

# DSBDL Assignment 03 - Descriptive Statistics: Measures of Central Tendency and Variability

## Part 2




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.

```
from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

```
import numpy as np
import pandas as pd
import seaborn as sns
```

```
ds = pd.read_csv('/content/drive/My Drive/DSBDL/Assignment3/iris.csv')
ds
```

	sepal_length	sepal_width	petal_length	petal_width	species	
0	5.1	3.5	1.4	0.2	setosa	
1	4.9	3.0	1.4	0.2	setosa	
2	4.7	3.2	1.3	0.2	setosa	
3	4.6	3.1	1.5	0.2	setosa	
4	5.0	3.6	1.4	0.2	setosa	
...	...	...	...	...	...	
145	6.7	3.0	5.2	2.3	virginica	
146	6.3	2.5	5.0	1.9	virginica	
147	6.5	3.0	5.2	2.0	virginica	
148	6.2	3.4	5.4	2.3	virginica	
149	5.9	3.0	5.1	1.8	virginica	

150 rows × 5 columns



Next steps:

[Generate code with ds](#)[View recommended plots](#)

```
ds.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
#   Column          Non-Null Count  Dtype
---  -
0   sepal_length    150 non-null   float64
1   sepal_width     150 non-null   float64
2   petal_length    150 non-null   float64
3   petal_width     150 non-null   float64
4   species         150 non-null   object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
```

```
ds.describe()
```



	sepal_length	sepal_width	petal_length	petal_width	
count	150.000000	150.000000	150.000000	150.000000	
mean	5.843333	3.054000	3.758667	1.198667	
std	0.828066	0.433594	1.764420	0.763161	
min	4.300000	2.000000	1.000000	0.100000	
25%	5.100000	2.800000	1.600000	0.300000	
50%	5.800000	3.000000	4.350000	1.300000	
75%	6.400000	3.300000	5.100000	1.800000	
max	7.900000	4.400000	6.900000	2.500000	

```
ds.isna().sum()
```



```
sepal_length    0
sepal_width     0
petal_length    0
petal_width     0
species         0
dtype: int64
```

```
iris = ds.groupby(ds['species'])
```



```
iris.mean()
```

	sepal_length	sepal_width	petal_length	petal_width	
species					
setosa	5.006	3.418	1.464	0.244	
versicolor	5.936	2.770	4.260	1.326	
virginica	6.588	2.974	5.552	2.026	



iris.median()

	sepal_length	sepal_width	petal_length	petal_width	
species					
setosa	5.0	3.4	1.50	0.2	
versicolor	5.9	2.8	4.35	1.3	
virginica	6.5	3.0	5.55	2.0	

iris.count()

	sepal_length	sepal_width	petal_length	petal_width	
species					
setosa	50	50	50	50	
versicolor	50	50	50	50	
virginica	50	50	50	50	

iris.min()

	sepal_length	sepal_width	petal_length	petal_width	
species					
setosa	4.3	2.3	1.0	0.1	
versicolor	4.9	2.0	3.0	1.0	
virginica	4.9	2.2	4.5	1.4	

iris.max()

sepal\_lengthsepal\_widthpetal\_lengthpetal\_width

iris.std()

sepal\_lengthsepal\_widthpetal\_lengthpetal\_width

species				
setosa	0.352490	0.381024	0.173511	0.107210
versicolor	0.516171	0.313798	0.469911	0.197753
virginica	0.635880	0.322497	0.551895	0.274650

iris.quantile()

sepal\_lengthsepal\_widthpetal\_lengthpetal\_width

species				
setosa	5.0	3.4	1.50	0.2
versicolor	5.9	2.8	4.35	1.3
virginica	6.5	3.0	5.55	2.0

Start coding or [generate](#) with AI.