Assignment -2

Python Programming

Assignment Date	25 September 2022
Student Name	Jaswanth Madela
Student Roll Number	720819106042
Maximum Marks	2 Marks

```
import pandas as pd import matplotlib.pyplot
as plt import
seaborn as sns import numpy as np
```

2.Loading the data Set

```
df=pd.read_csv("Churn_Modelling.csv")
df
```

	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender	Age	Tenure
0	1	15634602	Hargrave	619	France	Female	42	2
1	2	15647311	Hill	608	Spain	Female	41	1
2	3	15619304	Onio	502	France	Female	42	8
3	4	15701354	Boni	699	France	Female	39	1
4	5	15737888	Mitchell	850	Spain	Female	43	2
...	...	...	...	...	...	...	...	...
9995	9996	15606229	Obijiaku	771	France	Male	39	5
9996	9997	15569892	Johnstone	516	France	Male	35	10
9997	9998	15584532	Liu	709	France	Female	36	7

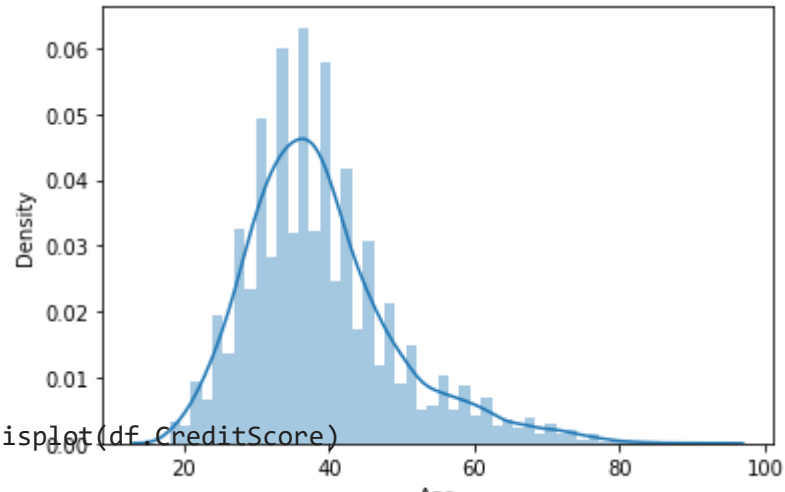
9998	9999	15682355	Sabbatini	772	Germany	Male	42	3
9999	10000	15628319	Walker	792	France	Female	28	4

100 0 14 [redacted]

3.Visulaizatoin 3.1

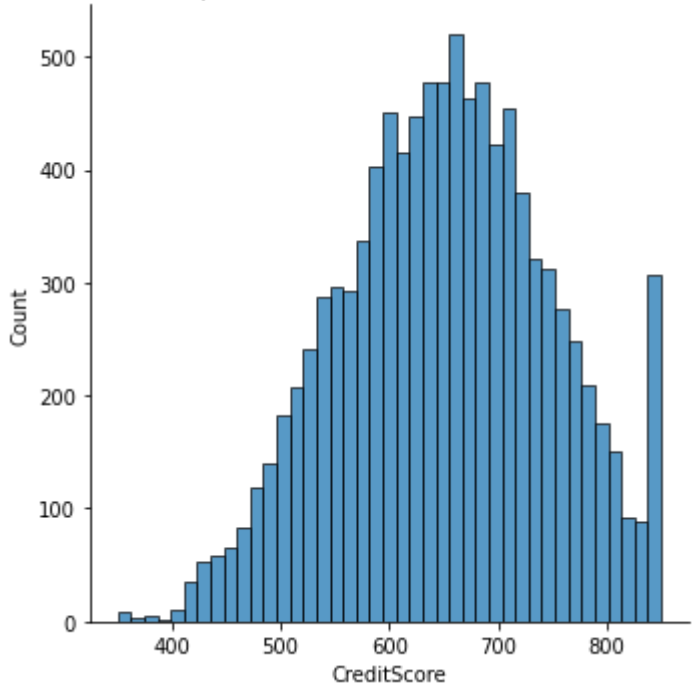
Univariate Analysis

```
sns.distplot(df.Age) C:\ProgramData\Anaconda3\lib\site-
packages\seaborn\distributions.py:2619: FutureWarni   warnings.warn(msg, FutureWarning)
     <AxesSubplot:xlabel='Age', ylabel='Density'>
```




```
sns.displot(df.CreditScore)
```

```
<seaborn.axisgrid.FacetGrid at 0x1d3d827f550>
```

