

ESSENTIALS OF DATA WAREHOUSING AND DATA MINING

LAB ASSIGNMENT - 12

NAME - Kaparotu Venkata Surya Tharani

USN - 22BTRAD018

BRANCH - AI & DE

1. Given the irish data. Except for the class attribute, take all attribute value in to a matrix A. Now calculate the corelation coefficiect (without using the direct correlation calculation function from any python library) between each sample (row vector - one irish flower) in matrix A with all the sample in matrix A. And for plotting and validating your calculations correctness, you can use the following code.

Code :-

```
import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

import numpy as np

f = pd.read_csv('Iris.csv')

print(f)

print(f.loc[:, 'PetalWidthCm'])

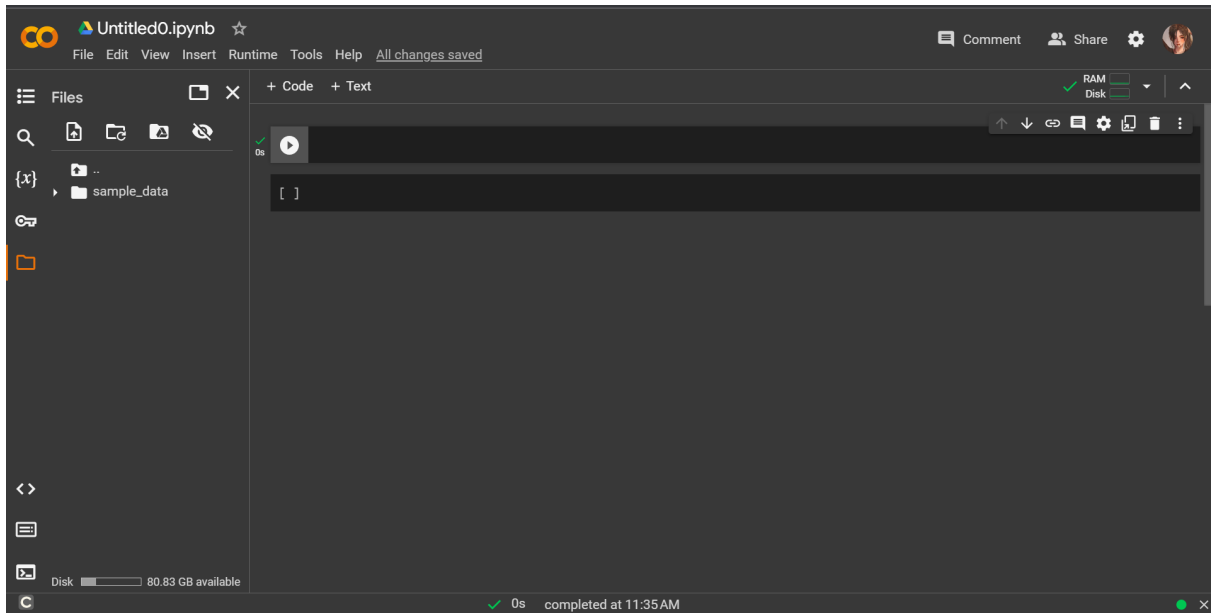
print(f.loc[:, 'PetalWidthCm'].T.corr().loc[:10, :10])

dcorr=f.loc[:, 'PetalWidthCm'].T.corr()

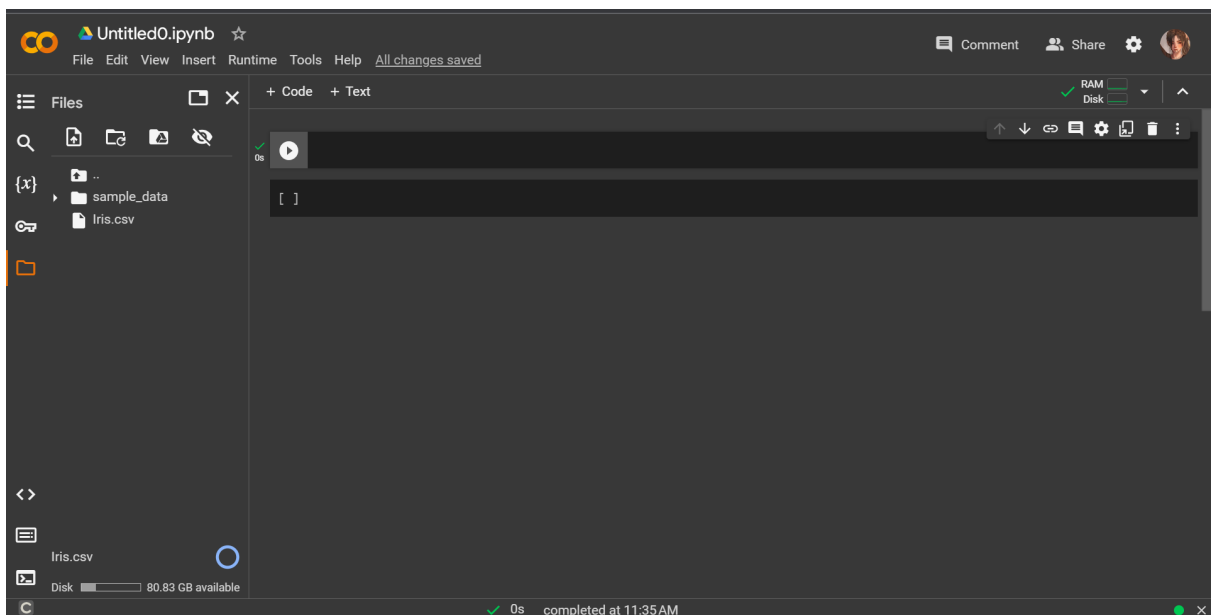
sns.heatmap(dcorr.corr().loc[:10, :10], annot = True, cmap = "coolwarm")
```

