```
import pandas as pd
df = pd.read csv("Churn Modelling (1).csv",sep = ",")
df
      RowNumber CustomerId Surname CreditScore Geography
                                                                Gender
Age \
              1
                   15634602
                              Hargrave
                                                 619
                                                        France
                                                                Female
42.0
1
              2
                   15647311
                                   Hill
                                                 225
                                                         Spain Female
41.0
2
              3
                   15619304
                                   Onio
                                                 629
                                                        France Female
42.0
3
              4
                   15701354
                                   Boni
                                                 699
                                                        France Female
39.0
              5
                              Mitchell
                   15737888
                                                 850
                                                           NaN Female
4
43.0
. . .
                                                 . . .
                                                            . . .
            . . .
                         . . .
9995
           9996
                              Obijiaku
                   15606229
                                                 771
                                                        France
                                                                   Male
39.0
9996
           9997
                   15569892
                             Johnstone
                                                 516
                                                        France
                                                                   Male
35.0
           9998
9997
                   15584532
                                    Liu
                                                 709
                                                        France Female
36.0
9998
           9999
                   15682355 Sabbatini
                                                 772
                                                       Germany
                                                                   Male
42.0
9999
          10000
                   15628319
                                 Walker
                                                 792
                                                        France Female
28.0
      Tenure
                Balance NumOfProducts HasCrCard IsActiveMember \
                   0.00
0
           2
                                      1
                                                 1
                                                                  1
1
           1
               83807.86
                                      1
                                                 0
                                                                  1
2
                                      3
                                                 1
                                                                  0
           8
              159660.80
3
                                      2
           1
                   0.00
                                                 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
                                      1
                                                                  1
                   0.00
                                                 0
           3
               75075.31
                                      2
                                                 1
                                                                  0
9998
9999
                                      1
                                                 1
                                                                  0
           4
              130142.79
      EstimatedSalary Exited
0
            101348.88
                             1
1
            112542.58
                            0
2
            113931.57
                             1
3
             93826.63
                            0
```

4

79084.10

0

9995 9996 9997 9998 9999	101 42 92	270.64 699.77 085.58 888.52 190.78	0 0 1 1 0			
[10000 row	s x 1	4 columns]				
df.shape						
(10000, 14)					
df.head()						
RowNumb	er C	ustomerId	Surname	CreditScore	Geography	Gender
Age \	1	15634602	Hargrave	619	France	Female
42.0 1	2	15647311	Hill	225	Spain	Female
41.0 2	3	15619304	Onio	629	France	Female
42.0 3	4	15701354	Boni	699	France	Female
39.0 4 43.0	5	15737888	Mitchell	850	NaN	Female
1 1 2 1 3	838 1596 1255	0.00 07.86 60.80 0.00 10.82 ary Exite .88 .58 .57	0fProducts 1 1 3 2 1 ed 1 0 1 0	. 1 . 0 . 1	IsActiveMe	ember \ 1 0 0 1
df.tail()	79004	. 10	O			
	umber	Customer	Id Surn	ame CreditS	core Geogra	phy Gender
Age \ 9995	9996	156062	29 Obiji		_	nce Male
39.0 9996	9997	155698	_			nce Male
35.0 9997	9998			Liu		nce Female

