

```
import numpy as np
import pandas as pd
df=pd.read_csv('/content/Churn_Modelling.csv')
df.head()
```

|   | RowNumber | CustomerId | Surname  | CreditScore | Geography | Gender | Age |
|---|-----------|------------|----------|-------------|-----------|--------|-----|
| 0 | 1         | 15634602   | Hargrave | 619         | France    | Female | 42  |
| 1 | 2         | 15647311   | Hill     | 608         | Spain     | Female | 41  |
| 2 | 3         | 15619304   | Onio     | 502         | France    | Female | 42  |
| 3 | 4         | 15701354   | Boni     | 699         | France    | Female | 39  |
| 4 | 5         | 15737888   | Mitchell | 850         | Spain     | Female | 43  |

|   | Tenure | Balance   | NumOfProducts | HasCrCard | IsActiveMember |  |
|---|--------|-----------|---------------|-----------|----------------|--|
| 0 | 2      | 0.00      | 1             | 1         | 1              |  |
| 1 | 1      | 83807.86  | 1             | 0         | 1              |  |
| 2 | 8      | 159660.80 | 3             | 1         | 0              |  |
| 3 | 1      | 0.00      | 2             | 0         | 0              |  |
| 4 | 2      | 125510.82 | 1             | 1         | 1              |  |

|   | EstimatedSalary | Exited |
|---|-----------------|--------|
| 0 | 101348.88       | 1      |
| 1 | 112542.58       | 0      |
| 2 | 113931.57       | 1      |
| 3 | 93826.63        | 0      |
| 4 | 79084.10        | 0      |

```
df
```

|       | RowNumber | CustomerId | Surname  | CreditScore | Geography | Gender |
|-------|-----------|------------|----------|-------------|-----------|--------|
| Age \ |           |            |          |             |           |        |
| 0     | 1         | 15634602   | Hargrave | 619         | France    | Female |
| 42    |           |            |          |             |           |        |
| 1     | 2         | 15647311   | Hill     | 608         | Spain     | Female |
| 41    |           |            |          |             |           |        |
| 2     | 3         | 15619304   | Onio     | 502         | France    | Female |
| 42    |           |            |          |             |           |        |
| 3     | 4         | 15701354   | Boni     | 699         | France    | Female |
| 39    |           |            |          |             |           |        |
| 4     | 5         | 15737888   | Mitchell | 850         | Spain     | Female |
| 43    |           |            |          |             |           |        |
| ...   | ...       | ...        | ...      | ...         | ...       | ...    |
| ...   |           |            |          |             |           |        |

|            |       |          |           |     |         |        |
|------------|-------|----------|-----------|-----|---------|--------|
| 9995<br>39 | 9996  | 15606229 | Obijiaku  | 771 | France  | Male   |
| 9996<br>35 | 9997  | 15569892 | Johnstone | 516 | France  | Male   |
| 9997<br>36 | 9998  | 15584532 | Liu       | 709 | France  | Female |
| 9998<br>42 | 9999  | 15682355 | Sabbatini | 772 | Germany | Male   |
| 9999<br>28 | 10000 | 15628319 | Walker    | 792 | France  | Female |

|      | Tenure | Balance   | NumOfProducts | HasCrCard | IsActiveMember | \ |
|------|--------|-----------|---------------|-----------|----------------|---|
| 0    | 2      | 0.00      | 1             | 1         | 1              |   |
| 1    | 1      | 83807.86  | 1             | 0         | 1              |   |
| 2    | 8      | 159660.80 | 3             | 1         | 0              |   |
| 3    | 1      | 0.00      | 2             | 0         | 0              |   |
| 4    | 2      | 125510.82 | 1             | 1         | 1              |   |
| ...  | ...    | ...       | ...           | ...       | ...            |   |
| 9995 | 5      | 0.00      | 2             | 1         | 0              |   |
| 9996 | 10     | 57369.61  | 1             | 1         | 1              |   |
| 9997 | 7      | 0.00      | 1             | 0         | 1              |   |
| 9998 | 3      | 75075.31  | 2             | 1         | 0              |   |
| 9999 | 4      | 130142.79 | 1             | 1         | 0              |   |

|      | EstimatedSalary | Exited |
|------|-----------------|--------|
| 0    | 101348.88       | 1      |
| 1    | 112542.58       | 0      |
| 2    | 113931.57       | 1      |
| 3    | 93826.63        | 0      |
| 4    | 79084.10        | 0      |
| ...  | ...             | ...    |
| 9995 | 96270.64        | 0      |
| 9996 | 101699.77       | 0      |
| 9997 | 42085.58        | 1      |
| 9998 | 92888.52        | 1      |
| 9999 | 38190.78        | 0      |

[10000 rows x 14 columns]

df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):
#   Column              Non-Null Count  Dtype
---  -
0   RowNumber           10000 non-null  int64
1   CustomerId          10000 non-null  int64
2   Surname             10000 non-null  object
3   CreditScore         10000 non-null  int64
```

```
4   Geography      10000 non-null object
5   Gender         10000 non-null object
6   Age            10000 non-null int64
7   Tenure         10000 non-null int64
8   Balance        10000 non-null float64
9   NumOfProducts  10000 non-null int64
10  HasCrCard      10000 non-null int64
11  IsActiveMember 10000 non-null int64
12  EstimatedSalary 10000 non-null float64
13  Exited         10000 non-null int64
```

```
dtypes: float64(2), int64(9), object(3)
```

```
memory usage: 1.1+ MB
```

```
df.isnull().sum()
```

```
RowNumber      0
CustomerId     0
Surname        0
CreditScore    0
Geography      0
Gender         0
Age            0
Tenure         0
Balance        0
NumOfProducts  0
HasCrCard      0
IsActiveMember 0
EstimatedSalary 0
Exited        0
```

