**Load the dataset.**

**import** numpy **as** np

**import** pandas **as** pd

**import** matplotlib.pyplot **as** plt

**import** seaborn **as** sns

df**=**pd**.**read\_csv("/content/Churn\_Modelling.csv")

df**.**head()

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Row** | **Cust** | **Sur** | **Credi** | **Geog** | **Ge** | **A** | **Te** | **Bala** | **NumOf** | **HasC** | **IsActiv** | **Estimat** | **Ex** |  |
| **Num** | **omer** | **na** | **tScor** | **raph** | **nd** | **g** | **nu** | **Produc** | **rCar** | **eMemb** | **edSalar** | **ite** |  |
| **nce** |  |
| **ber** | **Id** | **me** | **e** | **y** | **er** | **e** | **re** | **ts** | **d** | **er** | **y** | **d** |  |
|  |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | 1563 | Har |  | Fran | Fe |  |
| **0** | 1 | grav | 619 | mal |  |
| 4602 | ce |  |
|  |  | e |  | e |  |
|  |  |  |  |  |  |
|  |  | 1564 |  |  | Spai | Fe |  |
| **1** | 2 | Hill | 608 | mal |  |
| 7311 | n |  |
|  |  |  |  | e |  |
|  |  |  |  |  |  |  |
|  |  | 1561 | Oni |  | Fran | Fe |  |
| **2** | 3 | 502 | mal |  |
| 9304 | o | ce |  |
|  |  |  | e |  |
|  |  |  |  |  |  |  |
|  |  | 1570 | Bon |  | Fran | Fe |  |
| **3** | 4 | 699 | mal |  |
| 1354 | i | ce |  |
|  |  |  | e |  |
|  |  |  |  |  |  |  |
|  |  | 1573 | Mit |  | Spai | Fe |  |
| **4** | 5 | chel | 850 | mal |  |
| 7888 | n |  |
|  |  | l |  | e |  |
|  |  |  |  |  |  |
| 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 | | |  |  |

1. IsActiveMember 10000 non-null int64
2. EstimatedSalary 10000 non-null float64

4

2

4

1

4

2

3

9

4

3

2 0.00

1. 8380

7.86

1596

1. 60.8

0

1 0.00

1255

1. 10.8

2

1 1

1 0

3 1

1. 0
2. 1

101348.

1 1

88

112542.

1 0

58

113931.

0 1

57

93826.6

0 0

3

79084.1

1 0

0

13 Exited 10000 non-null int64

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

memory usage: 1.1+ MB

df**.**describe()

**RowN**

**umbe**

**r**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Custo** | **Credit** | **Age** | **Tenur** |  |
| **merId** | **Score** | **e** |  |
|  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Balanc** | **NumOf** | **HasC** | **IsActive** | **Estimat** |  |
| **Product** | **Membe** | **edSalarExited** |  |
| **e** | **rCard** |  |
| **s** | **r** | **y** |  |
|  |  |  |

**co**

**un**

**t**

**m**

**ea**

**n**

**st**

**d**

**mi**

**n**

**25**

**%**

**50**

**%**

**75**

**%**

**m**

**ax**

10000.

00000

5000.5

0000

2886.8

9568

1.0000

0

2500.7

5000

5000.5

0000

7500.2

5000

10000.

00000

1.0000

00e+0

4

1.5690

94e+0

7

7.1936

19e+0

4

1.5565

70e+0

7

1.5628

53e+0

7

1.5690

74e+0

7

1.5753

23e+0

7

1.5815

69e+0

7

10000.

00000

0

650.52

8800

96.653

299

350.00

0000

584.00

0000

652.00

0000

718.00

0000

850.00

0000

10000. 10000.

00000 00000

0 0

38.921 5.0128

800 00

10.487 2.8921

806 74

18.000 0.0000

000 00

32.000 3.0000

000 00

37.000 5.0000

000 00

44.000 7.0000

000 00

92.000 10.000

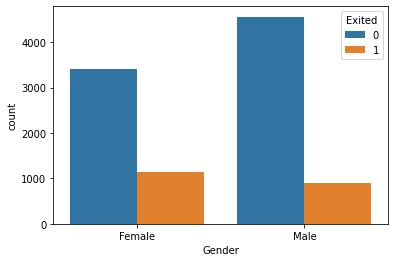
000 000

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 10000. | 10000.0 | 10000 | 10000.0 | 10000.0 | 10000. |  |
| .0000 | 00000 |  |
| 000000 | 00000 | 00000 | 00000 |  |
| 0 | 0 |  |
|  |  |  |  |  |
| 76485. | 1.53020 | 0.705 | 0.51510 | 100090. | 0.2037 |  |
| 889288 | 0 | 50 | 0 | 239881 | 00 |  |
| 62397. | 0.58165 | 0.455 | 0.49979 | 57510.4 | 0.4027 |  |
| 405202 | 4 | 84 | 7 | 92818 | 69 |  |
| 0.0000 | 1.00000 | 0.000 | 0.00000 | 11.5800 | 0.0000 |  |
| 00 | 0 | 00 | 0 | 00 | 00 |  |
| 0.0000 | 1.00000 | 0.000 | 0.00000 | 51002.1 | 0.0000 |  |
| 00 | 0 | 00 | 0 | 10000 | 00 |  |
| 97198. | 1.00000 | 1.000 | 1.00000 | 100193. | 0.0000 |  |
| 540000 | 0 | 00 | 0 | 915000 | 00 |  |
| 127644 | 2.00000 | 1.000 | 1.00000 | 149388. | 0.0000 |  |
| .24000 |  |
| 0 | 00 | 0 | 247500 | 00 |  |
| 0 |  |
|  |  |  |  |  |  |
| 250898 | 4.00000 | 1.000 | 1.00000 | 199992. | 1.0000 |  |
| .09000 |  |
| 0 | 00 | 0 | 480000 | 00 |  |
| 0 |  |
|  |  |  |  |  |  |