36.0

```
9998
           9999
                    15682355
                              Sabbatini
                                                  772
                                                         Germany
                                                                    Male
42.0
          10000
9999
                    15628319
                                 Walker
                                                  792
                                                          France
                                                                  Female
28.0
      Tenure
                Balance
                          NumOfProducts
                                          HasCrCard
                                                     IsActiveMember
                                                                      \
9995
           5
                    0.00
                                                  1
9996
          10
               57369.61
                                       1
                                                  1
                                                                   1
                                                                   1
                                       1
                                                  0
9997
           7
                    0.00
           3
               75075.31
                                       2
                                                  1
9998
                                                                   0
9999
           4
              130142.79
                                       1
                                                  1
                                                                   0
      EstimatedSalary
                        Exited
9995
             96270.64
                             0
9996
                             0
            101699.77
9997
             42085.58
                             1
9998
             92888.52
                             1
                             0
9999
             38190.78
df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 14 columns):
#
     Column
                       Non-Null Count
                                        Dtype
- - -
     _ _ _ _ _
     RowNumber
 0
                       10000 non-null
                                        int64
                       10000 non-null
 1
     CustomerId
                                        int64
 2
     Surname
                       10000 non-null
                                        object
 3
     CreditScore
                       10000 non-null
                                        int64
 4
                       9999 non-null
     Geography
                                        object
 5
     Gender
                       9999 non-null
                                        object
 6
                       9999 non-null
                                        float64
     Age
 7
                       10000 non-null
     Tenure
                                        int64
 8
     Balance
                       9999 non-null
                                        float64
 9
     NumOfProducts
                       10000 non-null
                                        int64
 10
    HasCrCard
                       10000 non-null
     IsActiveMember
                       10000 non-null
 11
                                        int64
                       10000 non-null
 12
     EstimatedSalary
                                       float64
 13
     Exited
                       10000 non-null int64
dtypes: float64(3), int64(8), object(3)
memory usage: 1.1+ MB
df.describe() # statical information of columns containing the numeric
```

values we will get

	RowNumber	CustomerId	CreditScore	Age
Tenure	\			
count	10000.00000	1.000000e+04	10000.00000	9999.000000
10000.	000000			
mean	5000.50000	1.569094e+07	650.48290	38.924892

```
5.012800
        2886.89568
                     7.193619e+04
                                       96.72014
                                                    10.486670
std
2.892174
           1.00000
                     1.556570e+07
                                      225.00000
                                                    18,000000
min
0.000000
25%
        2500.75000
                     1.562853e+07
                                      584.00000
                                                    32,000000
3.000000
50%
        5000.50000
                     1.569074e+07
                                      652.00000
                                                    37.000000
5.000000
75%
        7500.25000
                     1.575323e+07
                                      717.25000
                                                    44.000000
7.000000
                                      850.00000
                                                    92.000000
max
       10000.00000
                     1.581569e+07
10.000000
              Balance
                       NumOfProducts
                                          HasCrCard
                                                     IsActiveMember
         9999,000000
                        10000.000000
                                       10000.00000
                                                        10000.000000
count
        76483.536070
                                            0.70550
                                                            0.515100
mean
                             1.530200
        62400.186375
                             0.581654
                                            0.45584
                                                            0.499797
std
            0.000000
                             1.000000
                                            0.00000
                                                            0.000000
min
25%
            0.000000
                             1.000000
                                            0.00000
                                                            0.00000
        97188.620000
                             1.000000
                                            1.00000
                                                            1.000000
50%
       127646.040000
75%
                             2,000000
                                            1.00000
                                                            1.000000
       250898,090000
max
                             4.000000
                                            1.00000
                                                            1.000000
       EstimatedSalary
                                Exited
count
          10000.000000
                         10000.000000
         100090.239881
mean
                              0.203700
                              0.402769
std
          57510.492818
min
              11.580000
                              0.000000
25%
          51002.110000
                              0.00000
50%
         100193.915000
                              0.00000
75%
         149388.247500
                              0.00000
         199992.480000
                              1.000000
max
```

finding null values

```
x = df.head()
y = x.isnull()
y
```

	Number	CustomerId	Surname	CreditScore	Geography	Gender
Age \ 0 False	False	False	False	False	False	False
1	False	False	False	False	False	False
False	False	False	False	False	False	False
False 3 False	False	False	False	False	False	False

4 False	False	False	False	False	True F	alse
	ure Balar tedSalary		ducts H	asCrCard	IsActiveMember	
		Lse	False	False	False	
	lse Fal	Lse	False	False	False	
	lse Fal	Lse	False	False	False	
3 Fa	lse Fal	Lse	False	False	False	
False 4 Fa False	lse Fal	Lse	False	False	False	
0 Fa 1 Fa 2 Fa 3 Fa	ted lse lse lse lse					
y.sum() #head					
HasCrC IsActi Estima Exited	erId e Score phy e roducts ard veMember tedSalary int64	0 0 0 1 0 0 0 0 0				
	RowNumber	CustomerId	Surname	CreditSc	ore Geography	Gender
Age \ 0	False	False	False	Fa	lse False	False
False	False	Falco	Falco	Ea	lco Falco	Falco