```
dtype: int64
```

```
#Visualization
```

```
Univariate Analysis
```

```
import matplotlib.pyplot as plt
```

```
import seaborn as sns
```

```
df['EstimatedSalary'].mean()
```

```
100090.239881
```

```
df['EstimatedSalary'].median()
```

```
100193.915
```

```
df['EstimatedSalary'].std()
```

```
57510.49281769816
```

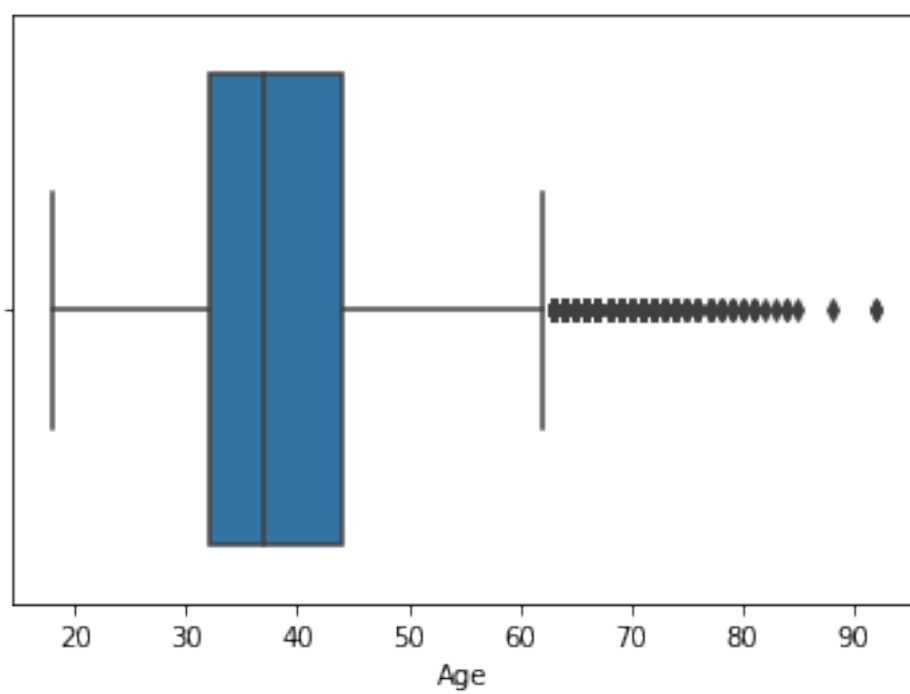
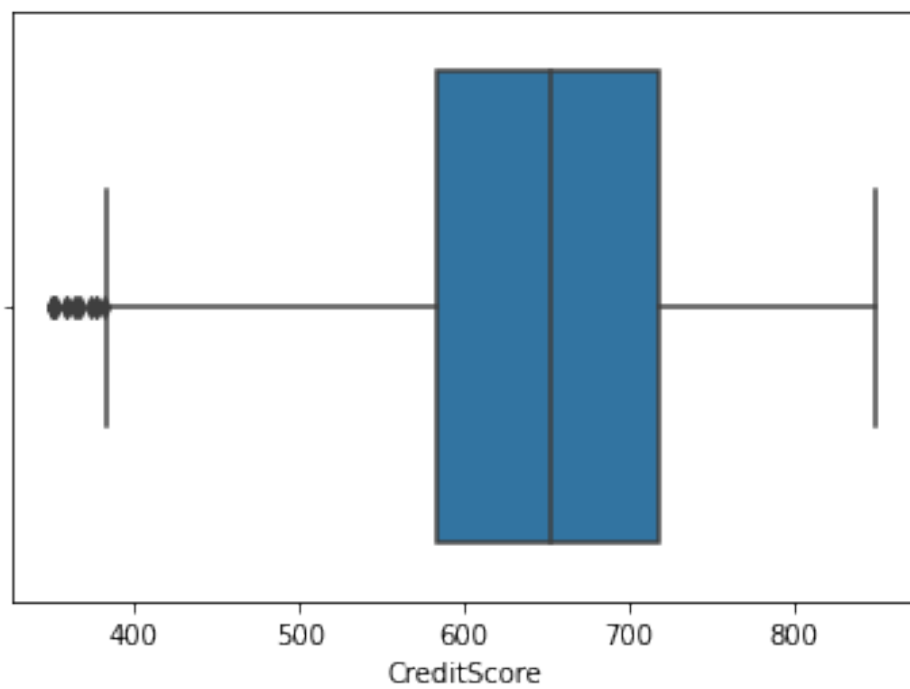
```
df=pd.read_csv('/content/Churn_Modelling.csv')
df=df.iloc[:,3:]
df
```

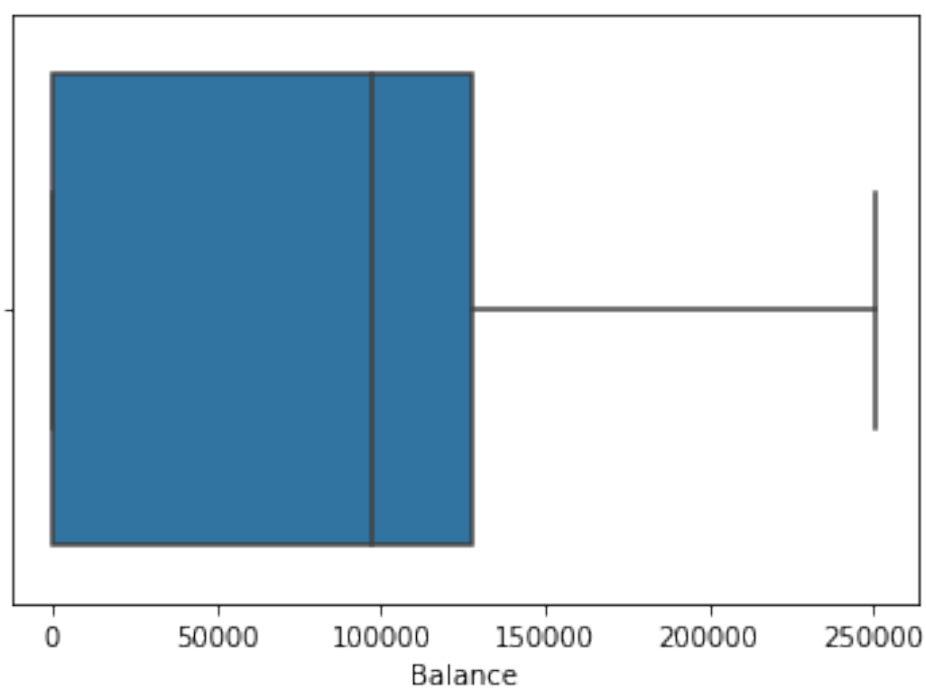
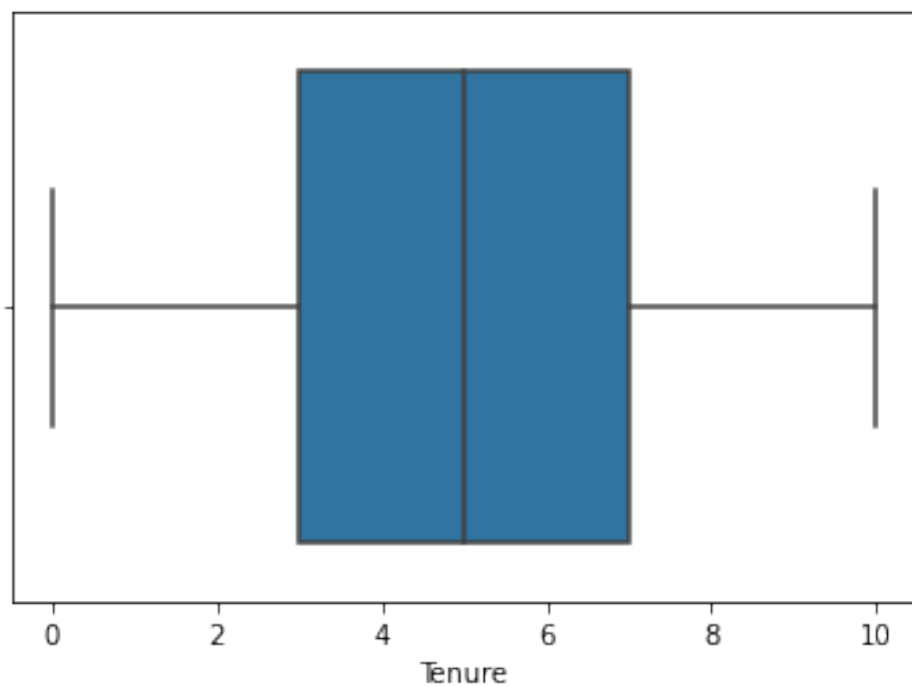
|      | CreditScore | Geography | Gender | Age | Tenure | Balance   |
|------|-------------|-----------|--------|-----|--------|-----------|
| 0    | 619         | France    | Female | 42  | 2      | 0.00      |
| 1    |             |           |        |     |        |           |
| 1    | 608         | Spain     | Female | 41  | 1      | 83807.86  |
| 1    |             |           |        |     |        |           |
| 2    | 502         | France    | Female | 42  | 8      | 159660.80 |
| 3    |             |           |        |     |        |           |
| 3    | 699         | France    | Female | 39  | 1      | 0.00      |
| 2    |             |           |        |     |        |           |
| 4    | 850         | Spain     | Female | 43  | 2      | 125510.82 |
| 1    |             |           |        |     |        |           |
| ...  | ...         | ...       | ...    | ... | ...    | ...       |
| ...  |             |           |        |     |        |           |
| 9995 | 771         | France    | Male   | 39  | 5      | 0.00      |
| 2    |             |           |        |     |        |           |
| 9996 | 516         | France    | Male   | 35  | 10     | 57369.61  |
| 1    |             |           |        |     |        |           |
| 9997 | 709         | France    | Female | 36  | 7      | 0.00      |
| 1    |             |           |        |     |        |           |
| 9998 | 772         | Germany   | Male   | 42  | 3      | 75075.31  |
| 2    |             |           |        |     |        |           |
| 9999 | 792         | France    | Female | 28  | 4      | 130142.79 |
| 1    |             |           |        |     |        |           |

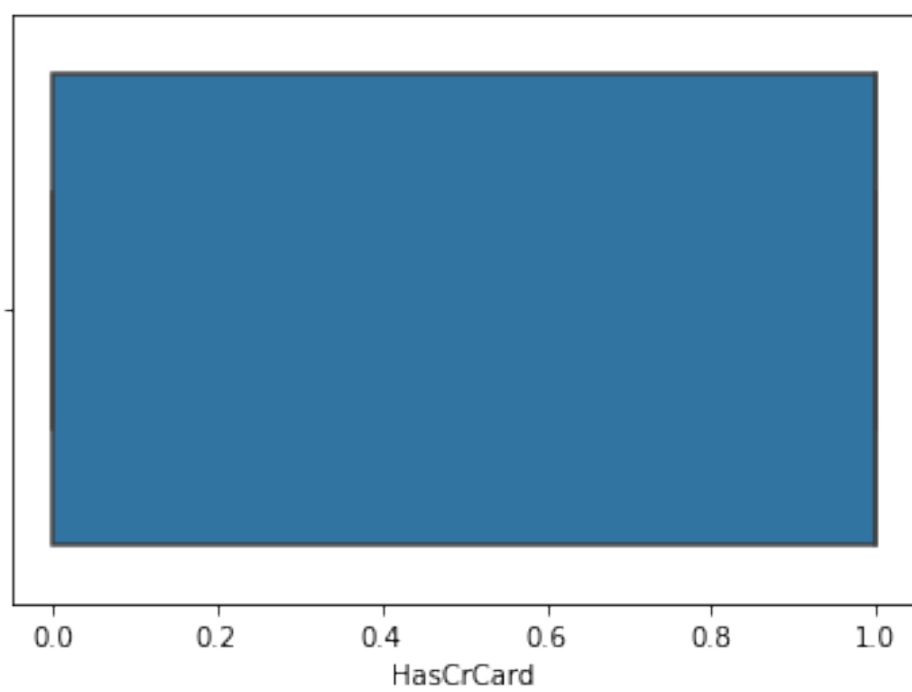
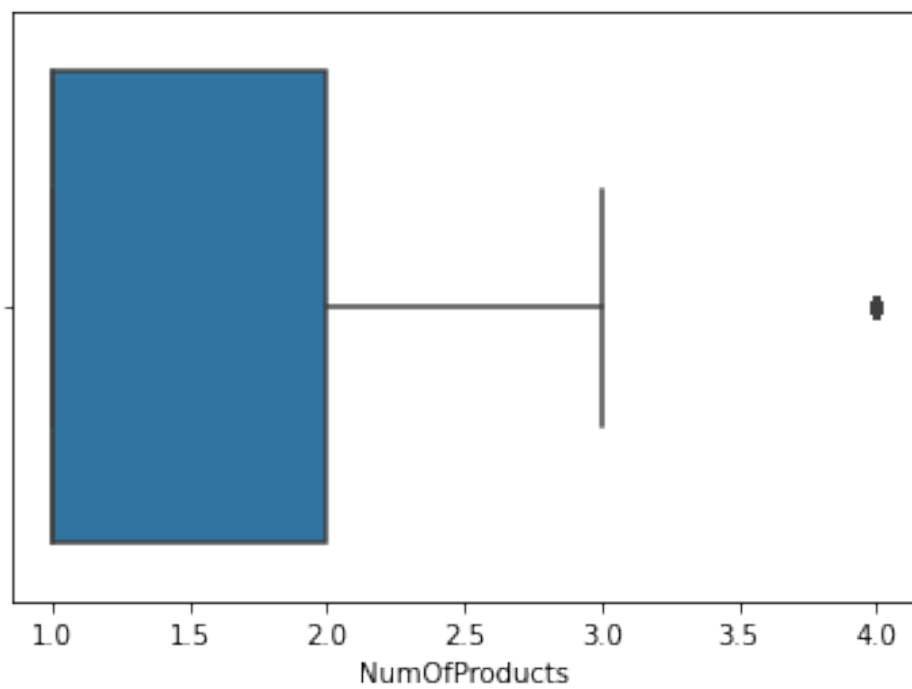
|      | HasCrCard | IsActiveMember | EstimatedSalary | Exited |
|------|-----------|----------------|-----------------|--------|
| 0    | 1         | 1              | 101348.88       | 1      |
| 1    | 0         | 1              | 112542.58       | 0      |
| 2    | 1         | 0              | 113931.57       | 1      |
| 3    | 0         | 0              | 93826.63        | 0      |
| 4    | 1         | 1              | 79084.10        | 0      |
| ...  | ...       | ...            | ...             | ...    |
| 9995 | 1         | 0              | 96270.64        | 0      |
| 9996 | 1         | 1              | 101699.77       | 0      |
| 9997 | 0         | 1              | 42085.58        | 1      |
| 9998 | 1         | 0              | 92888.52        | 1      |
| 9999 | 1         | 0              | 38190.78        | 0      |

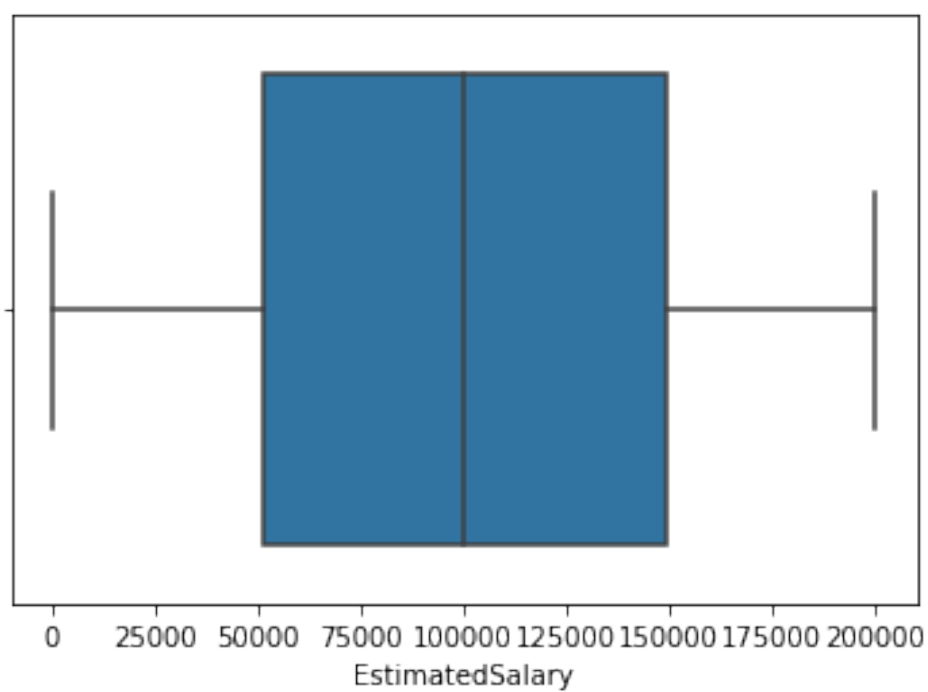
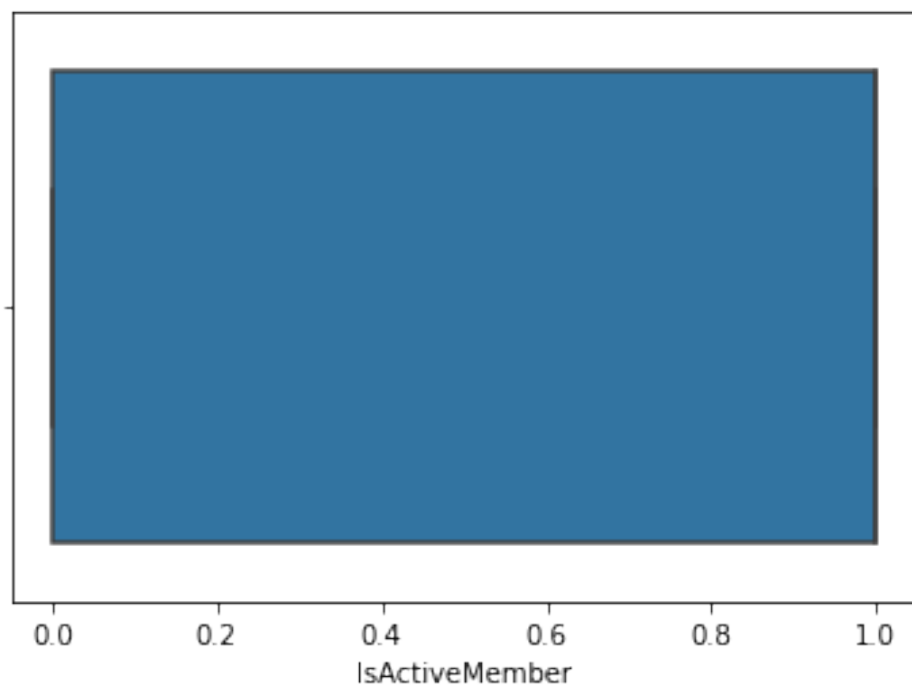
[10000 rows x 11 columns]

