Assignment No.3.

- Title:- Discriptive statistics - measures of central tendancy & variability.

∴ perform following operations on any open source dataset.

1) provide summary statistics for a dataset with numeric values grouped by one of the qualitative variable For example if your categorial variable is age group and quantitative variable is income, then provide summary statistics of income grouped by the age groups. create a list that contains numeric value for each response to the categorial variable.

2) write a python program to display some basic statistical details like percentile, mean, standard deviation etc. of the species of Iris-setosa, Iris-versicolor & Iris-versicolor of iris.csv dataset

= Objectives of assignment:- student should be able to perform statistical operations using python on any open-source dataset.

= prerequisite:

1) Basic of python programming.
2) concept of statistics such as mean, median, minimum, maximum, standard deviath, etc.

notebookd1e29aef84 | Kaggle

kaggle.com/jatharshveta/notebookd1e29aef84/edit

Gmail YouTube Maps

notebookd1e29aef84 Draft saved

File Edit View Run Add-ons Help

Share

Run All Code

Draft Session (28m)

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

[4]:
```
data=pd.read_csv('../input/student-feedback-survey-responses/student_feedback.csv')
```

[5]:
```
data
```

[5]:

	Unnamed: 0	Student ID	Well versed with the subject	Explains concepts in an understandable way	Use of presentations	Degree of difficulty of assignments	Solves doubts willingly	Structuring of the course	Provides support for students going above and beyond	Course recommendation based on relevance
0	0	340	5	2	7	6	9	2	1	8
1	1	253	6	5	8	6	2	1	2	9
2	2	680	7	7	6	5	4	2	3	1
3	3	806	9	6	7	1	5	9	4	6
4	4	632	8	10	8	4	6	6	9	9
...	...	...	...	...	...	...	...	...	...	...
996	996	55	8	7	6	2	5	7	7	9

Console

kaggle.com/jatharshveta/notebookd1e29aef84/edit

Apps Gmail YouTube Maps

notebookd1e29aef84 Draft saved

File Edit View Run Add-ons Help

Share

Run All Code

Draft Session (28m)

996	996	55	8	7	6	2	5	7	7	9
997	997	913	5	5	6	5	6	7	6	1
998	998	199	9	5	8	3	8	1	1	2
999	999	539	10	2	7	4	3	4	10	1
1000	1000	759	7	2	4	2	1	5	9	9

1001 rows × 10 columns

[6]:
```
data.mean()
```

[6]:
```
Unnamed: 0                                                500.000000
Student ID                                                500.000000
Well versed with the subject                                7.497502
Explains concepts in an understandable way                  6.081918
Use of presentations                                        5.942058
Degree of difficulty of assignments                         5.430569
Solves doubts willingly                                     5.474525
Structuring of the course                                   5.636364
Provides support for students going above and beyond        5.662338
Course recommendation based on relevance                    5.598402
dtype: float64
```

[7]:
```
data.loc[:,'Student ID'].mean()
```

Console

34°C Sunny

kaggle.com/jatharshveta/notebookd1e29aef84/edit

Apps Gmail YouTube Maps

notebookd1e29aef84 Draft saved

File Edit View Run Add-ons Help

Run All Code Draft Session (28m)

[8]:
```
0    38.0
1    29.3
2    71.7
3    85.6
dtype: float64
```

[9]:
```
data.median()
```

[9]:
```
Unnamed: 0                                          500.0
Student ID                                          500.0
Well versed with the subject                          8.0
Explains concepts in an understandable way            6.0
Use of presentations                                  6.0
Degree of difficulty of assignments                   5.0
Solves doubts willingly                               6.0
Structuring of the course                             6.0
Provides support for students going above and beyond  6.0
Course recommendation based on relevance              6.0
dtype: float64
```

[10]:
```
data.loc[:,'Student ID'].median()
```

[10]:
```
500.0
```

Console

notebookd1e29aef84 Draft saved

File Edit View Run Add-ons Help

Run All Code Share

Draft Session (29m)

[11]:
```
data.median(axis=1)[0:4]
```

[11]:
```
0    5.5
1    5.5
2    4.5
3    6.0
dtype: float64
```

[12]:
```
data.mode()
```

[12]:

