pandas-2

February 26, 2024

```
[]:
     import pandas as pd
[]: from google.colab import drive
     drive.mount('/content/drive')
    Drive already mounted at /content/drive; to attempt to forcibly remount, call
    drive.mount("/content/drive", force_remount=True).
[]: a=['jwalitha','ramya','durga','jahnvi','lahari','sunny','dhanush']
      r=pd.Series(a,index=[67,43,44,89,34,45,23])
     print(r)
    67
          jwalitha
    43
             ramya
    44
             durga
    89
             jahnvi
    34
             lahari
    45
             sunny
    23
           dhanush
    dtype: object
[]: df=pd.read_csv("/content/drive/MyDrive/DAP_TEAM 4/dataset/diabetcsv.csv")
     print(df)
                                  insu mass
               plas
                     pres
                            skin
                                                pedi
                                                      age
                                                                      class
    0
            6
                 148
                        72
                                         33.6
                                               0.627
                                                       50
                                                            tested_positive
                              35
                                      0
    1
             1
                 85
                        66
                              29
                                         26.6 0.351
                                                            tested_negative
                                      0
                                                       31
    2
            8
                                         23.3 0.672
                 183
                        64
                               0
                                      0
                                                       32
                                                            tested_positive
    3
                              23
             1
                 89
                        66
                                         28.1
                                               0.167
                                                            tested_negative
                                    94
                                                       21
    4
             0
                 137
                              35
                                         43.1
                        40
                                    168
                                               2.288
                                                       33
                                                            tested_positive
    . .
    763
           10
                 101
                        76
                              48
                                    180 32.9
                                               0.171
                                                       63
                                                           tested_negative
                                         36.8 0.340
    764
            2
                 122
                        70
                              27
                                      0
                                                       27
                                                            tested_negative
    765
            5
                 121
                        72
                              23
                                   112 26.2 0.245
                                                       30
                                                            tested_negative
    766
             1
                 126
                        60
                               0
                                      0
                                         30.1 0.349
                                                       47
                                                            tested positive
    767
             1
                  93
                                         30.4 0.315
                                                            tested_negative
                        70
                              31
                                                       23
```

[768 rows x 9 columns]

```
SEM2
     Names Initials
                        SEM1
                                      SEM3 Grade
                                                   Placed
0
        Joe
                    K
                         9.8
                               10.0
                                       9.9
                                               A+
1
    Rajesh
                    Μ
                         8.9
                                9.1
                                       9.3
                                                Α
                                                          1
                                                          0
2
    Kissan
                         9.9
                                9.8
                                      10.0
                                                Α
3
      Mary
                    N
                         7.7
                                8.0
                                       NaN
                                                В
                                                         0
4
       Jeen
                         9.8
                                9.1
                                       9.9
                    K
                                               A+
                                                          1
5
        Raj
                         8.9
                                9.1
                                       9.3
                                                Α
                                                          1
                    Μ
6
                    V
                         9.9
                                9.0
                                       9.2
                                                Α
                                                          1
    Hassan
7
      Mari
                         7.7
                                8.0
                                       7.1
                                                В
                                                          1
                    N
8
       Jess
                    K
                         {\tt NaN}
                                9.1
                                       9.9
                                               A+
                                                          1
                                                          0
9
    Rajini
                    Μ
                         {\tt NaN}
                                9.1
                                       9.3
                                                Α
10
     Kiran
                    ٧
                         {\tt NaN}
                                9.3
                                       9.2
                                                Α
                                                          0
      Maya
                         7.7
                                8.0
                                       7.1
                                                В
                                                          0
11
                    N
12
     Jolin
                    K
                         9.8
                                9.1
                                       9.9
                                               A+
                                                          1
13 Rajesh
                    Μ
                         8.9
                                9.1
                                       9.3
                                                Α
                                                          1
14
      Riya
                         9.3
                                9.9
                                      10.0
                                                Α
                                                          1
15
       Sana
                    ٧
                         9.9
                                9.3
                                       9.2
                                                Α
                                                         0
16
      Mark
                         7.7
                                8.0
                                       7.0
                                                В
                                                          0
```

```
[]: df=pd.read_excel("/content/drive/MyDrive/DAP_TEAM 4/datasets/grades_withnulls_u \( \text{\chi}(1).xlsx", sheet_name=0) \)
print(df)
```