False

1 False 2 False

3	Fal	se	False	False	False	e False	False
False 4 False	Fal	se	False	False	False	e True	False
9995 False	Fal	se	False	False	False	e False	False
9996 False	Fal	se	False	False	False	e False	False
9997 False	Fal	se	False	False	False	e False	False
9998 False	Fal	se	False	False	False	e False	False
9999 False	Fal	se	False	False	Fals	e False	False
0 1 2 3 4 9995 9996 9997 9998 9999	Tenure False	Balance False	NumOfPr	False False False False False False False False False	HasCrCard False	IsActiveMemb Fal Fal Fal Fal Fal Fal Fal	se se se se se se se se
0 1 2 3 4	Estimat	edSalary False False False False False	Exited False False False False				
9995 9996 9997 9998 9999		False False False False False	False False False False False				
[1000	0 rows x	14 colur	mns]				
df.is	null().s	um()					
RowNu Custo	merId	0 0 0					

0 0 0

Surname

```
CreditScore
                    0
Geography
                    1
Gender
                    1
Age
                    1
Tenure
                    0
Balance
                    1
NumOfProducts
                    0
HasCrCard
                    0
IsActiveMember
                    0
EstimatedSalary
                    0
Exited
                    0
dtype: int64
z = df.tail()
w = z.isnull()
                 CustomerId
      RowNumber
                              Surname CreditScore
                                                      Geography
                                                                 Gender
Age \
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 \
9995
       False
                False
                                 False
                                            False
                                                             False
9996
       False
                False
                                 False
                                            False
                                                             False
                False
9997
       False
                                 False
                                            False
                                                             False
9998
       False
                False
                                 False
                                            False
                                                             False
9999
       False
                False
                                 False
                                            False
                                                             False
      EstimatedSalary
                        Exited
9995
                False
                         False
9996
                 False
                         False
9997
                         False
                False
9998
                 False
                         False
9999
                False
                         False
w.sum() #tail
RowNumber
                    0
CustomerId
                    0
                    0
Surname
CreditScore
                    0
                    0
Geography
```

Gender 0 Age 0 Tenure 0 Balance 0 NumOfProducts 0 HasCrCard 0 IsActiveMember 0 EstimatedSalary 0 Exited 0 dtype: int64

Handle missing values
df.Age = df.Age.fillna(df.Age.median())

df							
۸۵۵	RowNumbe	er Custome	rId	Surname	CreditScore	Geography	Gender
Age 0	\	1 15634	602	Hargrave	619	France	Female
42.0		2 15647	311	Hill	225	Spain	Female
41.0		3 15619	304	Onio	629	France	Female
42.0		4 15701	354	Boni	699	France	Female
39.0		5 15737	888	Mitchell	850	NaN	Female
43.0							
9995	999	6 15606	229	0bijiaku	771	France	Male
39.0 9996	999	7 15569	892	Johnstone	516	France	Male
35.0 9997	999	98 15584	532	Liu	709	France	Female
36.0 9998	999	9 15682	355	Sabbatini	772	Germany	Male
42.0 9999 28.0	1000	00 15628	319	Walker	792	France	Female
	Tenure	Balance	Num	OfProducts		IsActiveMem	
0 1	2 1	0.00 83807.86		1 1	1 0		1 1
2	8	159660.80		3	1		0
3 4	1 2	0.00 125510.82		2 1	0 1		0 1
9995	5	0.00		2	1		0

9996 9997 9998 9999	10 7 3 4	57369.6 0.0 75075.3 130142.7	0 1	1 1 2 1	1 0 1 1	1 1 0 0
	Estimat	edSalary	Exited			
0		01348.88	1			
1	1	12542.58	0			
2	1	13931.57	1			
3		93826.63	0			
4		79084.10	0			
9995		96270.64	0			
9996	1	01699.77	0			
9997		42085.58	1			
9998		92888.52	1			
9999		38190.78	0			

