```
In [1]:
         import numpy as np
         import matplotlib.pyplot as plt
         %matplotlib inline
In [2]:
         import pandas as pd
         df = pd.read csv(r"D:\College\TE\SEM-2\Practical\DSBDA\2\AcademicPerformance.csv")
In [3]:
In [4]:
         print(df)
               gender race/ethnicity parental level of education
                                                                              lunch
         0
               female
                              group B
                                                  bachelor's degree
                                                                           standard
         1
               female
                                                       some college
                                                                           standard
                              group C
         2
               female
                              group B
                                                    master's degree
                                                                           standard
         3
                 male
                                                 associate's degree
                                                                      free/reduced
                              group A
         4
                 male
                              group C
                                                       some college
                                                                           standard
                  . . .
         2235
                  NaN
                                   NaN
                                                                 NaN
                                                                                NaN
         2236
                  NaN
                                   NaN
                                                                 NaN
                                                                                NaN
         2237
                  NaN
                                   NaN
                                                                 NaN
                                                                                NaN
         2238
                  NaN
                                   NaN
                                                                 NaN
                                                                                NaN
         2239
                  NaN
                                   NaN
                                                                 NaN
                                                                                NaN
              test preparation course
                                        Year Birth math score reading score \
         0
                                             1970.0
                                                             72.0
                                   none
                                             1961.0
         1
                             completed
                                                              NaN
                                                                              na
         2
                                             1958.0
                                                             90.0
                                                                              95
                                   none
         3
                                             1967.0
                                                              NaN
                                   none
                                                                             NaN
         4
                                   none
                                             1989.0
                                                             76.0
                                                                              78
                                    . . .
                                                 . . .
                                                              . . .
                                                                             . . .
         . . .
                                    NaN
                                                 NaN
                                                              NaN
         2235
                                                                             NaN
         2236
                                    NaN
                                                 NaN
                                                              NaN
                                                                             NaN
         2237
                                                              NaN
                                    NaN
                                                 NaN
                                                                             NaN
         2238
                                    NaN
                                                 NaN
                                                              NaN
                                                                             NaN
         2239
                                    NaN
                                                 NaN
                                                              NaN
                                                                             NaN
              writing score Dt_Admission College_Fees
         0
                          74
                                   6/16/14 $84,835.00
         1
                           Α
                                   6/15/14
                                            $57,091.00
         2
                          93
                                   5/13/14
                                            $67,267.00
         3
                          44
                               05-11-2014
                                            $32,474.00
         4
                          75
                               04-08-2014
                                            $21,474.00
                                       . . .
                         . . .
                                       NaN
         2235
                         NaN
                                                     NaN
         2236
                         NaN
                                       NaN
                                                     NaN
         2237
                         NaN
                                       NaN
                                                     NaN
         2238
                         NaN
                                       NaN
                                                     NaN
                                       NaN
                                                     NaN
         2239
                         NaN
```

