

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv("Customer_Churn.csv")

df.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

df["TotalCharges"] = df["TotalCharges"].replace(" ", "0")
df["TotalCharges"] = df["TotalCharges"].astype("float")

df.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	float64
20	Churn	7043	non-null	object

dtypes: float64(2), int64(2), object(17)

memory usage: 1.1+ MB

df.isnull().sum()

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

#converted 0 and 1 value to object yes or no in the case of Senior Citizen

```
df["SeniorCitizen"] = df["SeniorCitizen"].replace({1: "yes", 0: "no"})
df.head(30)
```

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure
0	7590-VHVEG	Female	no	Yes	No	1
1	5575-GNVDE	Male	no	No	No	34
2	3668-QPYBK	Male	no	No	No	2
3	7795-CF0CW	Male	no	No	No	45
4	9237-HQITU	Female	no	No	No	2
5	9305-CDSKC	Female	no	No	No	8
6	1452-KIOVK	Male	no	No	Yes	22
7	6713-OKOMC	Female	no	No	No	10
8	7892-P00KP	Female	no	Yes	No	28
9	6388-TABGU	Male	no	No	Yes	62
10	9763-GRSKD	Male	no	Yes	Yes	13
11	7469-LKBCI	Male	no	No	No	16
12	8091-TTVAX	Male	no	Yes	No	58
13	0280-XJGEX	Male	no	No	No	49
14	5129-JLPIS	Male	no	No	No	25
15	3655-SNQYZ	Female	no	Yes	Yes	69
16	8191-XWSZG	Female	no	No	No	52
17	9959-W0FKT	Male	no	No	Yes	71
18	4190-MFLUW	Female	no	Yes	Yes	10
19	4183-MYFRB	Female	no	No	No	21
20	8779-QRDMV	Male	yes	No	No	1
21	1680-VDCWW	Male	no	Yes	No	12

22	1066-JKSGK	Male	no	No	No	1
Yes						
23	3638-WEABW	Female	no	Yes	No	58
Yes						
24	6322-HRPFA	Male	no	Yes	Yes	49
Yes						
25	6865-JZNK0	Female	no	No	No	30
Yes						
26	6467-CHFZW	Male	no	Yes	Yes	47
Yes						
27	8665-UTDHz	Male	no	Yes	Yes	1
No						
28	5248-YGIJN	Male	no	Yes	No	72
Yes						
29	8773-HHU0Z	Female	no	No	Yes	17
Yes						

	MultipleLines	InternetService	OnlineSecurity	...	\
0	No phone service	DSL	No	...	
1	No	DSL	Yes	...	
2	No	DSL	Yes	...	
3	No phone service	DSL	Yes	...	
4	No	Fiber optic	No	...	
5	Yes	Fiber optic	No	...	
6	Yes	Fiber optic	No	...	
7	No phone service	DSL	Yes	...	
8	Yes	Fiber optic	No	...	
9	No	DSL	Yes	...	
10	No	DSL	Yes	...	
11	No	No	No internet service	...	
12	Yes	Fiber optic	No	...	
13	Yes	Fiber optic	No	...	
14	No	Fiber optic	Yes	...	
15	Yes	Fiber optic	Yes	...	
16	No	No	No internet service	...	
17	Yes	Fiber optic	Yes	...	
18	No	DSL	No	...	
19	No	Fiber optic	No	...	
20	No phone service	DSL	No	...	
21	No	No	No internet service	...	
22	No	No	No internet service	...	
23	Yes	DSL	No	...	
24	No	DSL	Yes	...	
25	No	DSL	Yes	...	
26	Yes	Fiber optic	No	...	
27	No phone service	DSL	No	...	
28	Yes	DSL	Yes	...	
29	No	DSL	No	...	

	DeviceProtection	TechSupport	StreamingTV	\
0	No	No	No	
1	Yes	No	No	
2	No	No	No	
3	Yes	Yes	No	
4	No	No	No	
5	Yes	No	Yes	
6	No	No	Yes	
7	No	No	No	
8	Yes	Yes	Yes	
9	No	No	No	
10	No	No	No	
11	No internet service	No internet service	No internet service	
12	Yes	No	Yes	
13	Yes	No	Yes	
14	Yes	Yes	Yes	
15	Yes	Yes	Yes	
16	No internet service	No internet service	No internet service	
17	Yes	No	Yes	
18	Yes	Yes	No	
19	Yes	No	No	
20	Yes	No	No	
21	No internet service	No internet service	No internet service	
22	No internet service	No internet service	No internet service	
23	No	Yes	No	
24	No	Yes	No	
25	No	No	No	
26	No	No	Yes	
27	No	No	No	
28	Yes	Yes	Yes	
29	No	No	Yes	
	StreamingMovies	Contract	PaperlessBilling	\
0	No	Month-to-month	Yes	
1	No	One year	No	
2	No	Month-to-month	Yes	
3	No	One year	No	
4	No	Month-to-month	Yes	
5	Yes	Month-to-month	Yes	
6	No	Month-to-month	Yes	
7	No	Month-to-month	No	
8	Yes	Month-to-month	Yes	
9	No	One year	No	
10	No	Month-to-month	Yes	
11	No internet service	Two year	No	
12	Yes	One year	No	
13	Yes	Month-to-month	Yes	
14	Yes	Month-to-month	Yes	
15	Yes	Two year	No	

