Assignment Date	28 September 2022
Student Name	M Hari Haran
Student Roll Number	913119104028
Maximum Marks	2 Marks

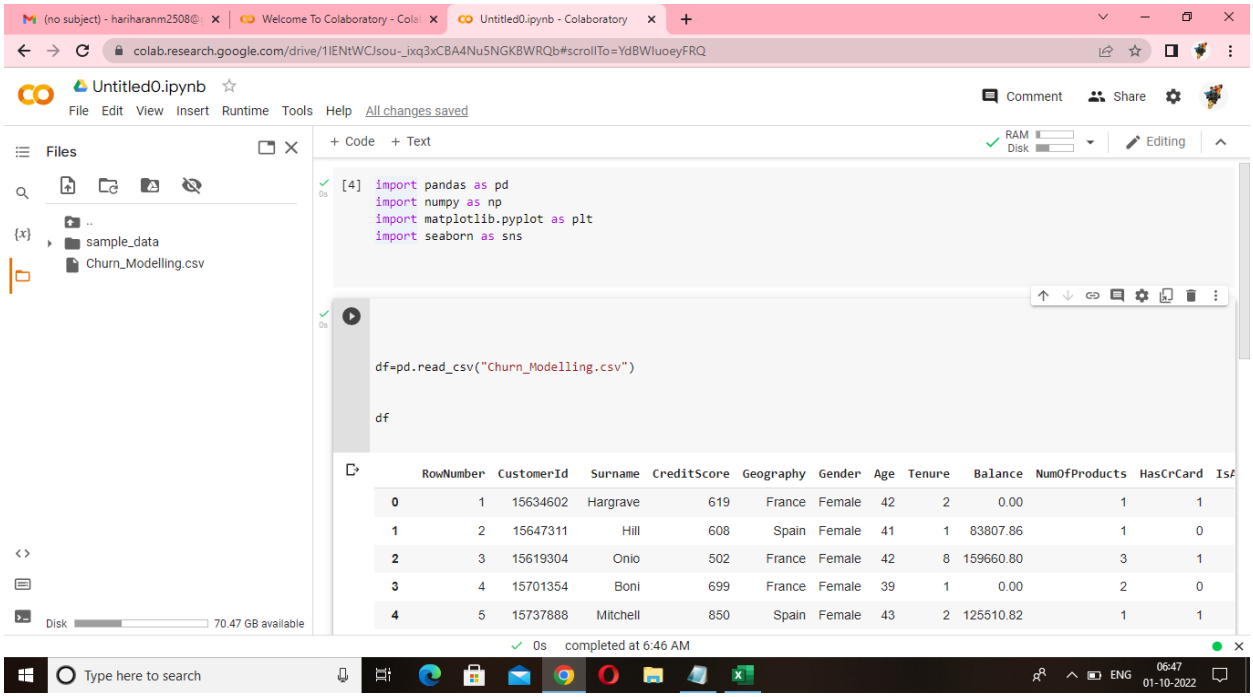
ASSIGNMENT 2

Question-1:

Download the dataset

Question-2:

Load the dataset :



	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure	Balance	NumOfProducts	HasCrCard	IsA
0	1	15634602	Hargrave	619	France	Female	42	2	0.00	1	1	
1	2	15647311	Hill	608	Spain	Female	41	1	83807.86	1	0	
2	3	15619304	Onio	502	France	Female	42	8	159660.80	3	1	
3	4	15701354	Boni	699	France	Female	39	1	0.00	2	0	
4	5	15737888	Mitchell	850	Spain	Female	43	2	125510.82	1	1	

Question-3:

Perform Below Visualizations.

Univariate Analysis

Bi – Variate Analysis

Mult – Variate Analysis

(no subject) - hariharanm2508@... | Welcome To Colaboratory - Cola... | Untitled0.ipynb - Colaboratory

colab.research.google.com/drive/1IENtWCJsou-_ixq3xCBA4Nu5NGKBWRQb#scrollTo=zFDYNcoMzIT9

Untitled0.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

Files

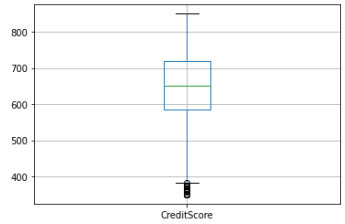
+ Code + Text

RAM Disk Editing

[6]

```
df.boxplot("CreditScore")
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f435da76a50>
```



sample_data

Churn_Modelling.csv

Disk 70.47 GB available

0s completed at 6:49 AM

Type here to search

ENG 06:49 01-10-2022

(no subject) - hariharanm2508@... | Welcome To Colaboratory - Cola... | Untitled0.ipynb - Colaboratory

colab.research.google.com/drive/1IENtWCJsou-_ixq3xCBA4Nu5NGKBWRQb#scrollTo=GwqZ1ZOs0Dz7

Untitled0.ipynb

File Edit View Insert Runtime Tools Help Saving...

