Assignment 2

9

9

5

9996

1560

6229

jiak

u

In []: #import libraries import numpy as np import pandas as pd import matplotlib.pyplot as plt import seaborn as sns %matplotlib inline #to view graph in colab itself In []: #load dataset df=pd.read csv("/content/Churn Modelling (1).csv") In []: df Out[]: Ex Row Cust Sur Cred Geo Ge A Te Bal NumO Has IsActiv Estima tedSal Num omer na itSco grap nd nu anc **fProdu** CrC eMem ite \mathbf{g} ber Id ard ber d hy me re er e re cts ary \mathbf{e} Har Fe 1563 4 101348 Fran 0 1 619 2 0.00 1 1 1 gra ma 4602 2 .88 ce ve le Fe 838 1564 112542 Spai 4 1 Hill 608 ma 07.8 1 0 0 7311 1 n .58 le 6 Fe 159 113931 1561 Oni Fran 4 2 3 502 3 8 1 660. 1 ma 9304 2 .57 o ce 80 le Fe 1570 Bon 3 93826. Fran 3 4 699 0.00 2 0 0 ma 1354 i 9 63 ce le 125 Mit Fe 1573 4 79084. Spai 4 5 chel 850 ma 510. 1 1 0 3 7888 n 10 le 1 82 9 Obi

3

0.00

2

1

Ma

le 9

Fran

ce

771

96270.

64

0

	Row Num ber	Cust omer Id	Sur na me	Cred itSco re	Geo grap hy	Ge nd er	A g e	Te nu re	Bal anc e	NumO fProdu cts	Has CrC ard	IsActiv eMem ber	Estima tedSal ary	Ex ite d
9 9 9 6	9997	1556 9892	Joh nsto ne	516	Fran ce	Ma le	3 5	10	573 69.6 1	1	1	1	101699 .77	0
9 9 9 7	9998	1558 4532	Liu	709	Fran ce	Fe ma le	3 6	7	0.00	1	0	1	42085. 58	1
9 9 9 8	9999	1568 2355	Sab bati ni	772	Ger man y	Ma le	4 2	3	750 75.3 1	2	1	0	92888. 52	1
9 9 9	1000	1562 8319	Wal ker	792	Fran ce	Fe ma le	2 8	4	130 142. 79	1	1	0	38190. 78	0
1000	0 rows	× 14 cc	olumns	5										
Perfo Anal		ow Vist	ualizat	ions.	Univar	iate A	naly	sis 🗨	Bi - Va	ariate Ana	ılysis ●	Multi - V	'ariate	
df.c	column	s												In []:
Inde	'G 'I	ender	', 'A	ge', ber',	'Tenur	ce',	'Ba	lanc	e', '	CreditS NumOfPr xited']	oduct		raphy'	
df["		Produc			ıe()									In []:
arra	ıy([1,	3, 2,	, 4])										C	Out[]:
df["	'NumOf	Produc	cts"]	.value	e_cour	nts()								In []:
1 2 3 4	5084 4590 266 60												C	Out[]:
	: Num ltypes	OfProd	ducts	, dtyr	pe: ir	nt64								ln []:
	Tumber			int(5 <i>4</i>								C)ut[]:
	omerI			int										

Cre Geo Gen Age Ten Bal Num Has IsA Est Exi	ename editSco ography der eure ance ofProc octivel imated rpe: ol	ducts d Member dSalar bject		obje int obje obje int int float int int int	64 ct ct 64 64 64 64 64 64									In []: Dut[]:
	Row Num ber	Cust omer Id	Sur na me	Cred itSco re	Geo grap hy	Ge nd er	A g e	Te nu re	Bal anc e	NumO fProdu cts	Has CrC ard	IsActiv eMemb er	Estima tedSala ry	Ex ite d
0	1	1563 4602	Har gra ve	619	Fran ce	Fe ma le	4 2	2	0.00	1	1	1	101348 .88	1
1	2	1564 7311	Hill	608	Spai n	Fe ma le	4	1	838 07.8 6	1	0	1	112542 .58	0
2	3	1561 9304	Oni o	502	Fran ce	Fe ma le	4 2	8	159 660. 80	3	1	0	113931 .57	1
3	4	1570 1354	Bon i	699	Fran ce	Fe ma le	3 9	1	0.00	2	0	0	93826. 63	0
4	5	1573 7888	Mit chel l	850	Spai n	Fe ma le	4 3	2	125 510. 82	1	1	1	79084. 10	0
df.	tail()												In []:
	Row Num ber	Cust omer Id	Sur na me	Cred itSco	Geo grap hy	Ge nd er	A g e	Te nu re	Bal anc	NumO fProdu cts	Has CrC ard	IsActiv eMem ber	Estima tedSal	Out[]: Ex ite d
	oer	10	me	re	ny	er	e	re	e	cts	ara	ber	ary	a