[10000 rows x 14 columns]

df.Balance=df.Balance.fillna(df.Balance.median())

df

۸۵۵	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender
Age 0	1	15634602	Hargrave	619	France	Female
42.0	2	15647311	Hill	225	Spain	Female
41.0	3	15619304	Onio	629	France	Female
42.0	4	15701354	Boni	699	France	Female
39.0 4 43.0	5	15737888	Mitchell	850	NaN	Female
9995	9996	15606229	0bijiaku	771	France	Male
39.0 9996	9997	15569892	Johnstone	516	France	Male
35.0 9997	9998	15584532	Liu	709	France	Female
36.0 9998	9999	15682355	Sabbatini	772	Germany	Male
42.0 9999 28.0	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.Geography = df.Geography.fillna(df.Geography.mode())

df

\ a a	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender
Age 0 42.0	1	15634602	Hargrave	619	France	Female
1 41.0	2	15647311	Hill	225	Spain	Female
2 42.0	3	15619304	Onio	629	France	Female
3 39.0	4	15701354	Boni	699	France	Female
4 43.0	5	15737888	Mitchell	850	NaN	Female
9995 39.0	9996	15606229	0bijiaku	771	France	Male
9996 35.0	9997	15569892	Johnstone	516	France	Male
9997 36.0	9998	15584532	Liu	709	France	Female

9998 42.0	99	99 15682	2355	Sabbatini	77	2 Germany	Male
9999 28.0	100	00 15628	3319	Walker	79	2 France	Female
0 1 2 3 4	Tenure 2 1 8 1 2	Balance 0.00 83807.86 159660.80 0.00 125510.82	Num(OfProducts 1 1 3 2 1	HasCrCard 1 0 1 0 1	IsActiveMen	1 1 0 0 1
9995 9996 9997 9998 9999	5 10 7 3 4	0.00 57369.61 0.00 75075.31 130142.79		2 1 1 2 1	1 1 0 1		0 1 1 0 0
0 1 2 3 4	1 1 1	edSalary B 01348.88 12542.58 13931.57 93826.63 79084.10		d 1 9 1 9			
9995 9996 9997 9998 9999	1	96270.64 01699.77 42085.58 92888.52 38190.78		9 9 1 1 9			
_	0 rows x null().s	14 columns	s]				
RowNun Custon Surnan Credi Geogra Gende Age Tenura Balan NumOff HasCra IsActa Estima Exite	mber merId me tScore aphy r e ce Products Card iveMembe atedSala	0 0 0 1 1 0 0 0 0					

```
df = df.fillna("unknown")
df
```

	RowNumber	CustomerId	Surname	CreditScore	Geography	Gender
Age 0 42.0 1 41.0 2 42.0 3 39.0 4 43.0 9995 39.0 9996 35.0 9997 36.0 9998 42.0 9999 28.0	1	15634602	Hargrave	619	France	Female
	2	15647311	Hill	225	Spain	Female
	3	15619304	Onio	629	France	Female
	4	15701354	Boni	699	France	Female
	5	15737888	Mitchell	850	unknown	Female
	9996	15606229	0bijiaku	771	France	Male
	9997	15569892	Johnstone	516	France	Male
	9998	15584532	Liu	709	France	Female
	9999	15682355	Sabbatini	772	Germany	Male
	10000	15628319	Walker	792	France	Female
0 1 2 3 4 9995 9996 9997 9998 9999	8 1 1 2 1 5 10 7 3	Balance Nu 0.00 83807.86 59660.80 0.00 25510.82 0.00 57369.61 0.00 75075.31 30142.79	mOfProducts 1 1 3 2 1 2 1 1 2 1	HasCrCard 1 0 1 0 1 1 1 0 1 1	IsActiveMen	nber \ 1
0 1 2 3 4 9995 9996	112 113 93 79 96	Salary Exit 348.88 542.58 931.57 826.63 084.10 270.64 699.77	ed 1 0 1 0 0 0			

```
9997
             42085.58
                              1
9998
                              1
              92888.52
              38190.78
9999
                              0
[10000 \text{ rows } \times 14 \text{ columns}]
df.isnull().sum() # it is useful when we dealing with large values
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
df.columns
Index(['RowNumber', 'CustomerId', 'Surname', 'CreditScore',
'Geography',
       'Gender', 'Age', 'Tenure', 'Balance', 'NumOfProducts',
'HasCrCard',
        'IsActiveMember', 'EstimatedSalary', 'Exited'],
      dtype='object')
Loc
df.loc[4:8]
   RowNumber CustomerId
                             Surname CreditScore Geography
                                                                Gender
Age \
           5
                 15737888
                           Mitchell
                                               850
                                                     unknown
                                                                Female
43.0
           6
                 15574012
                                 Chu
                                               645
                                                       Spain
                                                                  Male
44.0
           7
                 15592531
                           Bartlett
                                               619
                                                      France
                                                                  Male
6
50.0
           8
                 15656148
                              0binna
                                               376
                                                               unknown
7
                                                     Germany
29.0
           9
                 15792365
                                               501
                                                                  Male
                                  He
                                                      France
8
44.0
```

