

Assignment (Iris Data)

Download the Iris dataset from

<https://www.kaggle.com/datasets/uciml/iris>

and write a program that loads the CSV file and answers what is the average sepal length for each of three iris species.

Solution

1. Import pandas.

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

2. Download Iris dataset and read it.

```
[2]: data = pd.read_csv("Iris.csv")
data
```

```
[2]:
```

	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
...
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

150 rows × 6 columns

Peer Assignment

Sheikh Muhammad Sabih (2303.KHI.DEG.010)

M Humza Moeen (2303.KHI.DEG.019)

3. Find the **average Sepal Length** of each **Species**.

```
[3]: avg_length = data.groupby("Species")["SepalLengthCm"].mean()
```

```
[4]: avg_length
```

```
[4]: Species
Iris-setosa      5.006
Iris-versicolor  5.936
Iris-virginica   6.588
Name: SepalLengthCm, dtype: float64
```