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

data=pd.read_csv("Customer Churn.csv")

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]
```

```
# Rrtreive first 10 rows
data.head(10)
```

```
   customerID  gender  SeniorCitizen Partner Dependents  tenure
PhoneService  \
0  7590-VHVEG  Female              0     Yes         No       1
No
1  5575-GNVDE    Male              0      No         No      34
Yes
2  3668-QPYBK    Male              0      No         No       2
Yes
3  7795-CFOCW    Male              0      No         No      45
No
4  9237-HQITU  Female              0      No         No       2
Yes
5  9305-CDSKC  Female              0      No         No       8
Yes
6  1452-KIOVK    Male              0      No        Yes      22
Yes
7  6713-OKOMC  Female              0      No         No      10
No
8  7892-POOKP  Female              0     Yes         No      28
Yes
9  6388-TABGU    Male              0      No        Yes      62
Yes

      MultipleLines InternetService OnlineSecurity  ...
DeviceProtection  \
0  No phone service             DSL             No  ...
No
1                No             DSL            Yes  ...
Yes
2                No             DSL            Yes  ...
No
3  No phone service             DSL            Yes  ...
Yes
4                No     Fiber optic             No  ...
```

```
No
5               Yes     Fiber optic             No  ...
Yes
6               Yes     Fiber optic             No  ...
No
7  No phone service             DSL            Yes  ...
No
8               Yes     Fiber optic             No  ...
Yes
9                No             DSL            Yes  ...
No

  TechSupport StreamingTV StreamingMovies        Contract
PaperlessBilling  \
0          No          No              No  Month-to-month
Yes
1          No          No              No        One year
No
2          No          No              No  Month-to-month
Yes
3         Yes          No              No        One year
No
4          No          No              No  Month-to-month
Yes
5          No         Yes             Yes  Month-to-month
Yes
6          No         Yes              No  Month-to-month
Yes
7          No          No              No  Month-to-month
No
8         Yes         Yes             Yes  Month-to-month
Yes
9          No          No              No        One year
No

                PaymentMethod MonthlyCharges  TotalCharges Churn
0           Electronic check          29.85         29.85    No
1               Mailed check          56.95        1889.5    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         820.5   Yes
6    Credit card (automatic)          89.10        1949.4    No
7               Mailed check          29.75         301.9    No
8           Electronic check         104.80       3046.05   Yes
9  Bank transfer (automatic)          56.15       3487.95    No

[10 rows x 21 columns]
```

```
# Check how many null value persent in our dataset.
data.isnull()
```

```
      customerID  gender  SeniorCitizen  Partner  Dependents
tenure  \
0          False   False          False    False       False   False

1          False   False          False    False       False   False

2          False   False          False    False       False   False

3          False   False          False    False       False   False

4          False   False          False    False       False   False

...          ...     ...            ...      ...         ...     ...

7038       False   False          False    False       False   False

7039       False   False          False    False       False   False

7040       False   False          False    False       False   False

7041       False   False          False    False       False   False

7042       False   False          False    False       False   False


      PhoneService  MultipleLines  InternetService
OnlineSecurity  ...  \
0            False          False            False
False  ...
1            False          False            False
False  ...
2            False          False            False
False  ...
3            False          False            False
False  ...
4            False          False            False
False  ...
...            ...            ...              ...             ...  ..
.
7038         False          False            False
False  ...
7039         False          False            False
False  ...
7040         False          False            False
False  ...
7041         False          False            False
False  ...
7042         False          False            False
```

```
False  ...

      DeviceProtection  TechSupport  StreamingTV  StreamingMovies
Contract  \
0                False        False        False            False
False
1                False        False        False            False
False
2                False        False        False            False
False
3                False        False        False            False
False
4                False        False        False            False
False
...                ...          ...          ...              ...
...
7038             False        False        False            False
False
7039             False        False        False            False
False
7040             False        False        False            False
False
7041             False        False        False            False
False
7042             False        False        False            False
False

