Importing Libraries

In [107]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import warnings
import seaborn as sns
warnings.filterwarnings("ignore")
```

Exploratory Data Analysis

In [108]:
```
data=pd.read_csv('/home/placement/Desktop/TelecomCustomerChurn.csv')
print(data)
```

```
      customerID  gender  SeniorCitizen Partner Dependents  tenure  \
0     7590-VHVEG  Female              0     Yes         No       1
1     5575-GNVDE    Male              0      No         No      34
2     3668-QPYBK    Male              0      No         No       2
3     7795-CFOCW    Male              0      No         No      45
4     9237-HQITU  Female              0      No         No       2
...          ...     ...            ...     ...        ...     ...
7038  6840-RESVB    Male              0     Yes        Yes      24
7039  2234-XADUH  Female              0     Yes        Yes      72
7040  4801-JZAZL  Female              0     Yes        Yes      11
7041  8361-LTMKD    Male              1     Yes         No       4
7042  3186-AJIEK    Male              0      No         No      66

     PhoneService     MultipleLines InternetService OnlineSecurity  ...  \
0              No  No phone service             DSL             No  ...
1             Yes                No             DSL            Yes  ...
2             Yes                No             DSL            Yes  ...
3              No  No phone service             DSL            Yes  ...
4             Yes                No     Fiber optic             No  ...
...           ...               ...             ...            ...  ...
7038          Yes               Yes             DSL            Yes  ...
7039          Yes               Yes     Fiber optic             No  ...
7040           No  No phone service             DSL            Yes  ...
7041          Yes               Yes     Fiber optic             No  ...
7042          Yes                No     Fiber optic            Yes  ...

     DeviceProtection TechSupport StreamingTV StreamingMovies        Contract  \
0                  No          No          No              No  Month-to-month
1                 Yes          No          No              No        One year
2                  No          No          No              No  Month-to-month
3                 Yes         Yes          No              No        One year
4                  No          No          No              No  Month-to-month
...               ...         ...         ...             ...             ...
7038              Yes         Yes         Yes             Yes        One year
7039              Yes          No         Yes             Yes        One year
7040               No          No          No              No  Month-to-month
7041               No          No          No              No  Month-to-month
7042              Yes         Yes         Yes             Yes        Two year

     PaperlessBilling              PaymentMethod MonthlyCharges  TotalCharges  \
0                 Yes           Electronic check          29.85         29.85
```

```
1                 No              Mailed check          56.95        1889.5
2                Yes              Mailed check          53.85        108.15
3                 No  Bank transfer (automatic)          42.30       1840.75
4                Yes           Electronic check          70.70        151.65
...              ...                       ...            ...           ...
7038             Yes              Mailed check          84.80        1990.5
7039             Yes    Credit card (automatic)         103.20        7362.9
7040             Yes           Electronic check          29.60        346.45
7041             Yes              Mailed check          74.40         306.6
7042             Yes  Bank transfer (automatic)         105.65        6844.5
```

```
     Churn
0       No
1       No
2      Yes
3       No
4      Yes
...    ...
7038    No
7039    No
7040    No
7041   Yes
7042    No