[2240 rows x 11 columns]

```
In [5]: print(df['math score'])
         0
                 72.0
         1
                  NaN
         2
                 90.0
         3
                  NaN
         4
                 76.0
         2235
                  NaN
         2236
                  NaN
         2237
                  NaN
         2238
                  NaN
         2239
                  NaN
         Name: math score, Length: 2240, dtype: float64
In [6]: print(df['math score'].isnull())
         0
                 False
         1
                  True
         2
                 False
         3
                  True
         4
                 False
         2235
                  True
         2236
                  True
         2237
                  True
         2238
                  True
         2239
                  True
         Name: math score, Length: 2240, dtype: bool
In [7]:
        print(df['reading score'])
         0
                  72
         1
                  na
         2
                  95
         3
                 NaN
                  78
                 . . .
         2235
                 NaN
         2236
                 NaN
         2237
                 NaN
         2238
                 NaN
         2239
                 NaN
         Name: reading score, Length: 2240, dtype: object
In [8]: print(df['reading score'].isnull())
         0
                 False
         1
                 False
         2
                 False
         3
                  True
         4
                 False
         2235
                  True
         2236
                  True
         2237
                  True
         2238
                  True
         2239
                  True
         Name: reading score, Length: 2240, dtype: bool
```

```
In [9]:
          missing_values = ["n/a", "na", "--"]
          df = pd.read_csv(r"D:\College\TE\SEM-2\Practical\DSBDA\2\AcademicPerformance.csv", na_v
In [10]:
         print(df['reading score'])
                  72.0
          0
          1
                   NaN
          2
                  95.0
          3
                   NaN
          4
                  78.0
          2235
                   NaN
          2236
                   NaN
          2237
                   NaN
          2238
                   NaN
          2239
                   NaN
         Name: reading score, Length: 2240, dtype: float64
In [11]: | print(df['reading score'].isnull())
         0
                  False
          1
                   True
          2
                  False
          3
                   True
                  False
                  . . .
          2235
                   True
          2236
                   True
          2237
                   True
          2238
                   True
          2239
                   True
         Name: reading score, Length: 2240, dtype: bool
In [12]:
         dataset = [11,41,20,3,101,55,68,97,99,6]
In [13]:
         sorted(dataset)
Out[13]: [3, 6, 11, 20, 41, 55, 68, 97, 99, 101]
In [14]:
          quantile1, quantile3 = np.percentile(dataset, [25,75])
In [15]:
         print(quantile1, quantile3)
          13.25 89.75
In [16]:
         iqr_value = (quantile3 - quantile1)
In [17]:
         print(iqr_value)
          76.5
In [18]:
         lower_bound_value = quantile1 - (1.5*iqr_value)
In [19]:
          upper_bound_value = quantile3 + (1.5*iqr_value)
```

```
-101.5 204.5
In [21]:
            from datetime import date
            df['age'] = date.today().year - df['Year Birth']
In [22]:
            df['Year'] = pd.DatetimeIndex(df['Dt Admission']).year
            df['E_L'] = date.today().year - df['Year']
In [23]:
            df.head(5)
Out[23]:
                                         parental
                                                                       test
                                                                                                reading
                                                                                                         writing
                                                                                         math
                gender race/ethnicity
                                          level of
                                                         lunch
                                                                preparation
                                                                             Year_Birth
                                                                                                                  Dt_Admi
                                                                                         score
                                                                                                  score
                                                                                                           score
                                        education
                                                                    course
                                        bachelor's
            0
                                                                                 1970.0
                                                                                          72.0
                                                                                                   72.0
                                                                                                             74
                                                                                                                        6
                female
                              group B
                                                      standard
                                                                      none
                                           degree
                                            some
                female
                                                                                 1961.0
                                                                                          NaN
                                                                                                   NaN
                                                                                                              Α
                                                                                                                        6
                                                      standard
                                                                  completed
             1
                              group C
                                           college
                                         master's
            2
                                                      standard
                                                                                 1958.0
                                                                                          90.0
                                                                                                   95.0
                                                                                                              93
                                                                                                                        5,
                female
                              group B
                                                                      none
                                           degree
                                       associate's
             3
                                                                                 1967.0
                                                                                                                     05-11
                  male
                              group A
                                                   free/reduced
                                                                      none
                                                                                          NaN
                                                                                                   NaN
                                                                                                              44
                                           degree
                                            some
             4
                                                                                 1989.0
                                                                                          76.0
                                                                                                   78.0
                                                                                                             75
                                                                                                                     04-08
                  male
                              group C
                                                      standard
                                                                      none
                                           college
                                                                                                                       \triangleright
            df['Fees$'] = df['College_Fees'].str.replace(',', '').str.replace('$', '').str.replace(
In [24]:
            df['Fees M$'] = df['Fees$'].apply(lambda X:round(X/1000000))
In [25]:
            df.head(5)
Out[25]:
                                         parental
                                                                       test
                                                                                                         writing
                                                                                         math
                                                                                                reading
                gender race/ethnicity
                                          level of
                                                         lunch
                                                                preparation
                                                                             Year_Birth
                                                                                                                  Dt_Admi
                                                                                         score
                                                                                                  score
                                                                                                           score
                                        education
                                                                    course
                                        bachelor's
                female
                              group B
                                                      standard
                                                                      none
                                                                                 1970.0
                                                                                          72.0
                                                                                                   72.0
                                                                                                              74
                                                                                                                        6
                                           degree
                                            some
                female
                              group C
                                                      standard
                                                                  completed
                                                                                 1961.0
                                                                                          NaN
                                                                                                   NaN
                                                                                                              Α
                                                                                                                        6
                                           college
                                         master's
                                                                                                                        5,
             2
                female
                              group B
                                                      standard
                                                                                 1958.0
                                                                                          90.0
                                                                                                   95.0
                                                                                                             93
                                                                      none
                                           degree
                                       associate's
             3
                              group A
                                                   free/reduced
                                                                                 1967.0
                                                                                          NaN
                                                                                                   NaN
                                                                                                              44
                                                                                                                     05-11
                  male
                                                                      none
                                           degree
                                            some
             4
                  male
                              group C
                                                      standard
                                                                      none
                                                                                 1989.0
                                                                                          76.0
                                                                                                   78.0
                                                                                                              75
                                                                                                                     04-08
                                           college
 In [ ]:
```

In [20]:

print(lower bound value, upper bound value)