-14-110233 | BS(CS) | Adeel Ahmed | 17 | 0.0 | 5.0 | 5 | 8.0 | 10 |
| **18** | 022-14-110223 | BS(CS) | Madiha Jabeen | 16 | 5.0 | 2.0 | 5 | 8.0 | 10 |
| **29** | 022-14-110217 | BS(CS) | Muhammad Shahroz Khurshid | 17 | 7.0 | 1.0 | 7 | 6.0 | 0 |
| **47** | 022-14-110219 | BS(CS) | Waqar Ahmed | 11 | 9.0 | 5.0 | 9 | 5.0 | 7 |
| **3** | 022-14-19718 | BS(CS) | Ahmed Ali Raza | 14 | 7.0 | 2.0 | 7 | 0.0 | 2 |
| **38** | 022-14-110229 | BS(CS) | Sadaquat Rafique | 9 | 6.0 | 6.0 | 6 | 5.0 | 7 |
| **32** | 022-14-110035 | BS(CS) | Muhammad Wajahat Khan | 9 | 0.0 | 2.0 | 2 | 8.0 | 10 |
| **15** | 022-14-19916 | BS(CS) | Haseeb Sajid | 18 | 0.0 | 0.0 | 0 | 0.0 | 0 |

## 8. Display only the odd rows

In [26]:
```
# Display only the odd rows
df[df.index % 2 != 0]
```

Out[26]:

| | Student Code | Degree | Student Name | Mid | Quiz 1 | Quiz 2 | Best of Quizzes | Assignment 1 | Assignme |
|---|---|---|---|---|---|---|---|---|---|
| **19** | 022-14-110222 | BS(CS) | Mohammad Hunain | 27 | 9.0 | 6.0 | 9 | 8.0 | 10 |
| **45** | 022-14-110400 | BS(CS) | Syeda Sabika Raza | 27 | 9.0 | 6.0 | 9 | 9.0 | 0 |
| **43** | 022-14-110451 | BS(CS) | Syed Faizan Uddin | 28 | 9.0 | 4.0 | 9 | 6.0 | 8 |
| **37** | 022-14-110230 | BS(CS) | Sadaf Nosheen | 27 | 9.0 | 5.0 | 9 | 9.0 | 0 |
| **7** | 022-14-110388 | BS(CS) | Aneebullah Niazi | 26 | 9.0 | 6.0 | 9 | 8.0 | 10 |
| **27** | 022-14-110452 | BS(CS) | Muhammad Osama Khan | 27 | 8.0 | 6.0 | 8 | 8.0 | 10 |
| **25** | 022-14-110105 | BS(CS) | Muhammad Faraz | 26 | 9.0 | 6.0 | 9 | 8.0 | 10 |
| **9** | 022-14-110599 | BS(CS) | Arsalan | 28 | 8.0 | 6.0 | 8 | 8.0 | 0 |
| **5** | 022-14-110232 | BS(CS) | Ameer Hamza | 25 | 9.0 | 6.0 | 9 | 8.0 | 10 |
| **31** | 022-14-110231 | BS(CS) | Muhammad Taha Hasnain | 27 | 7.0 | 5.0 | 7 | 10.0 | 0 |
| **39** | 022-14-110107 | BS(CS) | Sania Iqbal | 28 | 4.0 | 4.0 | 4 | 9.0 | 11 |
| **13** | 022-14-110600 | BS(CS) | Hamza Abdul Jabbar | 24 | 8.0 | 4.0 | 8 | 8.0 | 10 |
| **23** | 022-14-19983 | BS(CS) | Muhammad Ali Iqbal | 26 | 5.0 | 0.0 | 5 | 8.0 | 10 |
| **11** | 022-14-110591 | BS(CS) | Habib Ullah | 28 | 8.0 | 5.0 | 8 | 5.0 | 0 |

