Pandas

pandas is a data manipulation package in Python for tabular data. That is, data in the form of rows and columns, also known as DataFrames.

pandas is used throughout the data analysis workflow. With pandas, you can: -Import datasets from databases, spreadsheets, comma-separated values (CSV) files, and more. -Clean datasets, for example, by dealing with missing values. -Tidy datasets by reshaping their structure into a suitable format for analysis. -Aggregate data by calculating summary statistics such as the mean of columns, correlation between them, and more. -Visualize datasets and uncover insights.

• We can access the elements with loc(starts from 0) or iloc(starts from 1)

```
#importing pandas
import pandas as pd
import pandas as pd
a = ["Jwalitha", 'Ramya'
'Durga', 'Jahnavi', 'Lahari', 'Sunny', 'Dhanush']
r = pd.Series(a, index = [67,43,44,89,34,45,23])
print(r)
67
      Jwalitha
43
         Ramya
44
         Durga
89
       Jahnavi
34
        Lahari
45
         Sunnv
       Dhanush
23
dtype: object
df = pd.read csv("D:\Jwalitha\Datasets\diabetcsv.csv")
print(df)
     preq
            plas
                  pres
                         skin
                               insu
                                      mass
                                             pedi
                                                    age
                                                                    class
0
             148
                    72
                           35
                                  0
                                      33.6
                                            0.627
                                                     50
                                                         tested positive
        6
1
              85
                           29
                                  0
                                     26.6
                                            0.351
                                                         tested negative
        1
                    66
                                                     31
2
        8
             183
                    64
                            0
                                  0
                                      23.3
                                            0.672
                                                     32
                                                         tested positive
3
        1
              89
                    66
                           23
                                 94
                                      28.1
                                            0.167
                                                     21
                                                         tested negative
4
        0
             137
                    40
                           35
                                      43.1
                                            2.288
                                                     33
                                                         tested positive
                                168
763
       10
             101
                    76
                           48
                                180
                                      32.9
                                            0.171
                                                     63
                                                         tested negative
764
        2
             122
                    70
                           27
                                  0
                                      36.8
                                            0.340
                                                     27
                                                         tested negative
765
        5
             121
                    72
                           23
                                112
                                      26.2
                                            0.245
                                                     30
                                                         tested negative
        1
766
             126
                    60
                            0
                                  0
                                      30.1
                                            0.349
                                                     47
                                                         tested positive
        1
                    70
                           31
                                  0
767
              93
                                      30.4
                                            0.315
                                                     23
                                                         tested negative
[768 rows x 9 columns]
```

```
df = pd.read csv("D:\Jwalitha\Datasets\grades.txt", sep=" ")
print(df)
                       SEM1
                              SEM2
                                    SEM3 Grade
     Names Initials
0
                        9.8
                              10.0
                                      9.9
       Joe
                   K
                                             A+
1
                               9.1
                                     9.3
    Raiesh
                        8.9
                   М
                                              Α
2
                    ٧
                        9.9
                               9.3
                                      9.2
                                              Α
    Kissan
3
                    N
                        7.7
                               8.0
                                     7.1
                                              В
      Mary
4
                    K
                        9.8
                                      9.9
      Jeen
                               9.1
                                             A+
5
                                      9.3
                   М
                        8.9
                               9.1
       Raj
                                              Α
6
                   ٧
                        9.9
                                     9.2
    Hassan
                               9.0
                                              Α
7
      Mari
                    N
                        7.7
                               8.0
                                      7.1
                                              В
8
                    K
                        9.8
                                     9.9
      Jess
                               9.1
                                             Α+
9
                        7.0
                               9.1
                                      9.3
    Rajini
                   М
                                              Α
10
                   ٧
                        9.9
                               9.3
                                      9.2
                                              Α
     Kiran
                        7.7
                                      7.1
                                              В
11
      Maya
                   Ν
                               8.0
12
     Jolin
                   K
                        9.8
                               9.1
                                      9.9
                                             A+
13
                               9.1
                                     9.3
      Riya
                   М
                        8.0
                                              Α
14
      Sana
                   ٧
                        9.9
                               9.3
                                      9.2
                                              Α
15
      Mark
                   N
                        7.7
                               8.0
                                     7.0
                                              В
df = pd.read excel("D:\Jwalitha\Datasets\diabetes.xlsx")
print(df)
            plas
                         skin
                                insu
                                               pedi
                                                                      class
     preq
                  pres
                                       mass
                                                     age
0
        6
             148
                     72
                           35
                                       33.6
                                             0.627
                                                      50
                                                           tested positive
                                   0
1
        1
              85
                     66
