1. Recalling Python Library and dataset

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sbn
df=pd.read_csv("/content/telecom_churn.csv")
df
```

	Churn	Account Weeks	ContractRenewal	Data Plan	DataUsage	CustServCalls	Day Mins	DayCalls	Monthly Charge	OverageFee	Roam Mins
0	0	128	1	1	2.70	1	265.1	110	89.0	9.87	10.0
1	0	107	1	1	3.70	1	161.6	123	82.0	9.78	13.7
2	0	137	1	0	0.00	0	243.4	114	52.0	6.06	12.2
3	0	84	0	0	0.00	2	299.4	71	57.0	3.10	6.6
4	0	75	0	0	0.00	3	166.7	113	41.0	7.42	10.1
...	...	...	...	...	...	...	...	...	...	...	...
3328	0	192	1	1	2.67	2	156.2	77	71.7	10.78	9.9
3329	0	68	1	0	0.34	3	231.1	57	56.4	7.67	9.6
3330	0	28	1	0	0.00	2	180.8	109	56.0	14.44	14.1
3331	0	184	0	0	0.00	2	213.8	105	50.0	7.98	5.0
3332	0	74	1	1	3.70	0	234.4	113	100.0	13.30	13.7

3333 rows × 11 columns

2. Checking Missing Values

```
Missing_values=df.isnull().sum()
Missing_values
Churn           0
AccountWeeks    0
ContractRenewal 0
DataPlan        0
DataUsage       0
CustServCalls   0
DayMins         0
DayCalls        0
MonthlyCharge   0
OverageFee      0
RoamMins        0
dtype: int64
```

⇒ There is no missing values.

3. Brief Description of Data

df.describe()

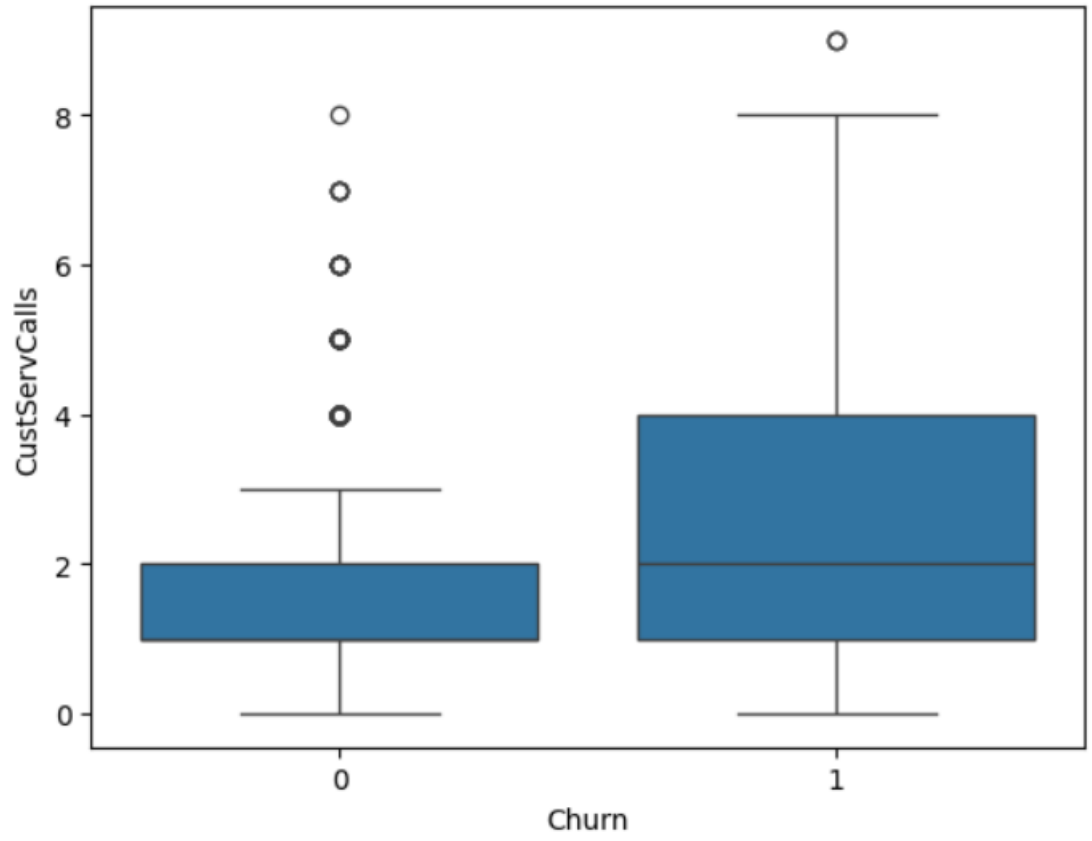
	Churn	AccountWeeks	ContractRenewal	DataPlan	DataUsage	CustServCalls	DayMins	DayCalls	MonthlyCharge	OverageFee	RoamMins