| | Student Code | Degree | Student Name | Mid | Quiz 1 | Quiz 2 | Best of Quizzes | Assignment 1 | Assignme |
|---|---|---|---|---|---|---|---|---|---|
| **21** | 022-14-110593 | BS(CS) | Muhammad Abdul Rehman Siddiqui | 20 | 9.0 | 6.0 | 9 | 9.0 | 11 |
| **17** | 022-14-110396 | BS(CS) | Khalid Anwer | 20 | 8.0 | 5.0 | 8 | 9.0 | 11 |
| **35** | 022-14-19923 | BS(CS) | Rabi Ahmed | 20 | 8.0 | 5.0 | 8 | 0.0 | 10 |
| **41** | 022-14-110225 | BS(CS) | Shariqa Ahmad | 20 | 7.0 | 6.0 | 7 | 8.0 | 10 |
| **33** | 022-14-19919 | BS(CS) | Muhammad Younus Baig | 20 | 4.0 | 2.0 | 4 | 8.0 | 10 |
| **1** | 022-14-110233 | BS(CS) | Adeel Ahmed | 17 | 0.0 | 5.0 | 5 | 8.0 | 10 |
| **29** | 022-14-110217 | BS(CS) | Muhammad Shahroz Khurshid | 17 | 7.0 | 1.0 | 7 | 6.0 | 0 |
| **47** | 022-14-110219 | BS(CS) | Waqar Ahmed | 11 | 9.0 | 5.0 | 9 | 5.0 | 7 |
| **3** | 022-14-19718 | BS(CS) | Ahmed Ali Raza | 14 | 7.0 | 2.0 | 7 | 0.0 | 2 |
| **15** | 022-14-19916 | BS(CS) | Haseeb Sajid | 18 | 0.0 | 0.0 | 0 | 0.0 | 0 |

In [27]:
```
#We can display all even rows
df[df.index%2 == 0]
```

Out[27]:

| | Student Code | Degree | Student Name | Mid | Quiz 1 | Quiz 2 | Best of Quizzes | Assignment 1 | Assignme |
|---|---|---|---|---|---|---|---|---|---|
| **10** | 022-14-110214 | BS(CS) | Fatima Haider Warsi | 30 | 8.0 | 7.0 | 8 | 8.0 | 0 |
| **40** | 022-14-110584 | BS(CS) | Sharif Taqi | 27 | 9.0 | 6.0 | 9 | 8.0 | 10 |
| **0** | 022-14-19987 | BS(CS) | Abdul Basit | 28 | 8.0 | 3.0 | 8 | 7.0 | 9 |
| **6** | 022-14-110588 | BS(CS) | Anas Ali Khan | 28 | 5.0 | 6.0 | 6 | 8.0 | 10 |
| **34** | 022-14-110413 | BS(CS) | Nazeer Bin Zafar | 25 | 9.0 | 6.0 | 9 | 8.0 | 10 |
| **24** | 022-14-110215 | BS(CS) | Muhammad Bilal | 28 | 5.0 | 4.0 | 5 | 9.0 | 11 |
| **42** | 022-14-110587 | BS(CS) | Sumbul Rehman | 28 | 5.0 | 6.0 | 6 | 8.0 | 10 |
| **46** | 022-14-19911 | BS(CS) | Usman Khan | 25 | 8.0 | 5.0 | 8 | 8.0 | 10 |
| **4** | 022-14-110648 | BS(CS) | Ahsan Ali Vohra | 27 | 7.0 | 6.0 | 7 | 7.0 | 9 |
| **36** | 022-14-110582 | BS(CS) | Rida Nasim | 25 | 9.0 | 5.0 | 9 | 9.0 | 0 |
| **26** | 022-14-110370 | BS(CS) | Muhammad Ghazali Faridi | 27 | 6.0 | 0.0 | 6 | 6.0 | 8 |
| **44** | 022-14-110589 | BS(CS) | Syed Sohaib | 25 | 7.0 | 5.0 | 7 | 9.0 | 0 |
| **30** | 022-14-110401 | BS(CS) | Muhammad Shozab | 23 | 8.0 | 6.0 | 8 | 9.0 | 0 |
| **12** | 022-15-110994 | BS(CS) | Hafiza Tooba Akbani | 23 | 7.0 | 5.0 | 7 | 8.0 | 10 |
| **20** | 022-14-110412 | BS(CS) | Muhammad Aamir | 24 | 7.0 | 6.0 | 7 | 7.0 | 9 |

| | Student Code | Degree | Student Name | Mid | Quiz 1 | Quiz 2 | Best of Quizzes | Assignment 1 | Assignme |
|---|---|---|---|---|---|---|---|---|---|
| **28** | 022-14-110387 | BS(CS) | Muhammad Saqib Intizar | 22 | 6.0 | 6.0 | 6 | 9.0 | 11 |
| **16** | 022-14-110596 | BS(CS) | Hassam Ahmed | 23 | 5.0 | 5.0 | 5 | 7.0 | 9 |
| **8** | 022-14-110601 | BS(CS) | Areesha Sohail | 19 | 9.0 | 4.0 | 9 | 7.0 | 9 |
| **14** | 022-14-110389 | BS(CS) | Hareem Afshan | 21 | 7.0 | 4.0 | 7 | 6.0 | 8 |
| **22** | 022-14-110398 | BS(CS) | Muhammad Abdullah | 20 | 7.0 | 6.0 | 7 | 7.0 | 0 |
| **2** | 022-14-110585 | BS(CS) | Afrah Zareen | 18 | 5.0 | 2.0 | 5 | 8.0 | 10 |
| **18** | 022-14-110223 | BS(CS) | Madiha Jabeen | 16 | 5.0 | 2.0 | 5 | 8.0 | 10 |
| **38** | 022-14-110229 | BS(CS) | Sadaquat Rafique | 9 | 6.0 | 6.0 | 6 | 5.0 | 7 |
| **32** | 022-14-110035 | BS(CS) | Muhammad Wajahat Khan | 9 | 0.0 | 2.0 | 2 | 8.0 | 10 |