[7043 rows x 21 columns]
```

In [109]:
```
data
```

3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...
...	...	...	...	...	...	...	...	...	...	...	...
7038	6840-RESVB	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	...
7039	2234-XADUH	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	...
7040	4801-JZAZL	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	...
7041	8361-LTMKD	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	...
7042	3186-AJIEK	Male	0	No	No	66	Yes	No	Fiber optic	Yes	...

7043 rows × 21 columns

In [110]:
```
data.head(7)
```

Out[110]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProte
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	
5	9305-CDSKC	Female	0	No	No	8	Yes	Yes	Fiber optic	No	...	
6	1452-KIOVK	Male	0	No	Yes	22	Yes	Yes	Fiber optic	No	...	

7 rows × 21 columns

In [111]:
```
data.describe()
```

Out[111]:

	SeniorCitizen	tenure	MonthlyCharges
count	7043.000000	7043.000000	7043.000000
mean	0.162147	32.371149	64.761692
std	0.368612	24.559481	30.090047
min	0.000000	0.000000	18.250000
25%	0.000000	9.000000	35.500000
50%	0.000000	29.000000	70.350000
75%	0.000000	55.000000	89.850000
max	1.000000	72.000000	118.750000

In [112]:
```
data.columns
```

Out[112]:
```
Index(['customerID', 'gender', 'SeniorCitizen', 'Partner', 'Dependents',
       'tenure', 'PhoneService', 'MultipleLines', 'InternetService',
       'OnlineSecurity', 'OnlineBackup', 'DeviceProtection', 'TechSupport',
       'StreamingTV', 'StreamingMovies', 'Contract', 'PaperlessBilling',
       'PaymentMethod', 'MonthlyCharges', 'TotalCharges', 'Churn'],
      dtype='object')
```

In [113]:
```
list(data)
```

Out[113]:
```
['customerID',
 'gender',
 'SeniorCitizen',
 'Partner',
 'Dependents',
 'tenure',
 'PhoneService',
 'MultipleLines',
 'InternetService',
 'OnlineSecurity',
 'OnlineBackup',
 'DeviceProtection',
 'TechSupport',
 'StreamingTV',
 'StreamingMovies',
 'Contract',
 'PaperlessBilling',
 'PaymentMethod',
 'MonthlyCharges',
 'TotalCharges',
 'Churn']
```

In [114]:
```
data.isna().sum()
```

Out[114]:
```
customerID          0
gender              0
SeniorCitizen       0
Partner             0
Dependents          0
tenure              0
PhoneService        0
MultipleLines       0
InternetService     0
OnlineSecurity      0
OnlineBackup        0
DeviceProtection    0
TechSupport         0
StreamingTV         0
StreamingMovies     0
Contract            0
PaperlessBilling    0
PaymentMethod       0
MonthlyCharges      0
TotalCharges        0
Churn               0
dtype: int64
```

In [115]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7043 entries, 0 to 7042
Data columns (total 21 columns):
 #   Column            Non-Null Count  Dtype  
---  ------            --------------  -----  
 0   customerID        7043 non-null   object 
 1   gender            7043 non-null   object 
 2   SeniorCitizen     7043 non-null   int64  
 3   Partner           7043 non-null   object 
 4   Dependents        7043 non-null   object 
 5   tenure            7043 non-null   int64  
 6   PhoneService      7043 non-null   object 
 7   MultipleLines     7043 non-null   object 
 8   InternetService   7043 non-null   object 
 9   OnlineSecurity    7043 non-null   object 
 10  OnlineBackup      7043 non-null   object 
 11  DeviceProtection  7043 non-null   object 
 12  TechSupport       7043 non-null   object 
 13  StreamingTV       7043 non-null   object 
 14  StreamingMovies   7043 non-null   object 
 15  Contract          7043 non-null   object 
 16  PaperlessBilling  7043 non-null   object 
 17  PaymentMethod     7043 non-null   object 
 18  MonthlyCharges    7043 non-null   float64
 19  TotalCharges      7043 non-null   object 
 20  Churn             7043 non-null   object 
dtypes: float64(1), int64(2), object(18)
memory usage: 1.1+ MB
```

changing the object type into float using numeric

In [116]:
```
data["TotalCharges"]=pd.to_numeric(data['TotalCharges'],errors='coerce')
```

In [117]:
```
data
```

Out[117]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DevicePr
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	
...	...	...	...	...	...	...	...	...	...	...	...	
7038	6840-RESVB	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	...	
7039	2234-XADUH	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	...	
7040	4801-JZAZL	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	...	
7041	8361-LTMKD	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	...	
7042	3186-AJIEK	Male	0	No	No	66	Yes	No	Fiber optic	Yes	...	

7043 rows × 21 columns

In [118]:
```
data.isna().sum()
```

Out[118]:
```
customerID          0
gender              0
SeniorCitizen       0
Partner             0
Dependents          0
tenure              0
PhoneService        0
MultipleLines       0
InternetService     0
OnlineSecurity      0
OnlineBackup        0
DeviceProtection    0
TechSupport         0
StreamingTV         0
StreamingMovies     0
Contract            0
PaperlessBilling    0
PaymentMethod       0
MonthlyCharges      0
TotalCharges       11
Churn               0
dtype: int64
```

Dropping unwanted columns

In [119]:
```
data1=data.drop(columns='customerID')
```

In [120]:
```
data1
```

Out[120]:

	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DevicePro
0	Female	0	Yes	No	1	No	No phone service	DSL	No	Yes	
1	Male	0	No	No	34	Yes	No	DSL	Yes	No	
2	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	
3	Male	0	No	No	45	No	No phone service	DSL	Yes	No	
4	Female	0	No	No	2	Yes	No	Fiber optic	No	No	
...	...	...	...	...	...	...	...	...	...	...	...
7038	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	No	
7039	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	Yes	
7040	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	No	
7041	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	No	
7042	Male	0	No	No	66	Yes	No	Fiber optic	Yes	No	

7043 rows × 20 columns

In [121]:
```
data2=data1.fillna(data1.median())
data2
```

Out[121]:

	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DevicePro
0	Female	0	Yes	No	1	No	No phone service	DSL	No	Yes	
1	Male	0	No	No	34	Yes	No	DSL	Yes	No	
2	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	
3	Male	0	No	No	45	No	No phone service	DSL	Yes	No	
4	Female	0	No	No	2	Yes	No	Fiber optic	No	No	
...	...	...	...	...	...	...	...	...	...	...	...
7038	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	No	
7039	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	Yes	
7040	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	No	
7041	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	No	
7042	Male	0	No	No	66	Yes	No	Fiber optic	Yes	No	

7043 rows × 20 columns

In [122]:
```
data2.isna().sum()
```

Out[122]:
```
gender              0
SeniorCitizen       0
Partner             0
Dependents          0
tenure              0
PhoneService        0
MultipleLines       0
InternetService     0
OnlineSecurity      0
OnlineBackup        0
DeviceProtection    0
TechSupport         0
StreamingTV         0
StreamingMovies     0
Contract            0
PaperlessBilling    0
PaymentMethod       0
MonthlyCharges      0
TotalCharges        0
Churn               0
dtype: int64
```

mapping the churn into 1 and 0

In [123]:
```
data2['Churn']=data2['Churn'].map({'Yes':1,'No':0})
```

In [124]:
```
data2
```

Out[124]:

	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DevicePro
0	Female	0	Yes	No	1	No	No phone service	DSL	No	Yes	
1	Male	0	No	No	34	Yes	No	DSL	Yes	No	
2	Male	0	No	No	2	Yes	No	DSL	Yes	Yes	
3	Male	0	No	No	45	No	No phone service	DSL	Yes	No	
4	Female	0	No	No	2	Yes	No	Fiber optic	No	No	
...	...	...	...	...	...	...	...	...	...	...	...
7038	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	No	
7039	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	Yes	
7040	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	No	
7041	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	No	
7042	Male	0	No	No	66	Yes	No	Fiber optic	Yes	No	

7043 rows × 20 columns

Getting Dummies

In [125]:
```
data3=pd.get_dummies(data2)
```

In [126]:
```
data3
```

Out[126]:

	SeniorCitizen	tenure	MonthlyCharges	TotalCharges	Churn	gender_Female	gender_Male	Partner_No	Partner_Yes	Dependents_No	...	S
0	0	1	29.85	29.85	0	1	0	0	1	1	...	
1	0	34	56.95	1889.50	0	0	1	1	0	1	...	
2	0	2	53.85	108.15	1	0	1	1	0	1	...	
3	0	45	42.30	1840.75	0	0	1	1	0	1	...	
4	0	2	70.70	151.65	1	1	0	1	0	1	...	
...	...	...	...	...	...	...	...	...	...	...	...	
7038	0	24	84.80	1990.50	0	0	1	0	1	0	...	
7039	0	72	103.20	7362.90	0	1	0	0	1	0	...	
7040	0	11	29.60	346.45	0	1	0	0	1	0	...	
7041	1	4	74.40	306.60	1	0	1	0	1	1	...	
7042	0	66	105.65	6844.50	0	0	1	1	0	1	...	

7043 rows × 46 columns

Corelation

In [127]:
```
cor=data3.corr()
cor
```

Out[127]:

	SeniorCitizen	tenure	MonthlyCharges	TotalCharges	Churn	gender_Female	gender_Male	Partner_No	Partner
SeniorCitizen	1.000000	0.016567	0.220173	0.102652	0.150889	0.001874	-0.001874	-0.016479	0.01
tenure	0.016567	1.000000	0.247900	0.825464	-0.352229	-0.005106	0.005106	-0.379697	0.37
MonthlyCharges	0.220173	0.247900	1.000000	0.650864	0.193356	0.014569	-0.014569	-0.096848	0.09
TotalCharges	0.102652	0.825464	0.650864	1.000000	-0.199037	0.000002	-0.000002	-0.318364	0.31
Churn	0.150889	-0.352229	0.193356	-0.199037	1.000000	0.008612	-0.008612	0.150448	-0.15
gender_Female	0.001874	-0.005106	0.014569	0.000002	0.008612	1.000000	-1.000000	-0.001808	0.00
gender_Male	-0.001874	0.005106	-0.014569	-0.000002	-0.008612	-1.000000	1.000000	0.001808	-0.00
Partner_No	-0.016479	-0.379697	-0.096848	-0.318364	0.150448	-0.001808	0.001808	1.000000	-1.00
Partner_Yes	0.016479	0.379697	0.096848	0.318364	-0.150448	0.001808	-0.001808	-1.000000	1.00
Dependents_No	0.211185	-0.159712	0.113890	-0.063593	0.164221	0.010517	-0.010517	0.452676	-0.45
Dependents_Yes	-0.211185	0.159712	-0.113890	0.063593	-0.164221	-0.010517	0.010517	-0.452676	0.45
PhoneService_No	-0.008576	-0.008448	-0.247398	-0.113013	-0.011942	-0.006488	0.006488	0.017706	-0.01
PhoneService_Yes	0.008576	0.008448	0.247398	0.113013	0.011942	0.006488	-0.006488	-0.017706	0.01
MultipleLines_No	-0.136213	-0.323088	-0.338314	-0.396377	-0.032569	-0.004476	0.004476	0.129929	-0.12
MultipleLines_No phone service	-0.008576	-0.008448	-0.247398	-0.113013	-0.011942	-0.006488	0.006488	0.017706	-0.01
MultipleLines_Yes	0.142948	0.331941	0.490434	0.468705	0.040102	0.008414	-0.008414	-0.142057	0.14
InternetService_DSL	-0.108322	0.013274	-0.160189	-0.052279	-0.124214	-0.006568	0.006568	0.000851	-0.00
InternetService_Fiber optic	0.255338	0.019720	0.787066	0.361045	0.308020	0.011286	-0.011286	-0.000304	0.00
InternetService_No	-0.182742	-0.039062	-0.763557	-0.374706	-0.227890	-0.006026	0.006026	-0.000615	0.00
OnlineSecurity_No	0.185532	-0.263746	0.360898	-0.063962	0.342637	-0.010429	0.010429	0.129936	-0.12
OnlineSecurity_No internet service	-0.182742	-0.039062	-0.763557	-0.374706	-0.227890	-0.006026	0.006026	-0.000615	0.00
OnlineSecurity_Yes	-0.038653	0.327203	0.296594	0.412091	-0.171226	0.017021	-0.017021	-0.143106	0.14
OnlineBackup_No	0.087952	-0.312694	0.210753	-0.177070	0.268005	-0.008191	0.008191	0.136058	-0.13
OnlineBackup_No internet service	-0.182742	-0.039062	-0.763557	-0.374706	-0.227890	-0.006026	0.006026	-0.000615	0.00