```
for col in df.columns:
    if(df.dtypes[col]=='int64' or df.dtypes[col]=='float64' ):
        sns.boxplot(x=df[col]).set( xlabel=col)
        plt.show()
```

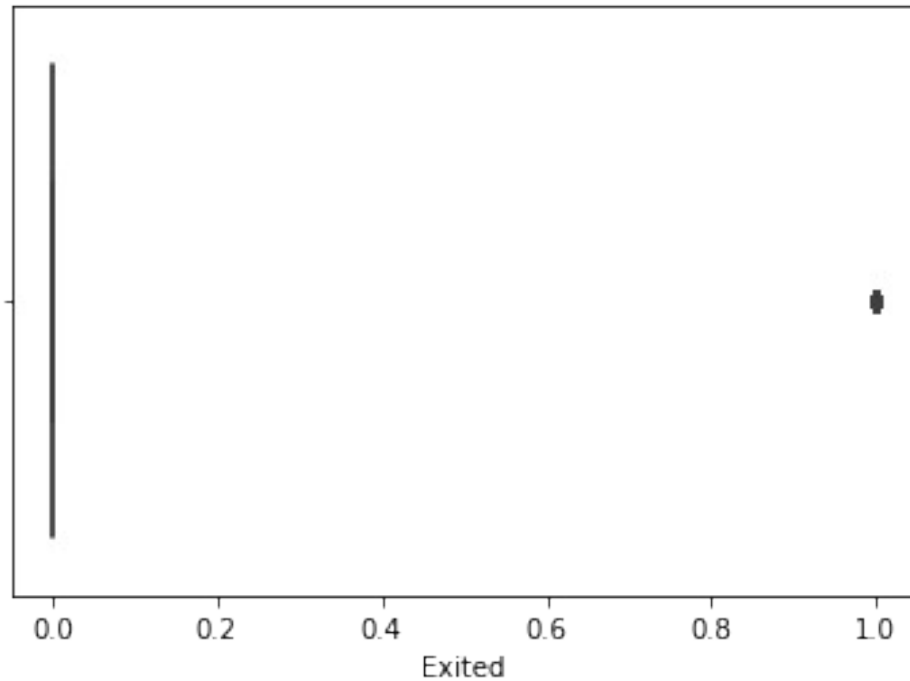








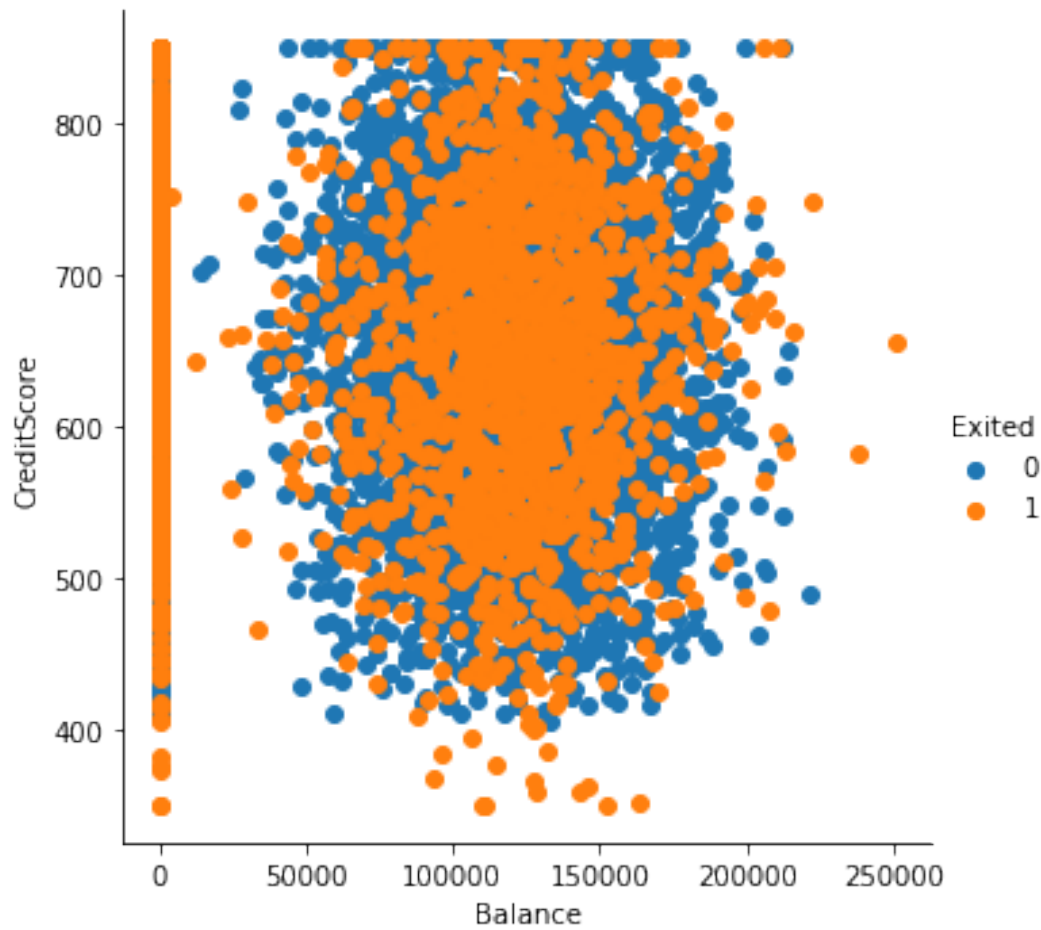




### Bivariate Analysis

```
sns.FacetGrid(df, hue='Exited', size=5).map(plt.scatter, "Balance", "CreditScore").add_legend()  
plt.show()
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/axisgrid.py:337:  
UserWarning: The `size` parameter has been renamed to `height`; please  
update your code.  
warnings.warn(msg, UserWarning)
```



Multivariate Analysis

```
sns.pairplot(df, hue='Exited', height=2)
```

```
<seaborn.axisgrid.PairGrid at 0x7f56d85415d0>
```



#Descriptive statistics

df.head()

|   | CreditScore | Geography | Gender | Age | Tenure | Balance   |
|---|-------------|-----------|--------|-----|--------|-----------|
| 0 | 619         | France    | Female | 42  | 2      | 0.00      |
| 1 | 608         | Spain     | Female | 41  | 1      | 83807.86  |
| 1 | 502         | France    | Female | 42  | 8      | 159660.80 |
| 2 | 699         | France    | Female | 39  | 1      | 0.00      |
| 3 | 850         | Spain     | Female | 43  | 2      | 125510.82 |

HasCrCard IsActiveMember EstimatedSalary Exited

|   |   |   |           |   |
|---|---|---|-----------|---|
| 0 | 1 | 1 | 101348.88 | 1 |
| 1 | 0 | 1 | 112542.58 | 0 |
| 2 | 1 | 0 | 113931.57 | 1 |
| 3 | 0 | 0 | 93826.63  | 0 |
| 4 | 1 | 1 | 79084.10  | 0 |

df.describe()

|                 | CreditScore  | Age          | Tenure       | Balance       |
|-----------------|--------------|--------------|--------------|---------------|
| NumOfProducts \ |              |              |              |               |
| count           | 10000.000000 | 10000.000000 | 10000.000000 | 10000.000000  |
| mean            | 650.528800   | 38.921800    | 5.012800     | 76485.889288  |
| std             | 96.653299    | 10.487806    | 2.892174     | 62397.405202  |
| min             | 350.000000   | 18.000000    | 0.000000     | 0.000000      |
| 25%             | 584.000000   | 32.000000    | 3.000000     | 0.000000      |
| 50%             | 652.000000   | 37.000000    | 5.000000     | 97198.540000  |
| 75%             | 718.000000   | 44.000000    | 7.000000     | 127644.240000 |
| max             | 850.000000   | 92.000000    | 10.000000    | 250898.090000 |

|       | HasCrCard    | IsActiveMember | EstimatedSalary | Exited       |
|-------|--------------|----------------|-----------------|--------------|
| count | 10000.000000 | 10000.000000   | 10000.000000    | 10000.000000 |
| mean  | 0.70550      | 0.515100       | 100090.239881   | 0.203700     |
| std   | 0.45584      | 0.499797       | 57510.492818    | 0.402769     |
| min   | 0.000000     | 0.000000       | 11.580000       | 0.000000     |
| 25%   | 0.000000     | 0.000000       | 51002.110000    | 0.000000     |
| 50%   | 1.000000     | 1.000000       | 100193.915000   | 0.000000     |
| 75%   | 1.000000     | 1.000000       | 149388.247500   | 0.000000     |
| max   | 1.000000     | 1.000000       | 199992.480000   | 1.000000     |

df.describe(include='all')

|          | RowNumber    | CustomerId   | Surname | CreditScore  | Geography |
|----------|--------------|--------------|---------|--------------|-----------|
| Gender \ |              |              |         |              |           |
| count    | 10000.000000 | 1.000000e+04 | 10000   | 10000.000000 | 10000     |
| unique   | NaN          | NaN          | 2932    | NaN          | 3         |
| top      | NaN          | NaN          | Smith   | NaN          | France    |
| freq     | NaN          | NaN          | 32      | NaN          | 5014      |
| mean     | 5000.50000   | 1.569094e+07 | NaN     | 650.528800   | NaN       |

|     |             |              |     |            |     |
|-----|-------------|--------------|-----|------------|-----|
| std | 2886.89568  | 7.193619e+04 | NaN | 96.653299  | NaN |
| NaN |             |              |     |            |     |
| min | 1.000000    | 1.556570e+07 | NaN | 350.000000 | NaN |
| NaN |             |              |     |            |     |
| 25% | 2500.75000  | 1.562853e+07 | NaN | 584.000000 | NaN |
| NaN |             |              |     |            |     |
| 50% | 5000.50000  | 1.569074e+07 | NaN | 652.000000 | NaN |
| NaN |             |              |     |            |     |
| 75% | 7500.25000  | 1.575323e+07 | NaN | 718.000000 | NaN |
| NaN |             |              |     |            |     |
| max | 10000.00000 | 1.581569e+07 | NaN | 850.000000 | NaN |
| NaN |             |              |     |            |     |

|              |              |              |               |               |
|--------------|--------------|--------------|---------------|---------------|
|              | Age          | Tenure       | Balance       | NumOfProducts |
| HasCrCard \  |              |              |               |               |
| count        | 10000.000000 | 10000.000000 | 10000.000000  | 10000.000000  |
| 10000.000000 |              |              |               |               |
| unique       | NaN          | NaN          | NaN           | NaN           |
| NaN          |              |              |               |               |
| top          | NaN          | NaN          | NaN           | NaN           |
| NaN          |              |              |               |               |
| freq         | NaN          | NaN          | NaN           | NaN           |
| NaN          |              |              |               |               |
| mean         | 38.921800    | 5.012800     | 76485.889288  | 1.530200      |
| 0.70550      |              |              |               |               |
| std          | 10.487806    | 2.892174     | 62397.405202  | 0.581654      |
| 0.45584      |              |              |               |               |
| min          | 18.000000    | 0.000000     | 0.000000      | 1.000000      |
| 0.00000      |              |              |               |               |
| 25%          | 32.000000    | 3.000000     | 0.000000      | 1.000000      |
| 0.00000      |              |              |               |               |
| 50%          | 37.000000    | 5.000000     | 97198.540000  | 1.000000      |
| 1.00000      |              |              |               |               |
| 75%          | 44.000000    | 7.000000     | 127644.240000 | 2.000000      |
| 1.00000      |              |              |               |               |
| max          | 92.000000    | 10.000000    | 250898.090000 | 4.000000      |
| 1.00000      |              |              |               |               |

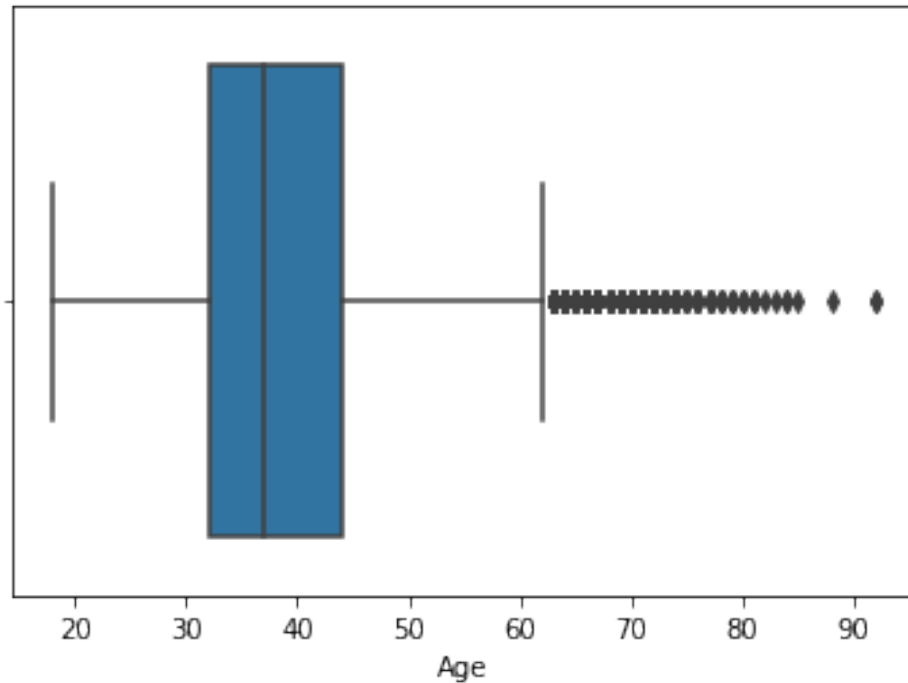
|        |                |                 |              |
|--------|----------------|-----------------|--------------|
|        | IsActiveMember | EstimatedSalary | Exited       |
| count  | 10000.000000   | 10000.000000    | 10000.000000 |
| unique | NaN            | NaN             | NaN          |
| top    | NaN            | NaN             | NaN          |
| freq   | NaN            | NaN             | NaN          |
| mean   | 0.515100       | 100090.239881   | 0.203700     |
| std    | 0.499797       | 57510.492818    | 0.402769     |
| min    | 0.000000       | 11.580000       | 0.000000     |
| 25%    | 0.000000       | 51002.110000    | 0.000000     |
| 50%    | 1.000000       | 100193.915000   | 0.000000     |
| 75%    | 1.000000       | 149388.247500   | 0.000000     |
| max    | 1.000000       | 199992.480000   | 1.000000     |