## 9. Display only those students who got failed in examination

In [28]:
```
df[df.Grade == "F"]
```

Out[28]:

| | Student Code | Degree | Student Name | Mid | Quiz 1 | Quiz 2 | Best of Quizzes | Assignment 1 | Assignme |
|---|---|---|---|---|---|---|---|---|---|
| **28** | 022-14-110387 | BS(CS) | Muhammad Saqib Intizar | 22 | 6.0 | 6.0 | 6 | 9.0 | 11 |
| **41** | 022-14-110225 | BS(CS) | Shariqa Ahmad | 20 | 7.0 | 6.0 | 7 | 8.0 | 10 |
| **33** | 022-14-19919 | BS(CS) | Muhammad Younus Baig | 20 | 4.0 | 2.0 | 4 | 8.0 | 10 |
| **22** | 022-14-110398 | BS(CS) | Muhammad Abdullah | 20 | 7.0 | 6.0 | 7 | 7.0 | 0 |
| **1** | 022-14-110233 | BS(CS) | Adeel Ahmed | 17 | 0.0 | 5.0 | 5 | 8.0 | 10 |
| **29** | 022-14-110217 | BS(CS) | Muhammad Shahroz Khurshid | 17 | 7.0 | 1.0 | 7 | 6.0 | 0 |
| **47** | 022-14-110219 | BS(CS) | Waqar Ahmed | 11 | 9.0 | 5.0 | 9 | 5.0 | 7 |
| **3** | 022-14-19718 | BS(CS) | Ahmed Ali Raza | 14 | 7.0 | 2.0 | 7 | 0.0 | 2 |
| **38** | 022-14-110229 | BS(CS) | Sadaquat Rafique | 9 | 6.0 | 6.0 | 6 | 5.0 | 7 |
| **32** | 022-14-110035 | BS(CS) | Muhammad Wajahat Khan | 9 | 0.0 | 2.0 | 2 | 8.0 | 10 |
| **15** | 022-14-19916 | BS(CS) | Haseeb Sajid | 18 | 0.0 | 0.0 | 0 | 0.0 | 0 |

## 10. Find out the basic statistical info about data

In [29]:
```
df.describe()
```

Out[29]:

| | Mid | Quiz 1 | Quiz 2 | Best of Quizzes | Assignment 1 | Assignment 2 | Assig |
|---|---|---|---|---|---|---|---|
| **count** | 48.000000 | 48.000000 | 48.000000 | 48.000000 | 48.000000 | 48.000000 | 48 |
| **mean** | 22.937500 | 6.833333 | 4.562500 | 7.020833 | 7.270833 | 7.083333 | 8 |
| **std** | 5.236558 | 2.336877 | 1.843981 | 1.973113 | 2.209646 | 4.365402 | 2 |
| **min** | 9.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0 |
| **25%** | 20.000000 | 5.750000 | 4.000000 | 6.000000 | 7.000000 | 1.500000 | 8 |
| **50%** | 25.000000 | 7.000000 | 5.000000 | 7.000000 | 8.000000 | 10.000000 | 10 |
| **75%** | 27.000000 | 9.000000 | 6.000000 | 9.000000 | 8.000000 | 10.000000 | 10 |
| **max** | 30.000000 | 9.000000 | 7.000000 | 9.000000 | 10.000000 | 11.000000 | 11 |

## 11. How many students got A, B, C, F?

In [30]:
```
grade_count = df['Grade'].value_counts(ascending=True)
grade_count
```

Out[30]:
```
A     8
F    11
C    13
B    16
Name: Grade, dtype: int64
```

## 12. What are the mean scores for students who got A, B, C, F?

In [31]:
```
grade_count_mean = df.groupby('Grade')['Total (100)'].mean()
```

In [32]:
```
grade_count_mean
```

Out[32]:
```
Grade
A    85.500000
B    73.187500
C    64.000000
F    44.454545
Name: Total (100), dtype: float64
```

# Done!

In [ ]:
```

```