      PaperlessBilling  PaymentMethod  MonthlyCharges  TotalCharges
Churn
0                False          False           False         False
False
1                False          False           False         False
False
2                False          False           False         False
False
3                False          False           False         False
False
4                False          False           False         False
False
...                ...            ...             ...           ...
...
7038             False          False           False         False
False
7039             False          False           False         False
False
7040             False          False           False         False
False
7041             False          False           False         False
False
```

```
7042              False          False           False         False
False

[7043 rows x 21 columns]
```

```
data.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
```

```
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
```

```
data.describe()
```

```
       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
```

```
data.shape
```

```
(7043, 21)
```

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

```
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
```

```
# Check duplicate value in datset
data.duplicated().sum()
```

```
0
```

```
# Chech duplicate  based on the unique columns.
data["customerID"].duplicated().sum()
```

```
0
```

```
# In this code we change value of senior citizen if 0 (NO) and  1(Yes)

def convert(value):
    if value==1:
        return "yes"
    else:
        return "No"

data["SeniorCitizen"]=data["SeniorCitizen"].apply(convert)
```

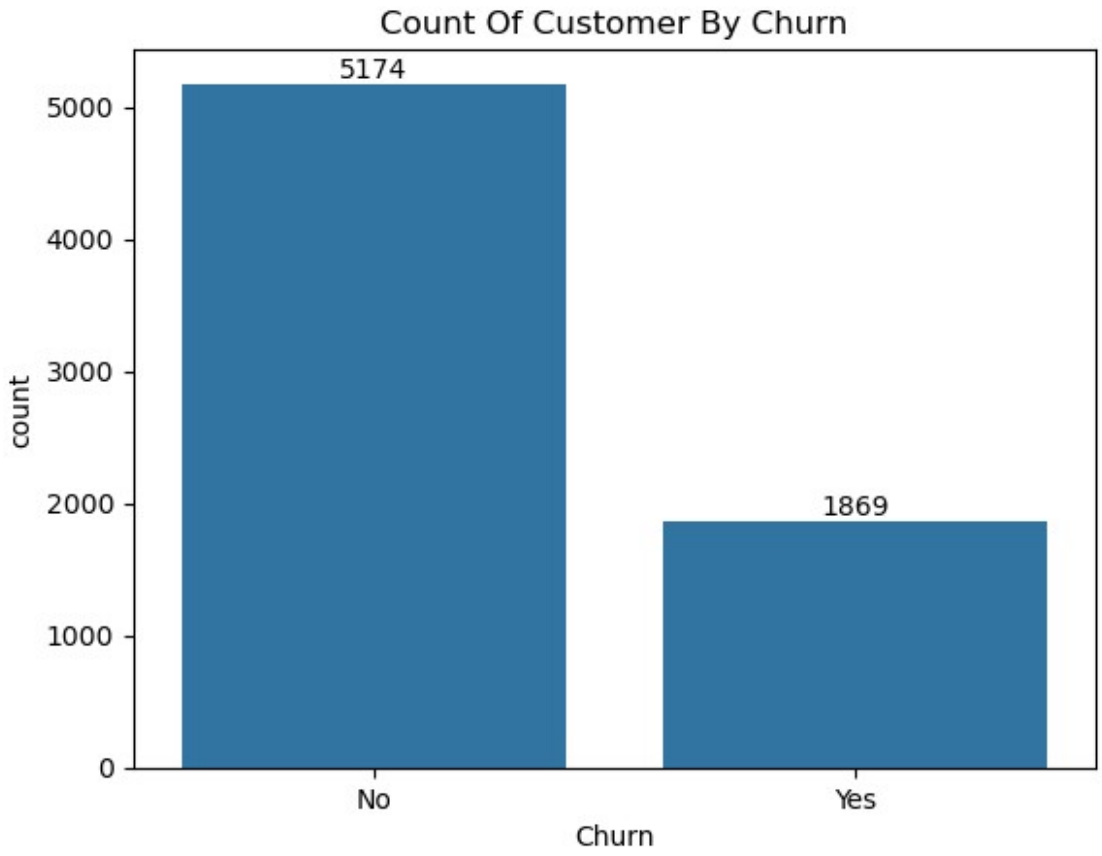
```
data
```