```
sns.boxplot(df['Age'])
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43:
FutureWarning: Pass the following variable as a keyword arg: x. From
version 0.12, the only valid positional argument will be `data`, and
passing other arguments without an explicit keyword will result in an
error or misinterpretation.
```

```
FutureWarning
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f56dee33e50>
```



```
df['Balance'].groupby(df['EstimatedSalary']).describe()
```

|                 | count | mean      | std | min       | 25%       | 50% \ |
|-----------------|-------|-----------|-----|-----------|-----------|-------|
| EstimatedSalary |       |           |     |           |           |       |
| 11.58           | 1.0   | 122917.71 | NaN | 122917.71 | 122917.71 |       |
| 122917.71       |       |           |     |           |           |       |
| 90.07           | 1.0   | 0.00      | NaN | 0.00      | 0.00      |       |
| 0.00            |       |           |     |           |           |       |
| 91.75           | 1.0   | 121669.93 | NaN | 121669.93 | 121669.93 |       |
| 121669.93       |       |           |     |           |           |       |
| 96.27           | 1.0   | 0.00      | NaN | 0.00      | 0.00      |       |
| 0.00            |       |           |     |           |           |       |
| 106.67          | 1.0   | 0.00      | NaN | 0.00      | 0.00      |       |
| 0.00            |       |           |     |           |           |       |
| ...             | ...   | ...       | ... | ...       | ...       | ..    |
| .               |       |           |     |           |           |       |

|           |     |           |     |           |           |
|-----------|-----|-----------|-----|-----------|-----------|
| 199909.32 | 1.0 | 0.00      | NaN | 0.00      | 0.00      |
| 0.00      |     |           |     |           |           |
| 199929.17 | 1.0 | 178755.84 | NaN | 178755.84 | 178755.84 |
| 178755.84 |     |           |     |           |           |
| 199953.33 | 1.0 | 153325.10 | NaN | 153325.10 | 153325.10 |
| 153325.10 |     |           |     |           |           |
| 199970.74 | 1.0 | 98635.77  | NaN | 98635.77  | 98635.77  |
| 98635.77  |     |           |     |           |           |
| 199992.48 | 1.0 | 0.00      | NaN | 0.00      | 0.00      |
| 0.00      |     |           |     |           |           |

|                 |           |           |
|-----------------|-----------|-----------|
|                 | 75%       | max       |
| EstimatedSalary |           |           |
| 11.58           | 122917.71 | 122917.71 |
| 90.07           | 0.00      | 0.00      |
| 91.75           | 121669.93 | 121669.93 |
| 96.27           | 0.00      | 0.00      |
| 106.67          | 0.00      | 0.00      |
| ...             | ...       | ...       |
| 199909.32       | 0.00      | 0.00      |
| 199929.17       | 178755.84 | 178755.84 |
| 199953.33       | 153325.10 | 153325.10 |
| 199970.74       | 98635.77  | 98635.77  |
| 199992.48       | 0.00      | 0.00      |

[9999 rows x 8 columns]

#Handle missing values

df

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

|            |       |          |           |     |         |        |
|------------|-------|----------|-----------|-----|---------|--------|
| 9998<br>42 | 9999  | 15682355 | Sabbatini | 772 | Germany | Male   |
| 9999<br>28 | 10000 | 15628319 | Walker    | 792 | France  | Female |

|      | Tenure | Balance   | NumOfProducts | HasCrCard | IsActiveMember | \   |
|------|--------|-----------|---------------|-----------|----------------|-----|
| 0    | 2      | 0.00      | 1             | 1         | 1              | 1   |
| 1    | 1      | 83807.86  | 1             | 0         | 1              | 1   |
| 2    | 8      | 159660.80 | 3             | 1         | 0              | 0   |
| 3    | 1      | 0.00      | 2             | 0         | 0              | 0   |
| 4    | 2      | 125510.82 | 1             | 1         | 1              | 1   |
| ...  | ...    | ...       | ...           | ...       | ...            | ... |
| 9995 | 5      | 0.00      | 2             | 1         | 0              | 0   |
| 9996 | 10     | 57369.61  | 1             | 1         | 1              | 1   |
| 9997 | 7      | 0.00      | 1             | 0         | 1              | 1   |
| 9998 | 3      | 75075.31  | 2             | 1         | 0              | 0   |
| 9999 | 4      | 130142.79 | 1             | 1         | 0              | 0   |

|      | EstimatedSalary | Exited |
|------|-----------------|--------|
| 0    | 101348.88       | 1      |
| 1    | 112542.58       | 0      |
| 2    | 113931.57       | 1      |
| 3    | 93826.63        | 0      |
| 4    | 79084.10        | 0      |
| ...  | ...             | ...    |
| 9995 | 96270.64        | 0      |
| 9996 | 101699.77       | 0      |
| 9997 | 42085.58        | 1      |
| 9998 | 92888.52        | 1      |
| 9999 | 38190.78        | 0      |

[10000 rows x 14 columns]

df.isnull()

|       | RowNumber | CustomerId | Surname | CreditScore | Geography | Gender |
|-------|-----------|------------|---------|-------------|-----------|--------|
| Age \ |           |            |         |             |           |        |
| 0     | False     | False      | False   | False       | False     | False  |
| False |           |            |         |             |           |        |
| 1     | False     | False      | False   | False       | False     | False  |
| False |           |            |         |             |           |        |
| 2     | False     | False      | False   | False       | False     | False  |
| False |           |            |         |             |           |        |
| 3     | False     | False      | False   | False       | False     | False  |
| False |           |            |         |             |           |        |
| 4     | False     | False      | False   | False       | False     | False  |
| False |           |            |         |             |           |        |
| ...   | ...       | ...        | ...     | ...         | ...       | ...    |
| ...   |           |            |         |             |           |        |
| 9995  | False     | False      | False   | False       | False     | False  |



|       |       |       |       |       |       |       |
|-------|-------|-------|-------|-------|-------|-------|
| False |       |       |       |       |       |       |
| 9996  | False | False | False | False | False | False |
| False |       |       |       |       |       |       |
| 9997  | False | False | False | False | False | False |
| False |       |       |       |       |       |       |
| 9998  | False | False | False | False | False | False |
| False |       |       |       |       |       |       |
| 9999  | False | False | False | False | False | False |
| False |       |       |       |       |       |       |

|      | Tenure | Balance | NumOfProducts | HasCrCard | IsActiveMember | \ |
|------|--------|---------|---------------|-----------|----------------|---|
| 0    | False  | False   | False         | False     | False          |   |
| 1    | False  | False   | False         | False     | False          |   |
| 2    | False  | False   | False         | False     | False          |   |
| 3    | False  | False   | False         | False     | False          |   |
| 4    | False  | False   | False         | False     | False          |   |
| ...  | ...    | ...     | ...           | ...       | ...            |   |
| 9995 | False  | False   | False         | False     | False          |   |
| 9996 | False  | False   | False         | False     | False          |   |
| 9997 | False  | False   | False         | False     | False          |   |
| 9998 | False  | False   | False         | False     | False          |   |
| 9999 | False  | False   | False         | False     | False          |   |

|      | EstimatedSalary | Exited |
|------|-----------------|--------|
| 0    | False           | False  |
| 1    | False           | False  |
| 2    | False           | False  |
| 3    | False           | False  |
| 4    | False           | False  |
| ...  | ...             | ...    |
| 9995 | False           | False  |
| 9996 | False           | False  |
| 9997 | False           | False  |
| 9998 | False           | False  |
| 9999 | False           | False  |

[10000 rows x 14 columns]

```
df.isnull().sum()
```

|               |   |
|---------------|---|
| RowNumber     | 0 |
| CustomerId    | 0 |
| Surname       | 0 |
| CreditScore   | 0 |
| Geography     | 0 |
| Gender        | 0 |
| Age           | 0 |
| Tenure        | 0 |
| Balance       | 0 |
| NumOfProducts | 0 |

```

HasCrCard      0
IsActiveMember 0
EstimatedSalary 0
Exited         0
dtype: int64

```

```
#outliers
```

```
df
```

|      | CreditScore | Geography | Gender | Age | Tenure | Balance   |
|------|-------------|-----------|--------|-----|--------|-----------|
| 0    | 619         | France    | Female | 42  | 2      | 0.00      |
| 1    |             |           |        |     |        |           |
| 1    | 608         | Spain     | Female | 41  | 1      | 83807.86  |
| 1    |             |           |        |     |        |           |
| 2    | 502         | France    | Female | 42  | 8      | 159660.80 |
| 3    |             |           |        |     |        |           |
| 3    | 699         | France    | Female | 39  | 1      | 0.00      |
| 2    |             |           |        |     |        |           |
| 4    | 850         | Spain     | Female | 43  | 2      | 125510.82 |
| 1    |             |           |        |     |        |           |
| ...  | ...         | ...       | ...    | ... | ...    | ...       |
| ...  |             |           |        |     |        |           |
| 9995 | 771         | France    | Male   | 39  | 5      | 0.00      |
| 2    |             |           |        |     |        |           |
| 9996 | 516         | France    | Male   | 35  | 10     | 57369.61  |
| 1    |             |           |        |     |        |           |
| 9997 | 709         | France    | Female | 36  | 7      | 0.00      |
| 1    |             |           |        |     |        |           |
| 9998 | 772         | Germany   | Male   | 42  | 3      | 75075.31  |
| 2    |             |           |        |     |        |           |
| 9999 | 792         | France    | Female | 28  | 4      | 130142.79 |
| 1    |             |           |        |     |        |           |

|      | HasCrCard | IsActiveMember | EstimatedSalary | Exited |
|------|-----------|----------------|-----------------|--------|
| 0    | 1         | 1              | 101348.88       | 1      |
| 1    | 0         | 1              | 112542.58       | 0      |
| 2    | 1         | 0              | 113931.57       | 1      |
| 3    | 0         | 0              | 93826.63        | 0      |
| 4    | 1         | 1              | 79084.10        | 0      |
| ...  | ...       | ...            | ...             | ...    |
| 9995 | 1         | 0              | 96270.64        | 0      |
| 9996 | 1         | 1              | 101699.77       | 0      |
| 9997 | 0         | 1              | 42085.58        | 1      |
| 9998 | 1         | 0              | 92888.52        | 1      |
| 9999 | 1         | 0              | 38190.78        | 0      |

```
[10000 rows x 11 columns]
```

```
df.shape
```

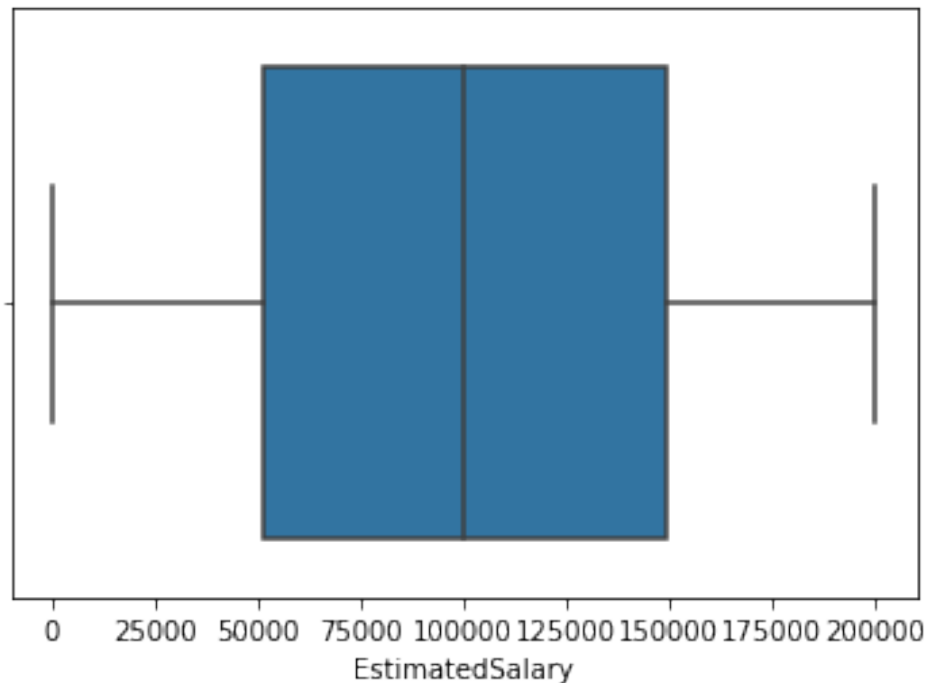
```
(10000, 14)
```

```
sns.boxplot(df['EstimatedSalary'])
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43:  
FutureWarning: Pass the following variable as a keyword arg: x. From  
version 0.12, the only valid positional argument will be `data`, and  
passing other arguments without an explicit keyword will result in an  
error or misinterpretation.
```