	SeniorCitizen	tenure	MonthlyCharges	TotalCharges	Churn	gender_Female	gender_Male	Partner_No	Partner
OnlineBackup_Yes	0.066572	0.360277	0.441780	0.509607	-0.082255	0.013773	-0.013773	-0.141498	0.14
DeviceProtection_No	0.094810	-0.312740	0.171836	-0.188911	0.252481	0.002988	-0.002988	0.147692	-0.14
DeviceProtection_No internet service	-0.182742	-0.039062	-0.763557	-0.374706	-0.227890	-0.006026	0.006026	-0.000615	0.00
DeviceProtection_Yes	0.059428	0.360653	0.482692	0.522374	-0.066160	0.002105	-0.002105	-0.153786	0.15
TechSupport_No	0.205620	-0.262143	0.322076	-0.083705	0.337281	-0.003397	0.003397	0.109443	-0.10
TechSupport_No internet service	-0.182742	-0.039062	-0.763557	-0.374706	-0.227890	-0.006026	0.006026	-0.000615	0.00
TechSupport_Yes	-0.060625	0.324221	0.338304	0.432329	-0.164674	0.009212	-0.009212	-0.119999	0.11
StreamingTV_No	0.049062	-0.245039	0.016951	-0.196622	0.128916	-0.003267	0.003267	0.124357	-0.12
StreamingTV_No internet service	-0.182742	-0.039062	-0.763557	-0.374706	-0.227890	-0.006026	0.006026	-0.000615	0.00
StreamingTV_Yes	0.105378	0.279756	0.629603	0.515279	0.063228	0.008393	-0.008393	-0.124666	0.12
StreamingMovies_No	0.034210	-0.252220	0.018075	-0.202387	0.130845	-0.005374	0.005374	0.117529	-0.11
StreamingMovies_No internet service	-0.182742	-0.039062	-0.763557	-0.374706	-0.227890	-0.006026	0.006026	-0.000615	0.00
StreamingMovies_Yes	0.120176	0.286111	0.627429	0.519884	0.061382	0.010487	-0.010487	-0.117412	0.11
Contract_Month-to-month	0.138360	-0.645561	0.060165	-0.445619	0.405103	0.003386	-0.003386	0.280865	-0.28
Contract_One year	-0.046262	0.202570	0.004904	0.170649	-0.177820	-0.008026	0.008026	-0.082783	0.08
Contract_Two year	-0.117000	0.558533	-0.074681	0.356226	-0.302253	0.003695	-0.003695	-0.248091	0.24
PaperlessBilling_No	-0.156530	-0.006152	-0.352150	-0.158055	-0.191825	-0.011754	0.011754	-0.014877	0.01
PaperlessBilling_Yes	0.156530	0.006152	0.352150	0.158055	0.191825	0.011754	-0.011754	0.014877	-0.01
PaymentMethod_Bank transfer (automatic)	-0.016159	0.243510	0.042812	0.186025	-0.117937	0.016024	-0.016024	-0.110706	0.11
PaymentMethod_Credit card (automatic)	-0.024135	0.233006	0.030550	0.182745	-0.134302	-0.001215	0.001215	-0.082029	0.08
PaymentMethod_Electronic check	0.171718	-0.208363	0.271625	-0.059971	0.301919	-0.000752	0.000752	0.083852	-0.08

	SeniorCitizen	tenure	MonthlyCharges	TotalCharges	Churn	gender_Female	gender_Male	Partner_No	Partner
PaymentMethod_Mailed check	-0.153477	-0.233852	-0.377437	-0.294814	-0.091683	-0.013744	0.013744	0.095125	-0.09

46 rows × 46 columns

In [128]:
```
y=data3['Churn']
x=data3.drop(columns='Churn')
```

In [129]:
```
x
```

Out[129]:

	SeniorCitizen	tenure	MonthlyCharges	TotalCharges	gender_Female	gender_Male	Partner_No	Partner_Yes	Dependents_No	Dependents_
0	0	1	29.85	29.85	1	0	0	1	1	
1	0	34	56.95	1889.50	0	1	1	0	1	
2	0	2	53.85	108.15	0	1	1	0	1	
3	0	45	42.30	1840.75	0	1	1	0	1	
4	0	2	70.70	151.65	1	0	1	0	1	
...	...	...	...	...	...	...	...	...	...	
7038	0	24	84.80	1990.50	0	1	0	1	0	
7039	0	72	103.20	7362.90	1	0	0	1	0	
7040	0	11	29.60	346.45	1	0	0	1	0	
7041	1	4	74.40	306.60	0	1	0	1	1	
7042	0	66	105.65	6844.50	0	1	1	0	1	

7043 rows × 45 columns

In [130]:
```
y
```

Out[130]:
```
0       0
1       0
2       1
3       0
4       1
       ..
7038    0
7039    0
7040    0
7041    1
7042    0
Name: Churn, Length: 7043, dtype: int64
```