	Names	Initials	SEM1	SEM2	SEM3	Grade	Placed
0	Joe	K	9.8	10.0	9.9	A+	1
1	Rajesh	M	8.9	9.1	9.3	Α	1
2	Kissan	V	9.9	9.8	10.0	Α	0
3	Mary	N	7.7	8.0	NaN	В	0
4	Jeen	K	9.8	9.1	9.9	A+	1
5	Raj	M	8.9	9.1	9.3	Α	1
6	Hassan	V	9.9	9.0	9.2	Α	1
7	Mari	N	7.7	8.0	7.1	В	1
8	Jess	K	NaN	9.1	9.9	A+	1
9	Rajini	M	NaN	9.1	9.3	Α	0
10	Kiran	V	NaN	9.3	9.2	Α	0
11	Maya	N	7.7	8.0	7.1	В	0
12	Jolin	K	9.8	9.1	9.9	A+	1
13	Rajesh	M	8.9	9.1	9.3	Α	1
14	Riya	М	9.3	9.9	10.0	Α	1
15	Sana	V	9.9	9.3	9.2	Α	0
16	Mark	N	7.7	8.0	7.0	В	0

```
[]: df=pd.read_excel("/content/drive/MyDrive/DAP_TEAM 4/datasets/diabetcsv.
      →xlsx", sheet_name=0)
     print(df)
                                                                        class
                plas
                      pres
                             skin
                                   insu
                                          mass
                                                 pedi
                                                        age
    0
                                          33.6
                                                             tested positive
             6
                 148
                         72
                               35
                                                0.627
                                                         50
    1
             1
                  85
                         66
                               29
                                          26.6
                                                0.351
                                                         31
                                                             tested_negative
    2
             8
                 183
                                0
                                          23.3 0.672
                                                         32
                                                             tested_positive
                         64
                                      0
    3
             1
                  89
                         66
                               23
                                     94
                                          28.1
                                                0.167
                                                         21
                                                             tested_negative
    4
             0
                 137
                         40
                               35
                                    168 43.1
                                                2.288
                                                         33
                                                             tested_positive
    763
                                          32.9
                                              0.171
            10
                 101
                        76
                               48
                                    180
                                                         63
                                                             tested_negative
    764
                                          36.8 0.340
             2
                 122
                         70
                               27
                                      0
                                                         27
                                                             tested_negative
    765
             5
                 121
                         72
                               23
                                    112
                                          26.2
                                                0.245
                                                         30
                                                             tested_negative
    766
             1
                 126
                                0
                                      0
                                          30.1 0.349
                                                         47
                                                             tested_positive
                         60
    767
                  93
                                          30.4 0.315
             1
                         70
                               31
                                                         23
                                                             tested_negative
    [768 rows x 9 columns]
[]: df=pd.read_csv("/content/drive/MyDrive/DAP_TEAM_4/datasets/grades_withnulls.
      ocsv")
     print(df.head())
     print(df.tail())
        Names Initials
                         SEM1
                                SEM2
                                      SEM3 Grade
                                                   Placed
    0
           Joe
                      K
                           9.8
                                10.0
                                        9.9
                                               A+
                                                         1
    1
       Rajesh
                      Μ
                           8.9
                                 9.1
                                        9.3
                                                Α
                                                         1
    2
                                 9.8 10.0
       Kissan
                      V
                           9.9
                                                Α
                                                         0
    3
         Mary
                      N
                           7.7
                                 8.0
                                        NaN
                                                В
                                                         0
    4
                      K
                                        9.9
                                                         1
         Jeen
                           9.8
                                 9.1
                                               A+
                           SEM1
                                 SEM2
                                       SEM3 Grade
         Names Initials
                                                    Placed
                                  9.1
                                         9.9
    12
         Jolin
                            9.8
                                                A+
                       K
                            8.9
                                  9.1
                                         9.3
    13
        Rajesh
                                                 Α
                                                          1
    14
           Riya
                       Μ
                            9.3