```
      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-CFOCW    Male            No      No         No      45
4     9237-HQITU  Female            No      No         No       2
...          ...     ...           ...     ...        ...     ...
7038  6840-RESVB    Male            No     Yes        Yes      24
7039  2234-XADUH  Female            No     Yes        Yes      72
7040  4801-JZAZL  Female            No     Yes        Yes      11
7041  8361-LTMKD    Male           yes     Yes         No       4
7042  3186-AJIEK    Male            No      No         No      66

     PhoneService     MultipleLines InternetService
OnlineSecurity  ...  \
0              No  No phone service             DSL
No  ...
1             Yes                No             DSL
Yes  ...
```

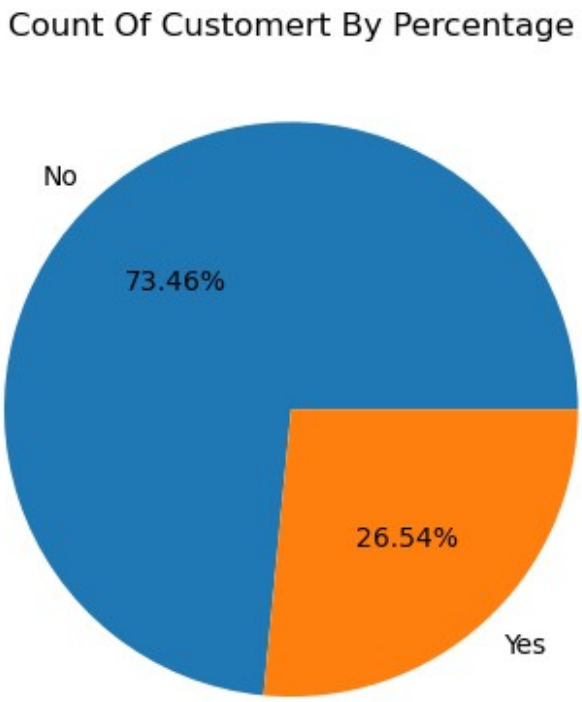
```
2             Yes                No             DSL
Yes  ...
3              No  No phone service             DSL
Yes  ...
4             Yes                No     Fiber optic
No  ...
...           ...               ...             ...             ...  ..
.
7038          Yes               Yes             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]
```

```
ax=sns.countplot(x = "Churn",data=data)

ax.bar_label(ax.containers[0])
plt.title("Count Of Customer By Churn")
plt.show()
```

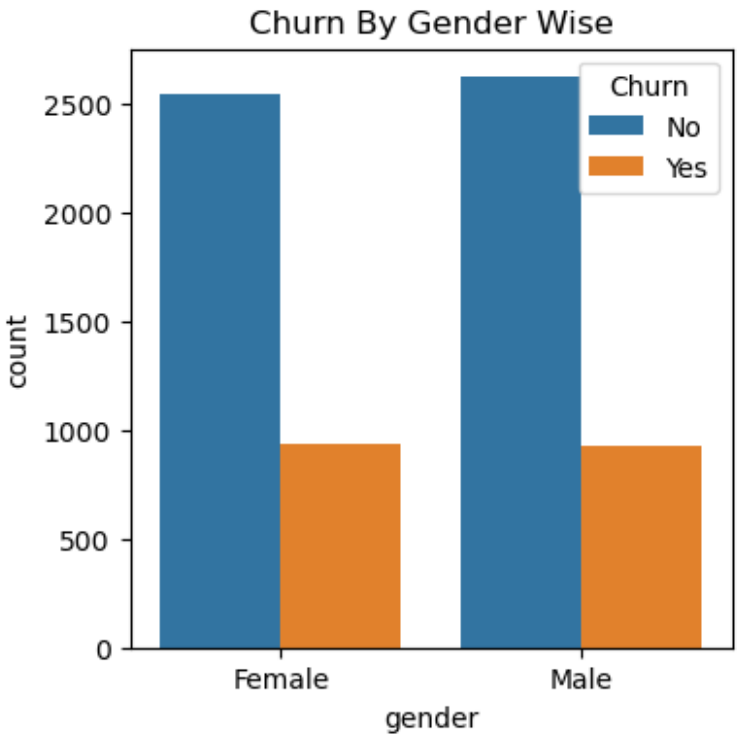