0

Obi

jiak u

1560

6229

9996

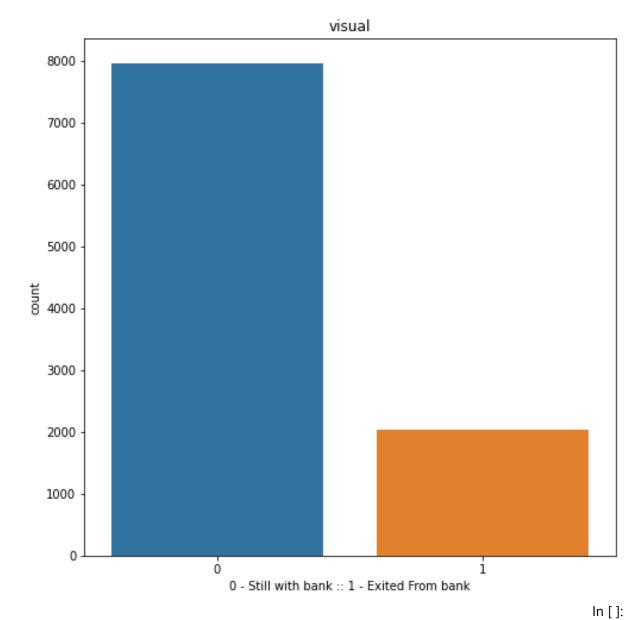
	Row Num ber	Cust omer Id	Sur na me	Cred itSco re	Geo grap hy	Ge nd er	A g e	Te nu re	Bal anc e	NumO fProdu cts	Has CrC ard	IsActiv eMem ber	Estin tedS a		Ex ite d
9 5															
9 9 9 6	9997	1556 9892	Joh nsto ne	516	Fran ce	Ma le	3 5	10	573 69.6 1	1	1	1	1016	99 77	0
9 9 9 7	9998	1558 4532	Liu	709	Fran ce	Fe ma le	3 6	7	0.00	1	0	1	4208	35. 58	1
9 9 9 8	9999	1568 2355	Sab bati ni	772	Ger man y	Ma le	4 2	3	750 75.3 1	2	1	0	9288	38. 52	1
9 9 9 9	1000	1562 8319	Wal ker	792	Fran ce	Fe ma le	2 8	4	130 142. 79	1	1	0	3819	90. 78	0
df.d	lescrik	pe()												I	n []:
	RowN umbe r	Custo merId			Age	Tenur e	В	Salanc e	NumO Produc		Men		timat Salar y		ut[]: ited
co un t	10000 .0000 0	1.0000 00e+0 4	000		0000. 00000 0	10000. 00000 0		10000. 00000	10000. 0000		1000		000.0		000. 0000 0
m ea n	5000. 50000	1.5690 94e+0 7	650	.52 38 300	8.921 800	5.0128		76485. 89288	1.5302	0 0.705 0 50			0090. 39881	0.2	0037
st d	2886. 89568	7.1936 19e+0 4	96.6	553 10 299	0.487 806	2.8921 74		52397. 05202	0.5816	5 0.455 4 84			510.4 92818	0.4	027 69
mi n	1.000	1.5565 70e+0 7	350	.00 18 000	8.000 000	0.0000	C	0000.000	1.0000	0 0.000 0 00		000 11 0	.5800	0.0	0000