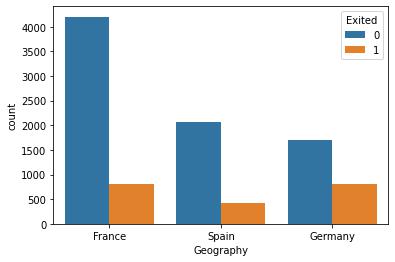
**Perform Below Visualizations.**

**Univariate Analysis**

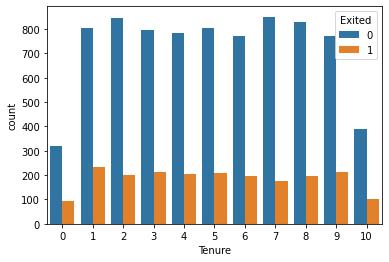
sns**.**countplot(x**=**"Gender",data**=**df,hue**=**"Exited") <matplotlib.axes.\_subplots.AxesSubplot at 0x7efc783e5450>



sns**.**countplot(x**=**"Geography",data**=**df,hue**=**"Exited") <matplotlib.axes.\_subplots.AxesSubplot at 0x7efc783590d0>



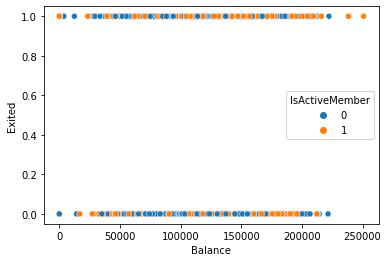
sns**.**countplot(x**=**"Tenure",data**=**df,hue**=**"Exited") <matplotlib.axes.\_subplots.AxesSubplot at 0x7efc77e375d0>



**Bi - Variate Analysis**

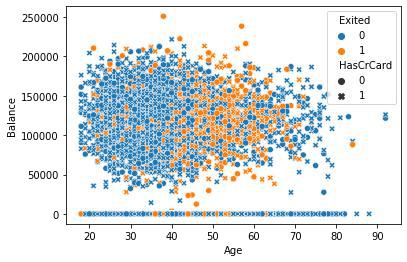
sns**.**scatterplot(x**=**"Balance", y**=**"Exited",data**=**df,hue**=**"IsActiveMember")

<matplotlib.axes.\_subplots.AxesSubplot at 0x7efc77e2a590>



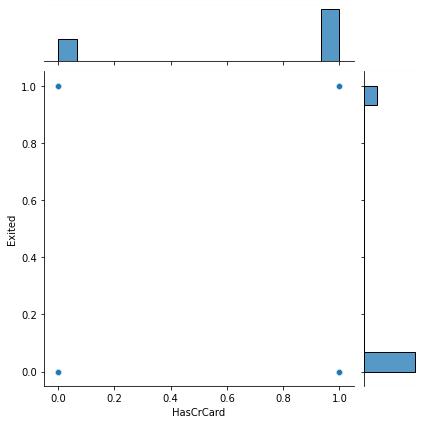
sns**.**scatterplot(x**=**"Age", y**=**"Balance",data**=**df,hue**=**"Exited" ,style**=**"HasCrCard")

<matplotlib.axes.\_subplots.AxesSubplot at 0x7efc77dec410>

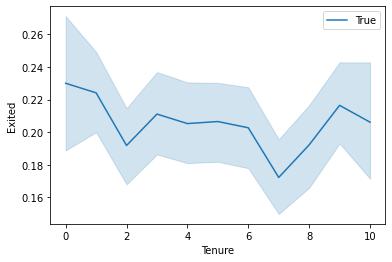


sns**.**jointplot(x**=**"HasCrCard",y**=**"Exited",data**=**df)

<seaborn.axisgrid.JointGrid at 0x7efc77d00650>



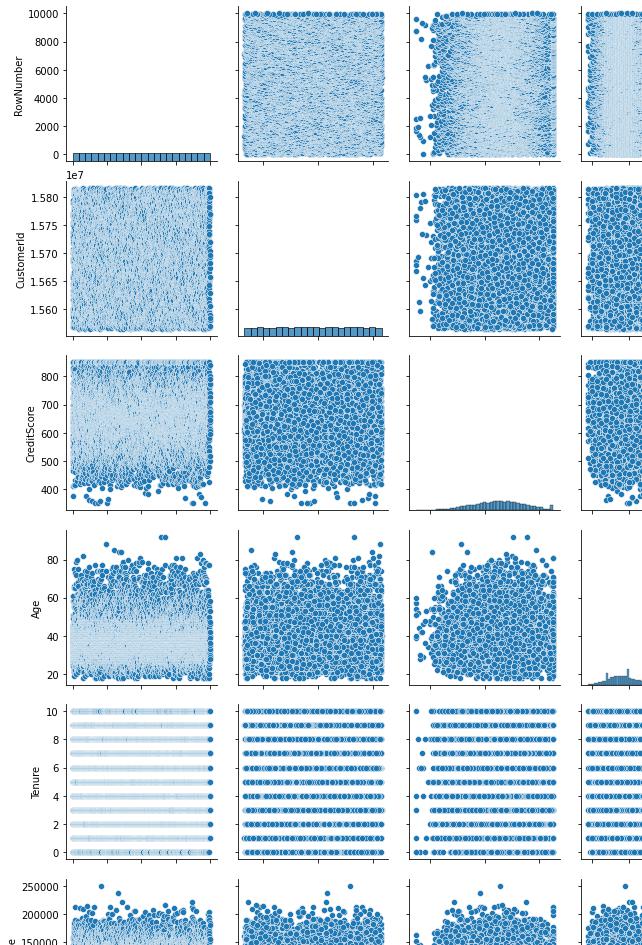
sns**.**lineplot(x**=**"Tenure", y**=**"Exited",data**=**df,hue**=True**) <matplotlib.axes.\_subplots.AxesSubplot at 0x7efc77b4d4d0>



**Multi - Variate Analysis**

sns**.**pairplot(data**=**df)

<seaborn.axisgrid.PairGrid at 0x7efc77b16e10>



**Perform descriptive statistics on the dataset.**

df**.**describe()