16	No internet service	One year	No
17	Yes	Two year	No
18	No	Month-to-month	No
19	Yes	Month-to-month	Yes
20	Yes	Month-to-month	Yes
21	No internet service	One year	No
22	No internet service	Month-to-month	No
23	No	Two year	Yes
24	No	Month-to-month	No
25	No	Month-to-month	Yes
26	Yes	Month-to-month	Yes
27	No	Month-to-month	No
28	Yes	Two year	Yes
29	Yes	Month-to-month	Yes

	PaymentMethod	MonthlyCharges	TotalCharges	Churn
0	Electronic check	29.85	29.85	No
1	Mailed check	56.95	1889.50	No
2	Mailed check	53.85	108.15	Yes
3	Bank transfer (automatic)	42.30	1840.75	No
4	Electronic check	70.70	151.65	Yes
5	Electronic check	99.65	820.50	Yes
6	Credit card (automatic)	89.10	1949.40	No
7	Mailed check	29.75	301.90	No
8	Electronic check	104.80	3046.05	Yes
9	Bank transfer (automatic)	56.15	3487.95	No
10	Mailed check	49.95	587.45	No
11	Credit card (automatic)	18.95	326.80	No
12	Credit card (automatic)	100.35	5681.10	No
13	Bank transfer (automatic)	103.70	5036.30	Yes
14	Electronic check	105.50	2686.05	No
15	Credit card (automatic)	113.25	7895.15	No
16	Mailed check	20.65	1022.95	No
17	Bank transfer (automatic)	106.70	7382.25	No
18	Credit card (automatic)	55.20	528.35	Yes
19	Electronic check	90.05	1862.90	No
20	Electronic check	39.65	39.65	Yes
21	Bank transfer (automatic)	19.80	202.25	No
22	Mailed check	20.15	20.15	Yes
23	Credit card (automatic)	59.90	3505.10	No
24	Credit card (automatic)	59.60	2970.30	No
25	Bank transfer (automatic)	55.30	1530.60	No
26	Electronic check	99.35	4749.15	Yes
27	Electronic check	30.20	30.20	Yes
28	Credit card (automatic)	90.25	6369.45	No
29	Mailed check	64.70	1093.10	Yes

[30 rows x 21 columns]