RowN umbe r	Custo merId	Credit Score	Age	Tenur e	Balanc e	NumOf Product s	HasC rCar d	IsActive Membe r	Estimat edSalar y	Exited
2500. 75000	1.5628 53e+0 7	584.00 0000	32.000 000	3.0000	0.0000	1.00000	0.000	0.00000	51002.1 10000	0.0000
5000. 50000	1.5690 74e+0 7	652.00 0000	37.000 000	5.0000	97198. 540000	1.00000	1.000	1.00000	100193. 915000	0.0000
7500. 25000	1.5753 23e+0 7	718.00 0000	44.000 000	7.0000	127644 .24000 0	2.00000	1.000	1.00000	149388. 247500	0.0000
10000 .0000 0	1.5815 69e+0 7	850.00 0000	92.000 000	10.000 000	250898 .09000 0	4.00000	1.000	1.00000	199992. 480000	1.0000
figure	(fiasi	ze = (8.	3))							In []:
_	_			a=df)						
_					1 - Ex:	ited Fro	m banl	ζ")		
	2500. 75000 5000. 50000 7500. 25000 10000 0 figure countp	umbe r Custo merId 2500. 1.5628 75000 53e+0 7 5000. 1.5690 74e+0 7 7500. 23e+0 7 10000 1.5815 .0000 69e+0 0 7	umbe r Custo merId Credit Score 2500. 75000 1.5628 53e+0 0000 584.00 0000 5000. 50000 74e+0 0000 652.00 0000 7500. 23e+0 7 0000 7000 7000 10000 1.5815 0000 0000 850.00 0000 850.00 0000 figure (figsize= (8, 8 000) 0000 69e+0 0000 69e+0 0000 figure (figsize= (8, 8 000) 0000 69e+0 0000 69e+0 0000	umbe r Custo merId Credit Score Age 2500. 75000 1.5628 53e+0 0000 0000 53e+0 0000 000 32.000 000 5000. 50000 74e+0 0000 000 652.00 000 37.000 000 7500. 23e+0 0000 000 7 7 18.00 000 000 10000 1.5815 0000 000 000 850.00 000 000 92.000 000 figure (figsize=(8,8)) countplot (x='Exited', data	umbe r Custo merId Credit Score Age Tenur e 2500. 75000 1.5628 53e+0 0000 7000 584.00 0000 000 32.000 000 3.0000 000 5000. 50000 1.5690 74e+0 0000 000 652.00 000 37.000 000 5.0000 00 7500. 23e+0 0000 7 718.00 000 000 44.000 000 7.0000 00 10000 1.5815 0000 69e+0 0000 000 850.00 000 000 92.000 000 10.000 000 figure (figsize=(8,8)) countplot (x='Exited', data=df)	umbe r Custo merId Credit Score Age Tenur e Balanc e 2500. 75000 1.5628 53e+0 70000 53e+0 0000 32.000 000 3.0000 00 0.0000 5000. 50000 1.5690 74e+0 70000 652.00 000 37.000 00 5.0000 97198. 540000 7500. 25000 1.5753 718.00 000 44.000 7.0000 00 7.0000 00 127644 000 00 10000 1.5815 0000 000 000 850.00 000 000 92.000 000 000 10.000 000 000 250898 0000 .0000 69e+0 0000 000 0000 000 000 000 000 000 000 000 000 000 figure (figsize=(8,8)) countplot (x='Exited', data=df) 4000 000 000 000 000	umbe r Custo merId Credit Score Age Tenur e Balanc e Product s 2500. 75000 1.5628 584.00 0000 000 000 000 000 000 000 000 00	umbe r Custo merId Credit Score Age Tenur e Balanc e Product s rCar d 2500. 75000 1.5628 53e+0 70 0000 584.00 0000 32.000 000 3.0000 000 0.0000 000 1.00000 000 0.0000 0.0000 0.0000 0.0000 0.0000 1.00000 000 0.0000 1.0000 000 1.0000 1.0000 1.0000 1.0000 000 1.0000 1.0000 1.0000 000 1.0000 1.0000 000 </td <td>umbe r Custo merId Credit Score Age Tenur Balanc e Product s rCar d Membe r 2500. 75000 1.5628 53e+0 70 0000 584.00 0000 32.000 000 3.0000 000 0.0000 1.0000 0.0000 000 0.0000 0.0000 000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.0000 1.0000 1.00000 1.0000 1.00000 1.0000 1.0000 1.00000 1.0000 1.00000 1.0000 1.0000 1.00000 1.00000 1.0000 1.00000 1.00000 1.00000 1.00000 1.00000 1.00000 1.00000 1.00000 1.00000 0</td> <td>umbe r Custo merId Credit Score Age Tenure e Balance e Product s rCar d Membe d edSalar y 2500. 1.5628 53e+0 75000 584.00 000 000 32.000 000 000 3.0000 000 000 0.0000 000 0.0000 000 0.0000 000 0.0000 000 51002.1 0000 5000. 1.5690 74e+0 7 0000 652.00 000 37.000 000 5.0000 000 97198. 0.0000 000 1.00000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000</td>	umbe r Custo merId Credit Score Age Tenur Balanc e Product s rCar d Membe r 2500. 75000 1.5628 53e+0 70 0000 584.00 0000 32.000 000 3.0000 000 0.0000 1.0000 0.0000 000 0.0000 0.0000 000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.0000 1.0000 1.00000 1.0000 1.00000 1.0000 1.0000 1.00000 1.0000 1.00000 1.0000 1.0000 1.00000 1.00000 1.0000 1.00000 1.00000 1.00000 1.00000 1.00000 1.00000 1.00000 1.00000 1.00000 0	umbe r Custo merId Credit Score Age Tenure e Balance e Product s rCar d Membe d edSalar y 2500. 1.5628 53e+0 75000 584.00 000 000 32.000 000 000 3.0000 000 000 0.0000 000 0.0000 000 0.0000 000 0.0000 000 51002.1 0000 5000. 1.5690 74e+0 7 0000 652.00 000 37.000 000 5.0000 000 97198. 0.0000 000 1.00000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000 1.0000 000

pl sr plt.ylabel("count")

plt.title("visual")

plt.show()



