In [24]:
```
import pandas as pd
import warnings
warnings.filterwarnings("ignore")
```

In [25]:
```
data=pd.read_csv("/home/placement/Desktop/divyasri/TelecomCustomerChurn.csv")
```

In [26]:
```
data.describe()
```

Out[26]:

	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 [27]:
```
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
```

In [28]:
```
list(data)
```

Out[28]:
```
['customerID',
 'gender',
 'SeniorCitizen',
 'Partner',
 'Dependents',
 'tenure',
 'PhoneService',
 'MultipleLines',
 'InternetService',
 'OnlineSecurity',
 'OnlineBackup',
 'DeviceProtection',
 'TechSupport',
 'StreamingTV',
 'StreamingMovies',
 'Contract',
 'PaperlessBilling',
 'PaymentMethod',
 'MonthlyCharges',
 'TotalCharges',
 'Churn']
```

In [29]:
```
data['TotalCharges']=pd.to_numeric(data['TotalCharges'],errors='coerce')
```

In [30]:
```
data.dtypes
```

Out[30]:
```
customerID          object
gender              object
SeniorCitizen        int64
Partner             object
Dependents          object
tenure               int64
PhoneService        object
MultipleLines       object
InternetService     object
OnlineSecurity      object
OnlineBackup        object
DeviceProtection    object
TechSupport         object
StreamingTV         object
StreamingMovies     object
Contract            object
PaperlessBilling    object
PaymentMethod       object
MonthlyCharges     float64
TotalCharges       float64
Churn               object
dtype: object
```

In [31]:
```
data.isna().sum()
```

Out[31]:
```
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
```

In [32]:
```
data
```

Out[32]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DevicePro
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 [33]:
```
list(data)
```

Out[33]:
```
['customerID',
 'gender',
 'SeniorCitizen',
 'Partner',
 'Dependents',
 'tenure',
 'PhoneService',
 'MultipleLines',
 'InternetService',
 'OnlineSecurity',
 'OnlineBackup',
 'DeviceProtection',
 'TechSupport',
 'StreamingTV',
 'StreamingMovies',
 'Contract',
 'PaperlessBilling',
 'PaymentMethod',
 'MonthlyCharges',
 'TotalCharges',
 'Churn']
```

In [36]:
```
data=data.drop(['customerID','SeniorCitizen','OnlineSecurity','OnlineBackup','DeviceProtection','TechSupport
```

In [39]:
```
data
```

Out[39]:

	gender	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	Contract	MonthlyCharges	TotalCharges	Churn
0	Female	Yes	No	1	No	No phone service	DSL	Month-to-month	29.85	29.85	No
1	Male	No	No	34	Yes	No	DSL	One year	56.95	1889.50	No
2	Male	No	No	2	Yes	No	DSL	Month-to-month	53.85	108.15	Yes
3	Male	No	No	45	No	No phone service	DSL	One year	42.30	1840.75	No
4	Female	No	No	2	Yes	No	Fiber optic	Month-to-month	70.70	151.65	Yes
...	...	...	...	...	...	...	...	...	...	...	...
7038	Male	Yes	Yes	24	Yes	Yes	DSL	One year	84.80	1990.50	No
7039	Female	Yes	Yes	72	Yes	Yes	Fiber optic	One year	103.20	7362.90	No
7040	Female	Yes	Yes	11	No	No phone service	DSL	Month-to-month	29.60	346.45	No
7041	Male	Yes	No	4	Yes	Yes	Fiber optic	Month-to-month	74.40	306.60	Yes
7042	Male	No	No	66	Yes	No	Fiber optic	Two year	105.65	6844.50	No

