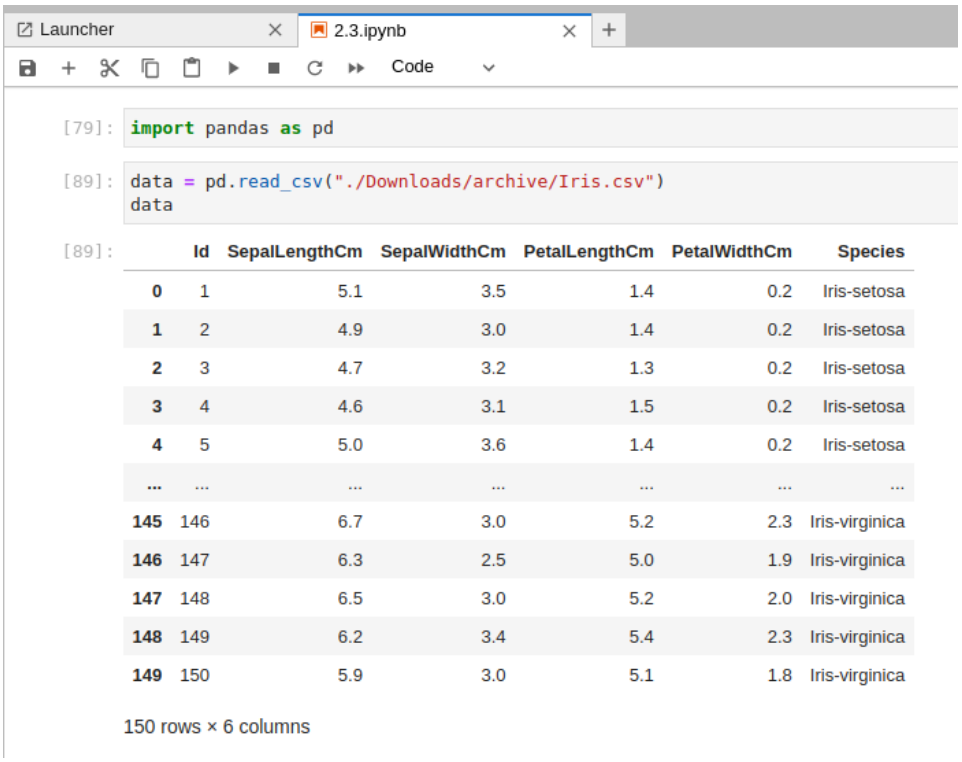
Talha Khan (2303.009.KHI.DEG)

Muhammad Moiz Khan (2303.022.KHI.DEG)

ASSIGNMNET 2.3



	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

- The first line imports the '**pandas**' library and makes it available under the alias '**pd**'.
- The second line reads the Iris dataset CSV file into a pandas DataFrame using the **read_csv()** method. The file path **./Downloads/archive/Iris.csv** assumes that the file is located in the **Downloads/archive** directory relative to the current working directory.

```
[81]: data.columns

[81]: Index(['Id', 'SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm', 'PetalWidthCm',
       'Species'],
      dtype='object')

[109]: avg = data.groupby('Species')['SepalLengthCm'].mean()

[108]: print("Average sepal length for :", avg)

Average sepal length for : Species
Iris-setosa        5.006
Iris-versicolor    5.936
Iris-virginica     6.588
Name: SepalLengthCm, dtype: float64

[ ]:
```

- The third line prints the column names of the DataFrame using the **columns** attribute. This step is optional and is not necessary for calculating the average sepal length.
- The fourth line calculates the average sepal length for each species using the **groupby()** method. This method groups the DataFrame by the **Species** column and calculates the mean of the **SepalLengthCm** column for each group.
- Finally, the fifth line prints the average sepal length for each species using the **print()** function. The output will be a pandas Series object that lists the average sepal length for each species.