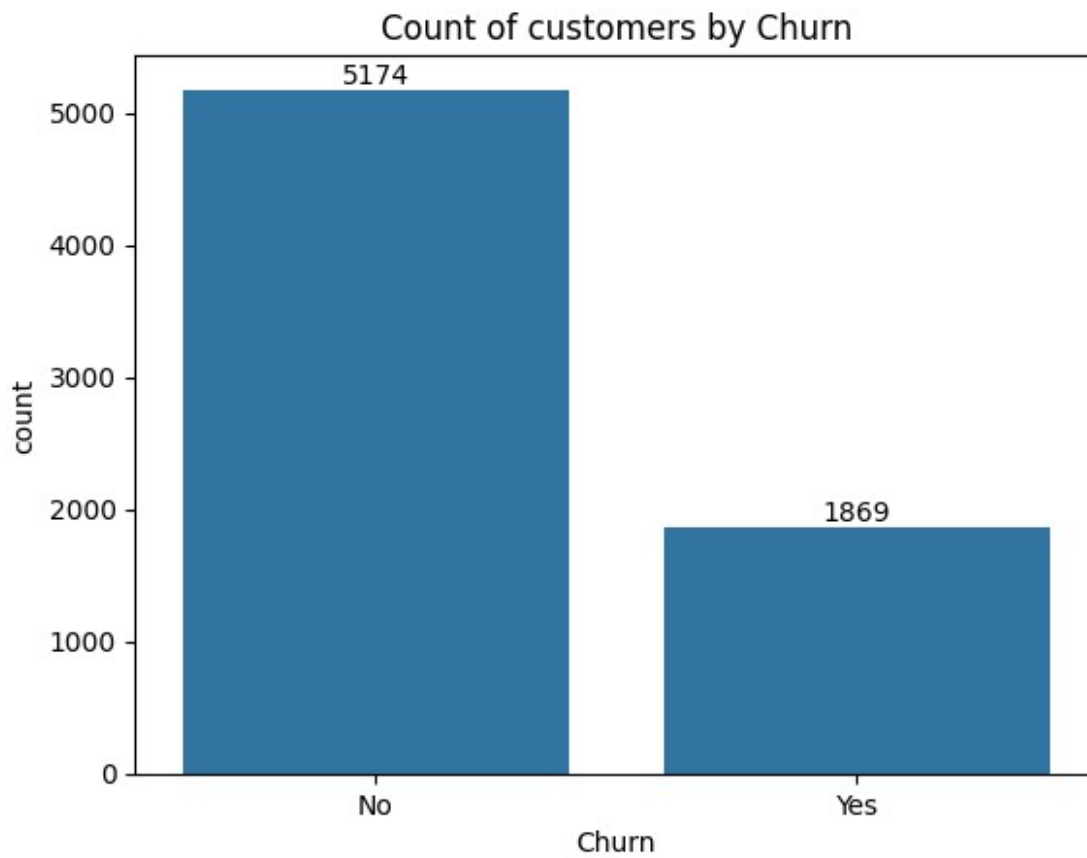
df.info()

```

<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   object
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   float64
20  Churn                  7043 non-null   object
dtypes: float64(2), int64(1), object(18)
memory usage: 1.1+ MB

ax = sns.countplot(x = df['Churn'], data = df)
ax.bar_label(ax.containers[0])
plt.title("Count of customers by Churn")
plt.show()

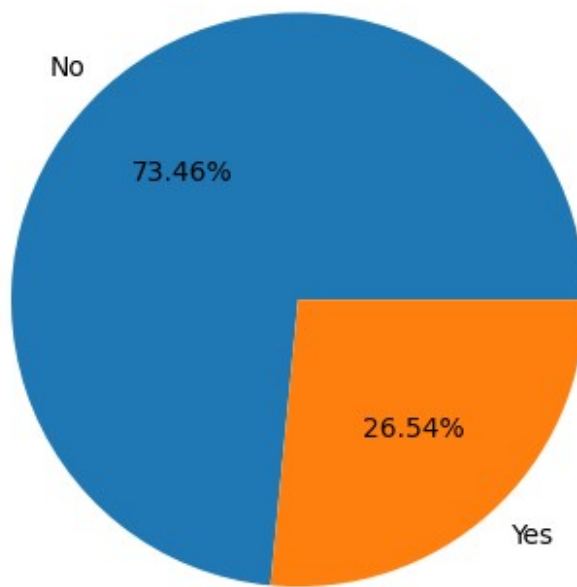
```



describing through pie chart and percentage involved

```
gb = df.groupby("Churn").agg({'Churn': "count"})
gb
plt.pie(gb['Churn'], labels = gb.index, autopct = "%1.2f%%")
plt.title("Percentage of Churned Customers")
plt.show()
```