Balance NumOfProducts HasCrCard IsActiveMember \

Tenure

5 8 1 6 7 7 4 1	.25510.82 .13755.78 0.00 .15046.74 .42051.07	1 2 2 4 2	1 1 1 1 0	1 0 1 0 1					
5 149 6 10 7 119	Salary Exit 084.10 756.71 062.80 346.88	ed 0 1 0 1							
df.loc[df.Age>40]									
RowNum		rId Surname	CreditScore	Geography					
Gender Age	1 15634	.602 Hargrave	619	France					
Female 42.0	2 15647	311 Hill	. 225	Spain					
Female 41.0	3 15619	304 Onio	629	France					
Female 42.0) 5 15737	888 Mitchell	. 850	unknown					
Female 43.0) 6 15574	012 Chu	ı 645	Spain					
Male 44.0									
		754 6 1 1							
9981 9 Male 42.0	982 15672	754 Burbidge	498	Germany					
9982 9 Female 46.0	15768 1	163 Griffir	655	Germany					
9986 9	987 15581	736 Bartlett	673	Germany					
	992 15769	959 Ajuluchukwu	597	France					
) 1999 15682	355 Sabbatini	. 772	Germany					
Male 42.0									
Tenure 0 2 1 1 2 8 4 2 5 8	0.00 83807.86 159660.80 125510.82 113755.78	NumOfProducts 1 1 3 1 2	HasCrCard Isa 1 0 1 1 1	ActiveMember \ 1 1 0 1					
9981 3 9982 7 9986 1	152039.70 137145.12	1 1 2	1 1 0	1 0 1					

```
9991
               88381.21
           4
                                                  1
                                                                   0
9998
           3
                                       2
                                                  1
                                                                   0
               75075.31
      EstimatedSalary
                        Exited
0
            101348.88
                             1
1
            112542.58
                             0
2
                             1
            113931.57
4
                             0
             79084.10
5
            149756.71
                             1
. . .
                            . . .
             53445.17
9981
                             1
9982
            115146.40
                             1
9986
             34047.54
                             0
                             1
9991
             69384.71
9998
             92888.52
                             1
[3582 rows x 14 columns]
df.loc[0:3,['Age','CustomerId','RowNumber']]
         CustomerId RowNumber
    Age
  42.0
           15634602
0
                              1
1 41.0
           15647311
                              2
                              3
2
  42.0
           15619304
                              4
3
  39.0
           15701354
df.loc[(df.HasCrCard == 1) & (df.Gender == 'Female')]
      RowNumber CustomerId
                                  Surname CreditScore Geography
Gender
         Age
              1
                    15634602
                                 Hargrave
                                                    619
                                                            France
Female
        42.0
2
              3
                    15619304
                                      Onio
                                                    629
                                                            France
Female 42.0
              5
                    15737888
                                 Mitchell
                                                    850
                                                           unknown
Female 43.0
             13
                    15632264
                                                    476
                                                            France
12
                                       Kay
Female
        34.0
14
                                     Scott
                                                    635
             15
                    15600882
                                                             Spain
        35.0
Female
                                                     . . .
                         . . .
9976
           9977
                    15656062
                                  Azikiwe
                                                    637
                                                            France
Female 33.0
9977
           9978
                    15579969
                                  Mancini
                                                    683
                                                            France
Female
        32.0
9982
                    15768163
                                  Griffin
           9983
                                                    655
                                                           Germany
Female
        46.0
                    15769959 Ajuluchukwu
                                                    597
9991
           9992
                                                            France
Female
        53.0
          10000
                    15628319
                                   Walker
                                                    792
9999
                                                            France
```

```
Tenure
                 Balance NumOfProducts HasCrCard
                                                        IsActiveMember
0
            2
                     0.00
                                                     1
                                                                       1
2
            8
               159660.80
                                         3
                                                     1
                                                                      0
4
            2
               125510.82
                                         1
                                                     1
                                                                       1
12
                                         2
           10
                     0.00
                                                     1
                                                                      0
            7
                                         2
                                                     1
14
                     0.00
                                                                       1
. . .
          . . .
               103377.81
9976
            7
                                         1
                                                     1
                                                                      0
            9
                                        2
                                                     1
                                                                      1
9977
                     0.00
            7
9982
               137145.12
                                        1
                                                     1
                                                                      0
9991
            4
               88381.21
                                        1
                                                     1
                                                                      0
            4
                                        1
                                                                      0
9999
               130142.79
                                                     1
      EstimatedSalary Exited
0
             101348.88
                               1
2
                               1
             113931.57
4
              79084.10
                               0
12
              26260.98
                               0
14
                               0
              65951.65
9976
              84419.78
                               0
9977
              24991.92
                               0
                               1
9982
             115146.40
                               1
9991
              69384.71
9999
              38190.78
                               0
[3191 rows x 14 columns]
u = df[['CustomerId','Surname','CreditScore']]
u.head()
   CustomerId
                 Surname CreditScore
     15634602 Hargrave
0
                                    619
                     Hill
                                    225
1
     15647311
2
     15619304
                     Onio
                                    629
3
     15701354
                                    699
                     Boni
     15737888
               Mitchell
                                    850
u = df[['CustomerId','Surname','CreditScore']].values
array([[15634602, 'Hargrave', 619], [15647311, 'Hill', 225],
        [15619304, 'Onio', 629],
        [15584532, 'Liu', 709],
        [15682355, 'Sabbatini', 772],
        [15628319, 'Walker', 792]], dtype=object)
```