In [239]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.33,random_state=42)
```

In [240]:
```
x_train
```

Out[240]:

	SeniorCitizen	tenure	MonthlyCharges	TotalCharges	gender_Female	gender_Male	Partner_No	Partner_Yes	Dependents_No	Dependents_
298	0	40	74.55	3015.75	0	1	0	1	0	
3318	0	10	29.50	255.25	0	1	1	0	1	
5586	0	27	19.15	501.35	1	0	1	0	1	
6654	0	7	86.50	582.50	1	0	0	1	1	
5362	0	65	24.75	1715.10	0	1	0	1	0	
...	...	...	...	...	...	...	...	...	...	
3772	0	1	95.00	95.00	0	1	0	1	1	
5191	0	23	91.10	2198.30	1	0	0	1	0	
5226	0	12	21.15	306.05	0	1	0	1	0	
5390	1	12	99.45	1200.15	0	1	1	0	1	
860	0	26	19.80	457.30	0	1	1	0	1	

4718 rows × 45 columns

In [241]:
```
x_test
```

Out[241]:

	SeniorCitizen	tenure	MonthlyCharges	TotalCharges	gender_Female	gender_Male	Partner_No	Partner_Yes	Dependents_No	Dependents_
185	0	1	24.80	24.80	1	0	0	1	1	
2715	0	41	25.25	996.45	0	1	1	0	1	
3825	0	52	19.35	1031.70	1	0	0	1	0	
1807	0	1	76.35	76.35	1	0	1	0	1	
132	0	67	50.55	3260.10	0	1	1	0	1	
...	...	...	...	...	...	...	...	...	...	
4147	1	71	24.85	1901.00	0	1	0	1	1	
3542	0	29	55.35	1636.95	0	1	0	1	0	
3759	0	7	89.35	631.85	0	1	1	0	1	
1114	0	32	98.85	3145.90	0	1	0	1	1	
4958	0	59	94.75	5597.65	1	0	0	1	1	

2325 rows × 45 columns

In [242]:
```
y_train
```

Out[242]:
```
298     0
3318    1
5586    0
6654    1
5362    0
       ..
3772    1
5191    0
5226    0
5390    1
860     0
Name: Churn, Length: 4718, dtype: int64
```

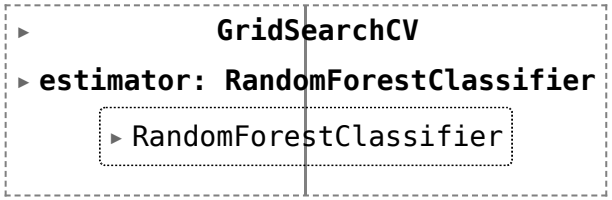
In [243]:
```
y_test
```

Out[243]:
```
185     1
2715    0
3825    0
1807    1
132     0
       ..
4147    0
3542    0
3759    1
1114    0
4958    0
Name: Churn, Length: 2325, dtype: int64
```

Random Forest Regression

In [244]:
```
from sklearn.model_selection import GridSearchCV     #GridSearchCV is for parameter tuning
from sklearn.ensemble import RandomForestClassifier
cls=RandomForestClassifier()
n_estimators=[25,50,75,100,125,150,175,200]          #number of decision trees in the forest, default = 10
criterion=['gini','entropy']                         #criteria for choosing nodes default = 'gini'
max_depth=[3,5,10]       #maximum number of nodes in a tree default = None (it will go till all possible n
parameters={'n_estimators': n_estimators,'criterion':criterion,'max_depth':max_depth} #this will undergo 8
RFC_cls = GridSearchCV(cls, parameters)
RFC_cls.fit(x_train,y_train)
```

Out[244]:



In [252]:
```
RFC_cls.best_params_
```

Out[252]:
```
{'criterion': 'entropy', 'max_depth': 10, 'n_estimators': 125}
```

In [253]:
```
y_pred=RFC_cls.predict(x_test)
y_pred
```

Out[253]:
```
array([1, 0, 0, ..., 1, 0, 0])
```

Confusion Matrix

In [254]:
```
from sklearn.metrics import confusion_matrix     #confusion matrix
confusion_matrix(y_test,y_pred)
```

Out[254]:
```
array([[1545,  152],
       [ 300,  328]])
```

Accuracy of the data

In [255]:
```
from sklearn.metrics import accuracy_score   ##accuracy of test data and predicted data
accuracy_score(y_test,y_pred)
```

Out[255]:
```
0.8055913978494623
```

Logistic Regression

In [256]:
```
from sklearn.linear_model import LogisticRegression
classifier=LogisticRegression()
classifier.fit(x_train,y_train)
```

Out[256]:
```
LogisticRegression
LogisticRegression()
```

In [257]:
```
y_pred=classifier.predict(x_test)
y_pred
```

Out[257]:
```
array([1, 0, 0, ..., 1, 0, 0])
```

In [258]:
```
from sklearn.metrics import confusion_matrix     #confusion matrix
confusion_matrix(y_test,y_pred)
```

Out[258]:
```
array([[1526,  171],
       [ 266,  362]])
```

In [259]:
```
from sklearn.metrics import accuracy_score   ##accuracy of test data and predicted data
accuracy_score(y_test,y_pred)
```

Out[259]:
```
0.8120430107526881
```

In [ ]:
```

```