Comment Share

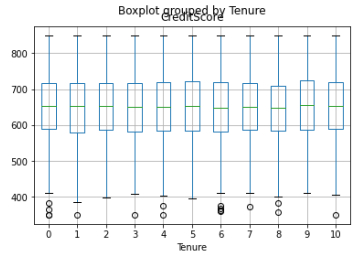
Files

+ Code + Text

RAM Disk Editing

```
df.boxplot("CreditScore","Tenure")
```

```
/usr/local/lib/python3.7/dist-packages/matplotlib/cbook/__init__.py:1376: VisibleDeprecationWarning: Creating an ndarray fr
  X = np.atleast_1d(X.T if isinstance(X, np.ndarray) else np.asarray(X))
<matplotlib.axes._subplots.AxesSubplot at 0x7f4373abb090>
```



sample_data

Churn_Modelling.csv

Disk 70.47 GB available

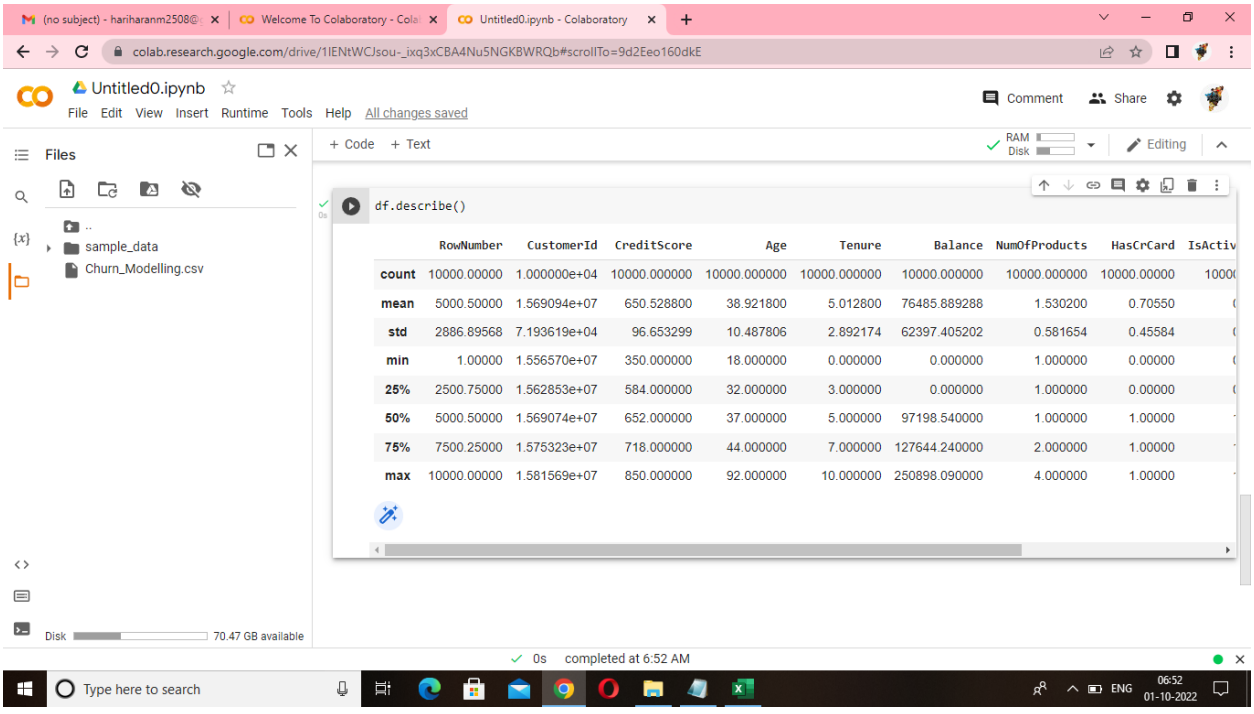
0s completed at 6:50 AM

Type here to search

ENG 06:50 01-10-2022

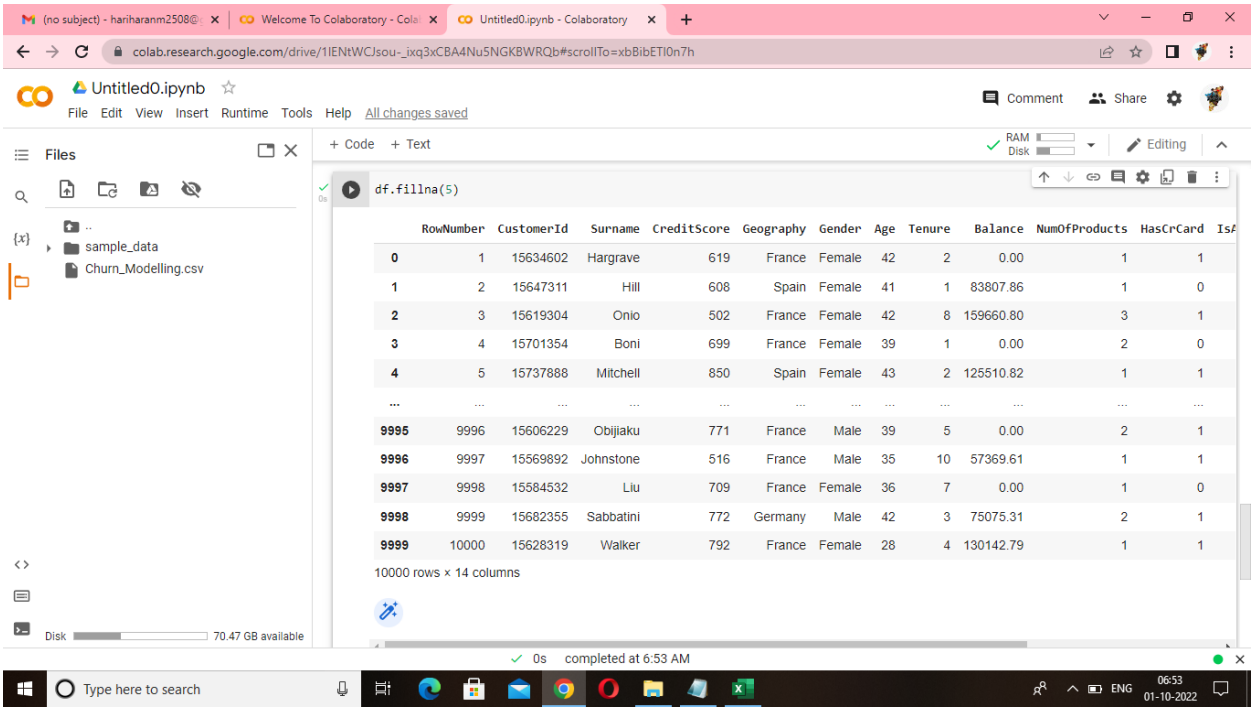
Question – 4

Perform descriptive statistics on the dataset