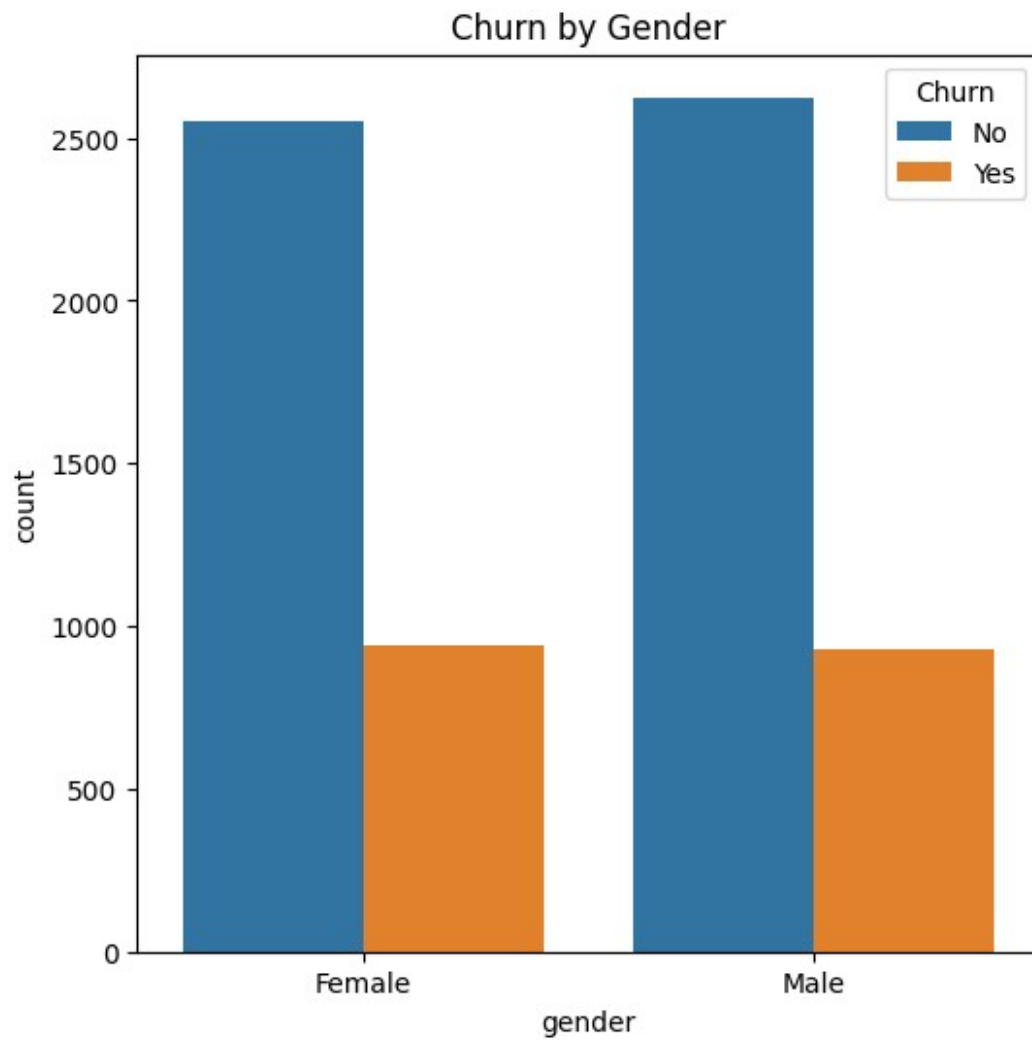

Percentage of Churned Customers



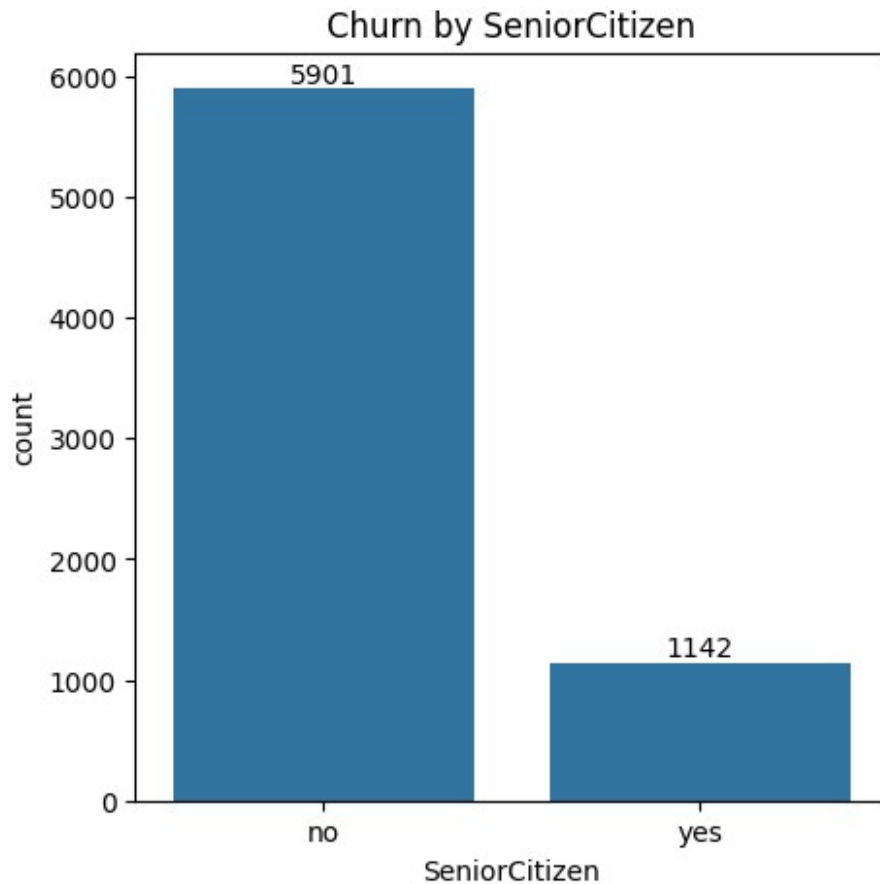
#from the given pie chart we can conclude that 26.54% of our customers have churned out

