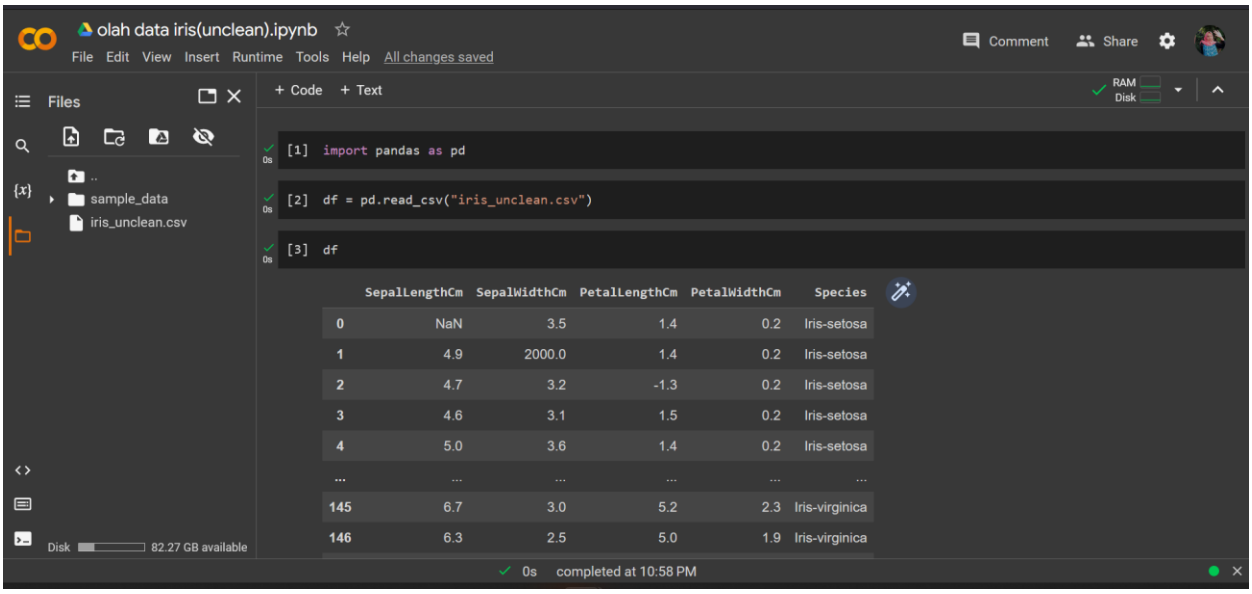
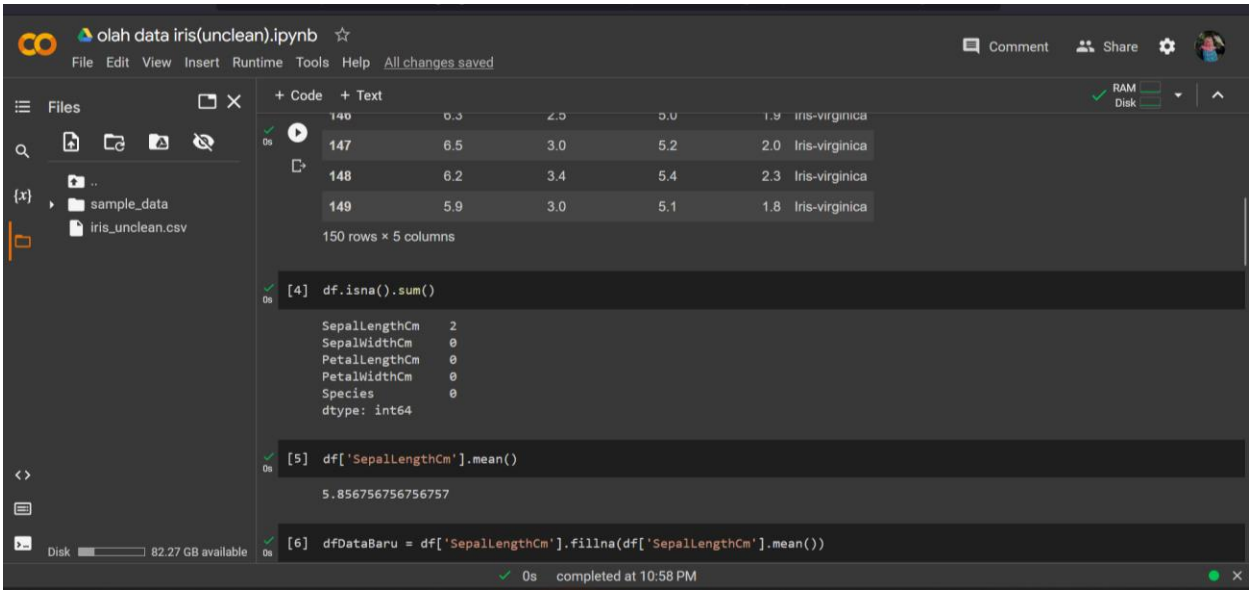