```
gb =data.groupby("Churn").agg({"Churn":"count"})
plt.pie(gb["Churn"],labels=gb.index ,autopct="%.2f%%")
plt.title("Count Of Customert By Percentage")
plt.show()
```



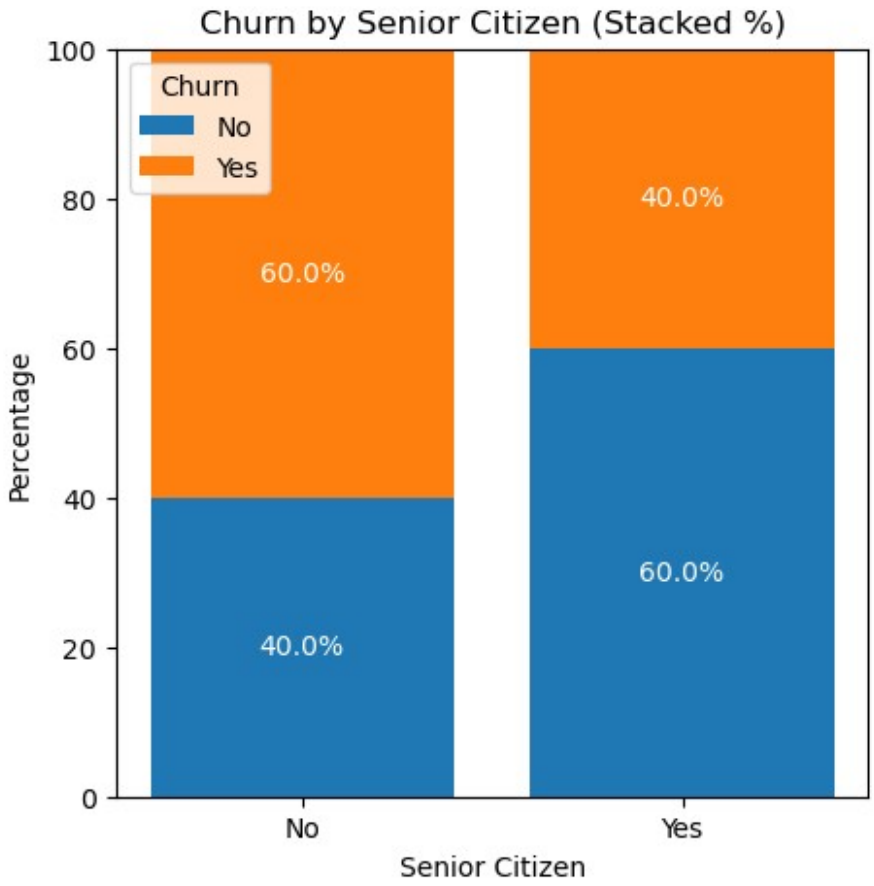
From the give pie chart we conclued 26.54 percentage of our customer have churned out.