#now lets explore the reason behind it

```
plt.figure(figsize= (6,6))  
sns.countplot(x = "gender", data = df, hue = "Churn")  
plt.title("Churn by Gender")  
plt.show()
```



```
plt.figure(figsize= (5,5))
ax = sns.countplot(x = "SeniorCitizen", data = df)
ax.bar_label(ax.containers[0])
plt.title("Churn by SeniorCitizen")
plt.show()
```



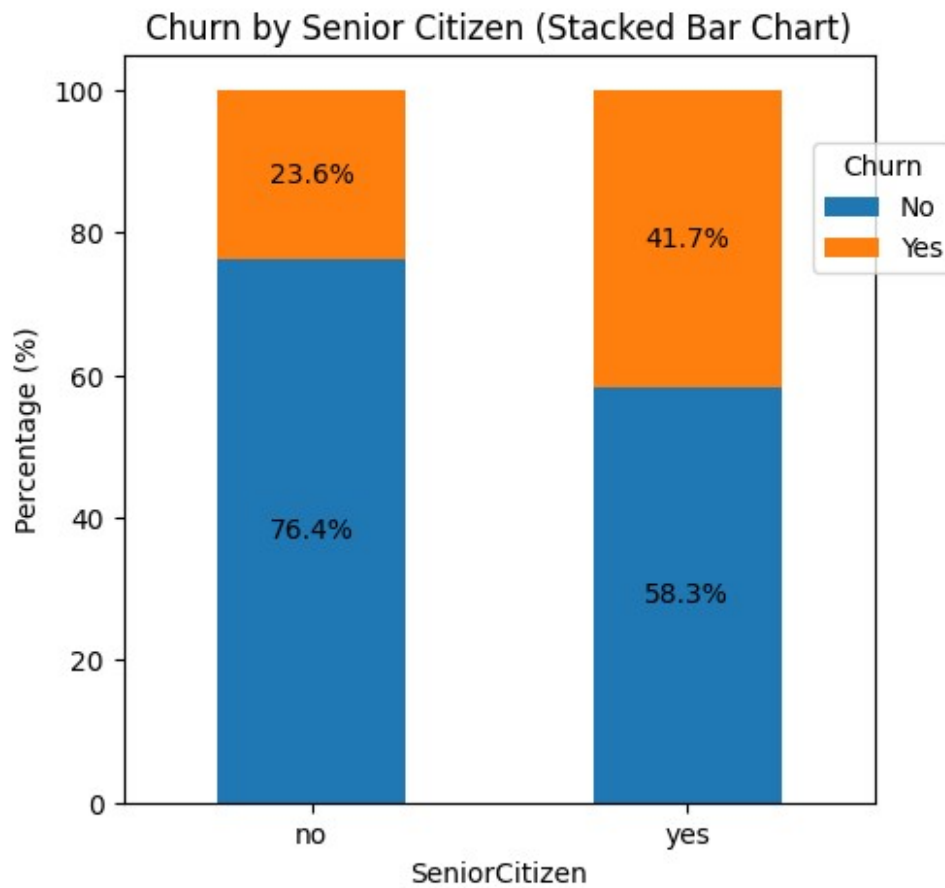
#comparatively a greated percentage of people in senior citizen category have churned

```
total_counts = df.groupby('SeniorCitizen')
['Churn'].value_counts(normalize=True).unstack() * 100

fig, ax = plt.subplots(figsize=(5,5))

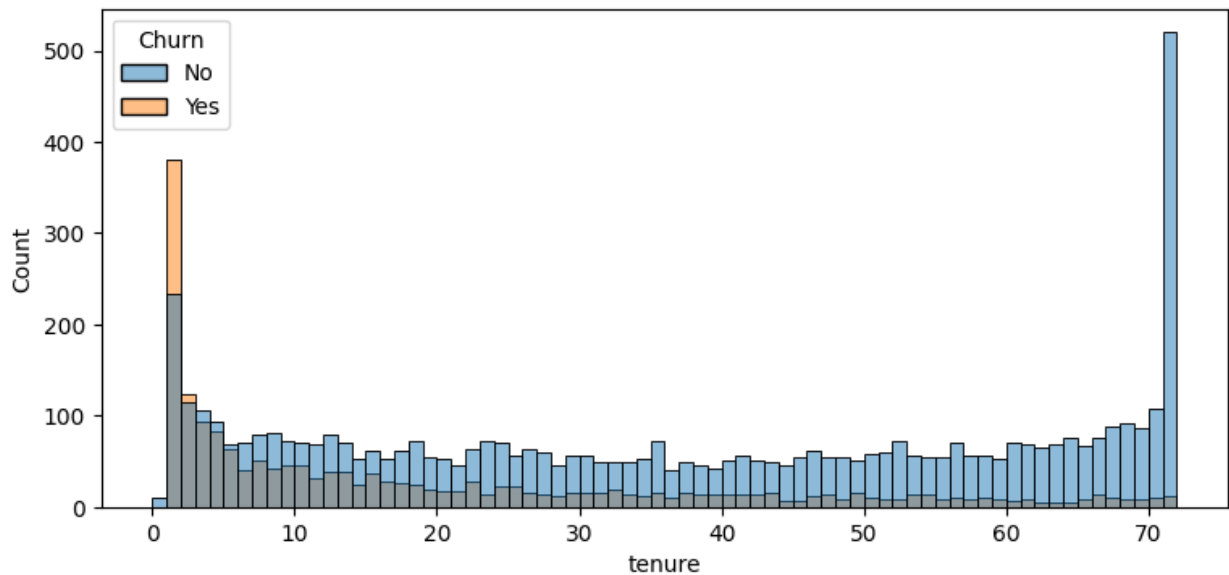
total_counts.plot(kind='bar', stacked=True, ax=ax,
color=['#1f77b4', '#ff7f0e'])
for p in ax.patches:
    width, height = p.get_width(), p.get_height()
    x,y = p.get_xy()
    ax.text(x + width/2, y + height / 2, f'{height:.1f}%',
ha='center', va='center')

plt.title('Churn by Senior Citizen (Stacked Bar Chart)')
plt.xlabel('SeniorCitizen')
plt.ylabel('Percentage (%)')
plt.xticks(rotation=0)
plt.legend(title="Churn", bbox_to_anchor=(0.9,0.9))
plt.show()
```