```
FutureWarning
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f56def47910>
```

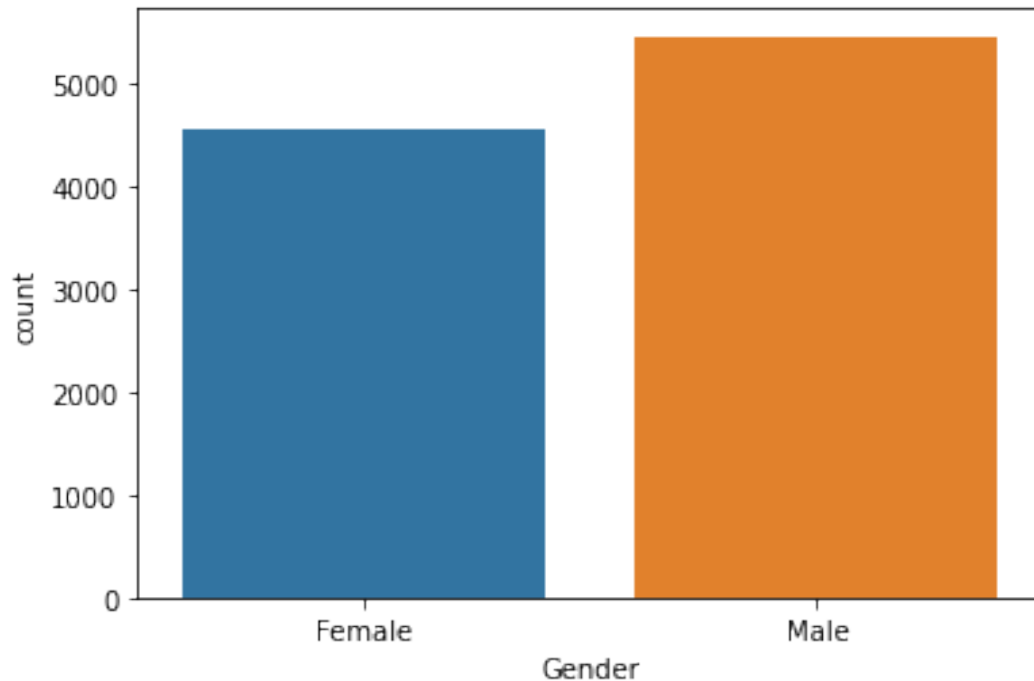


```
sns.countplot(df['Gender'])
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43:  
FutureWarning: Pass the following variable as a keyword arg: x. From  
version 0.12, the only valid positional argument will be `data`, and  
passing other arguments without an explicit keyword will result in an  
error or misinterpretation.
```

```
FutureWarning
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f56decc4e90>
```



IQR

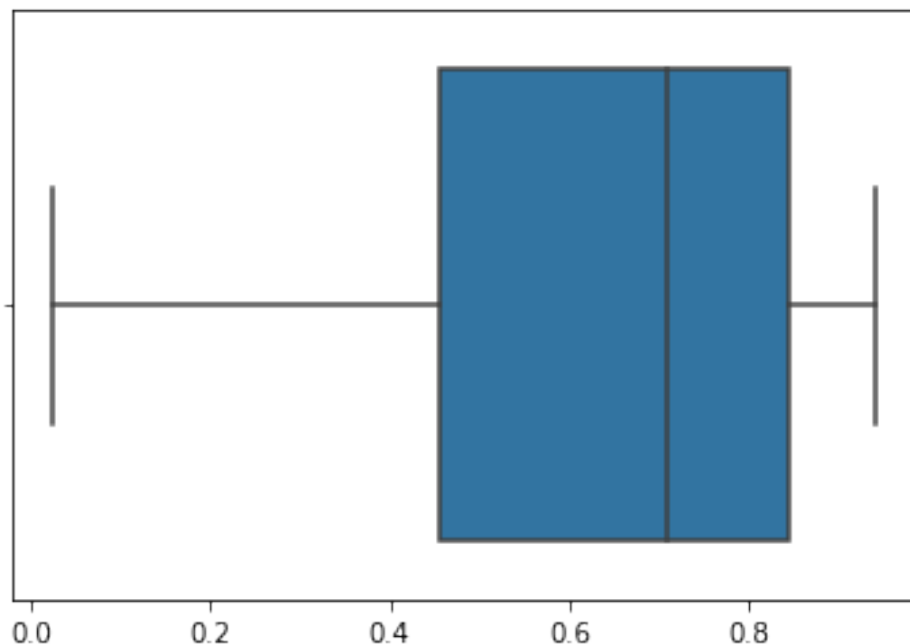
```
Q1=df.quantile(0.25)
Q3=df.quantile(0.75)
IQR=Q3-Q1
print(IQR)
```

```
RowNumber      4999.5000
CustomerId     124705.5000
CreditScore     134.0000
Age             12.0000
Tenure          4.0000
Balance        127644.2400
NumOfProducts   1.0000
HasCrCard       1.0000
IsActiveMember  1.0000
EstimatedSalary 98386.1375
Exited          0.0000
dtype: float64
```

```
sns.boxplot(np.random.rand(6))
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43:
FutureWarning: Pass the following variable as a keyword arg: x. From
version 0.12, the only valid positional argument will be `data`, and
passing other arguments without an explicit keyword will result in an
error or misinterpretation.
FutureWarning
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f56dbdf67d0>
```



```
df.corr()
```

|                 | RowNumber | CustomerId    | CreditScore | Age              |
|-----------------|-----------|---------------|-------------|------------------|
| Tenure \        |           |               |             |                  |
| RowNumber       | 1.000000  | 0.004202      | 0.005840    | 0.000783 -       |
| 0.006495        |           |               |             |                  |
| CustomerId      | 0.004202  | 1.000000      | 0.005308    | 0.009497 -       |
| 0.014883        |           |               |             |                  |
| CreditScore     | 0.005840  | 0.005308      | 1.000000    | -0.003965        |
| 0.000842        |           |               |             |                  |
| Age             | 0.000783  | 0.009497      | -0.003965   | 1.000000 -       |
| 0.009997        |           |               |             |                  |
| Tenure          | -0.006495 | -0.014883     | 0.000842    | -0.009997        |
| 1.000000        |           |               |             |                  |
| Balance         | -0.009067 | -0.012419     | 0.006268    | 0.028308 -       |
| 0.012254        |           |               |             |                  |
| NumOfProducts   | 0.007246  | 0.016972      | 0.012238    | -0.030680        |
| 0.013444        |           |               |             |                  |
| HasCrCard       | 0.000599  | -0.014025     | -0.005458   | -0.011721        |
| 0.022583        |           |               |             |                  |
| IsActiveMember  | 0.012044  | 0.001665      | 0.025651    | 0.085472 -       |
| 0.028362        |           |               |             |                  |
| EstimatedSalary | -0.005988 | 0.015271      | -0.001384   | -0.007201        |
| 0.007784        |           |               |             |                  |
| Exited          | -0.016571 | -0.006248     | -0.027094   | 0.285323 -       |
| 0.014001        |           |               |             |                  |
|                 | Balance   | NumOfProducts | HasCrCard   | IsActiveMember \ |
| RowNumber       | -0.009067 | 0.007246      | 0.000599    | 0.012044         |
| CustomerId      | -0.012419 | 0.016972      | -0.014025   | 0.001665         |

|                 |           |           |           |           |
|-----------------|-----------|-----------|-----------|-----------|
| CreditScore     | 0.006268  | 0.012238  | -0.005458 | 0.025651  |
| Age             | 0.028308  | -0.030680 | -0.011721 | 0.085472  |
| Tenure          | -0.012254 | 0.013444  | 0.022583  | -0.028362 |
| Balance         | 1.000000  | -0.304180 | -0.014858 | -0.010084 |
| NumOfProducts   | -0.304180 | 1.000000  | 0.003183  | 0.009612  |
| HasCrCard       | -0.014858 | 0.003183  | 1.000000  | -0.011866 |
| IsActiveMember  | -0.010084 | 0.009612  | -0.011866 | 1.000000  |
| EstimatedSalary | 0.012797  | 0.014204  | -0.009933 | -0.011421 |
| Exited          | 0.118533  | -0.047820 | -0.007138 | -0.156128 |

|                 |                 |           |
|-----------------|-----------------|-----------|
|                 | EstimatedSalary | Exited    |
| RowNumber       | -0.005988       | -0.016571 |
| CustomerId      | 0.015271        | -0.006248 |
| CreditScore     | -0.001384       | -0.027094 |
| Age             | -0.007201       | 0.285323  |
| Tenure          | 0.007784        | -0.014001 |
| Balance         | 0.012797        | 0.118533  |
| NumOfProducts   | 0.014204        | -0.047820 |
| HasCrCard       | -0.009933       | -0.007138 |
| IsActiveMember  | -0.011421       | -0.156128 |
| EstimatedSalary | 1.000000        | 0.012097  |
| Exited          | 0.012097        | 1.000000  |

replace outliers

```
CreditsMedian = df.loc[df['CreditScore']<400, 'CreditScore'].median()
ProdMedian = df.loc[df['NumOfProducts']>=3.5, 'NumOfProducts'].median()
```

```
df.loc[df.CreditScore < 400, 'CreditScore'] = np.nan
df.fillna(CreditsMedian,inplace=True)
df.loc[df.NumOfProducts > 3, 'NumOfProducts'] = np.nan
df.fillna(ProdMedian,inplace=True)
```

#Categorigal columns

df

|       | RowNumber | CustomerId | Surname  | CreditScore | Geography | Gender |
|-------|-----------|------------|----------|-------------|-----------|--------|
| Age \ |           |            |          |             |           |        |
| 0     | 1         | 15634602   | Hargrave | 619         | France    | Female |
| 42    |           |            |          |             |           |        |
| 1     | 2         | 15647311   | Hill     | 608         | Spain     | Female |
| 41    |           |            |          |             |           |        |
| 2     | 3         | 15619304   | Onio     | 502         | France    | Female |
| 42    |           |            |          |             |           |        |
| 3     | 4         | 15701354   | Boni     | 699         | France    | Female |
| 39    |           |            |          |             |           |        |
| 4     | 5         | 15737888   | Mitchell | 850         | Spain     | Female |
| 43    |           |            |          |             |           |        |
| ...   | ...       | ...        | ...      | ...         | ...       | ...    |
| ...   |           |            |          |             |           |        |

|            |       |          |           |     |         |        |
|------------|-------|----------|-----------|-----|---------|--------|
| 9995<br>39 | 9996  | 15606229 | Obijiaku  | 771 | France  | Male   |
| 9996<br>35 | 9997  | 15569892 | Johnstone | 516 | France  | Male   |
| 9997<br>36 | 9998  | 15584532 | Liu       | 709 | France  | Female |
| 9998<br>42 | 9999  | 15682355 | Sabbatini | 772 | Germany | Male   |
| 9999<br>28 | 10000 | 15628319 | Walker    | 792 | France  | Female |

|      | Tenure | Balance   | NumOfProducts | HasCrCard | IsActiveMember | \   |
|------|--------|-----------|---------------|-----------|----------------|-----|
| 0    | 2      | 0.00      | 1             | 1         | 1              |     |
| 1    | 1      | 83807.86  | 1             | 0         | 1              |     |
| 2    | 8      | 159660.80 | 3             | 1         | 0              |     |
| 3    | 1      | 0.00      | 2             | 0         | 0              |     |
| 4    | 2      | 125510.82 | 1             | 1         | 1              |     |
| ...  | ...    | ...       | ...           | ...       | ...            | ... |
| 9995 | 5      | 0.00      | 2             | 1         | 0              |     |
| 9996 | 10     | 57369.61  | 1             | 1         | 1              |     |
| 9997 | 7      | 0.00      | 1             | 0         | 1              |     |
| 9998 | 3      | 75075.31  | 2             | 1         | 0              |     |
| 9999 | 4      | 130142.79 | 1             | 1         | 0              |     |

|      | EstimatedSalary | Exited |
|------|-----------------|--------|
| 0    | 101348.88       | 1      |
| 1    | 112542.58       | 0      |
| 2    | 113931.57       | 1      |
| 3    | 93826.63        | 0      |
| 4    | 79084.10        | 0      |
| ...  | ...             | ...    |
| 9995 | 96270.64        | 0      |
| 9996 | 101699.77       | 0      |
| 9997 | 42085.58        | 1      |
| 9998 | 92888.52        | 1      |
| 9999 | 38190.78        | 0      |

[10000 rows x 14 columns]

Categorical Columns