                                  9.9
                                        10.0
                                                 Α
                                                          1
    15
                       V
                                         9.2
           Sana
                            9.9
                                  9.3
                                                 Α
                                                          0
    16
           Mark
                       N
                            7.7
                                  8.0
                                         7.0
                                                 В
                                                          0
[]: df=pd.read_csv("/content/drive/MyDrive/DAP_TEAM 4/datasets/diabetcsv.csv")
     print(df.describe)
    <bound method NDFrame.describe of</pre>
                                              preg plas pres skin insu mass
                                                                                     pedi
    age
                    class
    0
             6
                 148
                        72
                               35
                                       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
             1
                  89
                         66
                               23
                                          28.1 0.167
                                                         21
                                                             tested_negative
                                     94
    4
             0
                 137
                               35
                                         43.1
                                                2.288
                                                             tested_positive
                         40
                                    168
                                                         33
```

```
764
            2
                122
                             27
                                    0 36.8 0.340
                                                          tested_negative
                       70
                                                      27
            5
                                  112 26.2 0.245
    765
                121
                       72
                             23
                                                      30
                                                          tested_negative
    766
            1
                126
                       60
                              0
                                    0 30.1 0.349
                                                      47
                                                          tested_positive
    767
            1
                 93
                             31
                                       30.4 0.315
                                                          tested negative
                       70
                                                      23
    [768 rows x 9 columns]>
[]: df=pd.read_csv("/content/drive/MyDrive/DAP_TEAM 4/datasets/diabetcsv.csv")
     print(df.describe().T)
                                                                 50%
                                                                            75%
          count
                                             min
                                                       25%
                                    std
                       mean
    preg 768.0
                   3.845052
                               3.369578
                                           0.000
                                                   1.00000
                                                              3.0000
                                                                        6.00000
    plas 768.0
                 120.894531
                              31.972618
                                           0.000
                                                  99.00000
                                                            117.0000
                                                                      140.25000
    pres
         768.0
                  69.105469
                              19.355807
                                           0.000
                                                  62.00000
                                                             72.0000
                                                                       80.00000
         768.0
                  20.536458
                                           0.000
                                                   0.00000
                                                             23.0000
                                                                       32.00000
    skin
                              15.952218
    insu 768.0
                                           0.000
                                                             30.5000 127.25000
                  79.799479 115.244002
                                                   0.00000
    mass
          768.0
                  31.992578
                               7.884160
                                           0.000
                                                 27.30000
                                                             32.0000
                                                                       36.60000
          768.0
                                           0.078
                                                              0.3725
    pedi
                   0.471876
                               0.331329
                                                   0.24375
                                                                        0.62625
          768.0
                  33.240885
                              11.760232 21.000
                                                 24.00000
                                                             29.0000
                                                                       41.00000
    age
             max
           17.00
    preg
    plas 199.00
    pres 122.00
    skin
           99.00
    insu 846.00
           67.10
    mass
            2.42
    pedi
           81.00
    age
[]: df=pd.read_csv("/content/drive/MyDrive/DAP_TEAM_4/datasets/diabetcsv.csv")
     print(df.shape)
     print(df.shape[0])
     print(df.shape[1])
    (768, 9)
    768
    9
[]: df=pd.read_csv("/content/drive/MyDrive/DAP_TEAM 4/datasets/diabetcsv.csv")
     print(df.columns)
    Index(['preg', 'plas', 'pres', 'skin', 'insu', 'mass', 'pedi', 'age', 'class'],
    dtype='object')
[]: df2=df.copy()
     df2.loc[2:5, 'SEM1']=None
```