**RowN**

**umbe**

**r**

**co** 10000.

**un** 00000 **t**

1. 5000.5

**ea**0000

**n**

**st** 2886.8

1. 9568

**mi** 1.0000

1. 0
2. 2500.7

**%**5000

1. 5000.5

**%**0000

1. 7500.2

**%**5000

1. 10000. **ax** 00000

**Custo**

**merId**

1.0000

00e+0

4

1.5690

94e+0

7

7.1936

19e+0

4

1.5565

70e+0

7

1.5628

53e+0

7

1.5690

74e+0

7

1.5753

23e+0

7

1.5815

69e+0

7

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Credit** | **Age** | **Tenur** | **Balanc** |  |
| **Score** | **e** | **e** |  |
|  |  |
| 10000. | 10000. | 10000. | 10000. |  |
| 00000 | 00000 | 00000 |  |
| 000000 |  |
| 0 | 0 | 0 |  |
|  |  |
| 650.52 | 38.921 | 5.0128 | 76485. |  |
| 8800 | 800 | 00 | 889288 |  |
| 96.653 | 10.487 | 2.8921 | 62397. |  |
| 299 | 806 | 74 | 405202 |  |
| 350.00 | 18.000 | 0.0000 | 0.0000 |  |
| 0000 | 000 | 00 | 00 |  |
| 584.00 | 32.000 | 3.0000 | 0.0000 |  |
| 0000 | 000 | 00 | 00 |  |
| 652.00 | 37.000 | 5.0000 | 97198. |  |
| 0000 | 000 | 00 | 540000 |  |
| 718.00 | 44.000 | 7.0000 | 127644 |  |
| .24000 |  |
| 0000 | 000 | 00 |  |
| 0 |  |
|  |  |  |  |
| 850.00 | 92.000 | 10.000 | 250898 |  |
| .09000 |  |
| 0000 | 000 | 000 |  |
| 0 |  |
|  |  |  |  |

|  |  |  |
| --- | --- | --- |
| **NumOf** | **HasC** |  |
| **Product** |  |
| **rCard** |  |
| **s** |  |
|  |  |
| 10000.0 | 10000 |  |
| .0000 |  |
| 00000 |  |
| 0 |  |
|  |  |
| 1.53020 | 0.705 |  |
| 0 | 50 |  |
| 0.58165 | 0.455 |  |
| 4 | 84 |  |
| 1.00000 | 0.000 |  |
| 0 | 00 |  |
| 1.00000 | 0.000 |  |
| 0 | 00 |  |
| 1.00000 | 1.000 |  |
| 0 | 00 |  |
| 2.00000 | 1.000 |  |
| 0 | 00 |  |
| 4.00000 | 1.000 |  |
| 0 | 00 |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **IsActive** | **Estimat** |  |  |
| **Membe** | **edSalar** | **Exited** |  |
| **r** | **y** |  |  |
| 10000.0 | 10000.0 | 10000. |  |
| 00000 |  |
| 00000 | 00000 |  |
| 0 |  |
|  |  |  |
| 0.51510 | 100090. | 0.2037 |  |
| 0 | 239881 | 00 |  |
| 0.49979 | 57510.4 | 0.4027 |  |
| 7 | 92818 | 69 |  |
| 0.00000 | 11.5800 | 0.0000 |  |
| 0 | 00 | 00 |  |
| 0.00000 | 51002.1 | 0.0000 |  |
| 0 | 10000 | 00 |  |
| 1.00000 | 100193. | 0.0000 |  |
| 0 | 915000 | 00 |  |
| 1.00000 | 149388. | 0.0000 |  |
| 0 | 247500 | 00 |  |
| 1.00000 | 199992. | 1.0000 |  |
| 0 | 480000 | 00 |  |

**Handle the Missing values.**

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 | |

1. IsActiveMember 10000 non-null int64
2. 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

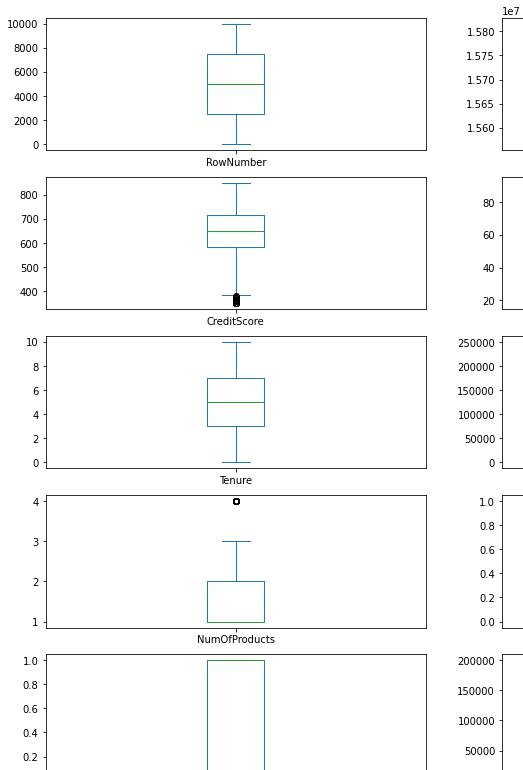
dtype: int64

There is no missing value

**Find the outliers and replace the outliers**

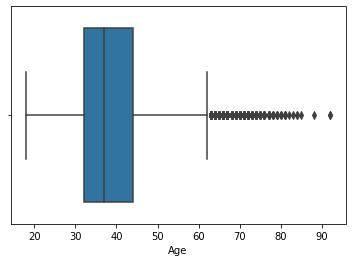
df**.**plot(kind**=**"box",subplots**=**"True",layout**=**(7,2),figsize**=**(15,20))

plt**.**show()



sns**.**boxplot(x**=**"Age",data**=**df)

<matplotlib.axes.\_subplots.AxesSubplot at 0x7efc70b494d0>



a**=** np**.**percentile(df**.**Age,[90])[0]

a

53.0

df[(df**.**Age**>**a)]