```
cat = list(df.select_dtypes(include='O').keys())
```

```
cat
```

```
['Surname', 'Geography', 'Gender']
```

Numerical Columns

```
num = list(df.select_dtypes(include=['int64', 'float64']).keys())
```

```
num
```

```
[ 'RowNumber',
  'CustomerId',
  'CreditScore',
  'Age',
  'Tenure',
  'Balance',
  'NumOfProducts',
  'HasCrCard',
  'IsActiveMember',
  'EstimatedSalary',
  'Exited']
```

#Encoding

One Hot Encoding(method 1)

df

|       | RowNumber | CustomerId | Surname   | CreditScore | Geography | Gender |
|-------|-----------|------------|-----------|-------------|-----------|--------|
| Age \ |           |            |           |             |           |        |
| 0     | 1         | 15634602   | Hargrave  | 619         | France    | 0      |
| 42    |           |            |           |             |           |        |
| 1     | 2         | 15647311   | Hill      | 608         | Spain     | 0      |
| 41    |           |            |           |             |           |        |
| 2     | 3         | 15619304   | Onio      | 502         | France    | 0      |
| 42    |           |            |           |             |           |        |
| 3     | 4         | 15701354   | Boni      | 699         | France    | 0      |
| 39    |           |            |           |             |           |        |
| 4     | 5         | 15737888   | Mitchell  | 850         | Spain     | 0      |
| 43    |           |            |           |             |           |        |
| ...   | ...       | ...        | ...       | ...         | ...       | ...    |
| ...   |           |            |           |             |           |        |
| 9995  | 9996      | 15606229   | Obijiaku  | 771         | France    | 1      |
| 39    |           |            |           |             |           |        |
| 9996  | 9997      | 15569892   | Johnstone | 516         | France    | 1      |
| 35    |           |            |           |             |           |        |
| 9997  | 9998      | 15584532   | Liu       | 709         | France    | 0      |
| 36    |           |            |           |             |           |        |
| 9998  | 9999      | 15682355   | Sabbatini | 772         | Germany   | 1      |
| 42    |           |            |           |             |           |        |
| 9999  | 10000     | 15628319   | Walker    | 792         | France    | 0      |
| 28    |           |            |           |             |           |        |

|      | Tenure | Balance   | NumOfProducts | HasCrCard | IsActiveMember | \   |
|------|--------|-----------|---------------|-----------|----------------|-----|
| 0    | 2      | 0.00      | 1             | 1         |                | 1   |
| 1    | 1      | 83807.86  | 1             | 0         |                | 1   |
| 2    | 8      | 159660.80 | 3             | 1         |                | 0   |
| 3    | 1      | 0.00      | 2             | 0         |                | 0   |
| 4    | 2      | 125510.82 | 1             | 1         |                | 1   |
| ...  | ...    | ...       | ...           | ...       |                | ... |
| 9995 | 5      | 0.00      | 2             | 1         |                | 0   |



|      |    |           |   |   |   |
|------|----|-----------|---|---|---|
| 9996 | 10 | 57369.61  | 1 | 1 | 1 |
| 9997 | 7  | 0.00      | 1 | 0 | 1 |
| 9998 | 3  | 75075.31  | 2 | 1 | 0 |
| 9999 | 4  | 130142.79 | 1 | 1 | 0 |

|      | EstimatedSalary | Exited |
|------|-----------------|--------|
| 0    | 101348.88       | 1      |
| 1    | 112542.58       | 0      |
| 2    | 113931.57       | 1      |
| 3    | 93826.63        | 0      |
| 4    | 79084.10        | 0      |
| ...  | ...             | ...    |
| 9995 | 96270.64        | 0      |
| 9996 | 101699.77       | 0      |
| 9997 | 42085.58        | 1      |
| 9998 | 92888.52        | 1      |
| 9999 | 38190.78        | 0      |

[10000 rows x 14 columns]

df['Gender']

|   |        |
|---|--------|
| 0 | Female |
| 1 | Female |
| 2 | Female |
| 3 | Female |
| 4 | Female |

|      |        |
|------|--------|
| ...  | ...    |
| 9995 | Male   |
| 9996 | Male   |
| 9997 | Female |
| 9998 | Male   |
| 9999 | Female |

Name: Gender, Length: 10000, dtype: object

Gender=pd.get\_dummies(df['Gender'])

Gender

|      | Female | Male |
|------|--------|------|
| 0    | 1      | 0    |
| 1    | 1      | 0    |
| 2    | 1      | 0    |
| 3    | 1      | 0    |
| 4    | 1      | 0    |
| ...  | ...    | ...  |
| 9995 | 0      | 1    |
| 9996 | 0      | 1    |
| 9997 | 1      | 0    |
| 9998 | 0      | 1    |
| 9999 | 1      | 0    |

```
[10000 rows x 2 columns]
```

```
Label Encoding(method 2)
```

```
from sklearn.preprocessing import LabelEncoder
from collections import Counter as count
```

```
count(df['Gender'])
```

```
Counter({'Female': 4543, 'Male': 5457})
```

```
le = LabelEncoder()
```

```
df['Gender'] = le.fit_transform(df['Gender'])
```

```
count(df['Gender'])
```

```
Counter({0: 4543, 1: 5457})
```

```
df
```

|       | RowNumber | CustomerId | Surname   | CreditScore | Geography | Gender |
|-------|-----------|------------|-----------|-------------|-----------|--------|
| Age \ |           |            |           |             |           |        |
| 0     | 1         | 15634602   | Hargrave  | 619         | France    | 0      |
| 42    |           |            |           |             |           |        |
| 1     | 2         | 15647311   | Hill      | 608         | Spain     | 0      |
| 41    |           |            |           |             |           |        |
| 2     | 3         | 15619304   | Onio      | 502         | France    | 0      |
| 42    |           |            |           |             |           |        |
| 3     | 4         | 15701354   | Boni      | 699         | France    | 0      |
| 39    |           |            |           |             |           |        |
| 4     | 5         | 15737888   | Mitchell  | 850         | Spain     | 0      |
| 43    |           |            |           |             |           |        |
| ...   | ...       | ...        | ...       | ...         | ...       | ...    |
| ...   |           |            |           |             |           |        |
| 9995  | 9996      | 15606229   | Obijiaku  | 771         | France    | 1      |
| 39    |           |            |           |             |           |        |
| 9996  | 9997      | 15569892   | Johnstone | 516         | France    | 1      |
| 35    |           |            |           |             |           |        |
| 9997  | 9998      | 15584532   | Liu       | 709         | France    | 0      |
| 36    |           |            |           |             |           |        |
| 9998  | 9999      | 15682355   | Sabbatini | 772         | Germany   | 1      |
| 42    |           |            |           |             |           |        |
| 9999  | 10000     | 15628319   | Walker    | 792         | France    | 0      |
| 28    |           |            |           |             |           |        |

|   | Tenure | Balance   | NumOfProducts | HasCrCard | IsActiveMember | \ |
|---|--------|-----------|---------------|-----------|----------------|---|
| 0 | 2      | 0.00      | 1             | 1         | 1              |   |
| 1 | 1      | 83807.86  | 1             | 0         | 1              |   |
| 2 | 8      | 159660.80 | 3             | 1         | 0              |   |
| 3 | 1      | 0.00      | 2             | 0         | 0              |   |

|      |     |           |     |     |     |
|------|-----|-----------|-----|-----|-----|
| 4    | 2   | 125510.82 | 1   | 1   | 1   |
| ...  | ... | ...       | ... | ... | ... |
| 9995 | 5   | 0.00      | 2   | 1   | 0   |
| 9996 | 10  | 57369.61  | 1   | 1   | 1   |
| 9997 | 7   | 0.00      | 1   | 0   | 1   |
| 9998 | 3   | 75075.31  | 2   | 1   | 0   |
| 9999 | 4   | 130142.79 | 1   | 1   | 0   |

|      | EstimatedSalary | Exited |
|------|-----------------|--------|
| 0    | 101348.88       | 1      |
| 1    | 112542.58       | 0      |
| 2    | 113931.57       | 1      |
| 3    | 93826.63        | 0      |
| 4    | 79084.10        | 0      |
| ...  | ...             | ...    |
| 9995 | 96270.64        | 0      |
| 9996 | 101699.77       | 0      |
| 9997 | 42085.58        | 1      |
| 9998 | 92888.52        | 1      |
| 9999 | 38190.78        | 0      |

[10000 rows x 14 columns]

Feature map(method 3)

count(df['Gender'])

Counter({0: 4543, 1: 5457})

df['Gender']=df['Gender'].replace(['No','Yes'],[0,1])

df

|       | RowNumber | CustomerId | Surname   | CreditScore | Geography | Gender |
|-------|-----------|------------|-----------|-------------|-----------|--------|
| Age \ |           |            |           |             |           |        |
| 0     | 1         | 15634602   | Hargrave  | 619         | France    | 0      |
| 42    |           |            |           |             |           |        |
| 1     | 2         | 15647311   | Hill      | 608         | Spain     | 0      |
| 41    |           |            |           |             |           |        |
| 2     | 3         | 15619304   | Onio      | 502         | France    | 0      |
| 42    |           |            |           |             |           |        |
| 3     | 4         | 15701354   | Boni      | 699         | France    | 0      |
| 39    |           |            |           |             |           |        |
| 4     | 5         | 15737888   | Mitchell  | 850         | Spain     | 0      |
| 43    |           |            |           |             |           |        |
| ...   | ...       | ...        | ...       | ...         | ...       | ...    |
| ...   |           |            |           |             |           |        |
| 9995  | 9996      | 15606229   | Obijiaku  | 771         | France    | 1      |
| 39    |           |            |           |             |           |        |
| 9996  | 9997      | 15569892   | Johnstone | 516         | France    | 1      |
| 35    |           |            |           |             |           |        |

|            |       |          |           |     |         |   |
|------------|-------|----------|-----------|-----|---------|---|
| 9997<br>36 | 9998  | 15584532 | Liu       | 709 | France  | 0 |
| 9998<br>42 | 9999  | 15682355 | Sabbatini | 772 | Germany | 1 |
| 9999<br>28 | 10000 | 15628319 | Walker    | 792 | France  | 0 |

|      | Tenure | Balance   | NumOfProducts | HasCrCard | IsActiveMember | \ |
|------|--------|-----------|---------------|-----------|----------------|---|
| 0    | 2      | 0.00      | 1             | 1         | 1              |   |
| 1    | 1      | 83807.86  | 1             | 0         | 1              |   |
| 2    | 8      | 159660.80 | 3             | 1         | 0              |   |
| 3    | 1      | 0.00      | 2             | 0         | 0              |   |
| 4    | 2      | 125510.82 | 1             | 1         | 1              |   |
| ...  | ...    | ...       | ...           | ...       | ...            |   |
| 9995 | 5      | 0.00      | 2             | 1         | 0              |   |
| 9996 | 10     | 57369.61  | 1             | 1         | 1              |   |
| 9997 | 7      | 0.00      | 1             | 0         | 1              |   |
| 9998 | 3      | 75075.31  | 2             | 1         | 0              |   |
| 9999 | 4      | 130142.79 | 1             | 1         | 0              |   |

|      | EstimatedSalary | Exited |
|------|-----------------|--------|
| 0    | 101348.88       | 1      |
| 1    | 112542.58       | 0      |
| 2    | 113931.57       | 1      |
| 3    | 93826.63        | 0      |
| 4    | 79084.10        | 0      |
| ...  | ...             | ...    |
| 9995 | 96270.64        | 0      |
| 9996 | 101699.77       | 0      |
| 9997 | 42085.58        | 1      |
| 9998 | 92888.52        | 1      |
| 9999 | 38190.78        | 0      |

[10000 rows x 14 columns]

#split the data into dependent and independent variables