763

10

101

76

48

180 32.9 0.171

63

tested_negative

```
df2.head(7)
[]:
         Names Initials
                         SEM1
                               SEM2
                                     SEM3 Grade
                                                 Placed
           Joe
                      K
                          9.8
                               10.0
                                      9.9
       Rajesh
                          8.9
                                9.1
                                      9.3
                                                      1
     1
                      Μ
                                              Α
                                9.8
                                     10.0
                                                      0
     2
       Kissan
                      V
                          NaN
                                              Α
     3
         Mary
                          NaN
                                8.0
                                      NaN
                                              В
                                                      0
                      N
     4
          Jeen
                                9.1
                                      9.9
                                             A+
                                                      1
                      K
                          NaN
     5
                                9.1
                                      9.3
                                                      1
           Raj
                      Μ
                          NaN
                                              Α
       Hassan
                          9.9
                                9.0
                                      9.2
                                                      1
                                              Α
[]: df2.isnull().head(7)
[]:
       Names
              Initials
                          SEM1
                                 SEM2
                                        SEM3
                                              Grade
                                                     Placed
     0 False
                  False
                         False
                                False False
                                              False
                                                      False
     1 False
                  False
                         False
                                False
                                       False
                                              False
                                                      False
     2 False
                                False False
                  False
                          True
                                              False
                                                      False
     3 False
                  False
                          True
                                False
                                        True
                                              False
                                                      False
     4 False
                  False
                          True
                                False False
                                              False
                                                      False
                                False False
     5 False
                  False
                          True
                                              False
                                                      False
     6 False
                  False False False
                                              False
                                                      False
[]: df.isnull().tail(7)
[]:
         Names
                Initials
                           SEM1
                                  SEM2
                                         SEM3
                                               Grade
                                                      Placed
     10 False
                   False
                           True
                                False
                                               False
                                                       False
                                        False
     11 False
                                 False
                   False
                         False
                                        False
                                               False
                                                       False
     12 False
                   False
                         False
                                False
                                        False
                                              False
                                                       False
                         False False
     13 False
                  False
                                        False False
                                                       False
     14 False
                  False False False False
                                                       False
                                                       False
     15 False
                  False False
                               False False False
                                                       False
     16 False
                  False False False False
    print(df2.isnull().sum())
    Names
                0
    Initials
                0
    SEM1
                7
    SEM2
                0
    SEM3
                1
    Grade
    Placed
    dtype: int64
[]: print(df2['SEM1'])
    0
          9.8
    1
          8.9
```

```
2
           NaN
     3
           {\tt NaN}
     4
           {\tt NaN}
     5
           {\tt NaN}
     6
           9.9
    7
           7.7
    8
           NaN
     9
           NaN
     10
           NaN
     11
           7.7
           9.8
     12
     13
           8.9
     14
           9.3
           9.9
     15
     16
            7.7
     Name: SEM1, dtype: float64
[]: print(df2[['SEM1','Names']])
         SEM1
                 Names
     0
          9.8
                    Joe
     1
          8.9
                Rajesh