```
plt.figure(figsize =(4,4))
sns.countplot(x="gender",data=data,hue="Churn")
plt.title("Churn By Gender Wise")
plt.show()
```

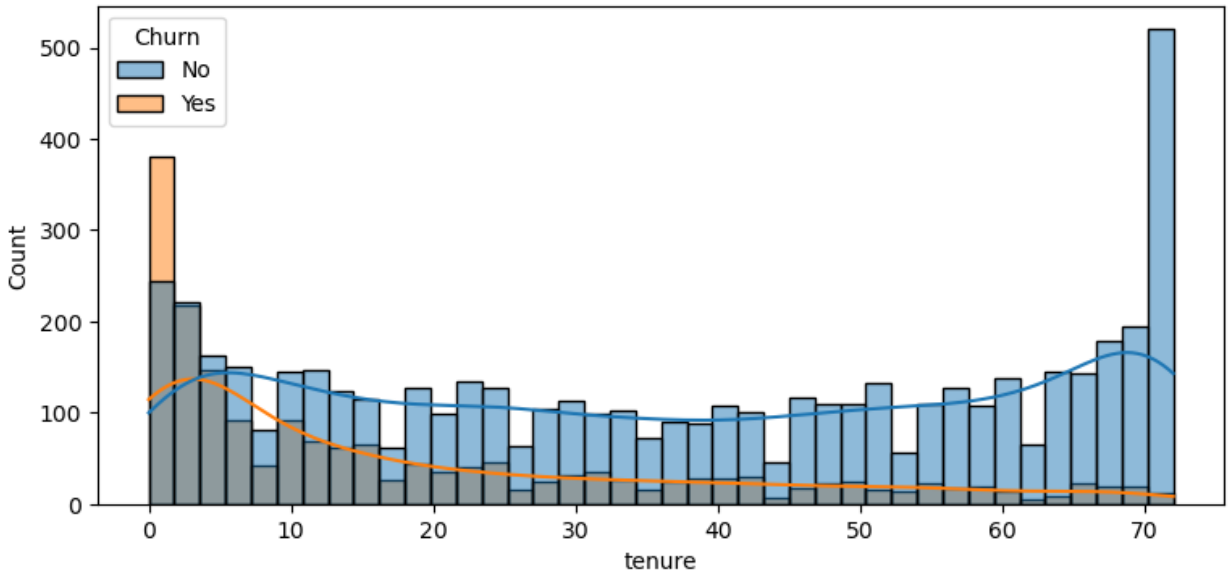



```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Sample Data
data = pd.DataFrame({
    "SeniorCitizen": ["Yes", "No", "Yes", "No", "Yes", "No", "Yes",
"No", "Yes", "No"],
    "Churn": ["Yes", "No", "No", "Yes", "No", "No", "Yes", "Yes",
"No", "Yes"]
})

# Create a cross-tabulation (counts)
ct = pd.crosstab(data["SeniorCitizen"], data["Churn"])

# Convert counts to percentages
ct_percent = ct.div(ct.sum(axis=1), axis=0) * 100

# Plot settings
fig, ax = plt.subplots(figsize=(5, 5))

# Define bar positions
bottom_values = [0] * len(ct_percent)  # Initial bottom positions

# Plot stacked bars
for churn_status in ct_percent.columns:
    bars = plt.bar(ct_percent.index, ct_percent[churn_status],
```

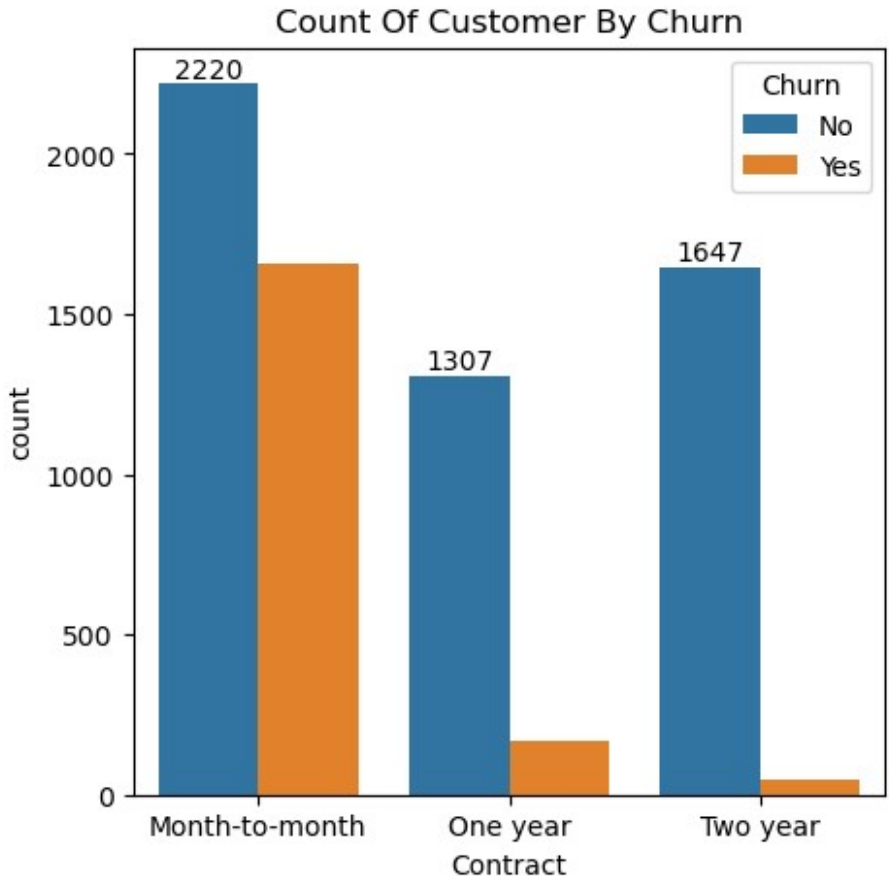
```
bottom=bottom_values, label=churn_status)