                           29
                                   0
                                       26.6
                                             0.351
                                                      31
                                                           tested negative
2
             183
        8
                     64
                            0
                                   0
                                       23.3
                                             0.672
                                                      32
                                                           tested positive
3
                                             0.167
        1
              89
                     66
                           23
                                  94
                                       28.1
                                                      21
                                                           tested negative
4
        0
             137
                     40
                           35
                                 168
                                       43.1
                                             2.288
                                                      33
                                                           tested positive
             . . .
                    . . .
                           . . .
                                 . . .
                                                     . . .
       . . .
763
       10
             101
                     76
                           48
                                 180
                                       32.9
                                             0.171
                                                      63
                                                           tested negative
        2
             122
                     70
                           27
                                       36.8
                                             0.340
                                                      27
                                                           tested negative
764
                                   0
        5
765
             121
                     72
                           23
                                 112
                                       26.2
                                             0.245
                                                      30
                                                           tested negative
766
        1
             126
                     60
                            0
                                   0
                                       30.1
                                             0.349
                                                      47
                                                           tested positive
        1
767
              93
                     70
                           31
                                   0
                                       30.4
                                             0.315
                                                      23
                                                           tested negative
[768 rows x 9 columns]
df = pd.read excel("D:\Jwalitha\Datasets\diabetes.xlsx",sheet name=1)
print(df)
#also you can use : df = pd.read json("diabetes.json") for loading
ison files
  Dead Alive
  yes
           no
1
  yes
           no
2
  yes
           no
3
  yes
           no
  yes
           no
```

```
df = pd.read csv("D:\Jwalitha\Datasets\grades.txt", sep=" ")
print(df.head())
print(df.tail(n=10))
    Names Initials
                      SEM1
                             SEM2
                                    SEM3 Grade
                                     9.9
0
                       9.8
                             10.0
      Joe
                   K
                                             A+
                       8.9
1
   Raiesh
                   М
                              9.1
                                     9.3
                                              Α
2
   Kissan
                   ٧
                       9.9
                              9.3
                                     9.2
                                              Α
3
     Mary
                   N
                       7.7
                              8.0
                                     7.1
                                              В
4
                       9.8
                              9.1
                                     9.9
     Jeen
                   K
                                             A+
                       SEM1
                                     SEM3 Grade
     Names Initials
                              SEM2
6
                         9.9
                               9.0
                                      9.2
    Hassan
                    ٧
                                               Α
7
                    N
                         7.7
                                      7.1
                                               В
      Mari
                               8.0
8
                         9.8
                               9.1
                                      9.9
      Jess
                    K
                                              A+
9
                    М
                         7.0
                               9.1
                                      9.3
    Rajini
                                               Α
10
                    ٧
     Kiran
                         9.9
                               9.3
                                      9.2
                                               Α
11
                    N
                         7.7
                               8.0
                                      7.1
                                               В
      Maya
12
                    K
                                      9.9
     Jolin
                         9.8
                               9.1
                                              A+
13
                    М
                         8.0
                               9.1
                                      9.3
                                               Α
      Riva
14
                    ٧
                         9.9
                               9.3
                                      9.2
                                               Α
      Sana
15
                         7.7
                                      7.0
                                               В
      Mark
                    N
                               8.0
df = pd.read_csv("D:\Jwalitha\Datasets\diabetcsv.csv")
print(df.describe)
<bound method NDFrame.describe of</pre>
                                            preg plas pres
                                                                skin insu
                               class
mass
        pedi
             age
                     72
                            35
0
         6
             148
                                    0
                                       33.6
                                              0.627
                                                       50
                                                            tested positive
1
         1
              85
                     66
                            29
                                    0
                                       26.6
                                              0.351
                                                       31
                                                            tested negative
2
         8
             183
                     64
                             0
                                    0
                                       23.3
                                              0.672
                                                       32
                                                            tested positive
3
                                       28.1
                            23
         1
              89
                     66
                                   94
                                              0.167
                                                       21
                                                            tested negative
4
         0
             137
                     40
                            35
                                  168
                                       43.1
                                              2.288
                                                       33
                                                            tested_positive
                     . . .