	Unnamed: 0	Student ID	Well versed with the subject	Explains concepts in an understandable way	Use of presentations	Degree of difficulty of assignments	Solves doubts willingly	Structuring of the course	Provides support for students going above and beyond	Course recommendation based on relevance
0	0	0	9.0	9.0	4.0	6.0	6.0	10.0	7.0	9.0
1	1	1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	2	2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	3	3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	4	4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...
996	996	996	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Console

notebookd1e29aef84 | Kaggle

kaggle.com/jatharshveta/notebookd1e29aef84/edit

Apps Gmail YouTube Maps

notebookd1e29aef84 Draft saved

File Edit View Run Add-ons Help

Share

Run All Code Draft Session (29m)

998	998	998	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
999	999	999	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
1000	1000	1000	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	

1001 rows × 10 columns

[13]:
```
data.loc[:,'Student ID'].mode()
```

[13]:
```
0          0
1          1
2          2
3          3
4          4
         ...
996      996
997      997
998      998
999      999
1000    1000
Length: 1001, dtype: int64
```

[14]:
```
data.min()
```

[14]:
```
Unnamed: 0                    0
Student ID                    0
```

Console

kaggle.com/jatharshveta/notebookd1e29aef84/edit

Apps Gmail YouTube Maps

notebookd1e29aef84 Draft saved

File Edit View Run Add-ons Help

Share

Run All Code

Draft Session (29m)

```
Student ID                                                  0
Well versed with the subject                                5
Explains concepts in an understandable way                  2
Use of presentations                                        4
Degree of difficulty of assignments                         1
Solves doubts willingly                                     1
Structuring of the course                                   1
Provides support for students going above and beyond        1
Course recommendation based on relevance                    1
dtype: int64
```