count	3333.000000	3333.000000	3333.000000	3333.000000	3333.000000	3333.000000	3333.000000	3333.000000	3333.000000	3333.000000	3333.000000
mean	0.144914	101.064806	0.903090	0.276628	0.816475	1.562856	179.775098	100.435644	56.305161	10.051488	10.237294
std	0.352067	39.822106	0.295879	0.447398	1.272668	1.315491	54.467389	20.069084	16.426032	2.535712	2.791840
min	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	14.000000	0.000000	0.000000
25%	0.000000	74.000000	1.000000	0.000000	0.000000	1.000000	143.700000	87.000000	45.000000	8.330000	8.500000
50%	0.000000	101.000000	1.000000	0.000000	0.000000	1.000000	179.400000	101.000000	53.500000	10.070000	10.300000

	Churn	AccountWeeks	ContractRenewal	DataPlan	DataUsage	CustServCalls	DayMins	DayCalls	MonthlyCharge	OverageFee	RoamMins
75%	0.000000	127.000000	1.000000	1.000000	1.780000	2.000000	216.400000	114.000000	66.200000	11.770000	12.100000
max	1.000000	243.000000	1.000000	1.000000	5.400000	9.000000	350.800000	165.000000	111.300000	18.190000	20.000000

4.1 Data Analysis

sbn.boxplot(x="Churn", y="CustServCalls", data=df)

```
<Axes: xlabel='Churn', ylabel='CustServCalls'>
```

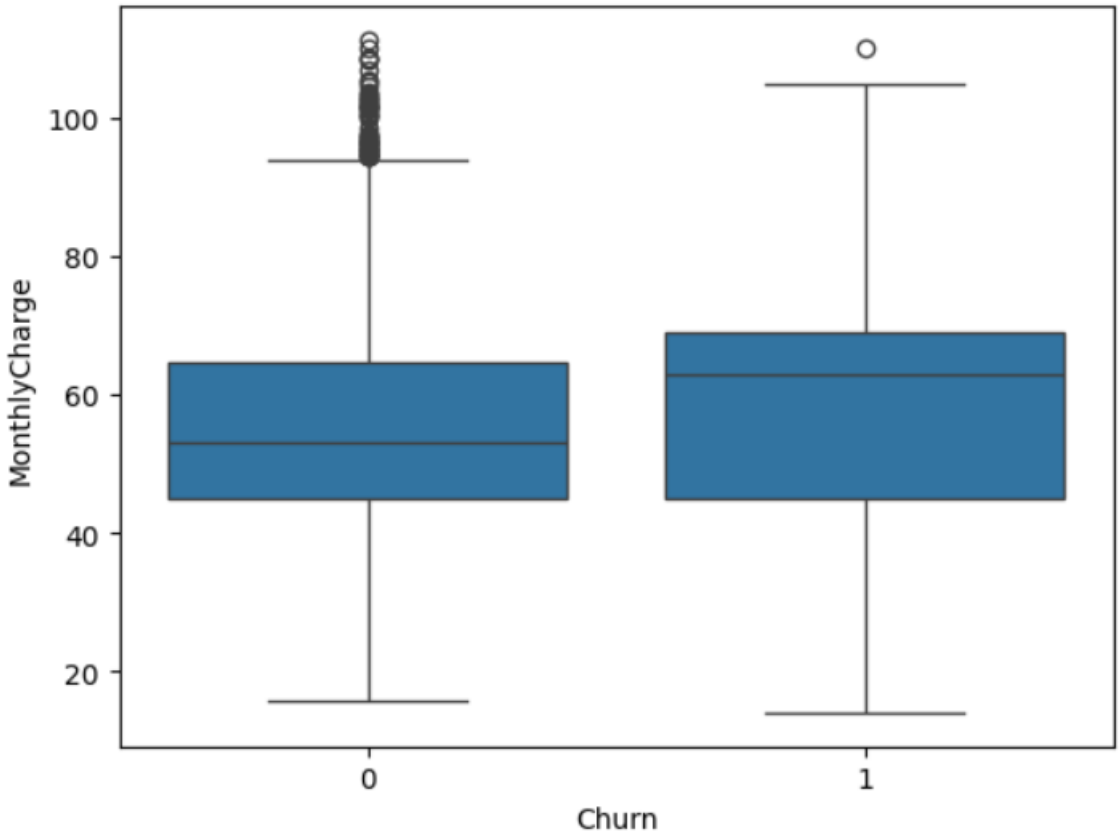


⇒ People having CustServCalls >=2 are leaving the network.