                                  . . .
                                                            tested negative
763
       10
             101
                     76
                            48
                                  180
                                       32.9
                                              0.171
                                                       63
764
         2
             122
                     70
                            27
                                    0
                                       36.8
                                              0.340
                                                       27
                                                            tested negative
         5
                                                            tested negative
765
             121
                     72
                            23
                                  112
                                       26.2
                                              0.245
                                                       30
766
         1
             126
                     60
                             0
                                    0
                                       30.1
                                              0.349
                                                       47
                                                            tested positive
              93
                     70
                            31
767
         1
                                    0
                                       30.4
                                              0.315
                                                       23
                                                            tested negative
[768 \text{ rows } x \text{ 9 columns}] >
df = pd.read csv("D:\Jwalitha\Datasets\grades.txt", sep=" ")
print(df.describe().T)
      count
                              std
                                    min
                                            25%
                                                   50%
                                                           75%
                                                                  max
                  mean
                                                  9.35
SEM1
        16.0
              8.90000
                         1.077652
                                    7.0
                                         7.700
                                                        9.825
                                                                  9.9
SEM2
        16.0
                                    8.0
                                                        9.150
              8.91250
                         0.589774
                                         8.750
                                                  9.10
                                                                10.0
SEM3
                                                        9.450
        16.0
              8.86875
                         1.104970
                                    7.0
                                         8.675
                                                  9.25
                                                                  9.9
```

df.info(show_counts=True, memory_usage=True, verbose=True) => also used to view and understand data

```
df = pd.read csv("D:\Jwalitha\Datasets\grades.txt", sep=" ")
print(df.shape) # Get the number of rows and columns
print(df.shape[0]) # Get the number of rows only
print(df.shape[1]) # Get the number of columns only
(16, 6)
16
6
#creating a list with all column names
df = pd.read csv("D:\Jwalitha\Datasets\grades.txt", sep=" ")
print(df.columns)
l1 = list(df.columns)
print(l1)
Index(['Names', 'Initials', 'SEM1', 'SEM2', 'SEM3', 'Grade'],
dtype='object')
['Names', 'Initials', 'SEM1', 'SEM2', 'SEM3', 'Grade']
#making null values
df = pd.read csv("D:\Jwalitha\Datasets\grades.txt", sep=" ")
df2 = df.copv()
df2.loc[2:5, 'SEM1'] = None
df2.head(7)
    Names Initials SEM1
                         SEM2 SEM3 Grade
0
                    9.8
                         10.0
                                9.9
      Joe
                K
                                       A+
                    8.9
                          9.1
                                9.3
1
   Raiesh
                М
                                        Α
2
  Kissan
                ٧
                    NaN
                          9.3
                                9.2
                                        Α
3
                                7.1
     Mary
                N
                    NaN
                          8.0
                                        В
4
                K
                    NaN
                          9.1
                                9.9
                                       A+
     Jeen
5
                          9.1
                                9.3
      Rai
                М
                    NaN
                                        Α
                ٧
                    9.9
                          9.0
                                9.2
  Hassan
                                        Α
df2.isnull().head(7)
   Names
         Initials
                    SEM1
                         SEM2
                                  SEM3
                                        Grade
   False
            False False
                          False False
                                        False
1
   False
            False False
                          False False
   False
            False
                    True
                          False False False
3
   False
            False
                    True
                          False False False
4
   False
            False
                    True
                          False False False
5
   False
                          False False
            False
                    True
   False
            False False False False
print(df2.isnull().sum())
Names
           0
Initials
           0
SEM1
           4
```

```
SEM2
            0
SEM3
            0
Grade
            0
dtype: int64
print(df2.isnull().sum().sum())
4
#extracting column
print(df['SEM1'])
0
      9.8
1
      8.9
2
      9.9
3
      7.7
4
      9.8
5
      8.9
6
      9.9
7
      7.7
8
      9.8
9
      7.0
10
      9.9
11
      7.7
12