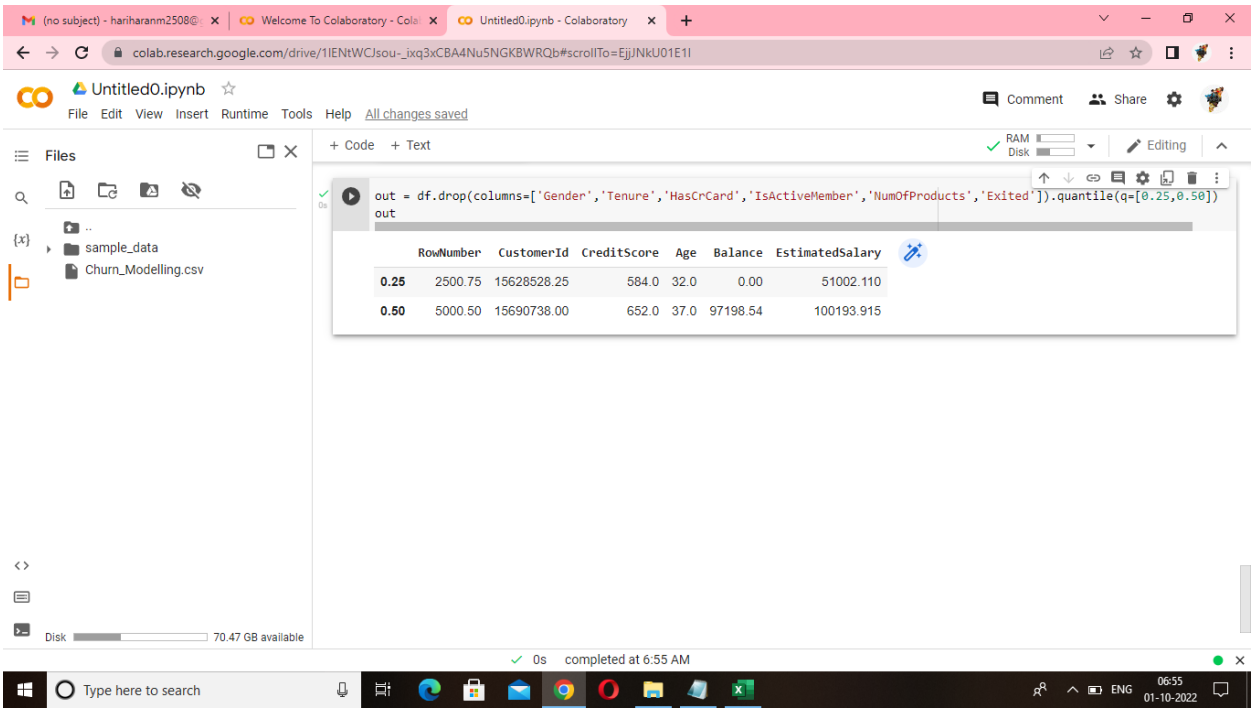
Question-5

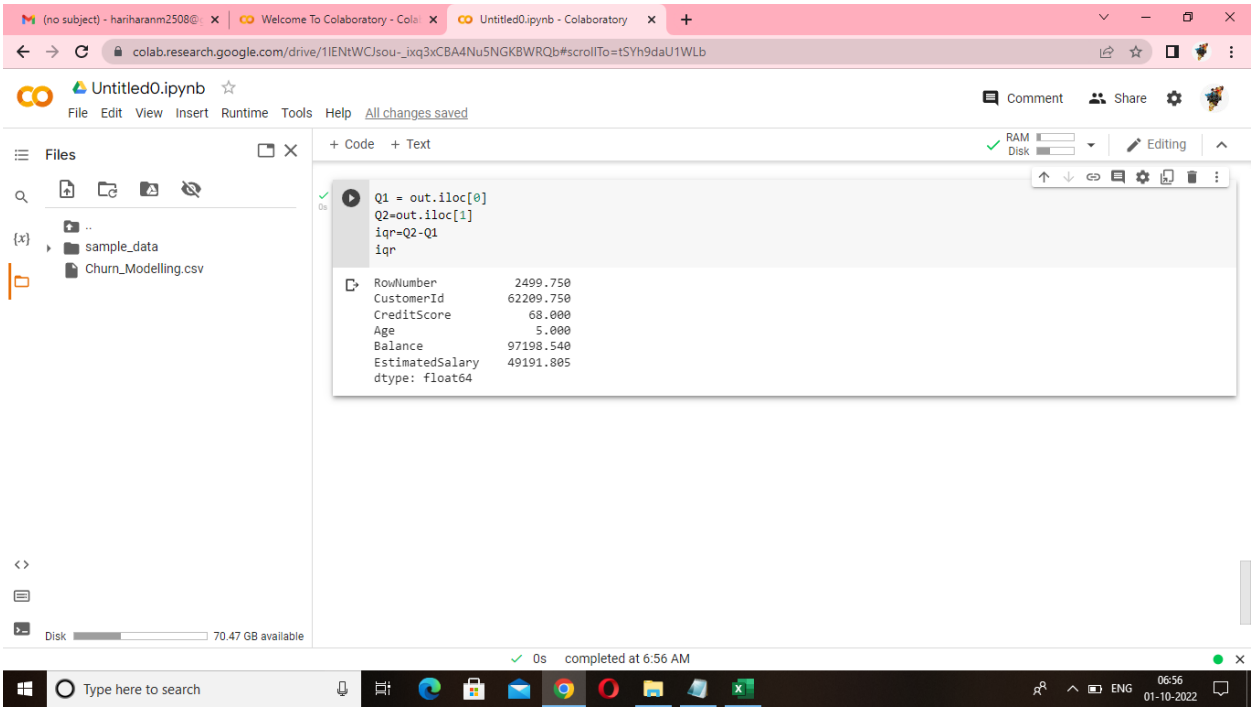
Handle the Missing Values

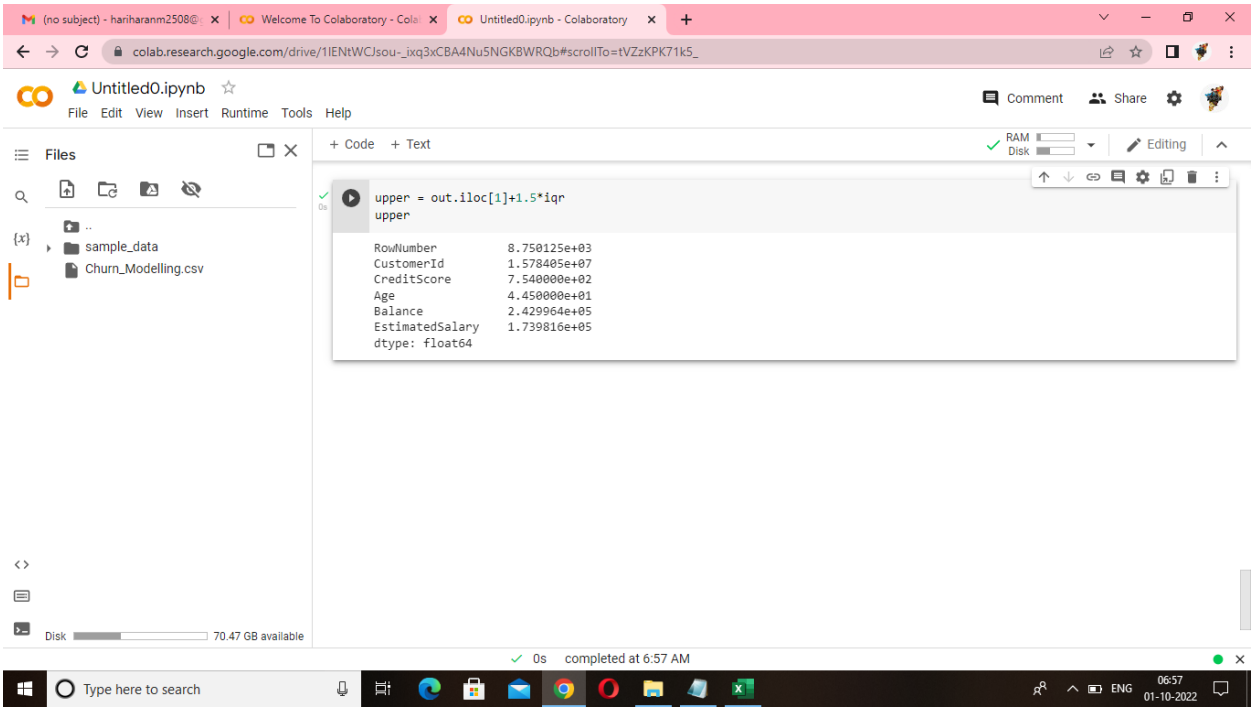


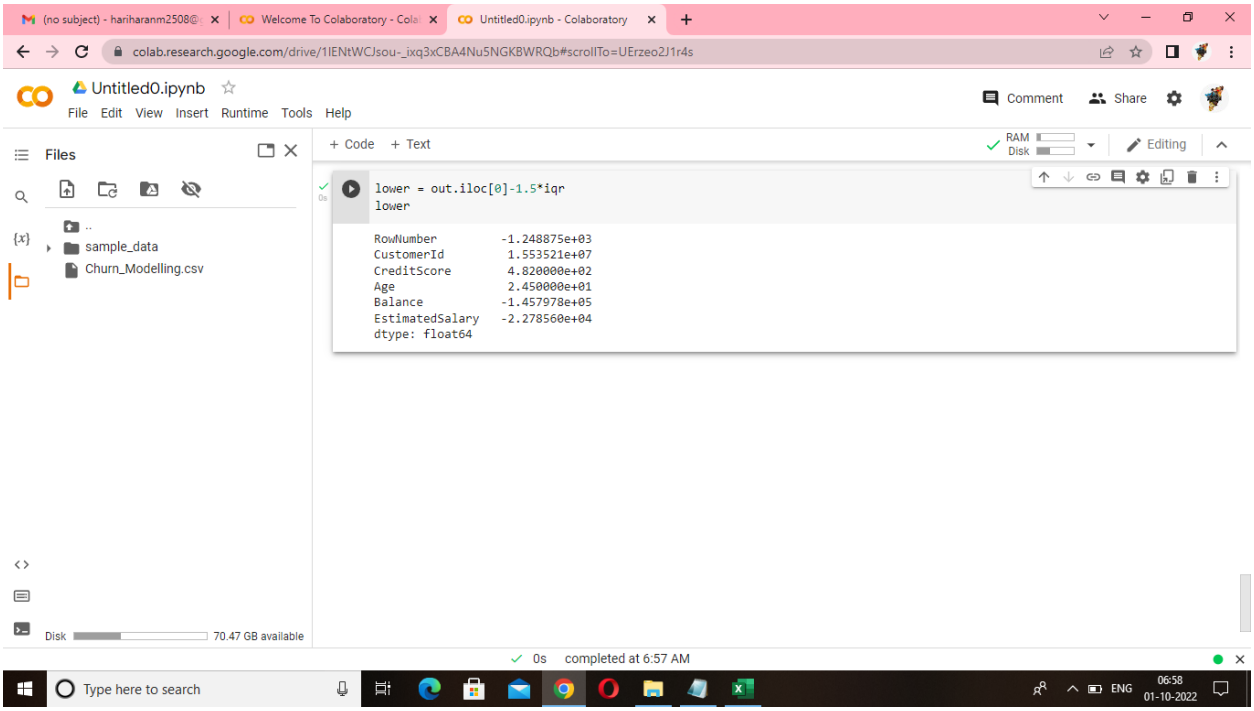
Question-6:

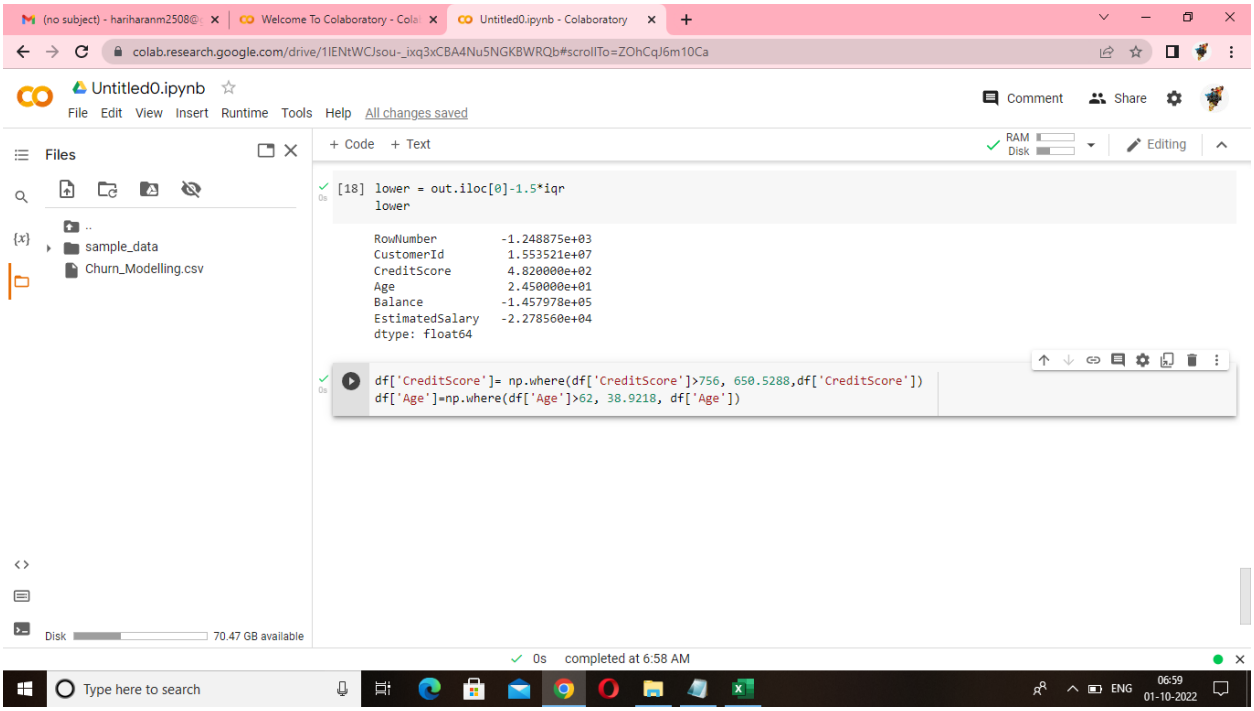
Find the outline & replace the outliners




Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1IENtWCJsou-_ixq3xCBA4Nu5NGKBWRQb#scrollTo=tSYh9daU1WLb
Untitled0.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
Editing
sample_data
Churn_Modelling.csv
Q1 = out.iloc[0]
Q2=out.iloc[1]
iqr=Q2-Q1
iqr
RowNumber 2499.750
CustomerId 62209.750
CreditScore 68.000
Age 5.000
Balance 97198.540
EstimatedSalary 49191.805
dtype: float64
Disk 70.47 GB available
0s completed at 6:56 AM
Type here to search
ENG
06:56
01-10-2022