     2
          NaN
                Kissan
     3
          {\tt NaN}
                  Mary
     4
          NaN
                  Jeen
     5
          {\tt NaN}
                   Raj
     6
          9.9
               Hassan
     7
          7.7
                  Mari
     8
          {\tt NaN}
                  Jess
     9
          NaN Rajini
     10
          {\tt NaN}
                 Kiran
          7.7
     11
                  Maya
     12
          9.8
                 Jolin
     13
          8.9
                Rajesh
     14
          9.3
                  Riya
     15
          9.9
                  Sana
          7.7
     16
                  Mark
[]: print(df2[df2.index==1])
         Names Initials
                           SEM1
                                  SEM2
                                         SEM3 Grade
                                                      Placed
     1 Rajesh
                             8.9
                                    9.1
                                           9.3
[]: print(df2[df2.index.isin(range(2,6))])
                                  SEM2
                                         SEM3 Grade Placed
         Names Initials SEM1
     2 Kissan
                        V
                             NaN
                                    9.8
                                         10.0
                                                    Α
                                                             0
     3
          Mary
                        N
                             NaN
                                    8.0
                                          {\tt NaN}
                                                    В
                                                             0
```

```
5
           Raj
                           NaN
                                 9.1
                                        9.3
                                                Α
                                                         1
                      М
[]: print(df2.loc[3])
    Names
                 Mary
    Initials
                    N
    SEM1
                  NaN
    SEM2
                  8.0
    SEM3
                  NaN
    Grade
                    В
                    0
    Placed
    Name: 3, dtype: object
[]: print(df2.loc[6:11])
         Names Initials
                           SEM1
                                 SEM2
                                       SEM3 Grade
                                                    Placed
                            9.9
                                  9.0
    6
        Hassan
                                         9.2
                                                 Α
                                                          1
    7
          Mari
                       N
                            7.7
                                  8.0
                                         7.1
                                                 В
                                                          1
    8
                                  9.1
                                         9.9
           Jess
                       K
                            NaN
                                                A+
                                                          1
    9
                                  9.1
                                         9.3
                                                          0
        Rajini
                       Μ
                            NaN
                                                 Α
         Kiran
                                  9.3
                                                          0
    10
                            NaN
                                         9.2
                                                 Α
    11
          Maya
                            7.7
                                  8.0
                                         7.1
                                                 В
                                                          0
[]: print(df2.loc[3]==None)
    Names
                 False
    Initials
                 False
    SEM1
                 False
    SEM2
                 False
    SEM3
                 False
    Grade
                 False
    Placed
                 False
    Name: 3, dtype: bool
[]: df2.loc[[10,2,7]]
                                        SEM3 Grade Placed
[]:
          Names Initials
                           SEM1
                                  SEM2
                                         9.2
     10
          Kiran
                            NaN
                                   9.3
                                                           0
     2
         Kissan
                            NaN
                                   9.8
                                        10.0
                                                           0
                                                  Α
     7
           Mari
                        N
                            7.7
                                   8.0
                                         7.1
                                                  В
                                                           1
[]: df2.loc[[1,2,3]]
[]:
         Names Initials
                          SEM1
                                SEM2
                                       SEM3 Grade
                                                   Placed
                                                          1
     1 Rajesh
                           8.9
                                  9.1
                                        9.3
                                                 Α
     2
        Kissan
                                      10.0
                                                         0
                       V
                           NaN
                                  9.8
                                                 Α
                                                 В
     3
          Mary
                           NaN
                                  8.0
                                        NaN
                                                         0
```