[15]:
```
data.loc[:,'Student ID'].min(skipna=False)
```

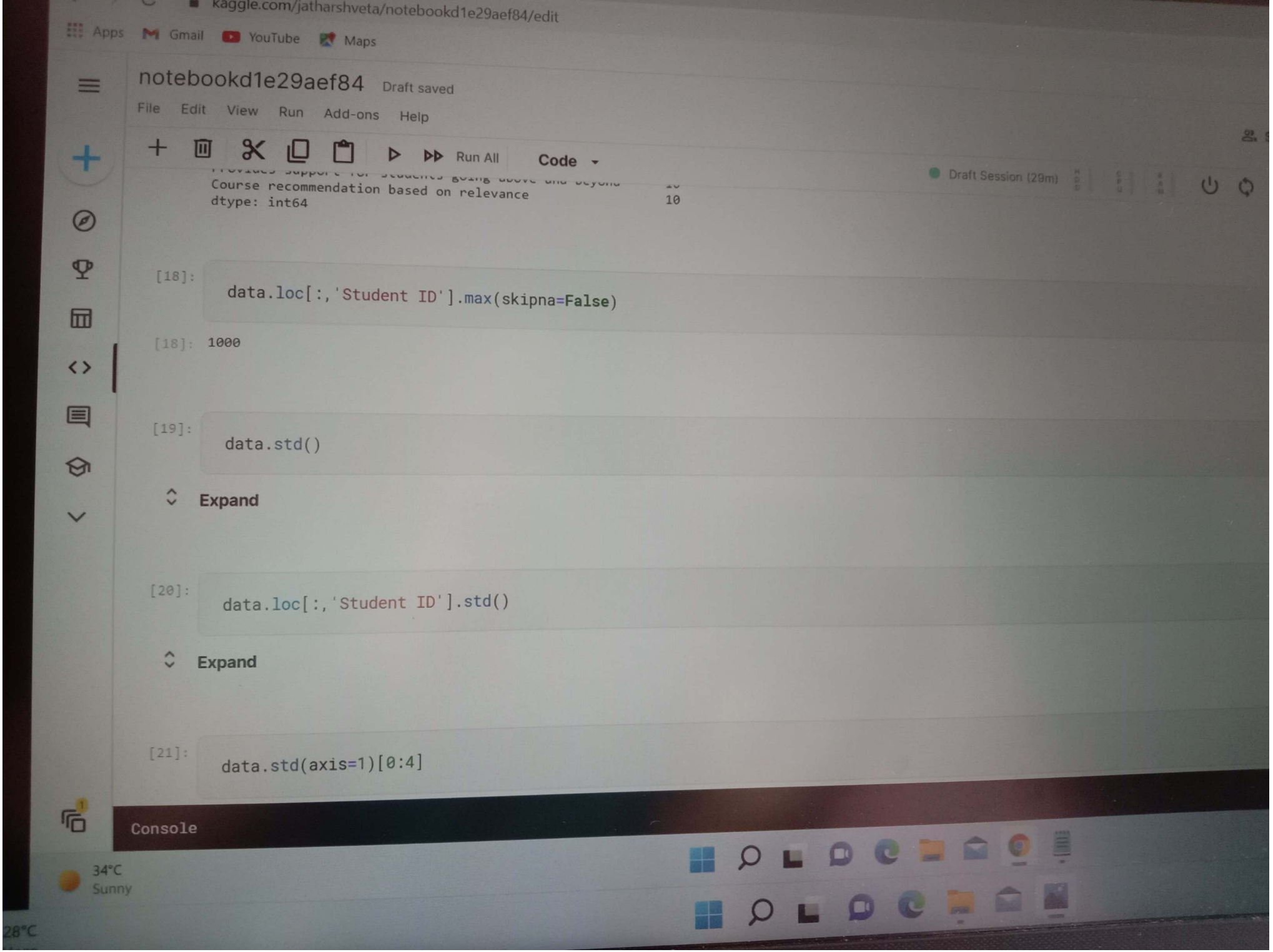
[15]:
```
0
```

[16]:
```
data.max()
```

[16]:
```
Unnamed: 0                                                  1000
Student ID                                                  1000
Well versed with the subject                                  10
Explains concepts in an understandable way                    10
Use of presentations                                           8
Degree of difficulty of assignments                           10
Solves doubts willingly                                       10
Structuring of the course                                     10
Provides support for students going above and beyond          10
Course recommendation based on relevance                      10
dtype: int64
```

Console

34°C



Apps Gmail YouTube Maps

notebookd1e29aef84 Draft saved

File Edit View Run Add-ons Help

Run All Code

Draft Session (30m)

[21]:
```
0    106.157116
1     78.652612
2    213.745461
3    253.135098
dtype: float64
```

[23]:
```
data.groupby(['Student ID'])['Use of presentations'].mean()
```

[23]:
```
Student ID
0       8.0
1       8.0
2       7.0
3       5.0
4       5.0
       ...
996     6.0
997     8.0
998     5.0
999     8.0
1000    5.0
Name: Use of presentations, Length: 1001, dtype: float64
```

[27]:
```
data_u=data.rename(columns={'Use of presentations':'Representation'},inplace=False)
```

Console

34°C Sunny

28°C Haze

notebookd1e29aef84 | Kaggle

kaggle.com/jatharshveta/notebookd1e29aef84/edit

Apps Gmail YouTube Maps

notebookd1e29aef84 Draft saved

File Edit View Run Add-ons Help

Run All Code

Draft Session (30m)

[29]:
```
data_u.groupby(['Student ID']).Representation.mean()
```

[29]:
```
Student ID
0       8.0
1       8.0
2       7.0
3       5.0
4       5.0
       ...
996     6.0
997     8.0
998     5.0
999     8.0
1000    5.0
Name: Representation, Length: 1001, dtype: float64
```

[33]:
```
from sklearn import preprocessing
enc=preprocessing.OneHotEncoder()
enc_data=pd.DataFrame(enc.fit_transform([['Student ID']]).toarray())
enc_data
```

[33]:

	0
0	1.0

Console

34°C Sunny

kaggle.com/jatharshveta/notebookd1e29aef84/edit

Apps Gmail YouTube Maps

notebookd1e29aef84 Draft saved

File Edit View Run Add-ons Help

Run All Code Share

Draft Session (30m)