    # Add percentage labels
    for bar in bars:
        height = bar.get_height()
        if height > 0:
            plt.text(
                bar.get_x() + bar.get_width() / 2,  # X Position
                bar.get_y() + height / 2,           # Y Position
                f"{height:.1f}%",                    # Label as
percentage
                ha="center", va="center", fontsize=10, color="white"
            )

    # Update bottom values for stacking
    bottom_values += ct_percent[churn_status]

# Title & Label

plt.title("Churn by Senior Citizen (Stacked %)")
plt.xlabel("Senior Citizen")
plt.ylabel("Percentage")
plt.legend(title="Churn")
plt.ylim(0, 100)  # Limit Y-axis to 100%
plt.show()
```




```
plt.figure(figsize=(9,4))
sns.histplot(x="tenure", data=data, bins=40, kde=True,
color="blue",hue="Churn")
plt.show()
```

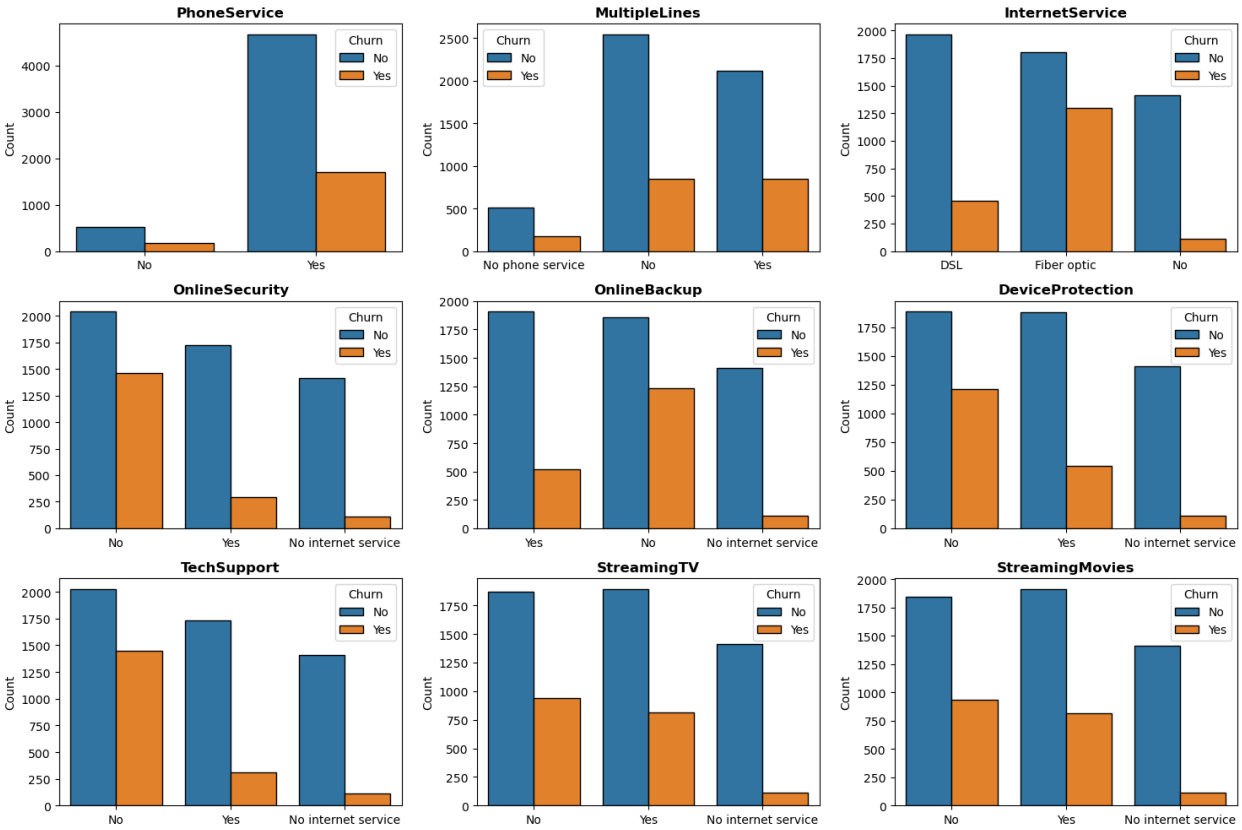