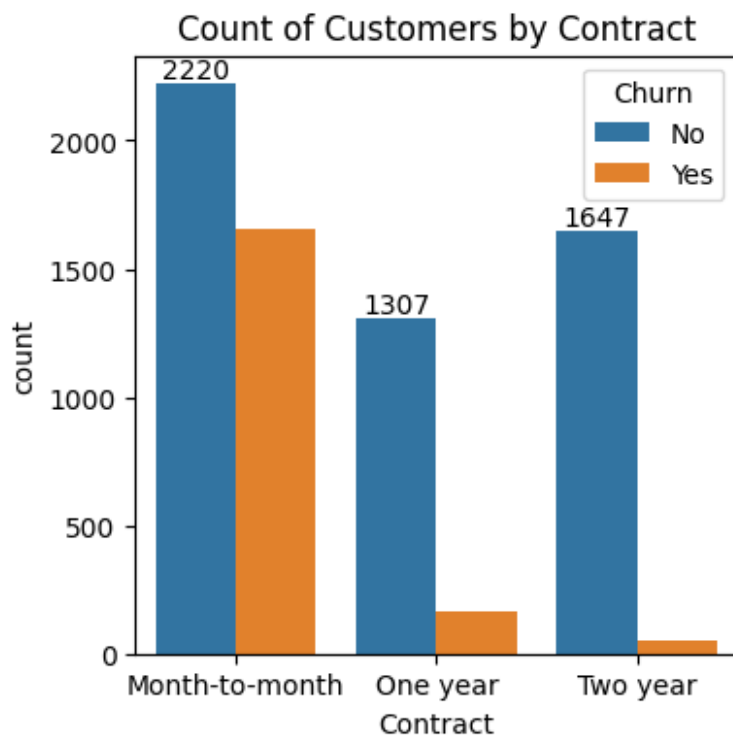
#comparative a greater percentage of people in senior citizen category have churned|

```
plt.figure(figsize = (9,4))
sns.histplot(x = "tenure", data = df, bins=72, hue= "Churn")
plt.show()
```



#people who have used our services for a long time have stayed and people who have used our services #1 or 2 months have churned

```
plt.figure(figsize=(4,4))
ax = sns.countplot(x= "Contract", data=df, hue="Churn")
ax.bar_label(ax.containers[0])
plt.title("Count of Customers by Contract")
plt.show()
```