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Row** | **Cust** | **Sur** | **Cred** | **Geo** | **Ge** | **A** | **Te** | **Bala** | **NumOf** | **Has** | **IsActiv** | **Estima** | **Ex** |  |
|  | **Num** | **omer** | **na** | **itSco** | **grap** | **nd** | **g** | **nu** | **Produc** | **CrC** | **eMemb** | **tedSala** | **ite** |  |
|  | **nce** |  |
|  | **ber** | **Id** | **me** | **re** | **hy** | **er** | **e** | **re** | **ts** | **ard** | **er** | **ry** | **d** |  |
|  |  |  |
| **1** |  | 1573 | Ro |  | Ger | Ma | 5 |  | 132 |  |  |  |  |  |  |
| 17 | 653 | man | 1 | 602. | 1 | 1 | 0 | 5097.67 | 1 |  |
| **6** | 7452 | meo | le | 8 |  |
|  |  | y |  | 88 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **4** |  | 1568 | Osb |  | Fran | Fe | 6 |  | 117 |  |  |  | 94153.8 |  |  |
| 43 | 556 | ma | 2 | 419. | 1 | 1 | 1 | 0 |  |
| **2** | 7946 | orne | ce | 1 | 3 |  |
|  |  | le |  | 35 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **4** |  | 1568 | Bia |  | Spai | Fe | 6 |  | 155 |  |  |  | 158338. |  |  |
| 45 | 660 | ma | 5 | 931. | 1 | 1 | 1 | 0 |  |
| **4** | 4171 | nchi | n | 1 | 39 |  |
|  |  | le |  | 11 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Row** | **Cust** | **Sur** | **Cred** | **Geo** | **Ge** | **A** | **Te** | **Bala** | **NumOf** | **Has** | **IsActiv** | **Estima** | **Ex** |  |
| **Num** | **omer** | **na** | **itSco** | **grap** | **nd** | **g** | **nu** | **Produc** | **CrC** | **eMemb** | **tedSala** | **ite** |  |
| **nce** |  |
| **ber** | **Id** | **me** | **re** | **hy** | **er** | **e** | **re** | **ts** | **ard** | **er** | **ry** | **d** |  |
|  |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **5** |  | 1562 | T'ie |  | Spai | Fe | 6 |  |
| 59 | 511 | ma |  |
| **8** | 3944 | n | n | 6 |  |
|  |  | le |  |
|  |  |  |  |  |  |  |  |
| **6** | 64 | 1575 | Piro | 684 | Spai | Ma | 5 |  |
| **3** | 1208 | zzi | n | le | 6 |  |
|  |  |  |
| **...** | ... | ... | ... | ... | ... | ... ... | |  |

|  |  |  |  |
| --- | --- | --- | --- |
| 4 | 0.00 | 1 | 1 |
|  | 787 |  |  |
| 8 | 07.1 | 1 | 1 |
|  | 6 |  |  |
| ... | ... | ... | ... |

0 1643.11 1

99398.3

1 0

6

... ... ...

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **9** |  |  |  |  |
| **9** | 9910 | 1577 | End |  |
| **0** | 3338 | rizzi |  |
|  |  |
| **9** |  |  |  |  |
| **9** |  |  |  |  |
| **9** | 9911 | 1578 | L? |  |
| **1** | 4042 |  |
|  |  |  |
| **0** |  |  |  |  |
| **9** |  |  |  |  |
| **9** | 9937 | 1565 | Par |  |
| **3** | 3037 | ks |  |
|  |  |
| **6** |  |  |  |  |
| **9** |  |  |  |  |
| **9** | 9940 | 1580 | Laj |  |
| **3** | 8971 | oie |  |
|  |  |
| **9** |  |  |  |  |
| **9** |  |  |  |  |
| **9** | 9980 | 1569 | Diri |  |
| **7** | 2664 | be |  |
|  |  |
| **9** |  |  |  |  |

966 rows × 14 columns

b**=**df["Age"]**.**mean()

b

38.9218

df**.**Age[(df**.**Age**>** a)] **=** b

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 739 | Fran | Ma | 5 |  |
| ce | le | 8 |  |
|  |  |
| 624 | Fran | Ma | 5 |  |
| ce | le | 5 |  |
|  |  |
| 609 | Fran | Ma | 7 |  |
| ce | le | 7 |  |
|  |  |
|  | Spai | Fe | 5 |  |
| 693 | ma |  |
| n | 7 |  |
|  | le |  |
|  |  |  |  |
|  | Fran | Fe | 5 |  |
| 677 | ma |  |
| ce | 8 |  |
|  | le |  |
|  |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
|  | 101 |  |  |
| 2 | 579. | 1 | 1 |
|  | 28 |  |  |
|  | 118 |  |  |
| 7 | 793. | 1 | 1 |
|  | 60 |  |  |
| 1 | 0.00 | 1 | 0 |
| 9 | 0.00 | 2 | 1 |
|  | 900 |  |  |
| 1 | 22.8 | 1 | 0 |
|  | 5 |  |  |

72168.5

1 0

3

95022.0

1 1

2

18708.7

1 0

6

135502.

1 0

77

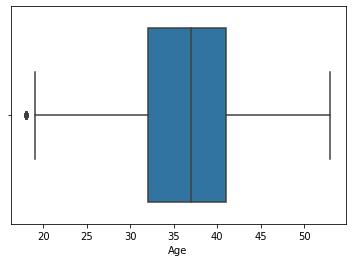
1 2988.28 0

df[(df**.**Age**>**a)]

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Row** | **Cust** | **Sur** | **Credi** | **Geog** | **Ge** | **A** | **Te** | **Bal** | **NumOf** | **HasC** | **IsActiv** | **Estimat** | **Ex** |
| **Num** | **omer** | **na** | **tScor** | **raph** | **nd** | **g** | **nu** | **anc** | **Produc** | **rCar** | **eMemb** | **edSalar** | **ite** |
| **ber** | **Id** | **me** | **e** | **y** | **er** | **e** | **re** | **e** | **ts** | **d** | **er** | **y** | **d** |

sns**.**boxplot(x**=**"Age",data**=**df)

<matplotlib.axes.\_subplots.AxesSubplot at 0x7efc702356d0>



c**=** np**.**percentile(df**.**Age,[1])[0]

c

21.0

df[(df**.**Age**<**19)]

