data-science-practicals-no-3-1

April 3, 2024

```
[41]: import pandas as pd
```

```
[42]: df=pd.read_csv("iris.csv")
```

```
[43]: df.head(2)
```

```
[43]:    Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm      Species
     0   1            5.1           3.5            1.4           0.2  Iris-setosa
     1   2            4.9           3.0            1.4           0.2  Iris-setosa
```

```
[44]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 6 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Id             150 non-null    int64
 1   SepalLengthCm  150 non-null    float64
 2   SepalWidthCm   150 non-null    float64
 3   PetalLengthCm  150 non-null    float64
 4   PetalWidthCm   150 non-null    float64
 5   Species        150 non-null    object
dtypes: float64(4), int64(1), object(1)
memory usage: 7.2+ KB
```

```
[45]: df.describe()
```

```
[45]:                Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
      count  150.000000     150.000000    150.000000     150.000000    150.000000
      mean    75.500000       5.843333      3.054000       3.758667      1.198667
      std     43.445368       0.828066      0.433594       1.764420      0.763161
      min      1.000000       4.300000      2.000000       1.000000      0.100000
      25%     38.250000       5.100000      2.800000       1.600000      0.300000
      50%     75.500000       5.800000      3.000000       4.350000      1.300000
      75%    112.750000       6.400000      3.300000       5.100000      1.800000
      max    150.000000       7.900000      4.400000       6.900000      2.500000
```

```
[46]: df.shape
```

```
[46]: (150, 6)
```

```
[47]: df.dtypes
```

```
[47]: Id                 int64
      SepalLengthCm    float64
      SepalWidthCm     float64
      PetalLengthCm    float64
      PetalWidthCm     float64
      Species           object
      dtype: object
```

```
[48]: # Filter data for each species
      setosa_data = df[df['Species'] == 'Iris-setosa']
      versicolor_data = df[df['Species'] == 'Iris-versicolor']
      virginica_data = df[df['Species'] == 'Iris-virginica']

      # Display the statistical details
      print("Statistics for Iris-setosa:")
      print(setosa_data.describe())
      print("\nStatistics for Iris-versicolor:")
      print(versicolor_data.describe())
      print("\nStatistics for Iris-virginica:")
      print(virginica_data.describe())
```

```
Statistics for Iris-setosa:
              Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
count   50.00000       50.00000     50.000000      50.000000      50.00000
mean    25.50000        5.00600      3.418000       1.464000       0.24400
std     14.57738        0.35249      0.381024       0.173511       0.10721
min      1.00000        4.30000      2.300000       1.000000       0.10000
25%     13.25000        4.80000      3.125000       1.400000       0.20000
50%     25.50000        5.00000      3.400000       1.500000       0.20000
75%     37.75000        5.20000      3.675000       1.575000       0.30000
max     50.00000        5.80000      4.400000       1.900000       0.60000

Statistics for Iris-versicolor:
              Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
count   50.00000      50.000000     50.000000      50.000000     50.000000
mean    75.50000       5.936000      2.770000       4.260000      1.326000
std     14.57738       0.516171      0.313798       0.469911      0.197753
min     51.00000       4.900000      2.000000       3.000000      1.000000
25%     63.25000       5.600000      2.525000       4.000000      1.200000
50%     75.50000       5.900000      2.800000       4.350000      1.300000
75%     87.75000       6.300000      3.000000       4.600000      1.500000
```

```
max      100.00000      7.000000      3.400000      5.100000      1.800000

Statistics for Iris-virginica:
            Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
count   50.00000      50.00000     50.000000      50.000000      50.00000
mean   125.50000       6.58800      2.974000       5.552000       2.02600
std     14.57738       0.63588      0.322497       0.551895       0.27465
min    101.00000       4.90000      2.200000       4.500000       1.40000
25%    113.25000       6.22500      2.800000       5.100000       1.80000
50%    125.50000       6.50000      3.000000       5.550000       2.00000
75%    137.75000       6.90000      3.175000       5.875000       2.30000
max    150.00000       7.90000      3.800000       6.900000       2.50000
```