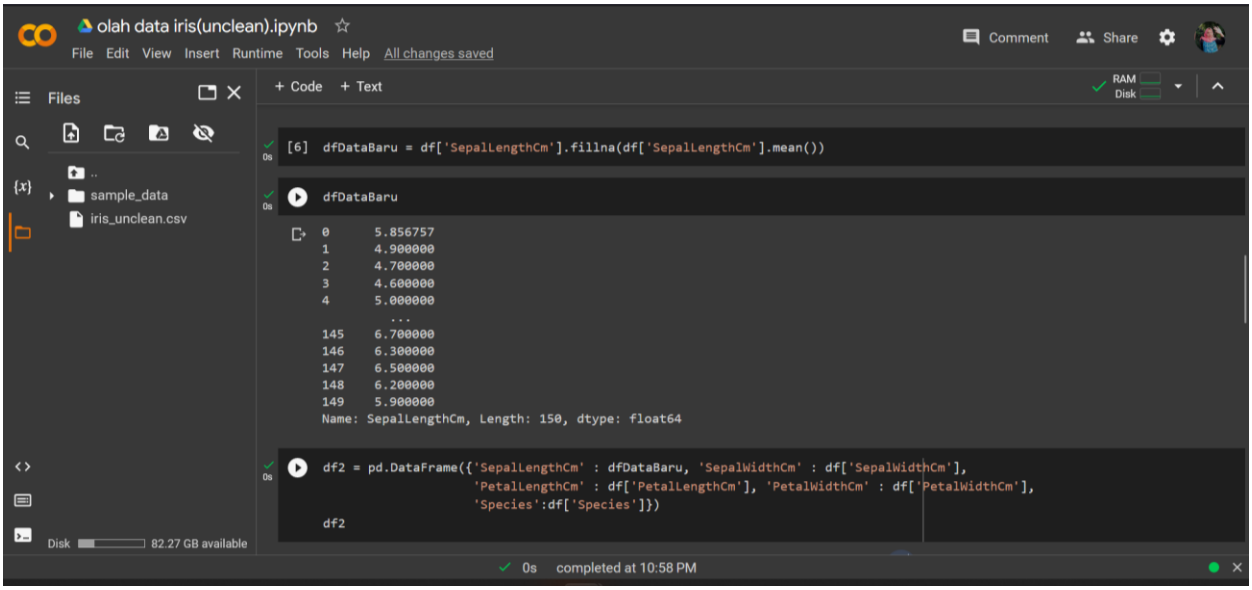
Nama : Jahrawati