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Row** | **Cust** | **Surna** | **Cred** | **Geo** | **Ge** | **A** | **Te** | **Bal** | **NumO** | **Has** | **IsActiv** | **Estima** | **Ex** |  |
|  | **Num** | **omer** | **itSco** | **grap** | **nd** | **g** | **nu** | **anc** | **fProdu** | **CrC** | **eMem** | **tedSal** | **ite** |  |
|  | **me** |  |
|  | **ber** | **Id** | **re** | **hy** | **er** | **e** | **re** | **e** | **cts** | **ard** | **ber** | **ary** | **d** |  |
|  |  |  |
| **7** |  | 1578 |  |  | Fran | Ma | 1 |  | 160 |  |  |  | 145936 |  |  |
| **4** | 747 | Hsieh | 844 | 8. | 2 | 980. | 1 | 0 | 0 | 0 |  |
| 7619 | ce | le | .28 |  |
| **6** |  |  |  | 0 |  | 03 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| **1** |  |  |  |  |  |  | 1 |  | 151 |  |  |  |  |  |  |
| **6** |  | 1577 | McDon |  | Fran | Ma |  |  |  |  | 127014 |  |  |
| 1620 | 656 | 8. | 10 | 762. | 1 | 0 | 1 | 0 |  |
| **1** | 0309 | ald | ce | le | .32 |  |
|  |  | 0 |  | 74 |  |  |  |  |  |
| **9** |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **1** |  |  |  |  |  | Fe | 1 |  | 827 |  |  |  |  |  |  |
| **6** |  | 1556 | Kharla |  | Fran |  |  |  |  | 71811. |  |  |
| 1679 | 570 | ma | 8. | 4 | 67.4 | 1 | 1 | 0 | 0 |  |
| **7** | 9178 | mov | ce | 90 |  |
|  |  | le | 0 |  | 2 |  |  |  |  |  |
| **8** |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**2**

**0**

**2**

**1**

**2**

**1**

**3**

**6**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Row** | **Cust** | **Surna** | **Cred** | **Geo** | **Ge** | **A** | **Te** | **Bal** | **NumO** | **Has** | **IsActiv** | **Estima** | **Ex** |  |
| **Num** | **omer** | **itSco** | **grap** | **nd** | **g** | **nu** | **anc** | **fProdu** | **CrC** | **eMem** | **tedSal** | **ite** |  |
| **me** |  |
| **ber** | **Id** | **re** | **hy** | **er** | **e** | **re** | **e** | **cts** | **ard** | **ber** | **ary** | **d** |  |
|  |  |
|  | 1579 | Vasilie |  | Ger | Fe | 1 |  | 128 |  |  |  | 197322 |  |  |
| 2022 | 716 | man | ma | 8. | 3 | 743. | 1 | 0 | 0 | 0 |  |
| 5519 | v | .13 |  |
|  |  | y | le | 0 |  | 80 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  | 1562 | Bellucc |  | Fran | Ma | 1 |  | 133 |  |  |  | 46941. |  |  |
| 2137 | 727 | 8. | 4 | 550. | 1 | 1 | 1 | 0 |  |
| 1893 | i | ce | le | 41 |  |
|  |  | 0 |  | 67 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |

**2**

**1**

**4**

**1**

**3**

**3**

**3**

**0**

**3**

**5**

**1**

**2**

**3**

**5**

**1**

**7**

**3**

**6**

**8**

**6**

**4**

**5**

**5**

**6**

**4**

**7**

**1**

**6**

**7**

**3**

**3**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1575 | Wallac |  | Fran | Ma | 1 |  |  |  |  |  |
| 2142 | 674 | 8. | 7 | 0.00 | 2 | 1 |  |
| 8372 | e | ce | le |  |
|  |  | 0 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  | 1565 |  |  | Fran | Ma | 1 |  |  |  |  |  |
| 3331 | Chao | 738 | 8. | 4 | 0.00 | 2 | 1 |  |
| 7439 | ce | le |  |
|  |  |  | 0 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  | 1565 |  |  | Spai | Ma | 1 |  |  |  |  |  |
| 3513 | Boylan | 806 | 8. | 3 | 0.00 | 2 | 1 |  |
| 7779 | n | le |  |
|  |  |  | 0 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  | 1575 | Burges |  | Spai | Ma | 1 |  |  |  |  |  |
| 3518 | 771 | 8. | 1 | 0.00 | 2 | 0 |  |
| 7821 | s | n | le |  |
|  |  | 0 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  | 1566 | Cattane |  | Fran | Ma | 1 |  | 176 |  |  |  |
| 3687 | 706 | 8. | 2 | 139. | 2 | 1 |  |
| 5327 | o | ce | le |  |
|  |  | 0 |  | 50 |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  | 1579 | Nwank |  | Fran | Fe | 1 |  | 988 |  |  |  |
| 4557 | 681 | ma | 8. | 1 | 94.3 | 1 | 1 |  |
| 6231 | wo | ce |  |
|  |  | le | 0 |  | 9 |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  | 1580 | Hallaha |  | Fran | Ma | 1 |  |  |  |  |  |
| 4717 | 646 | 8. | 10 | 0.00 | 2 | 0 |  |
| 5764 | n | ce | le |  |
|  |  | 0 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 7335 | 1575 | Vaguin | 616 | Fran | Ma | 1 | 6 | 0.00 | 2 | 1 |  |
| 9133 | e | ce | le |  |
|  |  | 8. |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |

55753.

1 1

12

47799.

1 0

15

86994.

1 0

54

41542.

0 0

95

129654

0 0

.22

9596.4

1 0

0

52795.

1 0

15

27308.