      9.8
13
      8.0
14
      9.9
15
      7.7
Name: SEM1, dtype: float64
print(df2[['SEM1','Names']])
    SEM1
           Names
0
     9.8
             Joe
     8.9
1
         Rajesh
2
          Kissan
     NaN
3
     NaN
            Mary
4
     NaN
            Jeen
5
     NaN
             Raj
6
         Hassan
     9.9
7
     7.7
            Mari
8
     9.8
            Jess
9
     7.0
         Rajini
10
     9.9
           Kiran
     7.7
11
           Maya
12
     9.8
           Jolin
13
     8.0
            Riya
14
     9.9
            Sana
15
     7.7
            Mark
```

```
#printing one row
print(df[df.index==1])
    Names Initials
                           SEM2
                                 SEM3 Grade
                    SEM1
1 Rajesh
                 М
                     8.9
                            9.1
                                  9.3
#printing multiple rows
print(df[df.index.isin(range(2,5))])
    Names Initials SEM1
                           SEM2
                                 SEM3 Grade
  Kissan
                     9.9
                            9.3
                                  9.2
                      7.7
                            8.0
                                  7.1
                                           В
3
     Mary
                 N
4
     Jeen
                 K
                     9.8
                            9.1
                                  9.9
                                         A+
#printing one row as series
print(df2.loc[1])
Names
            Rajesh
Initials
SEM1
               8.9
               9.1
SEM2
               9.3
SEM3
Grade
Name: 1, dtype: object
df2.loc[6:11]
     Names Initials
                     SEM1
                           SEM2
                                  SEM3 Grade
6
    Hassan
                  ٧
                      9.9
                             9.0
                                   9.2
                                            Α
7
                                            В
      Mari
                  N
                      7.7
                             8.0
                                   7.1
8
                  K
                       9.8
                             9.1
                                   9.9
      Jess
                                           A+
9
    Rajini
                  М
                       7.0
                             9.1
                                   9.3
                                           Α
10
                  ٧
                       9.9
                             9.3
                                   9.2
                                            Α
     Kiran
                       7.7
                                            В
11 Maya
                  Ν
                             8.0
                                   7.1
df2.loc[[10, 2, 7]]
     Names Initials
                            SEM2
                                  SEM3 Grade
                      SEM1
10
     Kiran
                       9.9
                             9.3
                                   9.2
                  ٧
                                            Α
                                            Α
    Kissan
                  ٧
                       NaN
                             9.3
                                   9.2
7
      Mari
                  N
                      7.7
                             8.0
                                   7.1
                                            В
df2.loc[10:15, ['Names', 'Grade',]]
    Names Grade
10
    Kiran
              Α
11
    Maya
              В
12
    Jolin
             A+
13
     Riya
              Α
14
     Sana
              Α
15
              В
     Mark
```

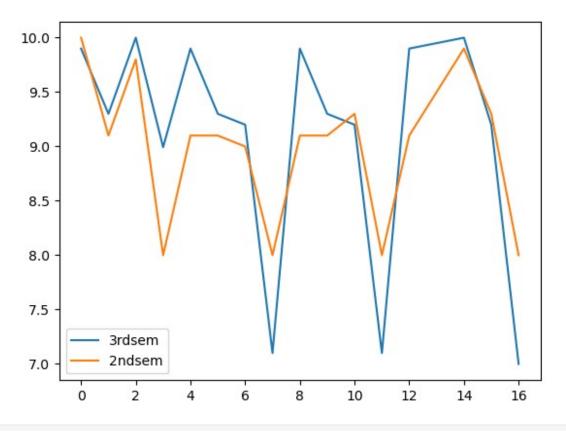
```
df2.iloc[10:14, :3] #[row,column]
   Names Initials SEM1
10
   Kiran
                    9.9
11
   Maya
                    7.7
12 Jolin
                Κ
                    9.8
13
                М
                    8.0
    Riya
df2.loc[:7, ["Names"]] #[row,column]
   Names
0
     Joe
1
  Rajesh
2
  Kissan
3
    Mary
4
    Jeen
5
     Raj
6
  Hassan
7
    Mari
df2.loc[df['Names']=='Rajesh']
                         SEM2 SEM3 Grade
   Names Initials SEM1
1 Rajesh M 8.9 9.1 9.3 A
df[df.Grade == 'A']
df[df.SEM2 == 10]
                       SEM2 SEM3 Grade
 Names Initials SEM1
0 Joe K 9.8 10.0 9.9 A+
df.loc[df['SEM1'] > 9.8, ['Names']] #printing names of students who
scored more than 9.8 in SEM1
    Names
2
   Kissan
6
   Hassan
10
    Kiran
14 Sana
df3 = df2.copy()
#df3 = df3.dropna()
df3
    Names Initials
                    SEM1 SEM2
                                SEM3 Grade
                          10.0
                                 9.9
0
      Joe
                 K
                     9.8
1
   Rajesh
                 М
                     8.9
                          9.1
                                 9.3
                                        Α
2
                 ٧
                           9.3
                                 9.2
                                        Α
   Kissan
                     NaN
3
     Marv
                 N
                     NaN
                           8.0
                                 7.1
                                        В
4
                 K
                     NaN
                           9.1
                                 9.9
                                       A+
     Jeen
5
      Raj
                 М
                     NaN
                           9.1
                                 9.3
                                        Α
6
                     9.9
                           9.0
                                 9.2
                                        Α
   Hassan
```

```
7
                               8.0
      Mari
                         7.7
                                      7.1
                                               В
                    N
8
                    K
                        9.8
                               9.1
                                      9.9
       Jess
                                              A+
9
    Rajini
                    М
                         7.0
                               9.1
                                      9.3
                                               Α
10
                                      9.2
     Kiran
                    ٧
                        9.9