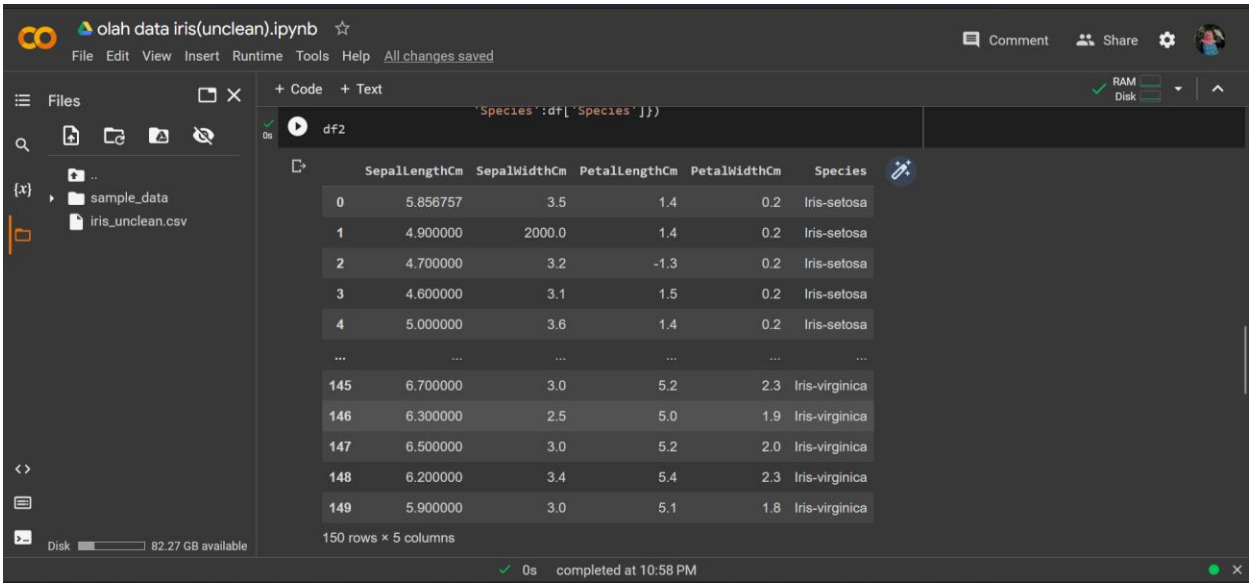
Nim : 20.01.013.050

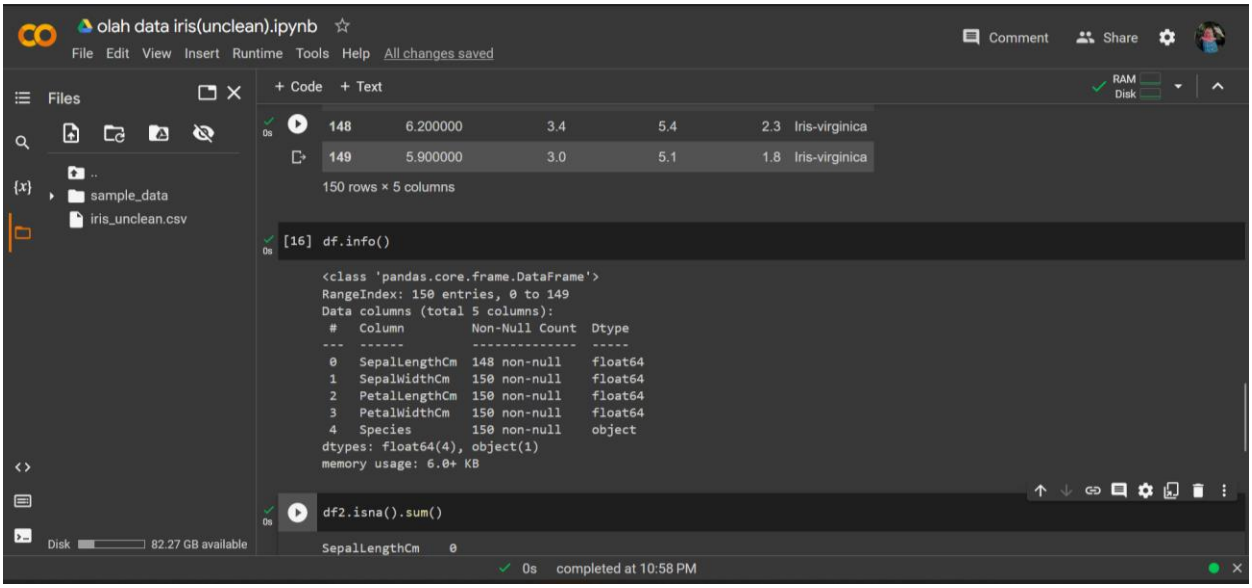
02. olah data iris (unclean)