4.2

sbn.boxplot(x="Churn", y="MonthlyCharge", data=df)

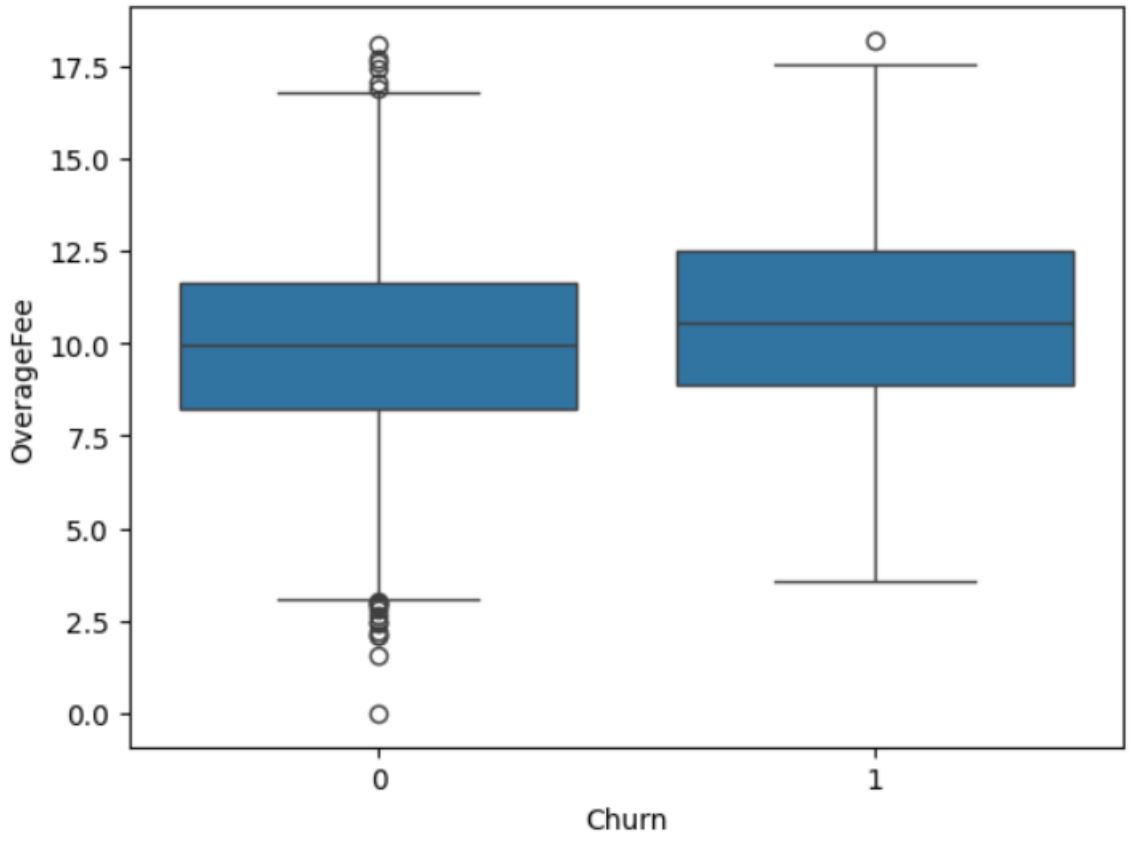
```
<Axes: xlabel='Churn', ylabel='MonthlyCharge'>
```



⇒ People having MonthlyCharge > $60 are leaving the system

4.3 sbn.boxplot(x="Churn", y="OverageFee", data=df)

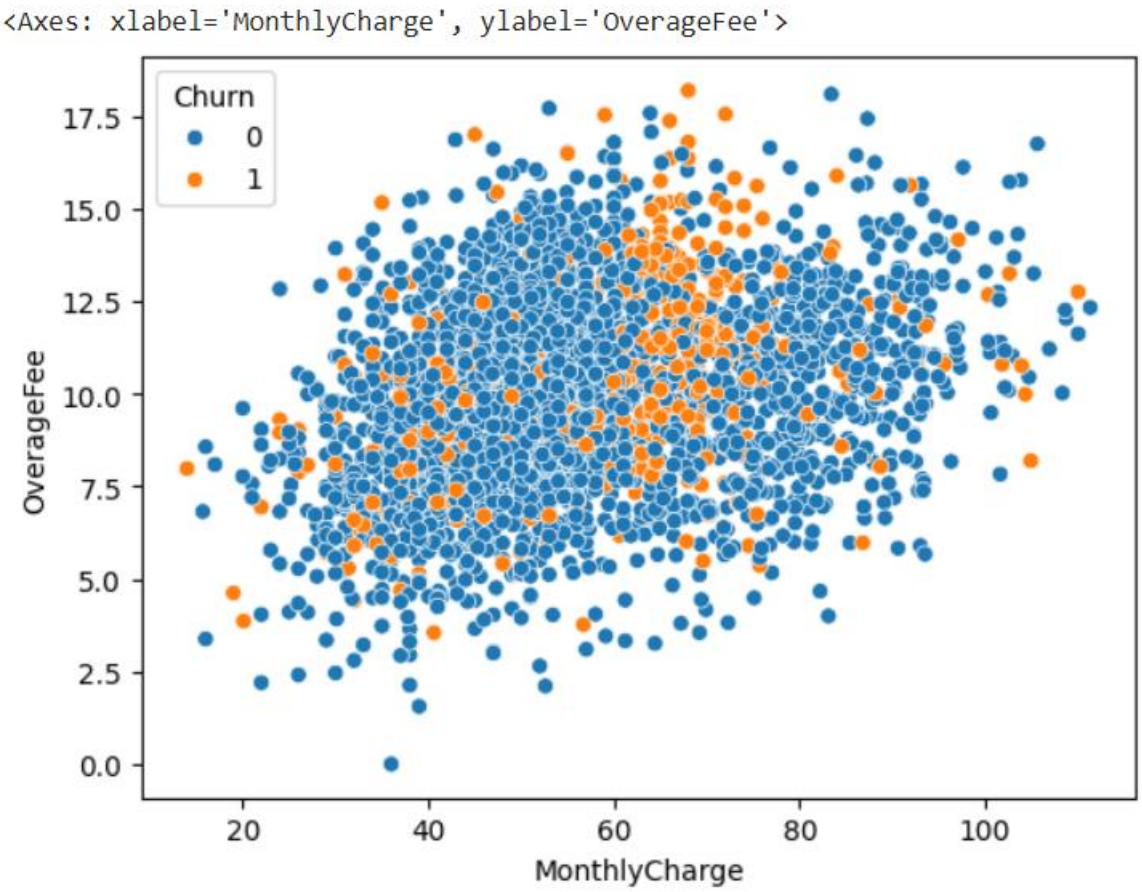
```