4

Jeen

K

NaN

9.1

9.9

A+

1

```
[]: df2.loc[10:15,['Names','Grade',]]
[]:
          Names Grade
     10
          Kiran
     11
                     В
           Maya
     12
           Jolin
                    Α+
     13
         Rajesh
                     Α
     14
           Riya
                     Α
     15
           Sana
                     Α
[]: df2.iloc[10:14,:3]
[]:
                            SEM1
          Names Initials
     10
          Kiran
                             NaN
     11
           Maya
                        N
                             7.7
                        K
     12
          Jolin
                             9.8
     13 Rajesh
                        М
                             8.9
[]: df2.loc[:7,["Names"]]
[]:
         Names
            Joe
     0
     1
        Rajesh
        Kissan
     2
     3
          Mary
     4
          Jeen
     5
           Raj
     6
       Hassan
     7
          Mari
[]: df2.loc[df2['Names']=='Rajesh']
[]:
                                         SEM3 Grade
          Names Initials
                            SEM1
                                   SEM2
                                                     Placed
                             8.9
                                          9.3
     1
         Rajesh
                        Μ
                                    9.1
                                                   Α
                                                            1
                                          9.3
     13 Rajesh
                        М
                             8.9
                                    9.1
                                                   Α
                                                            1
[]: df2.loc[df2['Grade']=='A']
[]:
          Names Initials
                            SEM1
                                   SEM2
                                         SEM3 Grade
                                                      Placed
     1
                             8.9
                                    9.1
                                          9.3
                                                   Α
         Rajesh
                                                            1
     2
         Kissan
                        V
                             NaN
                                    9.8
                                         10.0
                                                   Α
                                                            0
     5
            Raj
                        Μ
                             NaN
                                    9.1
                                          9.3
                                                   Α
                                                            1
     6
         Hassan
                        V
                             9.9
                                    9.0
                                          9.2
                                                   Α
                                                            1
     9
         Rajini
                        М
                             {\tt NaN}
                                    9.1
                                          9.3
                                                   Α
                                                            0
                        V
                                    9.3
                                          9.2
                                                            0
     10
          Kiran
                             {\tt NaN}
                                                   Α
     13
         Rajesh
                        М
                             8.9
                                    9.1
                                          9.3
                                                   Α
                                                            1
     14
           Riya
                             9.3
                                    9.9
                                         10.0
                                                   Α
                                                            1
                        Μ
```

```
15
           Sana
                            9.9
                                   9.3
                                         9.2
                                                 Α
                                                          0
[]: df2[df2.SEM2==9]
[]:
         Names Initials SEM1
                                SEM2
                                       SEM3 Grade Placed
     6 Hassan
                           9.9
                                  9.0
                                        9.2
[]: df2.loc[df2['SEM1']>9,['Names']]
          Names
[]:
     0
            Joe
     6
         Hassan
     12
          Jolin
     14
           Riya
     15
           Sana
[]: df3=df2.copy()
     df3.dropna(inplace=True,axis=1)
     df3
                           SEM2 Grade
[]:
          Names Initials
                                        Placed
                           10.0
     0
            Joe
                        K
                                    A+
                                              1
     1
         Rajesh
                            9.1
                                             1
                                     Α
         Kissan
                            9.8
                                             0
     2
                                     Α
     3
           Mary
                        N
                            8.0
                                     В
                                             0
     4
           Jeen
                        K
                            9.1
                                    A+
                                             1
     5
                            9.1
                                     Α
            Raj
                        Μ
                                              1
     6
                            9.0
         Hassan
                        V
                                     Α
                                             1
     7
                            8.0
                                     В
           Mari
                        N
                                              1
     8
           Jess
                        K
                            9.1
                                    A+
                                             1
         Rajini
                            9.1
                                             0