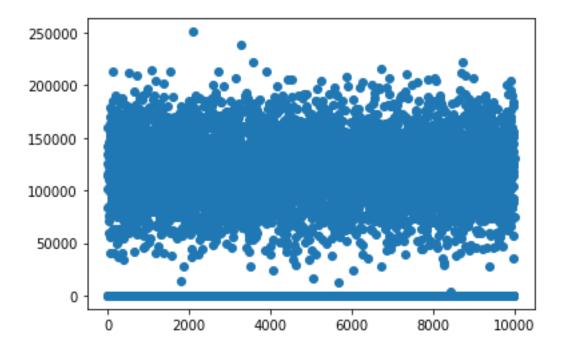
df.info()

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
                                                                            In []:
df.isna().any()
                                                                           Out[]:
            ra.
False
False
RowNumber
CustomerId
Surname
CreditScore False
Geography False
Geography
                  False
Gender
                  False
False
Age
Tenure
Balance
Balance False NumOfProducts False HasCrCard False
IsActiveMember False
EstimatedSalary False
Exited
                   False
dtype: bool
                                                                            In []:
df.isnull().sum()
                                                                           Out[]:
RowNumber
CustomerId
Surname
CreditScore
                  0
Geography
                   0
                   0
Gender
Age
                   0
                   0
Tenure
                   0
Balance
NumOfProducts 0
HasCrCard
IsActiveMember 0
EstimatedSalary
                    0
Exited
dtype: int64
                                                                            In []:
df1=df.copy()
                                                                            In []:
df1.shape
                                                                           Out[]:
(10000, 14)
                                                                            In []:
updated df=df.dropna(axis=1)
updated df.info()
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
```

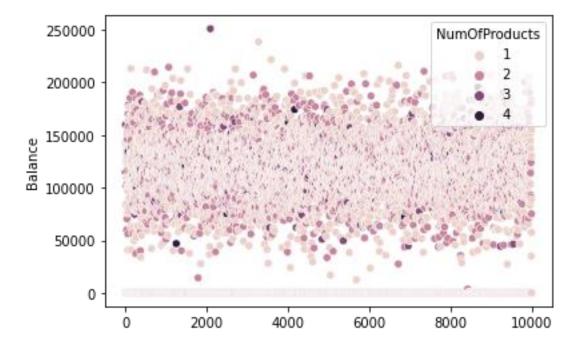
```
CreditScore 10000 non-null int64
Geography 10000 non-null object
Gender 10000 non-null object
 4
    Gender
 5
 6
    Age
                     10000 non-null int64
 7
    Tenure
                     10000 non-null int64
   Balance 10000 non-null float64
 8
    NumOfProducts 10000 non-null int64
 9
                      10000 non-null int64
 10 HasCrCard
 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
                                                                         In [ ]:
updated df['Balance']=updated df['Balance'].fillna(updated df['Balance'].me
an())
updated df.info()
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):
 # Column Non-Null Count Dtype
 0
   RowNumber
                     10000 non-null int64
                    10000 non-null int64
    CustomerId
 1
                     10000 non-null object
 2
    Surname
                    10000 non-null int64
 3 CreditScore
 4 Geography
                     10000 non-null object
 5
   Gender
                    10000 non-null object
                     10000 non-null int64
 6
    Age
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
                      10000 non-null int64
 13 Exited
dtypes: float64(2), int64(9), object(3)
memory usage: 1.1+ MB
                                                                         In []:
plt.scatter(df.index,df['Balance'])
plt.show()
```



sns.scatterplot(x=df.index,y=df['Balance'],hue=df['NumOfProducts'])

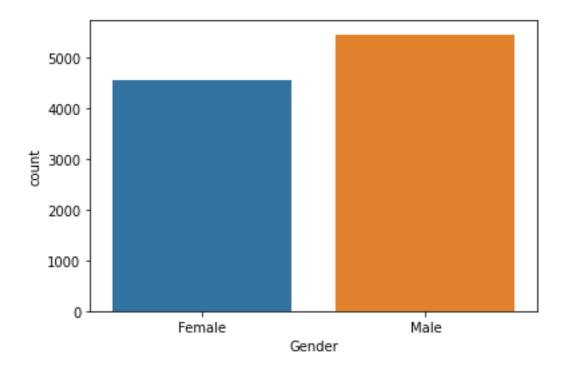