#people who have month to month contract are likely to churn then from those who have 1 or 2 years of contract

```
df.columns.values
array(['customerID', 'gender', 'SeniorCitizen', 'Partner',
      'Dependents',
      'tenure', 'PhoneService', 'MultipleLines', 'InternetService',
      'OnlineSecurity', 'OnlineBackup', 'DeviceProtection',
      'TechSupport', 'StreamingTV', 'StreamingMovies', 'Contract',
      'PaperlessBilling', 'PaymentMethod', 'MonthlyCharges',
      'TotalCharges', 'Churn'], dtype=object)

columns = ['PhoneService', 'MultipleLines', 'InternetService',
          'OnlineSecurity',
          'OnlineBackup', 'DeviceProtection', 'TechSupport',
          'StreamingTV', 'StreamingMovies']

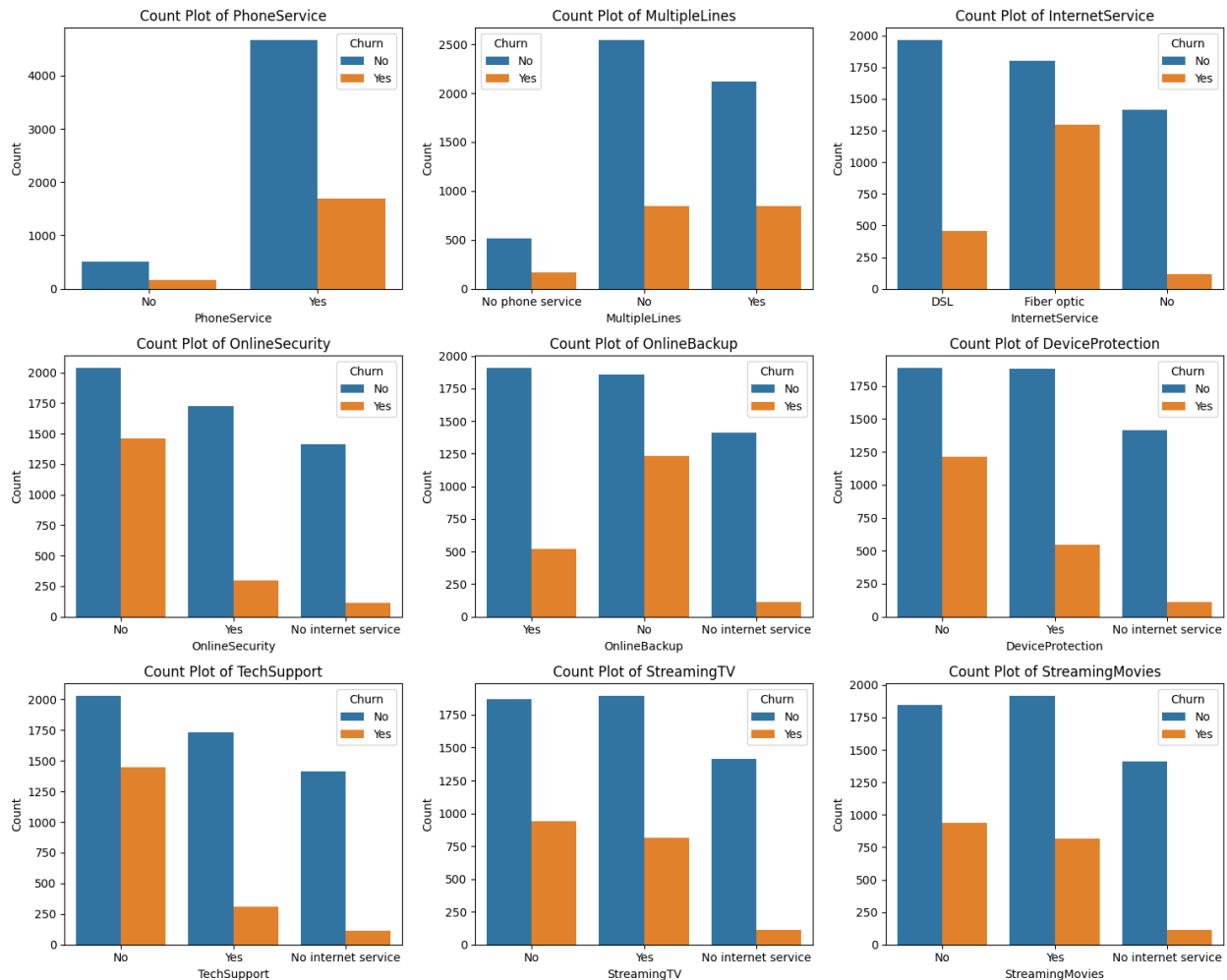
n_cols = 3
n_rows = (len(columns) + n_cols - 1) // n_cols

fig, axes = plt.subplots(n_rows, n_cols, figsize=(15, n_rows * 4))
axes = axes.flatten()

for i, col in enumerate(columns):
    sns.countplot(x=col, data=df, ax=axes[i], hue = df["Churn"])