ILOC

```
**1.used to slice the dataset in the machine learning
**2.iloc is uses for indexed based
**3.while loc is uses for label
#dataset.iloc[rows,columns]
df.iloc[0,4]
'France'
x = df.iloc[0:4,2:5]
Input and Output data
x1 = df.iloc[:,-1:]
x1
      Exited
0
            1
1
            0
2
            1
3
            0
4
            0
9995
            0
9996
            0
            1
9997
9998
            1
9999
            0
[10000 rows x 1 columns]
type(x1)
pandas.core.frame.DataFrame
x1 = df.iloc[:,-1]
x1
0
         1
         0
1
2
         1
3
         0
         0
9995
9996
         0
         1
9997
9998
```

```
9999
Name: Exited, Length: 10000, dtype: int64
type(x1)
pandas.core.series.Series
label encoding
df.columns
Index(['RowNumber', 'CustomerId', 'Surname', 'CreditScore',
'Geography',
       'Gender', 'Age', 'Tenure', 'Balance', 'NumOfProducts',
'HasCrCard',
       'IsActiveMember', 'EstimatedSalary', 'Exited'],
      dtype='object')
df.Geography.value counts()
France
           5014
Germany
           2509
Spain
           2476
unknown
              1
Name: Geography, dtype: int64
df.Age.value counts()
37.0
        479
38.0
        477
35.0
        474
36.0
        456
34.0
        447
92.0
          2
82.0
          1
88.0
          1
85.0
          1
83.0
          1
Name: Age, Length: 70, dtype: int64
# converting of categorical data to numerical data
Method 1
from sklearn.preprocessing import LabelEncoder
from collections import Counter as count
s = LabelEncoder()
print("Before lable encoder:",count(df['Geography']))
df['Geography']=s.fit transform(df['Geography'])
```

(10000, 10000)