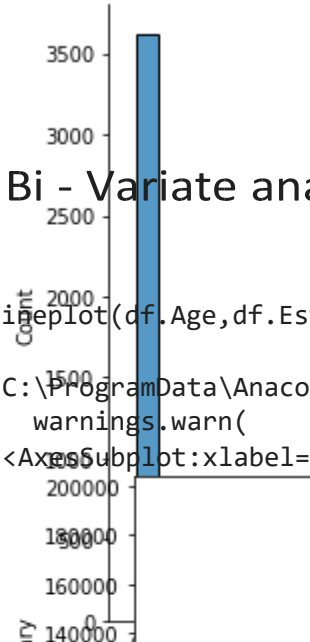



```
sns.displot(df.Balance)
```

```
<seaborn.axisgrid.FacetGrid at 0x1d3d40c6700>
```

```
sns.lineplot(df.Age,df.Balance)
```

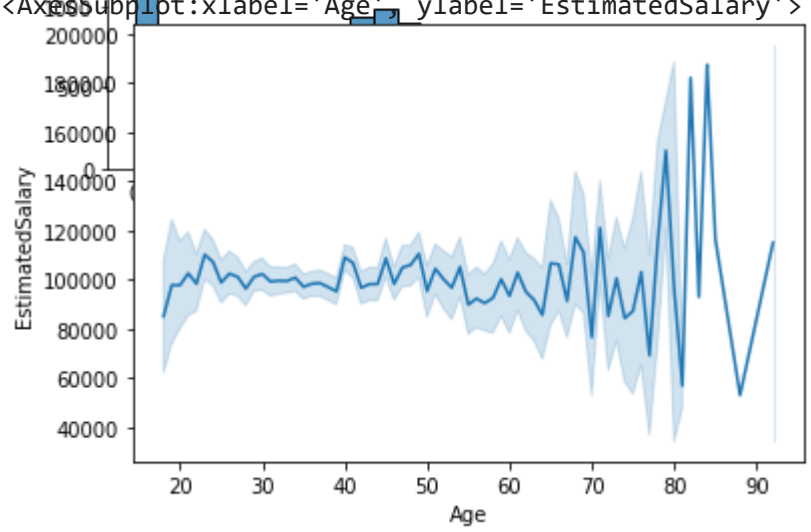
```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning:
```



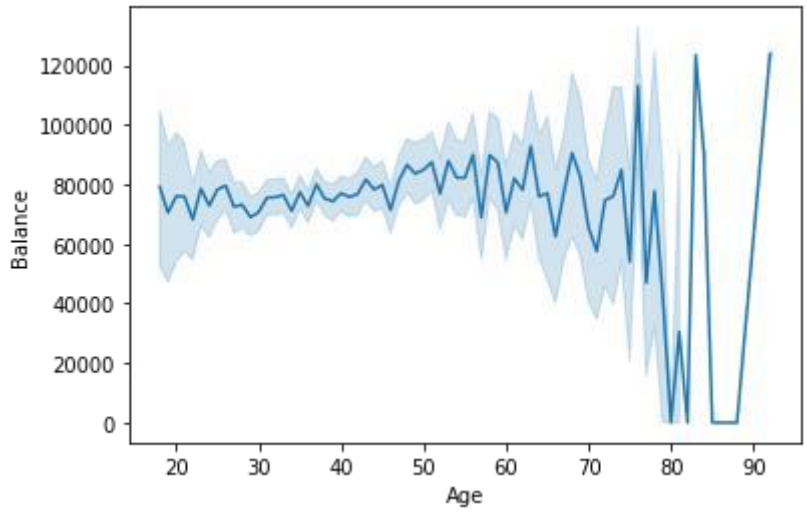
3.2 Bi - Variate analysis

```
sns.lineplot(df.Age,df.EstimatedSalary)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning:
  warnings.warn(
<AxesSubplot:xlabel='Age', ylabel='EstimatedSalary'>
```