    axes[i].set_title(f'Count Plot of {col}')
    axes[i].set_xlabel(col)
    axes[i].set_ylabel('Count')

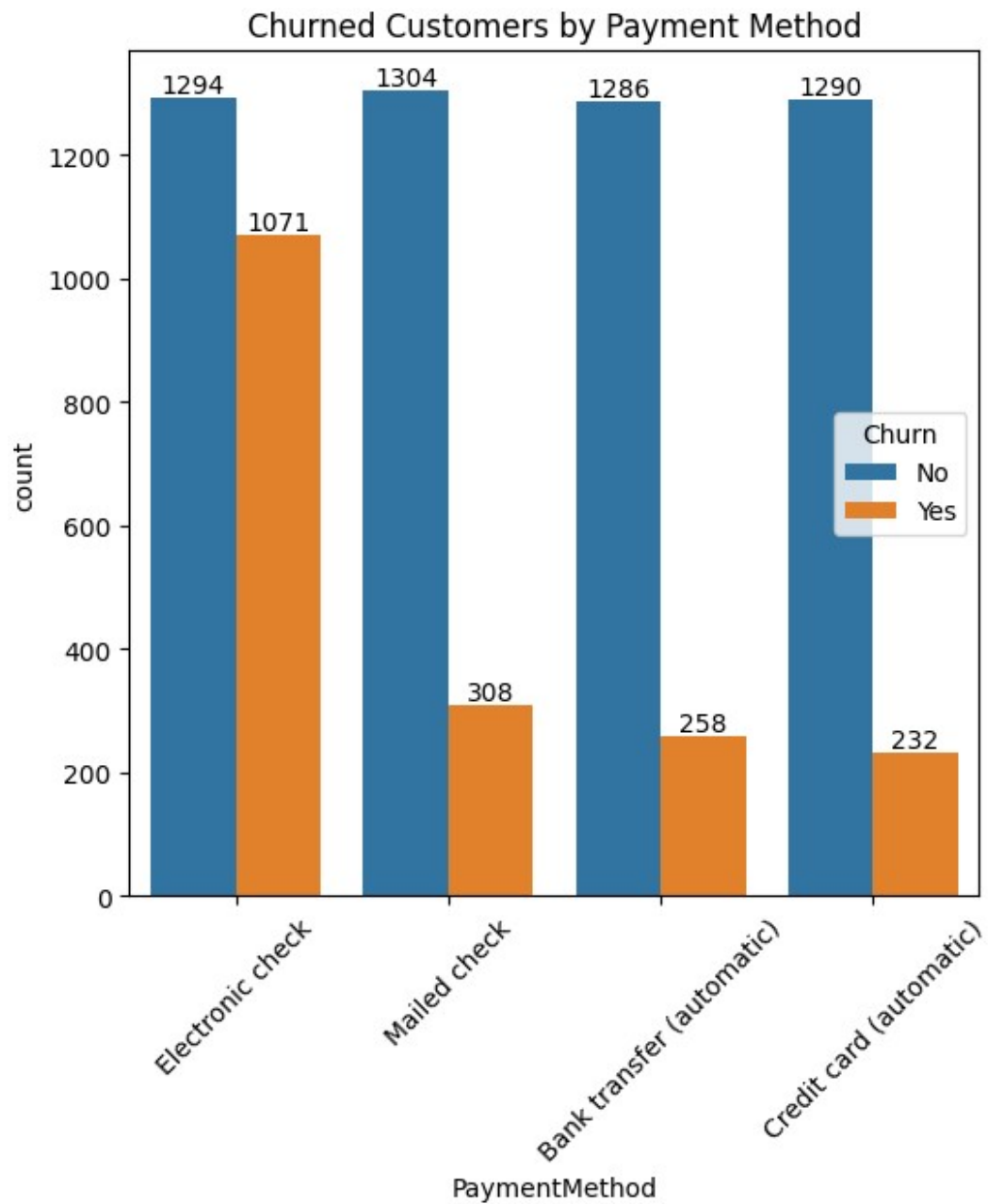
# Agar kuch empty subplots bach jaayein to unhe hata do
for j in range(i + 1, len(axes)):
    fig.delaxes(axes[j])

plt.tight_layout()
plt.show()
```



This visualization shows that customers without additional services like OnlineSecurity, TechSupport, or DeviceProtection tend to have a higher churn rate compared to those who use these services. Those with Fiber optic InternetService also exhibit more churn than DSL users. Overall, value-added services reduce churn, indicating that customers engaged with more bundled features are more likely to stay.

```
plt.figure(figsize=(6,6))
ax = sns.countplot(x = "PaymentMethod", data = df, hue = "Churn")
ax.bar_label(ax.containers[0])
ax.bar_label(ax.containers[1])
plt.title("Churned Customers by Payment Method")
plt.xticks(rotation = 45)
plt.show()
```



#Customer is likely to Churn When he is using electronic check as payment method