[34]:
```
data_encode=data_u.join(enc_data)
data_encode
```

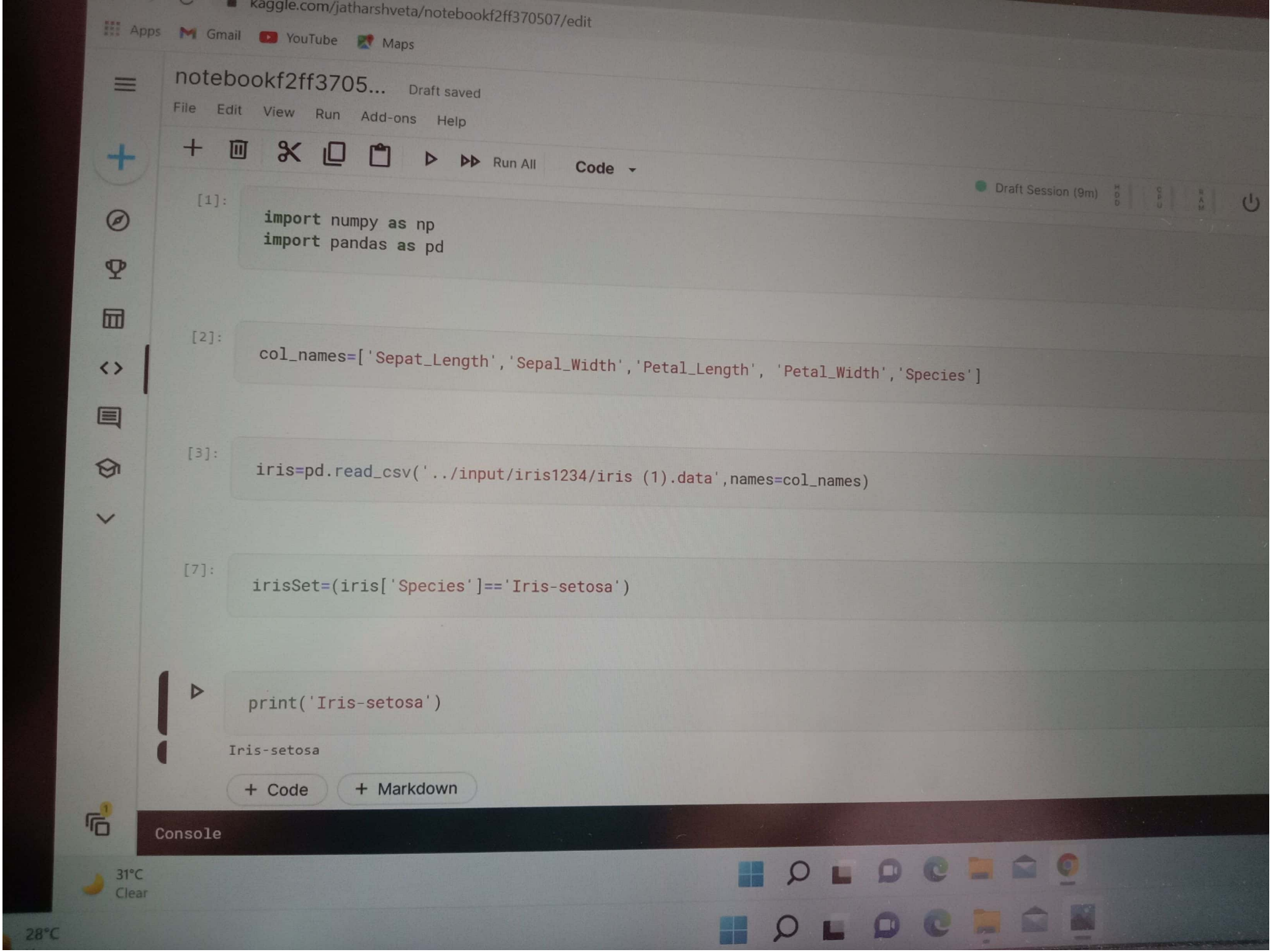
[34]:

	Unnamed: 0	Student ID	Well versed with the subject	Explains concepts in an understandable way	Representation	Degree of difficulty of assignments	Solves doubts willingly	Structuring of the course	Provides support for students going above and beyond	Course recommendation based on relevance	0
0	0	340	5	2	7	6	9	2	1	8	1.0
1	1	253	6	5	8	6	2	1	2	9	NaN
2	2	680	7	7	6	5	4	2	3	1	NaN
3	3	806	9	6	7	1	5	9	4	6	NaN
4	4	632	8	10	8	4	6	6	9	9	NaN
...	...	...	...	...	...	...	...	...	...	...	...
996	996	55	8	7	6	2	5	7	7	9	NaN
997	997	913	5	5	6	5	6	7	6	1	NaN
998	998	199	9	5	8	3	8	1	1	2	NaN
999	999	539	10	2	7	4	3	4	10	1	NaN
1000	1000	759	7	2	4	2	1	5	9	9	NaN