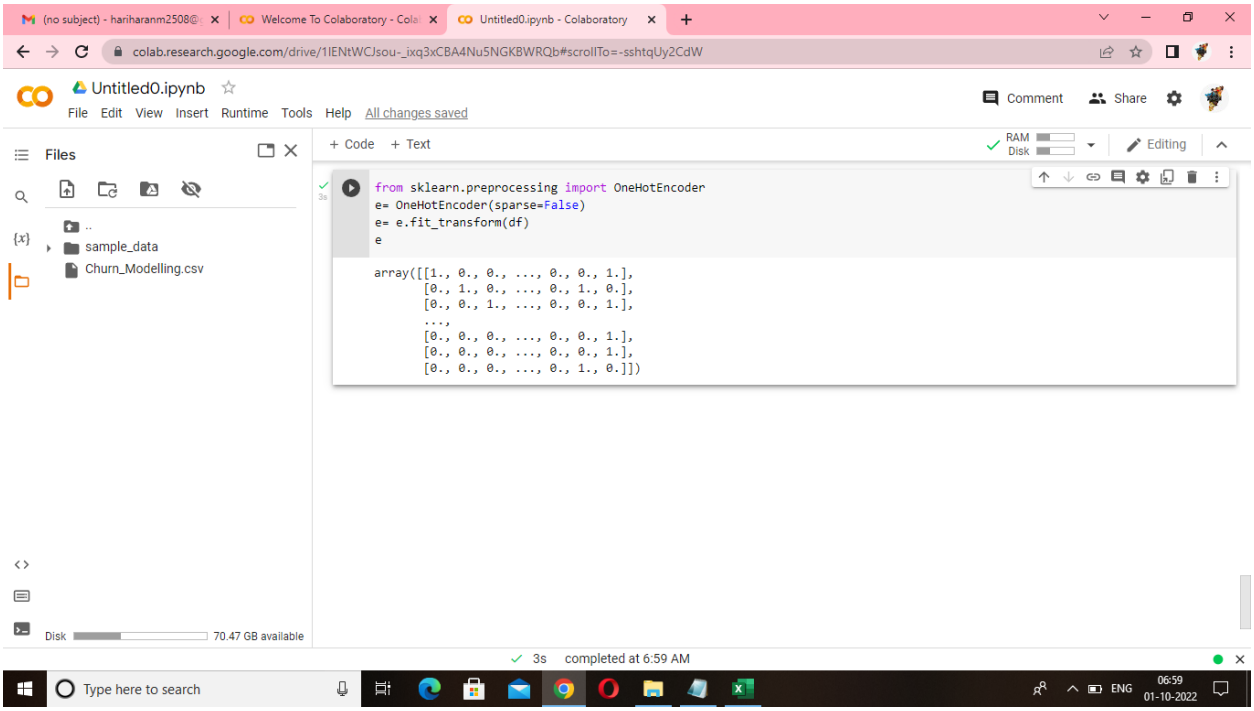

Untitled0.ipynb - Colaboratory
colab.research.google.com/drive/1IENtWCJsou-_ixq3xCBA4Nu5NGKBWRQb#scrollTo=tVZzKPK71k5_
Untitled0.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
Files
+ Code + Text
RAM
Disk
Editing
sample_data
Churn_Modelling.csv
upper = out.iloc[1]+1.5*iqr
upper
RowNumber 8.750125e+03
CustomerId 1.578405e+07
CreditScore 7.540000e+02
Age 4.450000e+01
Balance 2.429964e+05
EstimatedSalary 1.739816e+05
dtype: float64
Disk 70.47 GB available
0s completed at 6:57 AM
Type here to search
ENG
06:57
01-10-2022