                               9.3
                                               Α
11
      Maya
                    N
                        7.7
                               8.0
                                      7.1
                                               В
12
     Jolin
                    K
                        9.8
                               9.1
                                      9.9
                                              A+
13
                               9.1
                                      9.3
      Riya
                    М
                        8.0
                                               Α
14
       Sana
                    ٧
                         9.9
                               9.3
                                      9.2
                                               Α
15
                    N
                        7.7
                                      7.0
                                               В
      Mark
                               8.0
df3.dropna(inplace=True, axis=1) #inplace says not to reflect the
changes in the df
df3
     Names Initials
                       SEM2
                              SEM3 Grade
0
        Joe
                    K
                       10.0
                               9.9
                                       A+
    Rajesh
1
                               9.3
                    М
                        9.1
                                        Α
2
                        9.3
                               9.2
    Kissan
                    ٧
                                        Α
3
                    N
                        8.0
                               7.1
                                        В
      Mary
4
                    K
                         9.1
                               9.9
       Jeen
                                       A+
5
        Raj
                    М
                        9.1
                               9.3
                                        Α
6
                    ٧
                         9.0
                               9.2
                                        Α
    Hassan
7
                               7.1
      Mari
                    N
                        8.0
                                        В
8
      Jess
                    K
                        9.1
                               9.9
                                       A+
9
                        9.1
                               9.3
    Rajini
                    М
                                        Α
10
                    ٧
                        9.3
                               9.2
                                        Α
     Kiran
11
                        8.0
                               7.1
                                        В
      Maya
                    N
12
                    K
                               9.9
     Jolin
                        9.1
                                       A+
13
       Riya
                    М
                         9.1
                               9.3
                                        Α
                         9.3
14
                    ٧
                               9.2
      Sana
                                        Α
15
      Mark
                    N
                         8.0
                               7.0
                                        В
df3.dropna(inplace=True, how='all') #drops both rows and columns with
NaN
df3
     Names Initials
                       SEM2
                              SEM3 Grade
                               9.9
0
        Joe
                       10.0
                                       A+
                    K
1
    Raiesh
                    Μ
                        9.1
                               9.3
                                        Α
2
    Kissan
                    ٧
                        9.3
                               9.2
                                        Α
3
                         8.0
                               7.1
                                        В
      Mary
                    N
4
                               9.9
       Jeen
                    K
                        9.1
                                       A+
5
                    М
                        9.1
                               9.3
        Raj
                                        Α
6
    Hassan
                    ٧
                        9.0
                               9.2
                                        Α
7
                                        В
      Mari
                    N
                        8.0
                               7.1
8
                               9.9
       Jess
                    K
                        9.1
                                       A+
9
    Rajini
                    М
                        9.1
                               9.3
                                        Α
10
                    ٧
                         9.3
                               9.2
                                        Α
     Kiran
11
      Maya
                    N
                         8.0
                               7.1
                                        В
12
                    Κ
     Jolin
                         9.1
                               9.9
                                       A+
```

```
13
       Riva
                         9.1
                                 9.3
                                          Α
14
                         9.3
       Sana
                     ٧
                                 9.2
                                          Α
15
      Mark
                     N
                         8.0
                                 7.0
                                          В
dfn = pd.read csv("D:\Jwalitha\Datasets\grades withnulls.csv")
                        Sem1
                                                    Placed
     Names Initials
                               Sem2
                                      Sem3 Grade
0
        Joe
                     K
                         9.8
                               10.0
                                        9.9
                                                Α+
                                                           1
                                        9.3