Out[]:



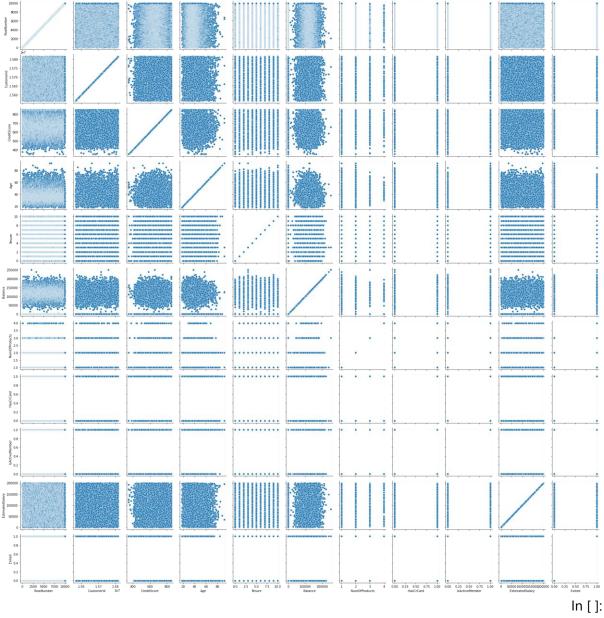
sns.barplot(x='Gender',y='Exited',data=df)
sns.countplot(x='Gender',data=df)

In []:

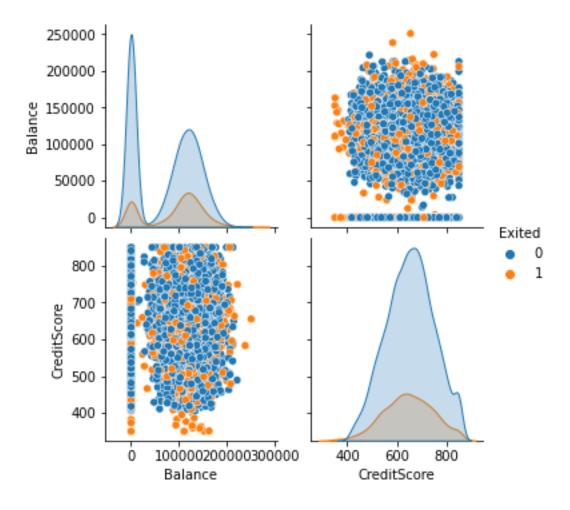


g=sns.PairGrid(df)
g.map(sns.scatterplot)

In []:



sns.pairplot(data=df[['Balance','CreditScore','Exited']],hue='Exited')



df.describe(include='all')

In []:

														Out[]:
	Row Nu mbe r	Cust ome rId	Su rn am e	Cre ditS core	Geo gra phy	G en de r	Age	Ten ure	Bala nce	Num OfPr oduct s	Has CrC ard	IsActi veMe mber	Estim atedS alary	Exit ed
co u nt	1000 0.00 000	1.00 0000 e+04	10 00 0	1000 0.00 0000	100 00	10 00 0	1000 0.00 0000	1000 0.00 0000	1000 0.000 000	10000 .0000 00	100 00.0 000 0	10000 .0000 00	10000 .0000 00	1000 0.00 0000
u ni q ue	NaN	NaN	29 32	NaN	3	2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
to p	NaN	NaN	Sm ith	NaN	Fra nce	M ale	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
fr eq	NaN	NaN	32	NaN	501 4	54 57	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

	Row Nu mbe r	Cust ome rId	Su rn am e	Cre ditS core	Geo gra phy	G en de r	Age	Ten ure	Bala nce	Num OfPr oduct s	Has CrC ard	IsActi veMe mber	Estim atedS alary	Exit ed
m ea n	5000 .500 00	1.56 9094 e+07	Na N	650. 5288 00	Na N	Na N	38.9 2180 0	5.01 2800	7648 5.889 288	1.530 200	0.70 550	0.515 100	10009 0.239 881	0.20 3700
st d	2886 .895 68	7.19 3619 e+04	Na N	96.6 5329 9	Na N	Na N	10.4 8780 6	2.89 2174	6239 7.405 202	0.581 654	0.45 584	0.499 797	57510 .4928 18	0.40 2769
m in	1.00 000	1.55 6570 e+07	Na N	350. 0000 00	Na N	Na N	18.0 0000 0	0.00	0.000	1.000	0.00 000	0.000	11.58 0000	0.00
25 %	2500 .750 00	1.56 2853 e+07	Na N	584. 0000 00	Na N	Na N	32.0 0000 0	3.00 0000	0.000	1.000	0.00	0.000	51002 .1100 00	0.00
50 %	5000 .500 00	1.56 9074 e+07	Na N	652. 0000 00	Na N	Na N	37.0 0000 0	5.00 0000	9719 8.540 000	1.000	1.00 000	1.000 000	10019 3.915 000	0.00
75 %	7500 .250 00	1.57 5323 e+07	Na N	718. 0000 00	Na N	Na N	44.0 0000 0	7.00 0000	1276 44.24 0000	2.000	1.00 000	1.000 000	14938 8.247 500	0.00
m ax	1000 0.00 000	1.58 1569 e+07	Na N	850. 0000 00	Na N	Na N	92.0 0000 0	10.0 0000 0	2508 98.09 0000	4.000 000	1.00 000	1.000 000	19999 2.480 000	1.00 0000

Find the outliers and replace the outliers

df[(df['NumOfProducts']>2) | (df['NumOfProducts']<3)]</pre>

In []:

													C)ut[]:
	Row Num ber	Cust omer Id	Sur na me	Cred itSco re	Geo grap hy	Ge nd er	A g e	Te nu re	Bal anc e	NumO fProdu cts	Has CrC ard	IsActiv eMem ber	Estima tedSal ary	Ex ite d
0	1	1563 4602	Har gra ve	619	Fran ce	Fe ma le	4 2	2	0.00	1	1	1	101348 .88	1
1	2	1564 7311	Hill	608	Spai n	Fe ma le	4 1	1	838 07.8 6	1	0	1	112542 .58	0

	Row Num ber	Cust omer Id	Sur na me	Cred itSco re	Geo grap hy	Ge nd er	A g e	Te nu re	Bal anc e	NumO fProdu cts	Has CrC ard	IsActiv eMem ber	Estima tedSal ary	Ex ite d
2	3	1561 9304	Oni o	502	Fran ce	Fe ma le	4 2	8	159 660. 80	3	1	0	113931 .57	1
3	4	1570 1354	Bon i	699	Fran ce	Fe ma le	3 9	1	0.00	2	0	0	93826. 63	0
4	5	1573 7888	Mit chel l	850	Spai n	Fe ma le	4 3	2	125 510. 82	1	1	1	79084. 10	0
9 9 9 5	9996	1560 6229	Obi jiak u	771	Fran ce	Ma le	3 9	5	0.00	2	1	0	96270. 64	0