     9
                        Μ
                                     Α
                            9.3
     10
          Kiran
                        V
                                     Α
                                             0
     11
                            8.0
                                     В
                                             0
           Maya
                        N
     12
          Jolin
                        K
                            9.1
                                    A+
                                             1
     13
        Rajesh
                        М
                            9.1
                                     Α
                                             1
                                     Α
     14
           Riya
                            9.9
                                              1
                        Μ
     15
           Sana
                        ٧
                                             0
                            9.3
                                     Α
     16
           Mark
                        N
                            8.0
                                     В
                                             0
[]: df3.dropna(inplace=True,how='all')
     df3
[]:
          Names Initials
                           SEM2 Grade
                                        Placed
            Joe
                           10.0
                                    Α+
                                             1
     0
                        K
     1
         Rajesh
                        М
                            9.1
                                     Α
                                             1
     2
                            9.8
                                     Α
                                             0
         Kissan
                        ٧
     3
                            8.0
                                     В
                                             0
           Mary
```

```
4
            Jeen
                             9.1
                                     A+
                         K
                                               1
     5
            Raj
                             9.1
                                      Α
                                               1
                         М
     6
                             9.0
         Hassan
                         V
                                      Α
                                               1
     7
                             8.0
           Mari
                         N
                                      В
                                               1
     8
            Jess
                         K
                             9.1
                                     A+
                                               1
     9
         Rajini
                             9.1
                                      Α
                                               0
                         М
     10
                             9.3
                                      Α
                                               0
          Kiran
                         ٧
     11
           Maya
                         N
                             8.0
                                      В
                                               0
     12
          Jolin
                             9.1
                                     A+
                                               1
                         K
     13
         Rajesh
                         М
                             9.1
                                      Α
                                               1
     14
           Riya
                             9.9
                                      Α
                         М
                                               1
     15
           Sana
                         V
                             9.3
                                      Α
                                               0
     16
           Mark
                                               0
                         N
                             8.0
                                      В
[]: df3['Names']
[]: 0
               Joe
     1
           Rajesh
     2
           Kissan
     3
              Mary
     4
              Jeen
     5
               Raj
     6
           Hassan
     7
              Mari
     8
              Jess
     9
           Rajini
     10
            Kiran
     11
              Maya
     12
             Jolin
     13
           Rajesh
     14
              Riya
     15
              Sana
     16
              Mark
     Name: Names, dtype: object
[]: mv = df2['SEM1'].mean()
     df2 = df2.fillna(mv)
     print(df2)
          Names Initials
                           SEM1
                                  SEM2
                                          SEM3 Grade
                                                      Placed
                           9.80
                                          9.90
    0
            Joe
                        K
                                  10.0
                                                  A+
                                                             1
    1
        Rajesh
                        M
                          8.90
                                   9.1
                                          9.30
                                                    Α
                                                            1
    2
        Kissan
                        V
                           8.96
                                   9.8
                                         10.00
                                                            0
                                                    Α
    3
                        N
                           8.96
                                   8.0
                                          8.96
                                                   В
                                                            0
           Mary
    4
           Jeen
                        K 8.96
                                   9.1
                                          9.90
                                                  A+
                                                            1
    5
            Raj
                        M 8.96
                                   9.1
                                          9.30
                                                    Α
                                                             1
```