                                                           1
1
    Rajesh
                     М
                         8.9
                                 9.1
                                                 Α
2
                                      10.0
    Kissan
                     V
                         9.9
                                 9.8
                                                 Α
                                                          0
3
                                                          0
                                                 В
      Mary
                     N
                         7.7
                                 8.0
                                       NaN
4
                                 9.1
                                        9.9
                                                           1
                     K
                         9.8
                                                Α+
       Jeen
5
                     М
                         8.9
                                 9.1
                                        9.3
                                                           1
        Raj
                                                 Α
6
                     ٧
                         9.9
                                 9.0
                                        9.2
                                                           1
    Hassan
                                                 Α
7
                                                           1
      Mari
                     N
                         7.7
                                 8.0
                                        7.1
                                                 В
                                                           1
8
       Jess
                     K
                         NaN
                                 9.1
                                        9.9
                                                Α+
9
                                 9.1
                                        9.3
                                                          0
    Rajini
                     М
                         NaN
                                                 Α
10
                                 9.3
                                        9.2
                                                          0
     Kiran
                     ٧
                         NaN
                                                 Α
11
                                 8.0
                                        7.1
                                                          0
      Maya
                     N
                         7.7
                                                 В
12
     Jolin
                     K
                         9.8
                                 9.1
                                        9.9
                                                Α+
                                                           1
                                                           1
13
                     М
                         8.9
                                 9.1
                                        9.3
    Rajesh
                                                 Α
                         9.3
                                                           1
14
       Riya
                     М
                                 9.9
                                      10.0
                                                 Α
                                9.3
                                                          0
15
                     ٧
                         9.9
                                        9.2
                                                 Α
       Sana
                     N
                         7.7
                                        7.0
                                                 В
16
      Mark
                                 8.0
                                                          0
mv = dfn['Sem1'].mean()
dfn = dfn.fillna(mv)
print(dfn)
      Names Initials
                                                 Sem3 Grade
                                                               Placed
                             Sem1
                                    Sem2
0
                        9.800000
                                    10.0
                                            9,900000
        Joe
                     K
                                                          A+
                                                                     1
1
                        8.900000
                                            9.300000
                                                           Α
                                                                     1
    Rajesh
                     М
                                     9.1
2
                                                                     0
    Kissan
                     ٧
                        9.900000
                                     9.8
                                           10.000000
                                                           Α
3
                     N
                        7,700000
                                     8.0
                                            8.992857
                                                           В
                                                                     0
      Mary
4
                     K
                                                                     1
       Jeen
                        9.800000
                                     9.1
                                            9.900000
                                                          A+
5
                        8.900000
                                     9.1
                                            9.300000
                                                           Α
                                                                     1
        Raj
                     М
6
                                                                     1
                     ٧
                                     9.0
                                                           Α
    Hassan
                        9.900000
                                            9.200000
7
      Mari
                        7.700000