1 0

58

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Row** | **Cust** | **Surna** | **Cred** | **Geo** | **Ge** | **A** | **Te** | **Bal** | **NumO** | **Has** | **IsActiv** | **Estima** | **Ex** |  |
| **Num** | **omer** | **itSco** | **grap** | **nd** | **g** | **nu** | **anc** | **fProdu** | **CrC** | **eMem** | **tedSal** | **ite** |  |
| **me** |  |
| **ber** | **Id** | **re** | **hy** | **er** | **e** | **re** | **e** | **cts** | **ard** | **ber** | **ary** | **d** |  |
|  |  |
| **4** |  |  |  |  |  | 0 |  |  |  |  |  |  |  |  |

**7**

**7**

**2**

**2**

**8**

**5**

**2**

**2**

**9**

**0**

**2**

**9**

**9**

**5**

**0**

**1**

**9**

**5**

**2**

**0**

**9**

**5**

**2**

**6**

**9**

**5**

**7**

**2**

**9**

**7**

**8**

**2**

**9**

**9**

**3**

**2**

7723

8523

9030

9502

9521

9527

9573

9783

9933

1557

0086

1561

9892

1572

2701

1563

4146

1567

3180

1566

5521

1564

1688

1572

8829

1581

3451

Lynch

Page

Bruno

Hou

Onyeka

ozulu

Chiaza

gomek

pele

Collier

Weigel

Fleetw

ood-

Smith

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Ger | Ma | 1 |  | 905 |  |  |  |
| 684 | man | 8. | 9 | 44.0 | 1 | 0 |  |
| le |  |
|  | y | 0 |  | 0 |  |  |  |
|  |  |  |  |  |  |
|  | Spai | Ma | 1 |  |  |  |  |  |
| 644 | 8. | 8 | 0.00 | 2 | 1 |  |
| n | le |  |
|  | 0 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  | Ger | Ma | 1 |  | 132 |  |  |  |
| 594 | man | 8. | 1 | 694. | 1 | 1 |  |
| le |  |
|  | y | 0 |  | 73 |  |  |  |
|  |  |  |  |  |  |
|  | Ger | Ma | 1 |  | 142 |  |  |  |
| 835 | man | 8. | 2 | 872. | 1 | 1 |  |
| le |  |
|  | y | 0 |  | 36 |  |  |  |
|  |  |  |  |  |  |
|  | Ger | Fe | 1 |  | 938 |  |  |  |
| 727 | man | ma | 8. | 2 | 16.7 | 2 | 1 |  |
|  | y | le | 0 |  | 0 |  |  |  |
|  | Ger | Ma | 1 |  | 111 |  |  |  |
| 642 | man | 8. | 5 | 183. | 2 | 0 |  |
| le |  |
|  | y | 0 |  | 53 |  |  |  |
|  |  |  |  |  |  |
|  | Spai | Ma | 1 |  |  |  |  |  |
| 644 | 8. | 7 | 0.00 | 1 | 0 |  |
| n | le |  |
|  | 0 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  | Fran | Ma | 1 |  | 102 |  |  |  |
| 509 | 8. | 7 | 983. | 1 | 1 |  |
| ce | le |  |
|  | 0 |  | 91 |  |  |  |
|  |  |  |  |  |  |  |
|  | Spai | Ma | 1 |  | 134 |  |  |  |
| 677 | 8. | 8 | 796. | 2 | 1 |  |
| n | le |  |
|  | 0 |  | 87 |  |  |  |
|  |  |  |  |  |  |  |

4777.2

1 0

3

59172.

0 0

42

167689

0 0

.56

117632

1 0

.63

126172

0 0

.11

10063.

1 0

75

59645.

1 1

24

171770

0 0

.58

114858

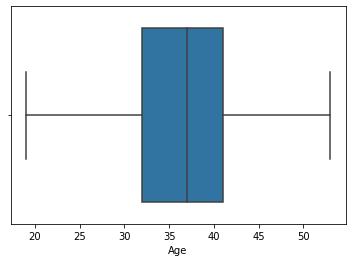
1 0

.90

df**.**Age[(df**.**Age**<**19)] **=** b

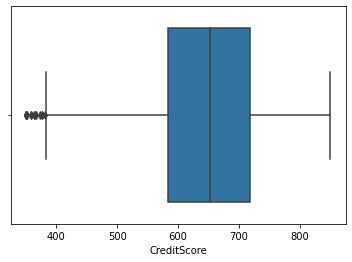
sns**.**boxplot(x**=**"Age",data**=**df)

<matplotlib.axes.\_subplots.AxesSubplot at 0x7efc701a34d0>



sns**.**boxplot(x**=**"CreditScore",data**=**df)

<matplotlib.axes.\_subplots.AxesSubplot at 0x7efc70173050>



d**=** np**.**percentile(df**.**CreditScore,[1])[0]

d

432.0

df[(df**.**CreditScore**<**d)]

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Row** | **Cust** | **Surn** | **Cred** | **Geo** | **Ge** | **A** | **Te** | **Bal** | **NumO** | **Has** | **IsActiv** | **Estima** | **Ex** |  |
| **Num** | **omer** | **itSco** | **grap** | **nd** | **g** | **nu** | **anc** | **fProdu** | **CrC** | **eMemb** | **tedSala** | **ite** |  |
| **ame** |  |
| **ber** | **Id** | **re** | **hy** | **er** | **e** | **re** | **e** | **cts** | **ard** | **er** | **ry** | **d** |  |
|  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **7** | 8 | 1565 | Obin |  |
| 6148 | na |  |
|  |  |  |
| **2** | 30 | 1565 | Lucc |  |
| **9** | 6300 | iano |  |
|  |  |
| **7** | 80 | 1580 | Postl |  |
| **9** | 3136 | e |  |
|  |  |
| **9** | 100 | 1563 | Fanu |  |
| **9** | 3059 | cci |  |
|  |  |
| **1** |  | 1579 | Harri |  |
| **4** | 150 |  |
| 4413 | s |  |
| **9** |  |  |
|  |  |  |  |

Ger

1. man y
2. Fran ce

Ger

1. man y
2. Fran ce
3. Fran ce

Fe

ma

le

Ma

le

Fe

ma

le

Ma

le

Ma

le

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 2 |  | 115 |  |  |  |
| 9. | 4 | 046. | 4 | 1 | 0 |
| 0 |  | 74 |  |  |  |
| 2 |  | 596 |  |  |  |
| 9. | 0 | 97.1 | 2 | 1 | 1 |
| 0 |  | 7 |  |  |  |
| 4 |  | 122 |  |  |  |
| 1. | 10 | 189. | 2 | 1 | 0 |
| 0 |  | 66 |  |  |  |
| 3 |  |  |  |  |  |
| 4. | 9 | 0.00 | 2 | 0 | 0 |
| 0 |  |  |  |  |  |
| 3 |  |  |  |  |  |
| 2. | 0 | 0.00 | 2 | 0 | 1 |
| 0 |  |  |  |  |  |

119346

.88

53483.

