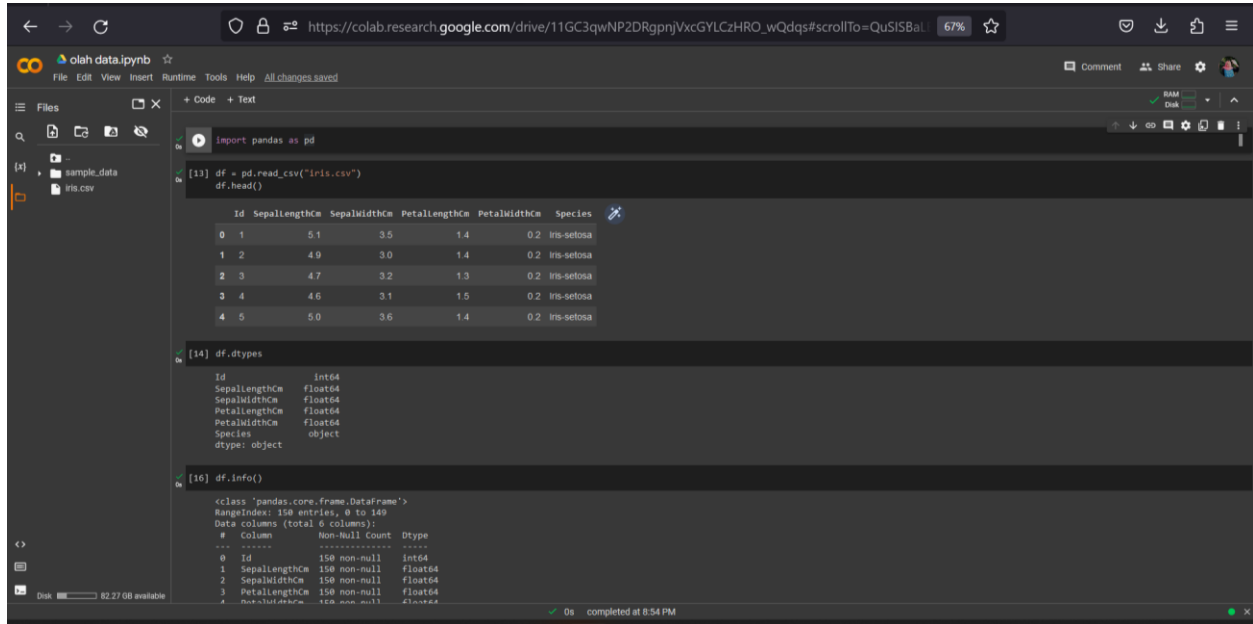


Nama : Jahrawati

Nim : 20.01.013.050

## 01. Olah data iris database



```
import pandas as pd

[13] df = pd.read_csv("iris.csv")
df.head()

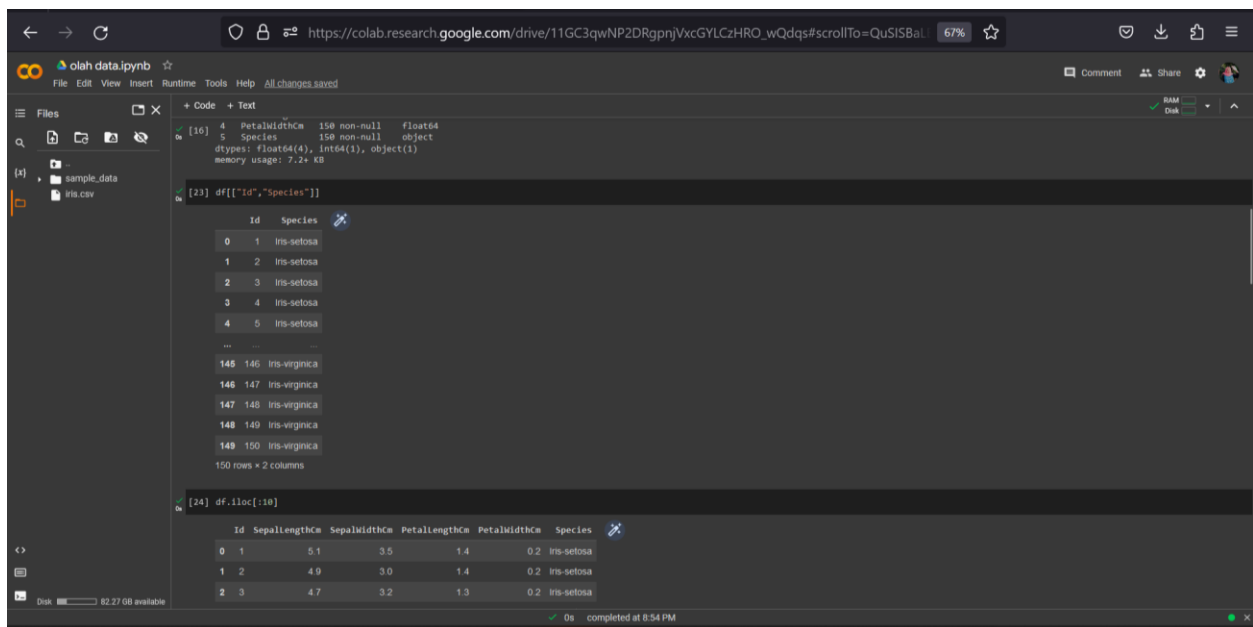
  Id  SepallengthCm  SepalwidthCm  PetallengthCm  PetalwidthCm  Species
0  1             5.1           3.5           1.4           0.2  Iris-setosa
1  2             4.9           3.0           1.4           0.2  Iris-setosa
2  3             4.7           3.2           1.3           0.2  Iris-setosa
3  4             4.6           3.1           1.5           0.2  Iris-setosa
4  5             5.0           3.6           1.4           0.2  Iris-setosa

[14] df.dtypes

Id                int64
SepallengthCm    float64
SepalwidthCm     float64
PetallengthCm    float64
PetalwidthCm     float64
Species          object
dtype: object

[16] 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: object (1), int64 (1), float64 (4)
memory usage: 7.2+ KB
```



```
[16] 4   PetalwidthCm         150 non-null   float64
      5   Species             150 non-null   object  
      dtypes: float64(4), int64(1), object(1)
      memory usage: 7.2+ KB

[23] df[["Id", "Species"]]

   Id  Species
0   1  Iris-setosa
1   2  Iris-setosa
2   3  Iris-setosa
3   4  Iris-setosa
4   5  Iris-setosa
...  ...
145 146  Iris-virginica
146 147  Iris-virginica
147 148  Iris-virginica
148 149  Iris-virginica
149 150  Iris-virginica
150 rows x 2 columns

[24] df.iloc[10]

   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
```

Colab interface showing a Jupyter Notebook with the following code and output:

```