9 9 9 6	9997	1556 9892	Joh nsto ne	516	Fran ce	Ma le	3 5	10	573 69.6 1	1	1	1	101699 .77	0
9 9 9 7	9998	1558 4532	Liu	709	Fran ce	Fe ma le	3 6	7	0.00	1	0	1	42085. 58	1
9 9 9 8	9999	1568 2355	Sab bati ni	772	Ger man y	Ma le	4 2	3	750 75.3 1	2	1	0	92888. 52	1
9 9 9	1000	1562 8319	Wal ker	792	Fran ce	Fe ma le	2 8	4	130 142. 79	1	1	0	38190. 78	0

10000 rows × 14 columns

df[(df['NumOfProducts']>2)]

In []:

	Row Num ber	Cust ome rId	Surna me	Cred itSco re	Geo grap hy	Ge nd er	A g e	Te nu re	Bal anc e	NumO fProdu cts	Has CrC ard	IsActiv eMem ber	Estima tedSal ary	Ex ite d
2	3	1561 9304	Onio	502	Fran ce	Fe ma le	4 2	8	159 660. 80	3	1	0	113931 .57	1
7	8	1565 6148	Obinn a	376	Ger man y	Fe ma le	2 9	4	115 046. 74	4	1	0	119346 .88	1
3 0	31	1558 9475	Aziki we	591	Spai n	Fe ma le	3 9	3	0.00	3	1	0	140469 .38	1
7 0	71	1570 3793	Konov alova	738	Ger man y	Ma le	5 8	2	133 745. 44	4	1	0	28373. 86	1
8	89	1562 2897	Sharp e	646	Fran ce	Fe ma le	4 6	4	0.00	3	1	0	93251. 42	1
•••														
9 7 3 7	9738	1574 1197	Calza da	710	Spai n	Ma le	2 2	8	0.00	3	1	0	107292 .91	0
9 7 4 7	9748	1577 5761	Iweob iegbun am	610	Ger man y	Fe ma le	6 9	5	860 38.2 1	3	0	0	192743 .06	1
9 8 0 0	9801	1564 0507	Li	762	Spai n	Fe ma le	3 5	3	119 349. 69	3	1	1	47114. 18	1
9 8 7 7	9878	1557 2182	Onwu amaez e	505	Ger man y	Fe ma le	3	3	106 506. 77	3	1	0	45445. 78	1
9	9896	1579 6764	Bruno	684	Ger man y	Fe ma le	5 6	3	127 585. 98	3	1	1	80593. 49	1

	Nι	ow ım oer	Cust ome rId	Su	rna me	Cred itSco)	Geo grap hy	Ge nd er	A g e	Te nu re	a	Bal nc e	Num fProd c		Has CrC ard		ctiv Iem ber		ima Sal ary	Ex ite d
9 5																					
326	ó row	s × 1	4 col	umns	S																
Che	eck fo	r Cat	egor	ical c	olum	ıns ar	nd p	erfo	rm en	codii	ng										
	['Ag		= df['Age	'].á	asty	pe ('flo	oat'))											n []:
ar.	· acy	рсь																		0	ut[]:
Rov	vNum.	ber				int	64														
	stom		d			int															
	rnam				(bje															
	edit ogra		ce		,	int bje															
	nder					bje. bje															
Age						Loat															
	nure					int															
	Lanc				f	Loat															
	nOfP sCrC		ıcts			int int															
	Acti		embe	r		int															
	ima				f	Loat															
Exi	ited					int	64														
dty	ype:	obj	ect																		
pd.	.get	_dum	nmie	s(df	, co.	Lumn	s = ['Tei	nure'	']).	hea	ıd()								l	n []:
																				0	ut[]:
	Ro	C	\mathbf{S}			G				Н		T	T	T	T	T	T	T	T	T	
	W	us	u	Cr	G	e			Nu	as		e	e	e	e	e	e	e	e	e	T
	Nu	to	r n	ed itS	eo gr	n	A g	al a	mO fPr	Cr		n u	n u	n u	n u	n u	n u	n u	n u	n u	en ur
	m be	m er	a	co	ap	d e	e	n	odu	C ar		re	re	re	re	re	re	re	re	re	e _
	r	Id	m e	re	hy	r		ce	cts	d		<u>-</u>	$\overline{2}$	$\overline{3}$	<u>-</u>	- 5	- 6	- 7	- 8	<u>-</u>	10
			·									•	_	J	•		Ū	,	Ü		
			Н																		
		15	ar		Fr	F e	4	0.													
0	1	63 46	gr	61 9	an	m	2	0.	1	1		0	1	0	0	0	0	0	0	0	0
		02	a v	9	ce	al	0	0													
		-	e			e															
						F															
		15 64	Н	60	Sp	e	4 1	8													
1	2	73	п ill	8	ai	m -1		3 8	1	0		1	0	0	0	0	0	0	0	0	0
		11			n	al e	0	0			•										
								7.													