<Axes: xlabel='Churn', ylabel='OverageFee'>
```



⇒ People having OverageFee > $10 are leaving the system

4.4

sbn.scatterplot(x="MonthlyCharge",y="OverageFee", data=df, hue="Churn")

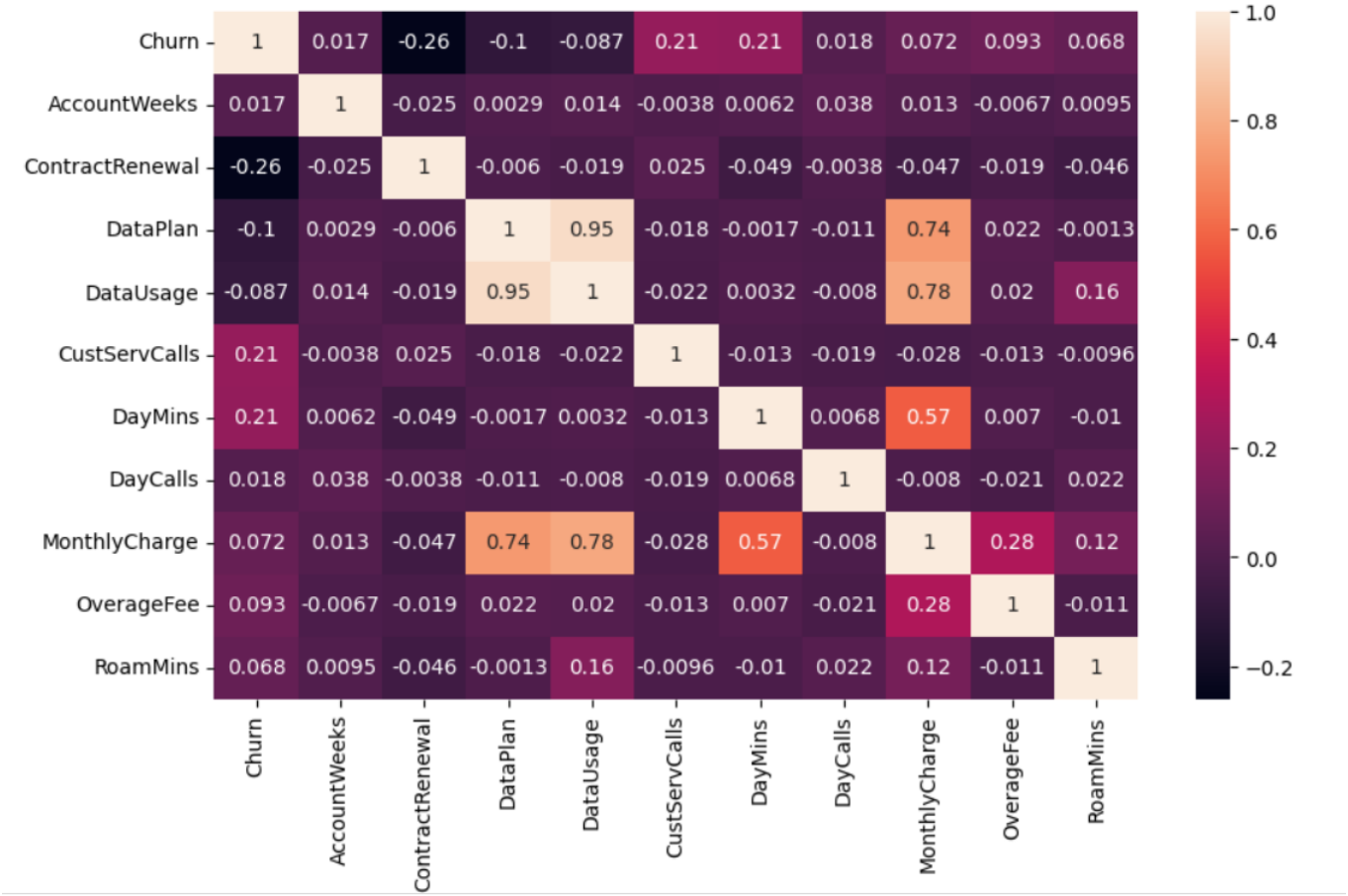


⇒ People having MonthlyCharge between $60 & $80 and OverageFee > $7.5 are leaving the network.

4.5

```
# Checking for Causation
plt.figure(figsize=(10,10))
sbn.heatmap(df.corr(),annot=True)
```

<Axes: >



⇒ **Causation:** The major effect on Customer Churn from the system is caused by 'CustServCalls'. The more CustServCalls may lead to Customer Churn.

Finally, CustServCalls has major impact on Customer Churn.