21

98301.

61

6534.1

8

878.87

1

0

0

0

0

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Row** | **Cust** | **Surn** | **Cred** | **Geo** | **Ge** | **A** | **Te** | **Bal** | **NumO** | **Has** | **IsActiv** | **Estima** | **Ex** |  |
| **Num** | **omer** | **itSco** | **grap** | **nd** | **g** | **nu** | **anc** | **fProdu** | **CrC** | **eMemb** | **tedSala** | **ite** |  |
| **ame** |  |
| **ber** | **Id** | **re** | **hy** | **er** | **e** | **re** | **e** | **cts** | **ard** | **er** | **ry** | **d** |  |
|  |  |

**...**

**9**

**3**

**5**

**7**

**9**

**4**

**0**

**7**

**9**

**5**

**2**

**2**

**9**

**6**

**2**

**4**

**9**

**9**

**3**

**0**

...

9358

9408

9523

9625

9931

...

1581

4405

1565

2835

1566

4504

1566

8309

1571

3604

...

Ches

noko

va

Lian

g

Beed

e

Masl

ow

Rossi

... ...

1. Fran ce
2. Spai n
3. Fran ce
4. Fran ce

Ger

1. man y

...

Fe

ma

le

Fe

ma

le

Ma

le

Fe

ma

le

Ma

le

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ... | ... | ... | ... | ... | ... |
| 4 |  |  |  |  |  |
| 6. | 9 | 0.00 | 1 | 1 | 1 |
| 0 |  |  |  |  |  |
| 2 |  | 121 |  |  |  |
| 7. | 2 | 580. | 1 | 0 | 1 |
| 0 |  | 42 |  |  |  |
| 3 |  |  |  |  |  |
| 5. | 7 | 0.00 | 2 | 1 | 1 |
| 0 |  |  |  |  |  |
| 4 |  | 111 |  |  |  |
| 0. | 0 | 098. | 1 | 1 | 1 |
| 0 |  | 85 |  |  |  |
| 4 |  | 166 |  |  |  |
| 0. | 9 | 776. | 2 | 0 | 1 |
| 0 |  | 60 |  |  |  |

...

81014.

50

134720

.51

88878.

15

172321

.21

172646

.88

...

1

0

0

1

0

99 rows × 14 columns

g**=**df["CreditScore"]**.**mean()

g

650.5288

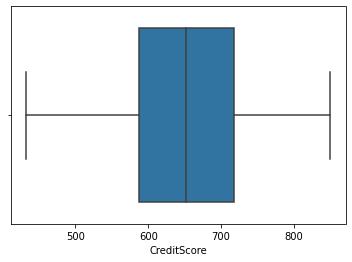
df**.**CreditScore[(df**.**CreditScore**<**d)] **=** g

df[(df**.**CreditScore**<**d)]

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Row** | **Cust** | **Sur** | **Credi** | **Geog** | **Ge** | **A** | **Te** | **Bal** | **NumOf** | **HasC** | **IsActiv** | **Estimat** | **Ex** |
| **Num** | **omer** | **na** | **tScor** | **raph** | **nd** | **g** | **nu** | **anc** | **Produc** | **rCar** | **eMemb** | **edSalar** | **ite** |
| **ber** | **Id** | **me** | **e** | **y** | **er** | **e** | **re** | **e** | **ts** | **d** | **er** | **y** | **d** |

sns**.**boxplot(x**=**"CreditScore",data**=**df)

<matplotlib.axes.\_subplots.AxesSubplot at 0x7efc700eadd0>



**Check for Categorical columns and perform encoding.**

**del** df['Surname']

df **=** pd**.**get\_dummies(df,drop\_first**=True**)

df**.**head()

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Row** | **Cust** | **Cre** | **A** | **Te** | **Bal** | **Num** | **Has** | **IsActi** | **Estim** | **E** | **Geograp** | **Geogr** | **Gend** |
| **Num** | **ome** | **ditS** | **g** | **nu** | **anc** | **OfPro** | **CrC** | **veMe** | **atedS** | **xit** | **hy\_Ger** | **aphy\_ er\_M** | |
| **ber** | **rId** | **core** | **e** | **re** | **e** | **ducts** | **ard** | **mber** | **alary** | **ed** | **many** | **Spain** | **ale** |

**0**

**1**

**2**

**3**

4

1 1563 619. 2

4602 0 .

0

4

2 1564 608. 1

7311 0 .

0

4

3 1561 502. 2

9304 0 .

0

3

4 1570 699. 9

1354 0 .

0

2

1

8

1

0.0

0

838

07.

86

159

660

.80

0.0

0

1 1

1 0

1. 1
2. 0

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 10134 | 1 | 0 | 0 | 0 |  |
| 8.88 |  |
|  |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 11254 | 0 | 0 | 1 | 0 |  |
| 2.58 |  |
|  |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 0 | 11393 | 1 | 0 | 0 | 0 |  |
| 1.57 |  |
|  |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 0 | 93826. | 0 | 0 | 0 | 0 |  |
| 63 |  |
|  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Row** | **Cust** | **Cre** | **A** | **Te** | **Bal** | **Num** | **Has** | **IsActi** | **Estim** | **E** | **Geograp** | **Geogr** | **Gend** |  |
|  | **Num** | **ome** | **ditS** | **g** | **nu** | **anc** | **OfPro** | **CrC** | **veMe** | **atedS** | **xit** | **hy\_Ger** | **aphy\_ er\_M** | |  |
|  | **ber** | **rId** | **core** | **e** | **re** | **e** | **ducts** | **ard** | **mber** | **alary** | **ed** | **many** | **Spain** | **ale** |  |
|  |  |  |  | 4 |  | 125 |  |  |  |  |  |  |  |  |  |
|  |  | 1573 | 850. | 3 |  |  |  |  | 79084. |  |  |  |  |  |
| **4** | 5 | 2 | 510 | 1 | 1 | 1 | 0 | 0 | 1 | 0 |  |
| 7888 | 0 | . | 10 |  |
|  |  |  | .82 |  |  |  |  |  |  |  |  |
|  |  |  |  | 0 |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**del** df['CustomerId']