```
x = df.iloc[:, :-1]
```

x

|       | RowNumber | CustomerId | Surname  | CreditScore | Geography | Gender |
|-------|-----------|------------|----------|-------------|-----------|--------|
| Age \ |           |            |          |             |           |        |
| 0     | 1         | 15634602   | Hargrave | 619         | France    | 0      |
| 42    |           |            |          |             |           |        |
| 1     | 2         | 15647311   | Hill     | 608         | Spain     | 0      |
| 41    |           |            |          |             |           |        |
| 2     | 3         | 15619304   | Onio     | 502         | France    | 0      |
| 42    |           |            |          |             |           |        |
| 3     | 4         | 15701354   | Boni     | 699         | France    | 0      |
| 39    |           |            |          |             |           |        |

|      |       |          |           |     |         |     |
|------|-------|----------|-----------|-----|---------|-----|
| 4    | 5     | 15737888 | Mitchell  | 850 | Spain   | 0   |
| 43   |       |          |           |     |         |     |
| ...  | ...   | ...      | ...       | ... | ...     | ... |
| ...  |       |          |           |     |         |     |
| 9995 | 9996  | 15606229 | Obijiaku  | 771 | France  | 1   |
| 39   |       |          |           |     |         |     |
| 9996 | 9997  | 15569892 | Johnstone | 516 | France  | 1   |
| 35   |       |          |           |     |         |     |
| 9997 | 9998  | 15584532 | Liu       | 709 | France  | 0   |
| 36   |       |          |           |     |         |     |
| 9998 | 9999  | 15682355 | Sabbatini | 772 | Germany | 1   |
| 42   |       |          |           |     |         |     |
| 9999 | 10000 | 15628319 | Walker    | 792 | France  | 0   |
| 28   |       |          |           |     |         |     |

|      | Tenure | Balance   | NumOfProducts | HasCrCard | IsActiveMember | \ |
|------|--------|-----------|---------------|-----------|----------------|---|
| 0    | 2      | 0.00      | 1             | 1         | 1              |   |
| 1    | 1      | 83807.86  | 1             | 0         | 1              |   |
| 2    | 8      | 159660.80 | 3             | 1         | 0              |   |
| 3    | 1      | 0.00      | 2             | 0         | 0              |   |
| 4    | 2      | 125510.82 | 1             | 1         | 1              |   |
| ...  | ...    | ...       | ...           | ...       | ...            |   |
| 9995 | 5      | 0.00      | 2             | 1         | 0              |   |
| 9996 | 10     | 57369.61  | 1             | 1         | 1              |   |
| 9997 | 7      | 0.00      | 1             | 0         | 1              |   |
| 9998 | 3      | 75075.31  | 2             | 1         | 0              |   |
| 9999 | 4      | 130142.79 | 1             | 1         | 0              |   |

|      | EstimatedSalary |
|------|-----------------|
| 0    | 101348.88       |
| 1    | 112542.58       |
| 2    | 113931.57       |
| 3    | 93826.63        |
| 4    | 79084.10        |
| ...  | ...             |
| 9995 | 96270.64        |
| 9996 | 101699.77       |
| 9997 | 42085.58        |
| 9998 | 92888.52        |
| 9999 | 38190.78        |

[10000 rows x 13 columns]

```
y=df.iloc[:, -1:]
y
```

|   | Exited |
|---|--------|
| 0 | 1      |
| 1 | 0      |
| 2 | 1      |

```

3          0
4          0
...      ...
9995       0
9996       0
9997       1
9998       1
9999       0

```

```
[10000 rows x 1 columns]
```

```
x=df.iloc[:, :-1].values
```

```
x
```

```

array([[1, 15634602, 'Hargrave', ..., 1, 1, 101348.88],
       [2, 15647311, 'Hill', ..., 0, 1, 112542.58],
       [3, 15619304, 'Onio', ..., 1, 0, 113931.57],
       ...,
       [9998, 15584532, 'Liu', ..., 0, 1, 42085.58],
       [9999, 15682355, 'Sabbatini', ..., 1, 0, 92888.52],
       [10000, 15628319, 'Walker', ..., 1, 0, 38190.78]],
      dtype=object)

```

```
y=df.iloc[:, -1:].values
```

```
y
```

```

array([[1],
       [0],
       [1],
       ...,
       [1],
       [1],
       [0]])

```

```
#Scale the independent variables
```

```

from sklearn.preprocessing import StandardScaler
from sklearn.datasets import load_iris

```

```

scaler = StandardScaler()
data = load_iris()

```

```

x_data=data.data
target = data.target

```

```

scaled_data = scaler.fit_transform(x_data)
scaled_data

```

```

array([[ -9.00681170e-01,  1.01900435e+00, -1.34022653e+00,
        -1.31544430e+00],
       [-1.14301691e+00, -1.31979479e-01, -1.34022653e+00,
        -1.31544430e+00],

```

[-1.38535265e+00, 3.28414053e-01, -1.39706395e+00,  
-1.31544430e+00],  
[-1.50652052e+00, 9.82172869e-02, -1.28338910e+00,  
-1.31544430e+00],  
[-1.02184904e+00, 1.24920112e+00, -1.34022653e+00,  
-1.31544430e+00],  
[-5.37177559e-01, 1.93979142e+00, -1.16971425e+00,  
-1.05217993e+00],  
[-1.50652052e+00, 7.88807586e-01, -1.34022653e+00,  
-1.18381211e+00],  
[-1.02184904e+00, 7.88807586e-01, -1.28338910e+00,  
-1.31544430e+00],  
[-1.74885626e+00, -3.62176246e-01, -1.34022653e+00,  
-1.31544430e+00],  
[-1.14301691e+00, 9.82172869e-02, -1.28338910e+00,  
-1.44707648e+00],  
[-5.37177559e-01, 1.47939788e+00, -1.28338910e+00,  
-1.31544430e+00],  
[-1.26418478e+00, 7.88807586e-01, -1.22655167e+00,  
-1.31544430e+00],  
[-1.26418478e+00, -1.31979479e-01, -1.34022653e+00,  
-1.44707648e+00],  
[-1.87002413e+00, -1.31979479e-01, -1.51073881e+00,  
-1.44707648e+00],  
[-5.25060772e-02, 2.16998818e+00, -1.45390138e+00,  
-1.31544430e+00],  
[-1.73673948e-01, 3.09077525e+00, -1.28338910e+00,  
-1.05217993e+00],  
[-5.37177559e-01, 1.93979142e+00, -1.39706395e+00,  
-1.05217993e+00],  
[-9.00681170e-01, 1.01900435e+00, -1.34022653e+00,  
-1.18381211e+00],  
[-1.73673948e-01, 1.70959465e+00, -1.16971425e+00,  
-1.18381211e+00],  
[-9.00681170e-01, 1.70959465e+00, -1.28338910e+00,  
-1.18381211e+00],  
[-5.37177559e-01, 7.88807586e-01, -1.16971425e+00,  
-1.31544430e+00],  
[-9.00681170e-01, 1.47939788e+00, -1.28338910e+00,  
-1.05217993e+00],  
[-1.50652052e+00, 1.24920112e+00, -1.56757623e+00,  
-1.31544430e+00],  
[-9.00681170e-01, 5.58610819e-01, -1.16971425e+00,  
-9.20547742e-01],  
[-1.26418478e+00, 7.88807586e-01, -1.05603939e+00,  
-1.31544430e+00],  
[-1.02184904e+00, -1.31979479e-01, -1.22655167e+00,  
-1.31544430e+00],  
[-1.02184904e+00, 7.88807586e-01, -1.22655167e+00,  
-1.05217993e+00],

[-7.79513300e-01, 1.01900435e+00, -1.28338910e+00,  
-1.31544430e+00],  
[-7.79513300e-01, 7.88807586e-01, -1.34022653e+00,  
-1.31544430e+00],  
[-1.38535265e+00, 3.28414053e-01, -1.22655167e+00,  
-1.31544430e+00],  
[-1.26418478e+00, 9.82172869e-02, -1.22655167e+00,  
-1.31544430e+00],  
[-5.37177559e-01, 7.88807586e-01, -1.28338910e+00,  
-1.05217993e+00],  
[-7.79513300e-01, 2.40018495e+00, -1.28338910e+00,  
-1.44707648e+00],  
[-4.16009689e-01, 2.63038172e+00, -1.34022653e+00,  
-1.31544430e+00],  
[-1.14301691e+00, 9.82172869e-02, -1.28338910e+00,  
-1.31544430e+00],  
[-1.02184904e+00, 3.28414053e-01, -1.45390138e+00,  
-1.31544430e+00],  
[-4.16009689e-01, 1.01900435e+00, -1.39706395e+00,  
-1.31544430e+00],  
[-1.14301691e+00, 1.24920112e+00, -1.34022653e+00,  
-1.44707648e+00],  
[-1.74885626e+00, -1.31979479e-01, -1.39706395e+00,  
-1.31544430e+00],  
[-9.00681170e-01, 7.88807586e-01, -1.28338910e+00,  
-1.31544430e+00],  
[-1.02184904e+00, 1.01900435e+00, -1.39706395e+00,  
-1.18381211e+00],  
[-1.62768839e+00, -1.74335684e+00, -1.39706395e+00,  
-1.18381211e+00],  
[-1.74885626e+00, 3.28414053e-01, -1.39706395e+00,  
-1.31544430e+00],  
[-1.02184904e+00, 1.01900435e+00, -1.22655167e+00,  
-7.88915558e-01],  
[-9.00681170e-01, 1.70959465e+00, -1.05603939e+00,  
-1.05217993e+00],  
[-1.26418478e+00, -1.31979479e-01, -1.34022653e+00,  
-1.18381211e+00],  
[-9.00681170e-01, 1.70959465e+00, -1.22655167e+00,  
-1.31544430e+00],  
[-1.50652052e+00, 3.28414053e-01, -1.34022653e+00,  
-1.31544430e+00],  
[-6.58345429e-01, 1.47939788e+00, -1.28338910e+00,  
-1.31544430e+00],  
[-1.02184904e+00, 5.58610819e-01, -1.34022653e+00,  
-1.31544430e+00],  
[ 1.40150837e+00, 3.28414053e-01, 5.35408562e-01,  
2.64141916e-01],  
[ 6.74501145e-01, 3.28414053e-01, 4.21733708e-01,  
3.95774101e-01],



[ 1.28034050e+00, 9.82172869e-02, 6.49083415e-01,  
3.95774101e-01],  
[-4.16009689e-01, -1.74335684e+00, 1.37546573e-01,  
1.32509732e-01],  
[ 7.95669016e-01, -5.92373012e-01, 4.78571135e-01,  
3.95774101e-01],  
[-1.73673948e-01, -5.92373012e-01, 4.21733708e-01,  
1.32509732e-01],  
[ 5.53333275e-01, 5.58610819e-01, 5.35408562e-01,  
5.27406285e-01],  
[-1.14301691e+00, -1.51316008e+00, -2.60315415e-01,  
-2.62386821e-01],  
[ 9.16836886e-01, -3.62176246e-01, 4.78571135e-01,  
1.32509732e-01],  
[-7.79513300e-01, -8.22569778e-01, 8.07091462e-02,  
2.64141916e-01],  
[-1.02184904e+00, -2.43394714e+00, -1.46640561e-01,  
-2.62386821e-01],  
[ 6.86617933e-02, -1.31979479e-01, 2.51221427e-01,  
3.95774101e-01],  
[ 1.89829664e-01, -1.97355361e+00, 1.37546573e-01,  
-2.62386821e-01],  
[ 3.10997534e-01, -3.62176246e-01, 5.35408562e-01,  
2.64141916e-01],  
[-2.94841818e-01, -3.62176246e-01, -8.98031345e-02,  
1.32509732e-01],  
[ 1.03800476e+00, 9.82172869e-02, 3.64896281e-01,  
2.64141916e-01],  
[-2.94841818e-01, -1.31979479e-01, 4.21733708e-01,  
3.95774101e-01],  
[-5.25060772e-02, -8.22569778e-01, 1.94384000e-01,  
-2.62386821e-01],  
[ 4.32165405e-01, -1.97355361e+00, 4.21733708e-01,  
3.95774101e-01],  
[-2.94841818e-01, -1.28296331e+00, 8.07091462e-02,  
-1.30754636e-01],  
[ 6.86617933e-02, 3.28414053e-01, 5.92245988e-01,  
7.90670654e-01],  
[ 3.10997534e-01, -5.92373012e-01, 1.37546573e-01,  
1.32509732e-01],  
[ 5.53333275e-01, -1.28296331e+00, 6.49083415e-01,  
3.95774101e-01],  
[ 3.10997534e-01, -5.92373012e-01, 5.35408562e-01,  
8.77547895e-04],  
[ 6.74501145e-01, -3.62176246e-01, 3.08058854e-01,  
1.32509732e-01],  
[ 9.16836886e-01, -1.31979479e-01, 3.64896281e-01,  
2.64141916e-01],  
[ 1.15917263e+00, -5.92373012e-01, 5.92245988e-01,  
2.64141916e-01],

[ 1.03800476e+00, -1.31979479e-01, 7.05920842e-01,  
6.59038469e-01],  
[ 1.89829664e-01, -3.62176246e-01, 4.21733708e-01,  
3.95774101e-01],  
[-1.73673948e-01, -1.05276654e+00, -1.46640561e-01,  
-2.62386821e-01],  
[-4.16009689e-01, -1.51316008e+00, 2.38717193e-02,  
-1.30754636e-01],  
[-4.16009689e-01, -1.51316008e+00, -3.29657076e-02,  
-2.62386821e-01],  
[-5.25060772e-02, -8.22569778e-01, 8.07091462e-02,  
8.77547895e-04],  
[ 1.89829664e-01, -8.22569778e-01, 7.62758269e-01,  
5.27406285e-01],  
[-5.37177559e-01, -1.31979479e-01, 4.21733708e-01,  
3.95774101e-01],  
[ 1.89829664e-01, 7.88807586e-01, 4.21733708e-01,  
5.27406285e-01],  
[ 1.03800476e+00, 9.82172869e-02, 5.35408562e-01,  
3.95774101e-01],  
[ 5.53333275e-01, -1.74335684e+00, 3.64896281e-01,  
1.32509732e-01],  
[-2.94841818e-01, -1.31979479e-01, 1.94384000e-01,  
1.32509732e-01],  
[-4.16009689e-01, -1.28296331e+00, 1.37546573e-01,  
1.32509732e-01],  
[-4.16009689e-01, -1.05276654e+00, 3.64896281e-01,  
8.77547895e-04],  
[ 3.10997534e-01, -1.31979479e-01, 4.78571135e-01,  
2.64141916e-01],  
[-5.25060772e-02, -1.05276654e+00, 1.37546573e-01,  
8.77547895e-04],  
[-1.02184904e+00, -1.74335684e+00, -2.60315415e-01,  
-2.62386821e-01],  
[-2.94841818e-01, -8.22569778e-01, 2.51221427e-01,  
1.32509732e-01],  
[-1.73673948e-01, -1.31979479e-01, 2.51221427e-01,  
8.77547895e-04],  
[-1.73673948e-01, -3.62176246e-01, 2.51221427e-01,  
1.32509732e-01],  
[ 4.32165405e-01, -3.62176246e-01, 3.08058854e-01,  
1.32509732e-01],  
[-9.00681170e-01, -1.28296331e+00, -4.30827696e-01,  
-1.30754636e-01],  
[-1.73673948e-01, -5.92373012e-01, 1.94384000e-01,  
1.32509732e-01],  
[ 5.53333275e-01, 5.58610819e-01, 1.27429511e+00,  
1.71209594e+00],  
[-5.25060772e-02, -8.22569778e-01, 7.62758269e-01,  
9.22302838e-01],

[ 1.52267624e+00, -1.31979479e-01, 1.21745768e+00,  
1.18556721e+00],  
[ 5.53333275e-01, -3.62176246e-01, 1.04694540e+00,  
7.90670654e-01],  
[ 7.95669016e-01, -1.31979479e-01, 1.16062026e+00,  
1.31719939e+00],  
[ 2.12851559e+00, -1.31979479e-01, 1.61531967e+00,  
1.18556721e+00],  
[-1.14301691e+00, -1.28296331e+00, 4.21733708e-01,  
6.59038469e-01],  
[ 1.76501198e+00, -3.62176246e-01, 1.44480739e+00,  
7.90670654e-01],  
[ 1.03800476e+00, -1.28296331e+00, 1.16062026e+00,  
7.90670654e-01],  
[ 1.64384411e+00, 1.24920112e+00, 1.33113254e+00,  
1.71209594e+00],  
[ 7.95669016e-01, 3.28414053e-01, 7.62758269e-01,  
1.05393502e+00],  
[ 6.74501145e-01, -8.22569778e-01, 8.76433123e-01,  
9.22302838e-01],  
[ 1.15917263e+00, -1.31979479e-01, 9.90107977e-01,  
1.18556721e+00],  
[-1.73673948e-01, -1.28296331e+00, 7.05920842e-01,  
1.05393502e+00],  
[-5.25060772e-02, -5.92373012e-01, 7.62758269e-01,  
1.58046376e+00],  
[ 6.74501145e-01, 3.28414053e-01, 8.76433123e-01,  
1.44883158e+00],  
[ 7.95669016e-01, -1.31979479e-01, 9.90107977e-01,  
7.90670654e-01],  
[ 2.24968346e+00, 1.70959465e+00, 1.67215710e+00,  
1.31719939e+00],  
[ 2.24968346e+00, -1.05276654e+00, 1.78583195e+00,  
1.44883158e+00],  
[ 1.89829664e-01, -1.97355361e+00, 7.05920842e-01,  
3.95774101e-01],  
[ 1.28034050e+00, 3.28414053e-01, 1.10378283e+00,  
1.44883158e+00],  
[-2.94841818e-01, -5.92373012e-01, 6.49083415e-01,  
1.05393502e+00],  
[ 2.24968346e+00, -5.92373012e-01, 1.67215710e+00,  
1.05393502e+00],  
[ 5.53333275e-01, -8.22569778e-01, 6.49083415e-01,  
7.90670654e-01],  
[ 1.03800476e+00, 5.58610819e-01, 1.10378283e+00,  
1.18556721e+00],  
[ 1.64384411e+00, 3.28414053e-01, 1.27429511e+00,  
7.90670654e-01],  
[ 4.32165405e-01, -5.92373012e-01, 5.92245988e-01,  
7.90670654e-01],