                                     8.0
                                            7.100000
                                                            В
                                                                     1
                     N
8
                     K
                        8.992857
                                     9.1
                                            9.900000
                                                                     1
       Jess
                                                          A+
9
                        8.992857
                                     9.1
                                            9.300000
                                                                     0
    Rajini
                     Μ
                                                            Α
10
                                     9.3
                     ٧
                        8.992857
                                            9.200000
                                                           Α
                                                                     0
     Kiran
11
                                                            В
                                                                     0
      Maya
                     N
                        7.700000
                                     8.0
                                            7.100000
12
     Jolin
                     K
                        9.800000
                                     9.1
                                            9.900000
                                                                     1
                                                          Α+
                                                                     1
13
                                     9.1
                                            9.300000
    Rajesh
                     М
                        8.900000
                                                            Α
14
       Riya
                     Μ
                        9.300000
                                     9.9
                                           10.000000
                                                            Α
                                                                     1
15
                                     9.3
                                                                     0
       Sana
                        9.900000
                                            9.200000
                                                            Α
                        7.700000
                                            7.000000
                                                            В
                                                                     0
16
      Mark
                     N
                                     8.0
```

```
dfn = dfn.drop duplicates()
print(dfn)
                                             Sem3 Grade
                                                          Placed
     Names Initials
                           Sem1
                                 Sem2
0
                      9.800000
                                 10.0
                                         9.900000
       Joe
                   K
                                                      Α+
1
                                                                1
    Raiesh
                      8.900000
                                  9.1
                                         9.300000
                   М
                                                       Α
2
    Kissan
                      9.900000
                                  9.8
                                        10.000000
                                                       Α
                                                                0
                   V
3
                      7.700000
                                  8.0
                                                       В
                                                                0
      Mary
                   N
                                         8.992857
4
                   K
                      9.800000
      Jeen
                                  9.1
                                         9.900000
                                                      A+
                                                                1
5
                   М
                      8.900000
                                  9.1
                                         9.300000
                                                       Α
                                                                1
       Raj
6
                                                                1
    Hassan
                      9.900000
                                  9.0
                                         9.200000
                                                       Α
7
      Mari
                      7.700000
                                  8.0
                                         7.100000
                                                       В
                                                                1
                   N
8
                                                                1
      Jess
                   K
                      8.992857
                                  9.1
                                         9.900000
                                                      Α+
9
                   M
                      8.992857
                                  9.1
                                                                0
    Rajini
                                         9.300000
                                                       Α
10
                   ٧
                      8.992857
                                  9.3
                                         9.200000
                                                       Α
                                                                0
     Kiran
11
                      7.700000
                                         7.100000
                                                       В
                                                                0
      Maya
                   N
                                  8.0
12
     Jolin
                   K
                      9.800000
                                  9.1