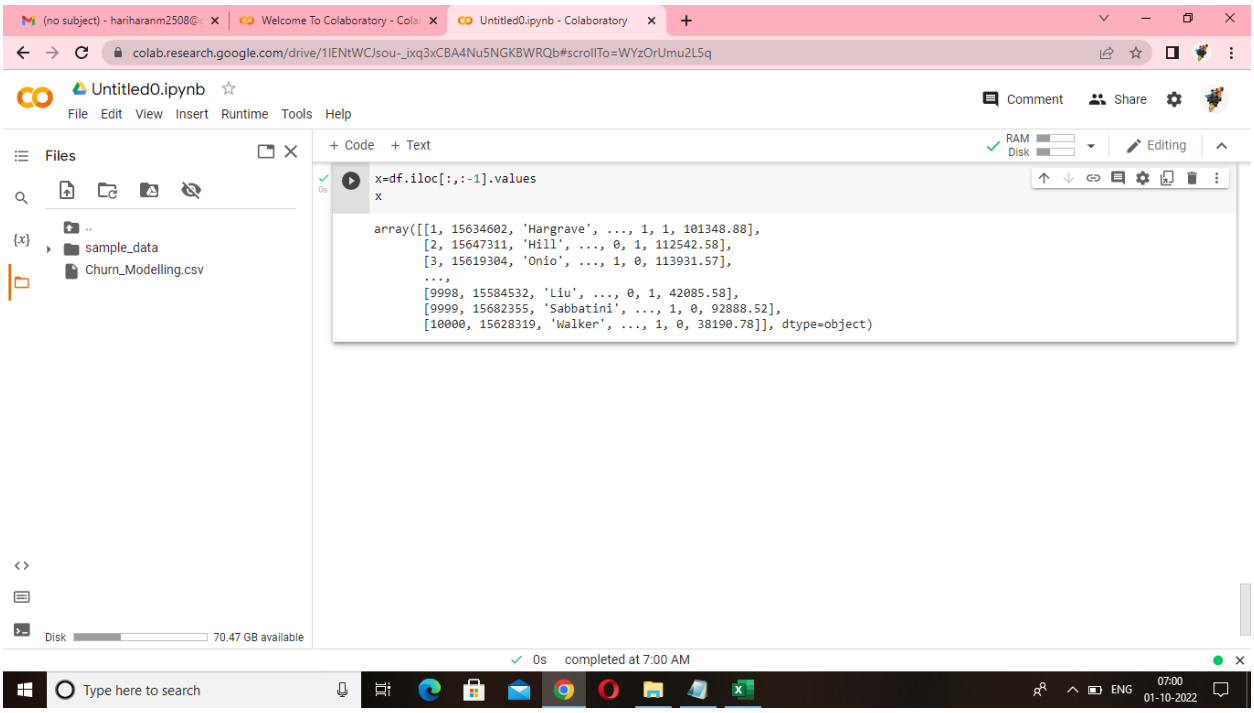
Question-7:

Check for Categorical columns and Perform encoding



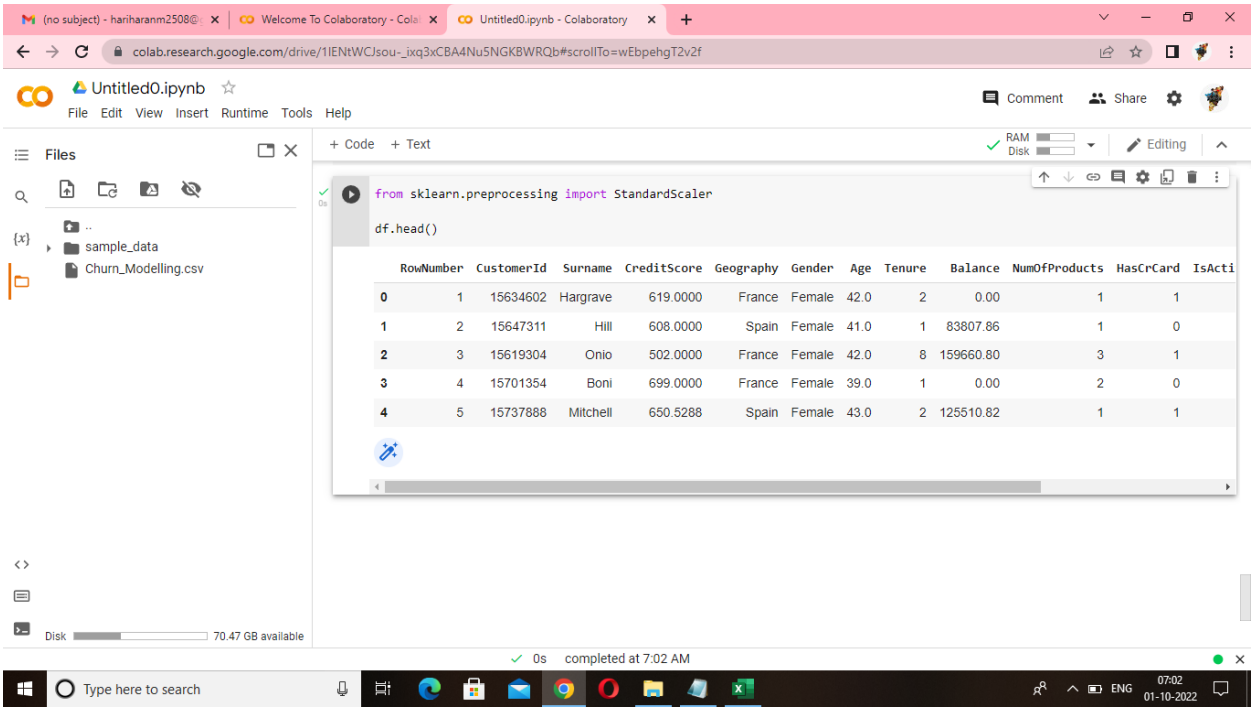
Question-8

Split the data into depentant and independent Variables.



Question-9

Scale the independent Variables



Question-10

Split the data into training & testing