```
\mathbf{C}
                                                               Н
    Ro
                                     G
                        Cr
                               G
                                                 В
                                                       Nu
                                                                                                                                   \mathbf{T}
                   u
                                                                           e
                                                                                 e
                                                                                                    e
                                                                                                                e
                                                                                                                             e
                                                                                       e
                                                                                             e
                                                                                                          \mathbf{e}
                                                                                                                      e
            us
                                                               as
     w
                                      \mathbf{e}
                                                      mО
                   r
                        ed
                               eo
                                           \mathbf{A}
                                                al
                                                                          n
                                                                                 n
                                                                                       n
                                                                                             n
                                                                                                   n
                                                                                                          n
                                                                                                                      n
                                                                                                                            n
                                                                                                                                  en
    Nu
                                                               \mathbf{Cr}
            to
                                      n
                   n
                        itS
                               gr
                                                 a
                                                       fPr
                                                                          u
                                                                                                                                  ur
                                           g
                                                                                       u
                                                                                                   u
                                                                                                                      u
                                      d
                                                               \mathbf{C}
    m
            m
                                                      odu
                                                                                            re
                   a
                        co
                               ap
                                           \mathbf{e}
                                                 n
                                                                          re
                                                                                re
                                                                                      re
                                                                                                  re
                                                                                                         re
                                                                                                               re
                                                                                                                      re
                                                                                                                            re
    be
            er
                                      e
                                                               ar
                               hy
                                                                                                                                  10
                  m
                        re
                                                        cts
                                                ce
            Id
                                                                d
      r
                                                                                                                             9
                                                                                                          <u>-</u>
                                                                                 \overline{2}
                                                                                                    5
                                                                                       3
                                                                                             4
                                                 8
                                                 6
                                                 1
                                                 5
                                      F
            15
                                                 9
                  O
                               Fr
                                      e
            61
                        50
                                           2
                                                 6
2
      3
                  ni
                                     m
                                                                                       0
                                                                                             0
                                                                                                   0
                               an
           93
                         2
                                                 6
                                     al
                   o
                               ce
                                           0
           04
                                                0.
                                      e
                                                 8
                                                 0
                                      F
            15
                                                0.
                  В
                               Fr
                                      e
            70
                        69
                                           9
3
                                                 0
                                                         2
                                                                           1
                                                                                 0
                                                                                       0
                                                                                             0
                                                                                                   0
                                                                                                          0
                                                                                                                0
                                                                                                                      0
                  o
                               an
                                     m
           13
                  ni
                               ce
                                     al
                                                 0
                                           0
                                                 1
                                                 2
                  M
                                      F
           15
                                                 5
                  it
                               Sp
                                      e
           73
                        85
                                           3
                                                 5
                   c
      5
                                     m
                                                                          0
                                                                                                   0
                                ai
           78
                                                 1
                  h
                         0
                                     al
                                           0
           88
                                                0.
                  el
                                      e
                   1
                                                 8
```

5 rows × 24 columns

Split the data into dependent and independent variables