                                         9.900000
                                                                1
                                                      A+
                                                                1
14
                                  9.9
      Riya
                   М
                      9.300000
                                        10.000000
                                                       Α
15
      Sana
                   ٧
                      9.900000
                                  9.3
                                         9.200000
                                                       Α
                                                                0
16
      Mark
                   N
                     7.700000
                                  8.0
                                         7.000000
                                                       В
                                                                0
dfn.rename(columns = {'Grade':'GPA'}, inplace = True)
dfn.head()
    Names Initials
                     Sem1
                            Sem2
                                        Sem3 GPA
                                                   Placed
0
                      9.8
                            10.0
                                    9.900000
                                              A+
      Joe
                  K
                                                        1
1
   Rajesh
                      8.9
                             9.1
                                   9.300000
                                               Α
                                                        1
                  М
2
                  ٧
                                                        0
   Kissan
                      9.9
                             9.8
                                  10.000000
                                               Α
3
     Mary
                  N
                      7.7
                             8.0
                                   8.992857
                                               В
                                                        0
4
     Jeen
                  K
                      9.8
                             9.1
                                   9.900000
                                              A+
                                                        1
dfn.columns = ['Student name', 'Mr/Ms', '1stsem', '2ndsem','3rdsem',
'F Outcome', 'Placed?']
dfn.head()
  Student name Mr/Ms
                        1stsem
                                2ndsem
                                            3rdsem F Outcome
                                                                Placed?
0
           Joe
                           9.8
                                   10.0
                                          9.900000
                                                                      1
                    K
                                                           Α+
                           8.9
                                                                      1
1
        Rajesh
                    М
                                    9.1
                                          9.300000
                                                            Α
2
                    ٧
                           9.9
                                                            Α
                                                                      0
        Kissan
                                    9.8
                                         10.000000
3
                           7.7
                                                                      0
          Mary
                    N
                                   8.0
                                          8.992857
                                                            В
                    K
                           9.8
                                          9.900000
                                                                      1
          Jeen
                                   9.1
                                                           Α+
dfn['Avg score'] = (dfn['1stsem']+dfn['2ndsem']+dfn['3rdsem'])/3
dfn.head()
                                            3rdsem F Outcome Placed?
  Student name Mr/Ms 1stsem 2ndsem
Avg_score
                                  10.0
                           9.8
            Joe
                    K
                                          9.900000
                                                           Α+
                                                                      1
9.900000
1
        Rajesh
                    Μ
                           8.9
                                   9.1
                                          9.300000
                                                            Α
9.100000
```

```
Kissan
                         9.9
                                 9.8 10.000000
                                                                 0
9.900000
3
          Mary
                         7.7
                                 8.0
                                       8.992857
                                                        В
                                                                 0
8.230952
                         9.8
          Jeen
                   K
                                 9.1 9.900000
                                                       Α+
                                                                 1
9.600000
dfn.groupby('Placed?').min()
        Student name Mr/Ms 1stsem 2ndsem 3rdsem F Outcome
Avg_score
Placed?
                               7.7
                                               7.0
               Kiran
                         M
                                       8.0
                                                           Α
7.566667
              Hassan
                         K
                               7.7
                                       8.0
                                               7.1
7.600000
dfn[['3rdsem', '2ndsem']].plot.line()
<Axes: >
```



#customize the color
#df[['BMI', 'Glucose']].plot.line(figsize=(20, 10), color={"BMI":
"red", "Glucose": "blue"})

```
dfn.plot.line(subplots=True)
array([<Axes: >, <Axes: >, <Axes: >, <Axes: >],
dtype=object)
```