```
[ 3.10997534e-01, -1.31979479e-01,  6.49083415e-01,
  7.90670654e-01],
[ 6.74501145e-01, -5.92373012e-01,  1.04694540e+00,
  1.18556721e+00],
[ 1.64384411e+00, -1.31979479e-01,  1.16062026e+00,
  5.27406285e-01],
[ 1.88617985e+00, -5.92373012e-01,  1.33113254e+00,
  9.22302838e-01],
[ 2.49201920e+00,  1.70959465e+00,  1.50164482e+00,
  1.05393502e+00],
[ 6.74501145e-01, -5.92373012e-01,  1.04694540e+00,
  1.31719939e+00],
[ 5.53333275e-01, -5.92373012e-01,  7.62758269e-01,
  3.95774101e-01],
[ 3.10997534e-01, -1.05276654e+00,  1.04694540e+00,
  2.64141916e-01],
[ 2.24968346e+00, -1.31979479e-01,  1.33113254e+00,
  1.44883158e+00],
[ 5.53333275e-01,  7.88807586e-01,  1.04694540e+00,
  1.58046376e+00],
[ 6.74501145e-01,  9.82172869e-02,  9.90107977e-01,
  7.90670654e-01],
[ 1.89829664e-01, -1.31979479e-01,  5.92245988e-01,
  7.90670654e-01],
[ 1.28034050e+00,  9.82172869e-02,  9.33270550e-01,
  1.18556721e+00],
[ 1.03800476e+00,  9.82172869e-02,  1.04694540e+00,
  1.58046376e+00],
[ 1.28034050e+00,  9.82172869e-02,  7.62758269e-01,
  1.44883158e+00],
[-5.25060772e-02, -8.22569778e-01,  7.62758269e-01,
  9.22302838e-01],
[ 1.15917263e+00,  3.28414053e-01,  1.21745768e+00,
  1.44883158e+00],
[ 1.03800476e+00,  5.58610819e-01,  1.10378283e+00,
  1.71209594e+00],
[ 1.03800476e+00, -1.31979479e-01,  8.19595696e-01,
  1.44883158e+00],
[ 5.53333275e-01, -1.28296331e+00,  7.05920842e-01,
  9.22302838e-01],
[ 7.95669016e-01, -1.31979479e-01,  8.19595696e-01,
  1.05393502e+00],
[ 4.32165405e-01,  7.88807586e-01,  9.33270550e-01,
  1.44883158e+00],
[ 6.86617933e-02, -1.31979479e-01,  7.62758269e-01,
  7.90670654e-01]])
```

#split the data into training and testing

```
from sklearn.model_selection import train_test_split
```

```
xtrain,xtest,ytrain,ytest =
train_test_split(x,y,test_size=0.3,random_state=0)
```

```
df
```

|       | RowNumber | CustomerId | Surname   | CreditScore | Geography | Gender |
|-------|-----------|------------|-----------|-------------|-----------|--------|
| Age \ |           |            |           |             |           |        |
| 0     | 1         | 15634602   | Hargrave  | 619         | France    | 0      |
| 42    |           |            |           |             |           |        |
| 1     | 2         | 15647311   | Hill      | 608         | Spain     | 0      |
| 41    |           |            |           |             |           |        |
| 2     | 3         | 15619304   | Onio      | 502         | France    | 0      |
| 42    |           |            |           |             |           |        |
| 3     | 4         | 15701354   | Boni      | 699         | France    | 0      |
| 39    |           |            |           |             |           |        |
| 4     | 5         | 15737888   | Mitchell  | 850         | Spain     | 0      |
| 43    |           |            |           |             |           |        |
| ...   | ...       | ...        | ...       | ...         | ...       | ...    |
| ...   |           |            |           |             |           |        |
| 9995  | 9996      | 15606229   | Obijiaku  | 771         | France    | 1      |
| 39    |           |            |           |             |           |        |
| 9996  | 9997      | 15569892   | Johnstone | 516         | France    | 1      |
| 35    |           |            |           |             |           |        |
| 9997  | 9998      | 15584532   | Liu       | 709         | France    | 0      |
| 36    |           |            |           |             |           |        |
| 9998  | 9999      | 15682355   | Sabbatini | 772         | Germany   | 1      |
| 42    |           |            |           |             |           |        |
| 9999  | 10000     | 15628319   | Walker    | 792         | France    | 0      |
| 28    |           |            |           |             |           |        |

|      | Tenure | Balance   | NumOfProducts | HasCrCard | IsActiveMember | \   |
|------|--------|-----------|---------------|-----------|----------------|-----|
| 0    | 2      | 0.00      | 1             | 1         | 1              |     |
| 1    | 1      | 83807.86  | 1             | 0         | 1              |     |
| 2    | 8      | 159660.80 | 3             | 1         | 0              |     |
| 3    | 1      | 0.00      | 2             | 0         | 0              |     |
| 4    | 2      | 125510.82 | 1             | 1         | 1              |     |
| ...  | ...    | ...       | ...           | ...       | ...            | ... |
| 9995 | 5      | 0.00      | 2             | 1         | 0              |     |
| 9996 | 10     | 57369.61  | 1             | 1         | 1              |     |
| 9997 | 7      | 0.00      | 1             | 0         | 1              |     |
| 9998 | 3      | 75075.31  | 2             | 1         | 0              |     |
| 9999 | 4      | 130142.79 | 1             | 1         | 0              |     |

|     | EstimatedSalary | Exited |
|-----|-----------------|--------|
| 0   | 101348.88       | 1      |
| 1   | 112542.58       | 0      |
| 2   | 113931.57       | 1      |
| 3   | 93826.63        | 0      |
| 4   | 79084.10        | 0      |
| ... | ...             | ...    |

```

9995      96270.64      0
9996     101699.77      0
9997      42085.58      1
9998      92888.52      1
9999      38190.78      0

```

```
[10000 rows x 14 columns]
```

```

X = df.iloc[:, :-1]
y = df.iloc[:, -1]

```

```

xtrain,xtest,ytrain,ytest =
train_test_split(x,y,test_size=0.3,random_state=0)

```

```
xtrain,xtest,ytrain,ytest
```

```

(array([[7682, 15633608, 'Black', ..., 1, 1, 55796.83],
       [9032, 15742323, 'Barese', ..., 1, 0, 19823.02],
       [3692, 15760244, 'Ives', ..., 0, 1, 13848.58],
       ...,
       [3265, 15574372, 'Hoolan', ..., 1, 0, 181429.87],
       [9846, 15664035, 'Parsons', ..., 1, 1, 148750.16],
       [2733, 15592816, 'Udokamma', ..., 1, 0, 118855.26]]),
dtype=object),
array([[9395, 15615753, 'Upchurch', ..., 1, 1, 192852.67],
       [899, 15654700, 'Fallaci', ..., 1, 0, 128702.1],
       [2399, 15633877, 'Morrison', ..., 1, 1, 75732.25],
       ...,
       [9308, 15680405, 'P'eng", ..., 1, 1, 167400.29],
       [8395, 15597983, 'Brown', ..., 1, 1, 70849.47],
       [5234, 15591286, 'Simmons', ..., 1, 1, 33759.41]]),
dtype=object),
7681      1
9031      0
3691      0
202       1
5625      0
..
9225      0
4859      0
3264      0
9845      0
2732      1
Name: Exited, Length: 7000, dtype: int64,
9394      0
898       1
2398      0
5906      0
2343      0
..

```

```
4004    0
7375    0
9307    0
8394    0
5233    1
Name:Exited, Length: 3000, dtype: int64)
```

```
xtrain.shape,xtest.shape
```

```
((7000, 13), (3000, 13))
```