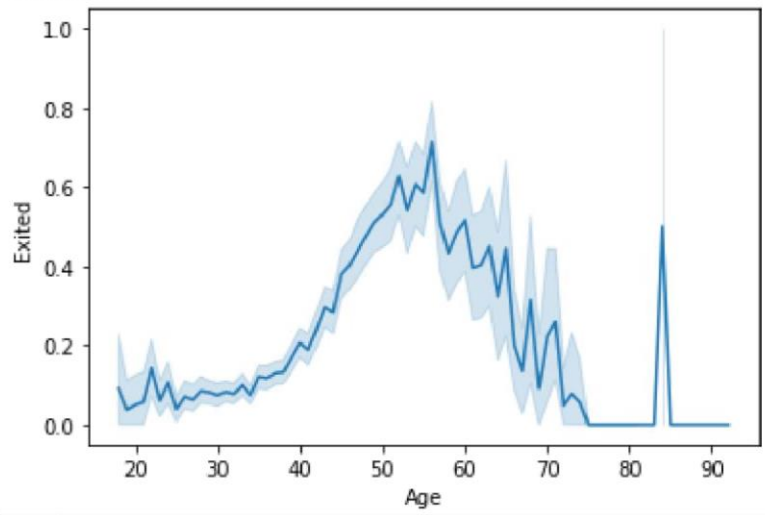



```
  warnings.warn( <AxesSubplot:xlabel='Age',
ylabel='Balance'>
```

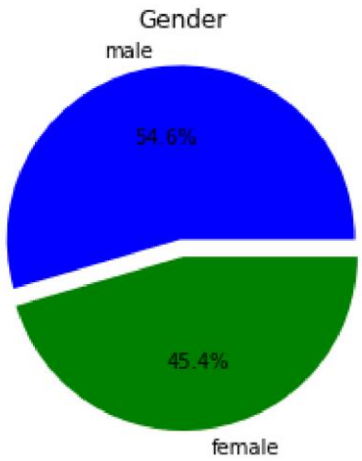



```
sns.lineplot(df.Age,df.Exited)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning:
  warnings.warn(
<AxesSubplot:xlabel='Age', ylabel='Exited'>
```

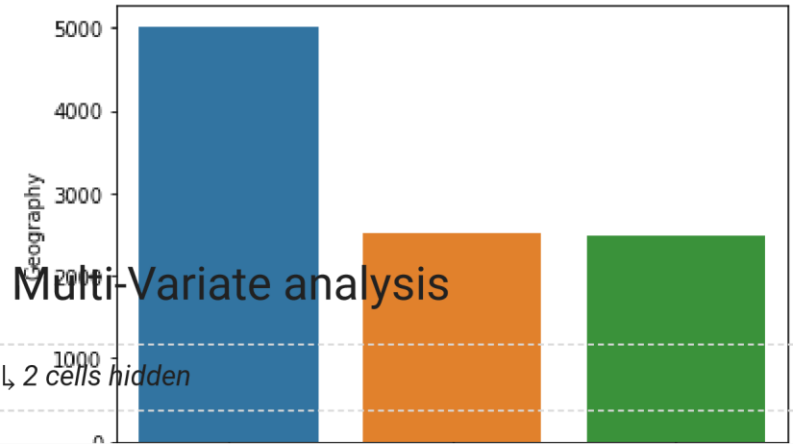



```
plt.pie(df.Gender.value_counts(),[0.1,0],labels=["male","female"],autopct="%1.1f%%",colors
plt.title("Gender")
plt.show()
```




```
sns.barplot(df.Geography.value_counts().index,df.Geography.value_counts())
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning:
       warnings.warn(
```

```
<AxesSubplot:ylabel='Geography'>
```



3.3 Multi-Variate analysis

[ ] ↳ *2 cells hidden*

4.Perform the discriptive Statistics

[ ] ↳ *1 cell hidden*

5.Handling the missing Values

[ ] ↳ *1 cell hidden*

6.Finding Outliers

[ ] ↳ *11 cells hidden*

7. Check for Categorical columns and perform encoding

[ ] ↳ *3 cells hidden*

8. Split the data into dependent and independent variables

[ ] ↳ *4 cells hidden*

9. Scale the independent variables

[ ] ↳ *2 cells hidden*

10. Split the data into training and testing

[ ] ↳ *2 cells hidden*