7043 rows × 11 columns

In [40]:
```
data=data.fillna(data.median())
```

In [41]:
```
data
```

Out[41]:

	gender	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	Contract	MonthlyCharges	TotalCharges	Churn
0	Female	Yes	No	1	No	No phone service	DSL	Month-to-month	29.85	29.85	No
1	Male	No	No	34	Yes	No	DSL	One year	56.95	1889.50	No
2	Male	No	No	2	Yes	No	DSL	Month-to-month	53.85	108.15	Yes
3	Male	No	No	45	No	No phone service	DSL	One year	42.30	1840.75	No
4	Female	No	No	2	Yes	No	Fiber optic	Month-to-month	70.70	151.65	Yes
...	...	...	...	...	...	...	...	...	...	...	...
7038	Male	Yes	Yes	24	Yes	Yes	DSL	One year	84.80	1990.50	No
7039	Female	Yes	Yes	72	Yes	Yes	Fiber optic	One year	103.20	7362.90	No
7040	Female	Yes	Yes	11	No	No phone service	DSL	Month-to-month	29.60	346.45	No
7041	Male	Yes	No	4	Yes	Yes	Fiber optic	Month-to-month	74.40	306.60	Yes
7042	Male	No	No	66	Yes	No	Fiber optic	Two year	105.65	6844.50	No

7043 rows × 11 columns

In [42]:
```
data['Churn']=data['Churn'].map({'Yes':1,'No':0})
data
```

Out[42]:

	gender	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	Contract	MonthlyCharges	TotalCharges	Churn
0	Female	Yes	No	1	No	No phone service	DSL	Month-to-month	29.85	29.85	0
1	Male	No	No	34	Yes	No	DSL	One year	56.95	1889.50	0
2	Male	No	No	2	Yes	No	DSL	Month-to-month	53.85	108.15	1
3	Male	No	No	45	No	No phone service	DSL	One year	42.30	1840.75	0
4	Female	No	No	2	Yes	No	Fiber optic	Month-to-month	70.70	151.65	1
...	...	...	...	...	...	...	...	...	...	...	...
7038	Male	Yes	Yes	24	Yes	Yes	DSL	One year	84.80	1990.50	0
7039	Female	Yes	Yes	72	Yes	Yes	Fiber optic	One year	103.20	7362.90	0
7040	Female	Yes	Yes	11	No	No phone service	DSL	Month-to-month	29.60	346.45	0
7041	Male	Yes	No	4	Yes	Yes	Fiber optic	Month-to-month	74.40	306.60	1
7042	Male	No	No	66	Yes	No	Fiber optic	Two year	105.65	6844.50	0

7043 rows × 11 columns

In [43]:
```
data=pd.get_dummies(data)
data
```

Out[43]:

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

7043 rows × 21 columns

In [44]:
```
y=data['Churn']
x=data.drop(['Churn'],axis=1)
```

In [45]:
```
y
```

Out[45]:
```
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 [46]:
```
x
```

Out[46]:

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

7043 rows × 20 columns

In [47]:
```
from sklearn.model_selection import train_test_split #spliting of training and testing
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.33,random_state=42)
```

In [53]:
```
y_train.head(5)
```

Out[53]:
```
298     0
3318    1
5586    0
6654    1
5362    0
Name: Churn, dtype: int64
```

In [54]:
```
x_train.head(5)
```

Out[54]:

	tenure	MonthlyCharges	TotalCharges	gender_Female	gender_Male	Partner_No	Partner_Yes	Dependents_No	Dependents_Yes	PhoneS
298	40	74.55	3015.75	0	1	0	1	0	1	
3318	10	29.50	255.25	0	1	1	0	1	0	
5586	27	19.15	501.35	1	0	1	0	1	0	
6654	7	86.50	582.50	1	0	0	1	1	0	
5362	65	24.75	1715.10	0	1	0	1	0	1	

In [55]:
```
y_test.head(5)
```

Out[55]:
```
185     1
2715    0
3825    0
1807    1
132     0
Name: Churn, dtype: int64
```

In [56]:
```
x_test.head(5)
```

Out[56]:

	tenure	MonthlyCharges	TotalCharges	gender_Female	gender_Male	Partner_No	Partner_Yes	Dependents_No	Dependents_Yes	PhoneServ
185	1	24.80	24.80	1	0	0	1	1	0	
2715	41	25.25	996.45	0	1	1	0	1	0	
3825	52	19.35	1031.70	1	0	0	1	0	1	
1807	1	76.35	76.35	1	0	1	0	1	0	
132	67	50.55	3260.10	0	1	1	0	1	0	

In [57]:
```
cor=data.corr()
cor
```

Out[57]:

	tenure	MonthlyCharges	TotalCharges	Churn	gender_Female	gender_Male	Partner_No	Partner_Yes	Dependents_No
tenure	1.000000	0.247900	0.825464	-0.352229	-0.005106	0.005106	-0.379697	0.379697	-0.159712
MonthlyCharges	0.247900	1.000000	0.650864	0.193356	0.014569	-0.014569	-0.096848	0.096848	0.113890
TotalCharges	0.825464	0.650864	1.000000	-0.199037	0.000002	-0.000002	-0.318364	0.318364	-0.063593
Churn	-0.352229	0.193356	-0.199037	1.000000	0.008612	-0.008612	0.150448	-0.150448	0.164221
gender_Female	-0.005106	0.014569	0.000002	0.008612	1.000000	-1.000000	-0.001808	0.001808	0.010517
gender_Male	0.005106	-0.014569	-0.000002	-0.008612	-1.000000	1.000000	0.001808	-0.001808	-0.010517
Partner_No	-0.379697	-0.096848	-0.318364	0.150448	-0.001808	0.001808	1.000000	-1.000000	0.452676
Partner_Yes	0.379697	0.096848	0.318364	-0.150448	0.001808	-0.001808	-1.000000	1.000000	-0.452676
Dependents_No	-0.159712	0.113890	-0.063593	0.164221	0.010517	-0.010517	0.452676	-0.452676	1.000000
Dependents_Yes	0.159712	-0.113890	0.063593	-0.164221	-0.010517	0.010517	-0.452676	0.452676	-1.000000
PhoneService_No	-0.008448	-0.247398	-0.113013	-0.011942	-0.006488	0.006488	0.017706	-0.017706	-0.001762
PhoneService_Yes	0.008448	0.247398	0.113013	0.011942	0.006488	-0.006488	-0.017706	0.017706	0.001762
MultipleLines_No	-0.323088	-0.338314	-0.396377	-0.032569	-0.004476	0.004476	0.129929	-0.129929	-0.023198
MultipleLines_No phone service	-0.008448	-0.247398	-0.113013	-0.011942	-0.006488	0.006488	0.017706	-0.017706	-0.001762
MultipleLines_Yes	0.331941	0.490434	0.468705	0.040102	0.008414	-0.008414	-0.142057	0.142057	0.024526
InternetService_DSL	0.013274	-0.160189	-0.052279	-0.124214	-0.006568	0.006568	0.000851	-0.000851	-0.052010
InternetService_Fiber optic	0.019720	0.787066	0.361045	0.308020	0.011286	-0.011286	-0.000304	0.000304	0.165818
InternetService_No	-0.039062	-0.763557	-0.374706	-0.227890	-0.006026	0.006026	-0.000615	0.000615	-0.139812
Contract_Month-to-month	-0.645561	0.060165	-0.445619	0.405103	0.003386	-0.003386	0.280865	-0.280865	0.231720
Contract_One year	0.202570	0.004904	0.170649	-0.177820	-0.008026	0.008026	-0.082783	0.082783	-0.068368
Contract_Two year	0.558533	-0.074681	0.356226	-0.302253	0.003695	-0.003695	-0.248091	0.248091	-0.204613

21 rows × 21 columns

In [49]:
```
from sklearn.linear_model import LogisticRegression #logistic regression
classifier=LogisticRegression()
classifier.fit(x_train,y_train)
```

Out[49]:
```
▾ LogisticRegression

LogisticRegression()
```

In [50]:
```
y_pred=classifier.predict(x_test)
y_pred
```

Out[50]:
```
array([1, 0, 0, ..., 1, 1, 0])
```

In [51]:
```
from sklearn.metrics import confusion_matrix #confusion_matrix
confusion_matrix(y_test,y_pred)
```

Out[51]:
```
array([[1514,  183],
       [ 272,  356]])
```

In [52]:
```
from sklearn.metrics import accuracy_score
accuracy_score(y_test,y_pred)              #accuracy value
```

Out[52]:
```
0.8043010752688172
```

In [ ]:
```

```