```
In []:
x=df.iloc[:,:-1].values #independent variables
y=df.iloc[:,-1].values #dependent variables
print(x,y)

[[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]] [1 0 1 ... 1 1 0]
Scale the independent variables

In []:
x=df.iloc[1:3,:-1].values
x
```

```
[3, 15619304, 'Onio', 502, 'France', 'Female', 42, 8, 159660.8, 3,
      1, 0, 113931.57]], dtype=object)
                                                             In []:
x=df[['Gender','Age']]
print(x)
     Gender
            Age
    Female 42.0
1
    Female 41.0
    Female 42.0
2
     Female 39.0
3
    Female 43.0
      9995
     Male 39.0
9996
     Male 35.0
9997 Female 36.0
9998
     Male 42.0
9999 Female 28.0
[10000 rows x 2 columns]
Split the data into training and testing
                                                             In [ ]:
from sklearn.model selection import train test split
                                                             In []:
training data, testing data=train test split(df, test size=1, random state=3)
print(training data, testing data)
     RowNumber CustomerId
                         Surname CreditScore Geography Gender
6555
        6556 15581505
                                             641
                                Bales
                                                   France Male
                               Diribe
                                             555 Germany Female
1448
        1449 15585367
                                             474 Germany Female
        3352 15792729
                              Holland
3351
                                                  France Male
231
         232 15627000
                              Freeman
                                              610
1204
        1205 15650098
                             Baranova
                                             630 France Female
         . . .
                                             . . .
                                                    ... ...
                . . .
                               . . .
        6401 15585907
6400
                              Collier
                                             676
                                                    Spain Female
                                             778 France
        9161
               15753679 Mullawirraburka
9160
                                                           Male
        9860
                                             678 Germany
9859
               15615430
                                Adams
                                                            Male
                                              601
                                Valdez
1688
        1689 15804610
                                                   France Female
        5995 15746065
5994
                              Lo Duca
                                             580 Germany Male
     Age Tenure Balance NumOfProducts HasCrCard IsActiveMember \
6555 35.0
         5
                           2
                  0.00
                                             1
              4 120392.99
1448 46.0
                                    1
                                              1
                                                            0
             9 176311.36
3351 34.0
                                    1
                                              1
                                                            0
             0.00
                                    2
231 40.0
                                              1
                                                            0
1204 40.0
             7
                    0.00
                                    2
                                             1
                                                            1
. . .
     . . .
             . . .
                     . . .
                                   . . .
                                  2
6400 30.0
                    0.00
             5
                                             0
                                                           0
                                    2
9160 24.0
              4
                                              1
                                                            1
                    0.00
                                   1
9859 55.0
             4 129646.91
                                                           1
                                             1
                                   2
             1 0.00
                                             0
1688 41.0
                                                           1
5994 35.0
            10 136281.41
                                   2
                                             1
                                                            1
     EstimatedSalary Exited
```

array([[2, 15647311, 'Hill', 608, 'Spain', 'Female', 41, 1, 83807.86, 1,

0, 1, 112542.58],

6555	93148.93	0				
1448	177719.88	1				
3351	160213.27	0				
231	62232.60	0				
1204	34453.17	0				
6400	179066.58	0				
9160	162809.20	0				
9859	184125.10	1				
1688	160607.06	0				
5994	24799.47	0				
-	ows x 14 columns] aphy Gender Age		r Customer	Id Surname	Cred	itScor
_	5877 155853		70	4 France	Male	39.0
Ψе	enure Balance N	JumOfProducts	HasCrCard	TsActiveMemb	er \	
5876	2 111525.02	1	1	18716 CT V CTTCHILD	0	
-						
	stimatedSalary Ex					
5876	199484.96	0				