Α

9.20

6

Hassan

9.90

9.0

```
7
                        N 7.70
           Mari
                                   8.0
                                          7.10
                                                    В
                                                             1
    8
           Jess
                        K
                           8.96
                                   9.1
                                          9.90
                                                             1
                                                   A+
    9
                           8.96
                                          9.30
                                                             0
         Rajini
                                   9.1
                                                    Α
                        М
    10
          Kiran
                        V
                           8.96
                                   9.3
                                          9.20
                                                    Α
                                                             0
                            7.70
           Maya
                                   8.0
                                          7.10
                                                    В
                                                             0
    11
                        N
    12
          Jolin
                        K
                            9.80
                                   9.1
                                          9.90
                                                   A+
                                                             1
    13
         Rajesh
                           8.90
                                   9.1
                                          9.30
                                                    Α
                                                             1
                            9.30
    14
           Riya
                        М
                                   9.9
                                         10.00
                                                    Α
                                                             1
    15
           Sana
                            9.90
                                   9.3
                                          9.20
                                                             0
                                                    Α
                           7.70
    16
           Mark
                                   8.0
                                          7.00
                                                             0
                        N
                                                    В
[]: df2 = df2.drop_duplicates()
     print(df2)
          Names Initials
                           SEM1
                                  SEM2
                                          SEM3 Grade
                                                       Placed
    0
            Joe
                        K
                           9.80
                                  10.0
                                          9.90
                                                   Α+
                                                             1
                           8.90
                                          9.30
    1
         Rajesh
                        M
                                   9.1
                                                             1
                                                    Α
    2
         Kissan
                        ٧
                           8.96
                                   9.8
                                         10.00
                                                    Α
                                                             0
    3
                            8.96
                                          8.96
                                                             0
                        N
                                   8.0
                                                    В
           Mary
    4
           Jeen
                        K
                           8.96
                                   9.1
                                          9.90
                                                   A+
                                                             1
    5
                           8.96
                                          9.30
            Raj
                                   9.1
                                                    Α
                                                             1
    6
         Hassan
                        V
                           9.90
                                   9.0
                                          9.20
                                                    Α
                                                             1
    7
           Mari
                        N
                            7.70
                                   8.0
                                          7.10
                                                    В
                                                             1
    8
                           8.96
                                          9.90
           Jess
                        K
                                   9.1
                                                   A+
                                                             1
    9
         Rajini
                           8.96
                                   9.1
                                          9.30
                                                    Α
                                                             0
                        Μ
    10
          Kiran
                        V
                           8.96
                                   9.3
                                          9.20
                                                    Α
                                                             0
                           7.70
                                                             0
    11
           Maya
                        N
                                   8.0
                                          7.10
                                                    В
    12
          Jolin
                           9.80
                                   9.1
                                          9.90
                                                             1
                        K
                                                   A+
    14
                           9.30
                                   9.9
                                         10.00
                                                             1
           Riya
                                                    Α
                            9.90
    15
           Sana
                        V
                                   9.3
                                          9.20
                                                    Α
                                                             0
    16
           Mark
                        N
                           7.70
                                   8.0
                                          7.00
                                                    В
[]: df2.rename(columns = {'Grade': 'GPA'}, inplace = True)
     df2.head()
[]:
         Names Initials
                           SEM1
                                  SEM2
                                          SEM3 GPA
                                                    Placed
            Joe
                           9.80
                                  10.0
                                          9.90
                                                A+
                        K
                        M 8.90
     1
        Rajesh
                                   9.1
                                          9.30
                                                 Α
                                                          1
     2
                           8.96
                                                          0
        Kissan
                        V
                                   9.8
                                        10.00
                                                 Α
     3
                                                          0
                           8.96
                                   8.0
                                          8.96
                                                 В
           Mary
                        N
     4
           Jeen
                           8.96
                                   9.1
                                          9.90
                                                          1
[]: df2['Avg_score'] = (df2['1stsem']+df2['2ndsem']+df2['3rdsem'])/3
     df2.head()
[]:
       Student_name Mr/Ms
                             1stsem
                                      2ndsem
                                               3rdsem F_Outcome
                                                                   Placed?
                                                                             Avg_score
     0
                          K
                                 9.8
                                         10.0
                                                   9.9
                                                               A+
                                                                          1
                                                                                    9.9
                 Joe
     1
                                 8.9
                                          9.1
                                                   9.3
                                                                          1
              Rajesh
                          М
                                                                Α
                                                                                    9.1
```

```
2
          Kissan
                         V
                                           9.8
                                                    10.0
                                                                                             {\tt NaN}
                                 NaN
                                                                     Α
3
             Mary
                         N
                                 NaN
                                           8.0
                                                      NaN
                                                                     В
                                                                                 0
                                                                                             NaN
4
             Jeen
                         K
                                 {\tt NaN}
                                           9.1
                                                      9.9
                                                                                             {\tt NaN}
                                                                    A+
```

[]: df2.groupby('Placed?').min()

```
Student_name Mr/Ms 1stsem 2ndsem 3rdsem F_Outcome Avg_score
[]:
     Placed?
     0
                                    7.7
                                            8.0
                                                    7.0
                                                                    7.566667
                   Kiran
                              М
                                                                    7.600000
     1
                   Hassan
                              K
                                    7.7
                                            8.0
                                                    7.1
```

```
[]: import pandas as pd df2.plot.line(subplots=True)
```

[]: array([<Axes: >, <Axes: >, <Axes: >, <Axes: >], dtype=object)

