

Name : Tanishq Thuse
 Div : SY CSAI - B
 Roll no. : 60
 PRN : 12310237
 Subject : DV Assignment - 4

Assignment -4

Write a Python program to display some basic statistical details like percentile, mean, standard deviation etc. of the species of 'Iris-setosa', 'Iris-versicolor' and 'Iris-versicolor' of iris.csv dataset

Dataset link : <https://www.kaggle.com/datasets/saurabh00007/iris.csv?resource=download>


Importing Libraries

```
import pandas as pd
```



Creating a dataframe & EDA

```
path = "/content/Iris.csv"
df = pd.read_csv(path)
```

```
df.head()
```



| | 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 |
| 2 | 3 | 4.7 | 3.2 | 1.3 | 0.2 | Iris-setosa |
| 3 | 4 | 4.6 | 3.1 | 1.5 | 0.2 | Iris-setosa |
| 4 | 5 | 5.0 | 3.6 | 1.4 | 0.2 | Iris-setosa |


Next steps:

[Generate code with df](#)



☒ [View recommended plots](#)

[New interactive sheet](#)

```
df.tail()
```



| | Id | SepalLengthCm | SepalWidthCm | PetalLengthCm | PetalWidthCm | Species |
|-----|-----|---------------|--------------|---------------|--------------|----------------|
| 145 | 146 | 6.7 | 3.0 | 5.2 | 2.3 | Iris-virginica |
| 146 | 147 | 6.3 | 2.5 | 5.0 | 1.9 | Iris-virginica |
| 147 | 148 | 6.5 | 3.0 | 5.2 | 2.0 | Iris-virginica |
| 148 | 149 | 6.2 | 3.4 | 5.4 | 2.3 | Iris-virginica |
| 149 | 150 | 5.9 | 3.0 | 5.1 | 1.8 | Iris-virginica |

```
#Function to display statistics in more detail
```

```
def display_statistics(species_name):
    species_data = df[df['Species'] == species_name]
    print(f"Statistics for {species_name}:")
    print(f"Mean:\n{species_data.mean(numeric_only=True)}\n")
    print(f"Standard Deviation:\n{species_data.std(numeric_only=True)}\n")
    print(f"25th Percentile:\n{species_data.quantile(0.25,numeric_only=True)}\n")
    print(f"50th Percentile (Median):\n{species_data.median(numeric_only=True)}\n")
    print(f"75th Percentile:\n{species_data.quantile(0.75,numeric_only=True)}\n")
```

```
#List for species
```

```
species_list = ['Iris-setosa', 'Iris-versicolor', 'Iris-virginica']
```

- ✓ Statistical details like percentile, mean, standard deviation etc. of the species of 'Iris-setosa'.

```
#Filtering the data for each species
setosa = df[df['Species'] == 'Iris-setosa']
versicolor = df[df['Species'] == 'Iris-versicolor']
virginica = df[df['Species'] == 'Iris-virginica']
```

```
#Calculating basic statistical details
print("Iris-setosa:")
print(setosa.describe())
```

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

```
#Display in more detail
display_statistics(species_list[0])
```

```
Statistics for Iris-setosa:
```

Mean:

```
Id          25.500
SepalLengthCm  5.006
SepalWidthCm   3.418
PetalLengthCm  1.464
PetalWidthCm   0.244
dtype: float64
```

Standard Deviation:

```
Id          14.577380
SepalLengthCm  0.352490
SepalWidthCm   0.381024
PetalLengthCm  0.173511
PetalWidthCm   0.107210
dtype: float64
```

25th Percentile:

```
Id          13.250
SepalLengthCm  4.800
SepalWidthCm   3.125
PetalLengthCm  1.400
PetalWidthCm   0.200
Name: 0.25, dtype: float64
```

50th Percentile (Median):

```
Id          25.5
SepalLengthCm  5.0
SepalWidthCm   3.4
PetalLengthCm  1.5
PetalWidthCm   0.2
dtype: float64
```

75th Percentile:

```
Id          37.750
SepalLengthCm  5.200
SepalWidthCm   3.675
PetalLengthCm  1.575
PetalWidthCm   0.300
Name: 0.75, dtype: float64
```

- ✓ Statistical details like percentile, mean, standard deviation etc. of the species of 'Iris-versicolor'.

```
print("\nIris-versicolor:")
print(versicolor.describe())
```

```
Iris-versicolor:
   count  Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
count  50.00000  50.000000  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 |

```
#Display in more detail
display_statistics(species_list[1])
```

Statistics for Iris-versicolor:

Mean:

| | |
|---------------|--------|
| Id | 75.500 |
| SepalLengthCm | 5.936 |
| SepalWidthCm | 2.770 |
| PetalLengthCm | 4.260 |
| PetalWidthCm | 1.326 |

dtype: float64

Standard Deviation:

| | |
|---------------|-----------|
| Id | 14.577380 |
| SepalLengthCm | 0.516171 |
| SepalWidthCm | 0.313798 |
| PetalLengthCm | 0.469911 |
| PetalWidthCm | 0.197753 |

dtype: float64

25th Percentile:

| | |
|---------------|--------|
| Id | 63.250 |
| SepalLengthCm | 5.600 |
| SepalWidthCm | 2.525 |
| PetalLengthCm | 4.000 |
| PetalWidthCm | 1.200 |

Name: 0.25, dtype: float64

50th Percentile (Median):

| | |
|---------------|-------|
| Id | 75.50 |
| SepalLengthCm | 5.90 |
| SepalWidthCm | 2.80 |
| PetalLengthCm | 4.35 |
| PetalWidthCm | 1.30 |

dtype: float64

75th Percentile:

| | |
|---------------|-------|
| Id | 87.75 |
| SepalLengthCm | 6.30 |
| SepalWidthCm | 3.00 |
| PetalLengthCm | 4.60 |
| PetalWidthCm | 1.50 |

Name: 0.75, dtype: float64

✓ Statistical details like percentile, mean, standard deviation etc. of the species of 'Iris-virginica'.

```
print("\nIris-virginica:")
print(virginica.describe())
```

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 |

```
#Display in more detail
display_statistics(species_list[2])
```

Statistics for Iris-virginica:

Mean:

| | |
|---------------|---------|
| Id | 125.500 |
| SepalLengthCm | 6.588 |
| SepalWidthCm | 2.974 |
| PetalLengthCm | 5.552 |

```
PetalWidthCm      2.026  
dtype: float64
```

```
Standard Deviation:  
Id                14.577380  
SepalLengthCm    0.635880  
SepalWidthCm     0.322497  
PetalLengthCm    0.551895  
PetalWidthCm     0.274650  
dtype: float64
```

```
25th Percentile:  
Id                113.250  
SepalLengthCm     6.225  
SepalWidthCm      2.800  
PetalLengthCm     5.100  
PetalWidthCm      1.800  
Name: 0.25, dtype: float64
```

```
50th Percentile (Median):  
Id                125.50  
SepalLengthCm     6.50  
SepalWidthCm      3.00  
PetalLengthCm     5.55  
PetalWidthCm      2.00  
dtype: float64
```

```
75th Percentile:  
Id                137.750  
SepalLengthCm     6.900  
SepalWidthCm      3.175  
PetalLengthCm     5.875  
PetalWidthCm      2.300  
Name: 0.75, dtype: float64
```