1001 rows × 11 columns

Console

34°C



Apps Gmail YouTube Maps

notebookf2ff3705... Draft saved

File Edit View Run Add-ons Help

Run All Code

Draft Session (9m)

[10]:
```
print(iris[irisSet].describe())
```

	Sepat_Length	Sepal_Width	Petal_Length	Petal_Width
count	50.00000	50.000000	50.000000	50.00000
mean	5.00600	3.418000	1.464000	0.24400
std	0.35249	0.381024	0.173511	0.10721
min	4.30000	2.300000	1.000000	0.10000
25%	4.80000	3.125000	1.400000	0.20000
50%	5.00000	3.400000	1.500000	0.20000
75%	5.20000	3.675000	1.575000	0.30000
max	5.80000	4.400000	1.900000	0.60000

[11]:
```
irisVer=(iris['Species']=='Iris-versicolor')
```

[12]:
```
print('Iris-versicolor')
```

```
Iris-versicolor
```

[13]:
```
print(iris[irisVer].describe())
```

```
       Sepat_Length  Sepal_Width  Petal_Length  Petal_Width
```

Console

31°C Clear

28°C Haze

kaggle.com/jatharshveta/notebookf2ff370507/edit

Apps Gmail YouTube Maps

notebookf2ff3705... Draft saved

File Edit View Run Add-ons Help

Run All Code

Draft Session (9m)

	Sepat_Length	Sepal_Width	Petal_Length	Petal_Width
count	50.000000	50.000000	50.000000	50.000000
mean	5.936000	2.770000	4.260000	1.326000
std	0.516171	0.313798	0.469911	0.197753
min	4.900000	2.000000	3.000000	1.000000
25%	5.600000	2.525000	4.000000	1.200000
50%	5.900000	2.800000	4.350000	1.300000
75%	6.300000	3.000000	4.600000	1.500000
max	7.000000	3.400000	5.100000	1.800000

[14]:
```
irisVir=(iris['Species']=='Iris-virginica')
```

Expand

Iris-virginica

[16]:
```
print(iris[irisVir].describe())
```

	Sepat_Length	Sepal_Width	Petal_Length	Petal_Width
count	50.00000	50.000000	50.000000	50.00000
mean	6.58800	2.974000	5.552000	2.02600
std	0.63588	0.322497	0.551895	0.27465
min	4.90000	2.200000	4.500000	1.40000
25%	6.22500	2.800000	5.100000	1.80000
50%	6.50000	3.000000	5.550000	2.00000

Console

31°C Clear

28°C

notebookf2ff370507 | Kaggle

(3) WhatsApp

kaggle.com/jatharshveta/notebookf2ff370507/edit

Apps Gmail YouTube Maps

notebookf2ff3705... Draft saved

File Edit View Run Add-ons Help

Run All Code

Draft Session (9m)

```
25%     5.600000    2.525000    4.000000    1.200000
50%     5.900000    2.800000    4.350000    1.300000
75%     6.300000    3.000000    4.600000    1.500000
max     7.000000    3.400000    5.100000    1.800000
```

[14]:

```
irisVir=(iris['Species']=='Iris-virginica')
```

Expand

```
Iris-virginica
```

[16]:

```
print(iris[irisVir].describe())
```

```
       Sepat_Length  Sepal_Width  Petal_Length  Petal_Width
count      50.00000    50.000000     50.000000     50.00000
mean        6.58800     2.974000      5.552000      2.02600
std         0.63588     0.322497      0.551895      0.27465
min         4.90000     2.200000      4.500000      1.40000
25%         6.22500     2.800000      5.100000      1.80000
50%         6.50000     3.000000      5.550000      2.00000
75%         6.90000     3.175000      5.875000      2.30000
max         7.90000     3.800000      6.900000      2.50000
```

Console

31°C Clear