**del** df['RowNumber'] df**.**head()

**Credi** **A** **Te**

**tScor** **g** **nur**

**e** **e** **e**

4

**0** 619.0 2. 2

0

4

**1** 608.0 1. 1

0

4

**2** 502.0 2. 8

0

3

**3** 699.0 9. 1

0

4

**4** 850.0 3. 2

0

**Bala**

**nce**

0.00

8380

7.86

1596

60.8

0

0.00

1255

10.8

2

|  |  |  |
| --- | --- | --- |
| **NumOf** | **HasC** | **IsActive** |
| **Product** | **rCar** | **Membe** |
| **s** | **d** | **r** |
| 1 | 1 | 1 |
| 1 | 0 | 1 |
| 3 | 1 | 0 |
| 2 | 0 | 0 |
| 1 | 1 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Estimat** | **Ex** | **Geograph** | **Geograp** | **Gende** |  |
| **edSalar** | **ite** | **y\_German** | **hy\_Spai** | **r\_Mal** |  |
| **y** | **d** | **y** | **n** | **e** |  |
| 101348. | 1 | 0 | 0 | 0 |  |
| 88 |  |
|  |  |  |  |  |
| 112542. | 0 | 0 | 1 | 0 |  |
| 58 |  |
|  |  |  |  |  |
| 113931. | 1 | 0 | 0 | 0 |  |
| 57 |  |
|  |  |  |  |  |
| 93826.6 | 0 | 0 | 0 | 0 |  |
| 3 |  |
|  |  |  |  |  |
| 79084.1 | 0 | 0 | 1 | 0 |  |
| 0 |  |
|  |  |  |  |  |

**Split the data into dependent and independent variables.**

x**=**df**.**loc[:,df**.**columns**!=**"Exited"]

y**=**df["Exited"]

x**.**head()

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Credit** | **A** | **Ten** | **Bala** | **NumOfP** | **HasC** | **IsActive** | **Estimate** | **Geography** | **Geograp** | **Gender** |  |
|  | **Score** | **ge** | **ure** | **nce** | **roducts** | **rCard** | **Member** | **dSalary** | **\_Germany** | **hy\_Spain** | **\_Male** |  |
|  |  | 4 |  |  |  |  |  | 101348.8 |  |  |  |  |
| **0** | 619.0 | 2. | 2 | 0.00 | 1 | 1 | 1 | 0 | 0 | 0 |  |
| 8 |  |
|  |  | 0 |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Credit** | **A** | **Ten** | **Bala** | **NumOfP** | **HasC** | **IsActive** | **Estimate** | **Geography** | **Geograp** | **Gender** |  |
|  | **Score** | **ge** | **ure** | **nce** | **roducts** | **rCard** | **Member** | **dSalary** | **\_Germany** | **hy\_Spain** | **\_Male** |  |
|  |  | 4 |  | 8380 |  |  |  | 112542.5 |  |  |  |  |
| **1** | 608.0 | 1. | 1 | 1 | 0 | 1 | 0 | 1 | 0 |  |
| 7.86 | 8 |  |
|  |  | 0 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 4 |  | 1596 |  |  |  | 113931.5 |  |  |  |  |
| **2** | 502.0 | 2. | 8 | 3 | 1 | 0 | 0 | 0 | 0 |  |
| 60.80 | 7 |  |
|  |  | 0 |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | 3 |  |  |  |  |  |  |  |  |  |  |
| **3** | 699.0 | 9. | 1 | 0.00 | 2 | 0 | 0 | 93826.63 | 0 | 0 | 0 |  |
|  |  | 0 |  |  |  |  |  |  |  |  |  |  |
|  |  | 4 |  | 1255 |  |  |  |  |  |  |  |  |
| **4** | 850.0 | 3. | 2 | 1 | 1 | 1 | 79084.10 | 0 | 1 | 0 |  |
| 10.82 |  |
|  |  | 0 |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |

y**.**head()

1. 1
2. 0
3. 1
4. 0
5. 0

Name: Exited, dtype: int64

**Scale the independent variables**

**from** sklearn **import** linear\_model

**from** sklearn.preprocessing **import** StandardScaler

scale **=** StandardScaler()

x**=**df[['Age','Tenure']]

scaledx **=** scale**.**fit\_transform(x)

scaledx

array([[ 0.75333656, -1.04175968],

[ 0.60998931, -1.38753759],

[ 0.75333656, 1.03290776],

...,

[-0.10674695, 0.68712986],

[ 0.75333656, -0.69598177],

[-1.25352496, -0.35020386]])

**Split the data into training and testing**

**from** sklearn.model\_selection **import** train\_test\_split

x\_train, x\_test, y\_train, y\_test**=**train\_test\_split(x,y,test\_size**=**0.2,random\_state**=**0)

x\_train**.**head()

**Age** **Tenure**

|  |  |  |
| --- | --- | --- |
|  | **Age** | **Tenure** |
| **7389** | 34.0000 | 5 |
| **9275** | 42.0000 | 1 |
| **2995** | 29.0000 | 2 |
| **5316** | 40.0000 | 5 |
| **356** | 38.9218 | 8 |

x\_test**.**head()

**Age** **Tenure**

|  |  |  |
| --- | --- | --- |
| **9394** | 35.0 | 8 |
| **898** | 40.0 | 2 |
| **2398** | 42.0 | 8 |
| **5906** | 32.0 | 4 |
| **2343** | 38.0 | 5 |
| y\_train**.**head() | |  |
| 7389 | 0 |  |
| 9275 | 0 |  |
| 2995 | 0 |  |
| 5316 | 0 |  |
| 356 | 0 |  |

Name: Exited, dtype: int64

y\_test**.**head()

9394 0

1. 1

2398 0

5906 0

2343 0

Name: Exited, dtype: int64

print('X Train shape:{},Y.Train Shape:{}'**.**format(x\_train**.**shape,y\_train**.**shape))

X Train shape:(8000, 2),Y.Train Shape:(8000,)

print('X Test Shape :{},Y Test Shape:{}'**.**format(x\_test**.**shape,y\_test**.**shape))

X Test Shape :(2000, 2),Y Test Shape:(2000,)