+ Code + Text
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
5 6 5.4 3.9 1.7 0.4 Iris-setosa
6 7 4.6 3.4 1.4 0.3 Iris-setosa
7 8 5.0 3.4 1.5 0.2 Iris-setosa
8 9 4.4 2.9 1.4 0.2 Iris-setosa
9 10 4.9 3.1 1.5 0.1 Iris-setosa
```

```
[25] df[["Id","Species"]].iloc[11:16]
```

Id	Species
11	Iris-setosa
12	Iris-setosa
13	Iris-setosa
14	Iris-setosa
15	Iris-setosa
16	Iris-setosa

```
[26] df.tail(3)
```

Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
147	148	6.5	3.0	5.2	Iris-virginica
148	149	6.2	3.4	5.4	Iris-virginica
149	150	5.9	3.0	5.1	Iris-virginica

```
[27] df.mean()
```

completed at 8:54 PM

Colab interface showing a Jupyter Notebook with the following code and output:

```
+ Code + Text
[27] df.mean()
clipython-input-27-c61fecf89b5:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric_only=None') is deprecated; in a future version this will raise TypeError. Select
df.mean()
Id 75.500000
SepalLengthCm 5.843333
SepalWidthCm 3.954000
PetalLengthCm 3.758667
PetalWidthCm 1.198667
dtype: float64
```

```
[34] df[["PetalLengthCm"]].mean()
```

3.7586666666666666

```
[37] df[["SepalWidthCm"]].min()
```

2.0

```
[38] df[["Species"]].value_counts()
```

Species	count
Iris-setosa	50
Iris-versicolor	50
Iris-virginica	50

```
dfValueCountsSpecies = df[["Species"]].value_counts().rename_axis("Species Value Counts").reset_index(name="Count")
```

Species Value Counts	Count
0 Iris-setosa	50
1 Iris-versicolor	50
2 Iris-virginica	50

```
[49] dfValueCountsSpecies = df[["PetalLengthCm"]].value_counts().rename_axis("PetalLengthCm Value Counts").reset_index(name="Count")
```

completed at 8:54 PM