SCREENSHOTS :-

1. Click the folder.



2. Click on upload button and select the iris dataset.



3. Write the python code and execute.

```
Untitled0.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
f = pd.read_csv('Iris.csv')
print(f)
print(f.loc[:, 'PetalWidthCm'])
print(f.loc[:, 'PetalWidthCm'].T.corr().loc[:10, :10])
dcorr=f.loc[:, 'PetalWidthCm'].T.corr()
sns.heatmap(dcorr.corr().loc[:10, :10], annot = True, cmap = "coolwarm")
```

Output :

```
Untitled0.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
0 1 2 3 4 5 ... 145 146 147 148 149 150
SepalLengthCm 5.1 4.9 4.7 4.6 5.0 ... 6.7 6.3 6.5 6.2 5.9
SepalWidthCm 3.5 3.0 3.2 3.1 3.6 ... 3.0 2.5 3.0 3.4 3.0
PetalLengthCm 1.4 1.4 1.3 1.5 1.4 ... 5.2 5.0 5.2 5.4 5.1
PetalWidthCm 0.2 0.2 0.2 0.2 0.2 ... 2.3 1.9 2.0 2.3 1.8
Species
0 Iris-setosa
1 Iris-setosa
2 Iris-setosa
3 Iris-setosa
4 Iris-setosa
...
145 Iris-virginica
146 Iris-virginica
147 Iris-virginica
148 Iris-virginica
149 Iris-virginica
```

```
Untitled0.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[150 rows x 6 columns]
0 1 2 3 4 5 6
0 1.000000 0.963372 0.868565 0.715549 0.634651 0.555049 0.327686
1 0.963372 1.000000 0.962732 0.868439 0.804075 0.743467 0.551987
2 0.868565 0.962732 1.000000 0.966743 0.933644 0.893975 0.752314
3 0.715549 0.868439 0.966743 1.000000 0.991214 0.976251 0.891561
4 0.634651 0.804075 0.933644 0.991214 1.000000 0.994978 0.938072
5 0.555049 0.743467 0.893975 0.976251 0.994978 1.000000 0.967724
6 0.327686 0.551987 0.752314 0.891561 0.938072 0.967724 1.000000
7 0.284741 0.517718 0.722297 0.872177 0.920587 0.955001 0.998184
8 0.127156 0.375376 0.601880 0.782933 0.845727 0.894810 0.977402
9 0.131254 0.380052 0.605077 0.785693 0.847347 0.896193 0.977629
10 0.141504 0.386859 0.613447 0.790759 0.854188 0.901678 0.981136
```