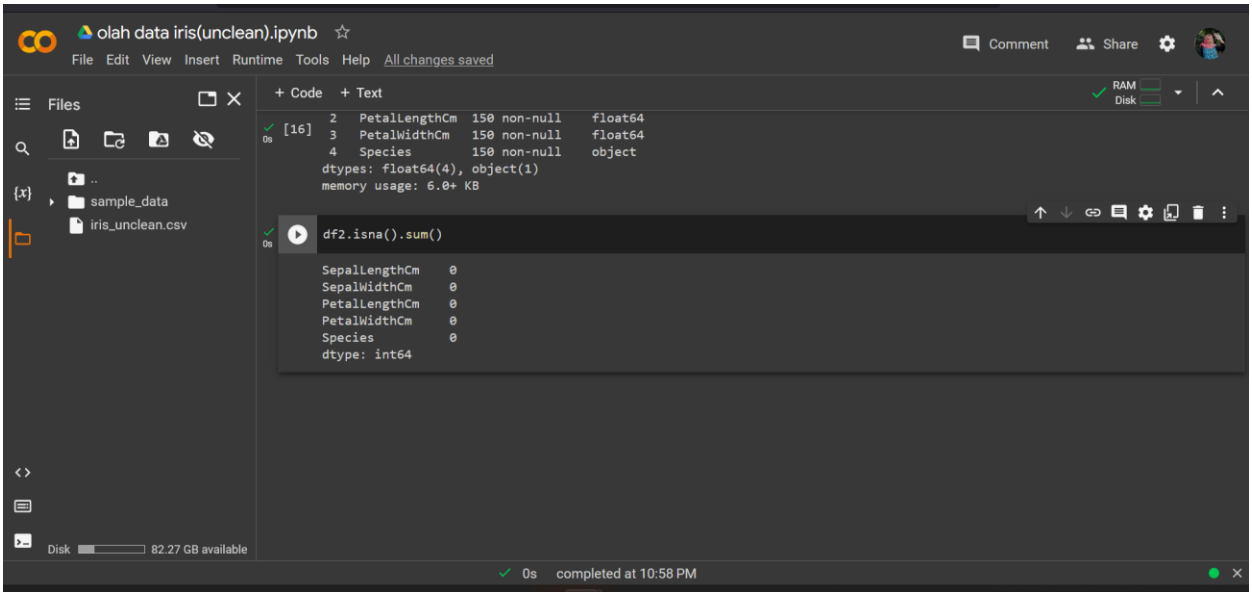





olah data iris(unclean).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code
+ Text
RAM
Disk
[6] dfDataBaru = df['SepalLengthCm'].fillna(df['SepalLengthCm'].mean())
dfDataBaru
0 5.856757
1 4.900000
2 4.700000
3 4.600000
4 5.000000
...
145 6.700000
146 6.300000
147 6.500000
148 6.200000
149 5.900000
Name: SepalLengthCm, Length: 150, dtype: float64
df2 = pd.DataFrame({'SepalLengthCm' : dfDataBaru, 'SepalWidthCm' : df['SepalWidthCm'],
'PetalLengthCm' : df['PetalLengthCm'], 'PetalWidthCm' : df['PetalWidthCm'],
'Species':df['Species']})
df2
..
sample_data
iris_unclean.csv
Disk
82.27 GB available
0s completed at 10:58 PM


olah data iris(unclean).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code
+ Text
RAM
Disk
'Species':df['Species']})
df2
SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm Species
0 5.856757 3.5 1.4 0.2 Iris-setosa
1 4.900000 2000.0 1.4 0.2 Iris-setosa
2 4.700000 3.2 -1.3 0.2 Iris-setosa
3 4.600000 3.1 1.5 0.2 Iris-setosa
4 5.000000 3.6 1.4 0.2 Iris-setosa
... ...
145 6.700000 3.0 5.2 2.3 Iris-virginica
146 6.300000 2.5 5.0 1.9 Iris-virginica
147 6.500000 3.0 5.2 2.0 Iris-virginica
148 6.200000 3.4 5.4 2.3 Iris-virginica
149 5.900000 3.0 5.1 1.8 Iris-virginica
150 rows × 5 columns
..
sample_data
iris_unclean.csv
Disk
82.27 GB available
0s completed at 10:58 PM


olah data iris(unclean).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code
+ Text
RAM
Disk
148 6.200000 3.4 5.4 2.3 Iris-virginica
149 5.900000 3.0 5.1 1.8 Iris-virginica
150 rows × 5 columns
[16] df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
Column Non-Null Count Dtype
0 SepalLengthCm 148 non-null float64
1 SepalWidthCm 150 non-null float64
2 PetalLengthCm 150 non-null float64
3 PetalWidthCm 150 non-null float64
4 Species 150 non-null object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
df2.isna().sum()
SepalLengthCm 0
sample_data
iris_unclean.csv
Disk 82.27 GB available
0s completed at 10:58 PM


olah data iris(unclean).ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code
+ Text
RAM
Disk
[16] 2 PetalLengthCm 150 non-null float64
3 PetalWidthCm 150 non-null float64
4 Species 150 non-null object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
df2.isna().sum()
SepalLengthCm 0
SepalWidthCm 0
PetalLengthCm 0
PetalWidthCm 0
Species 0
dtype: int64
sample_data
iris_unclean.csv
Disk 82.27 GB available
0s completed at 10:58 PM