```
plt.figure(figsize=(5,5))
ax=sns.countplot(x = "Contract",data=data,hue="Churn")

ax.bar_label(ax.containers[0])
plt.title("Count Of Customer By Churn")

plt.show()
```



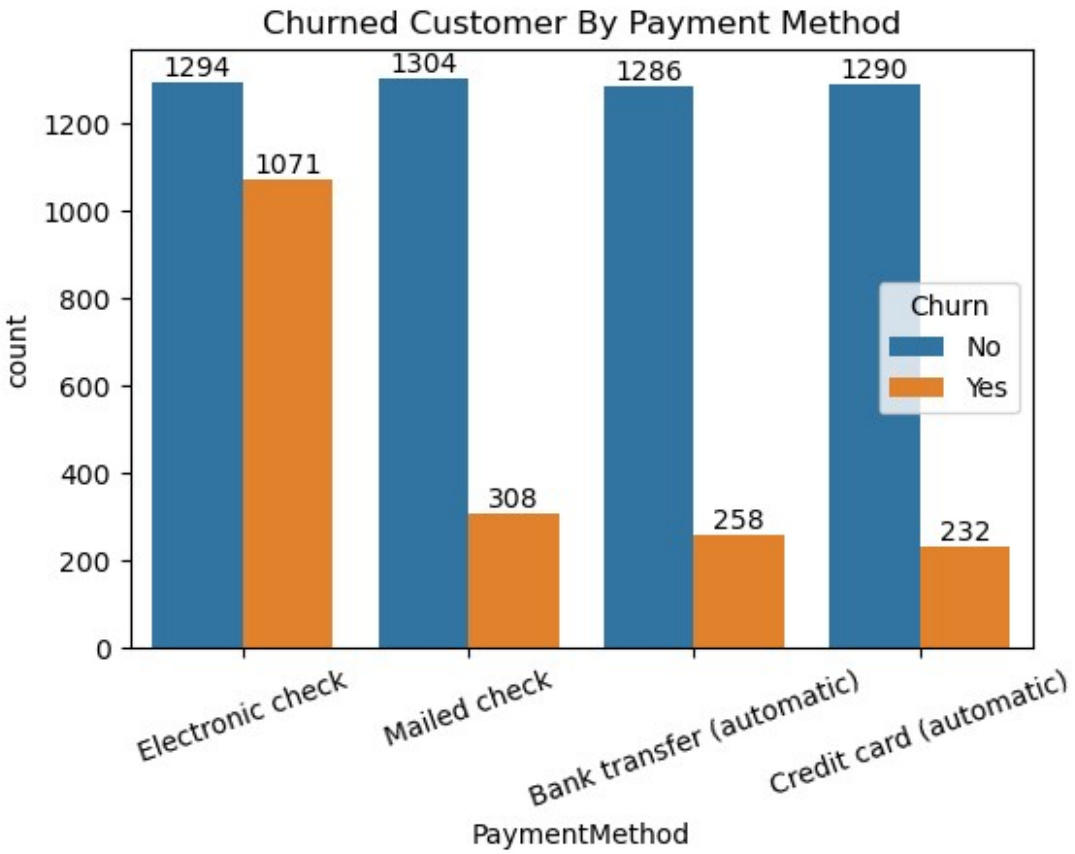

```
data.columns.values

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'
```

```
]

# Set up subplots
num_cols = 3  # Number of columns in subplot grid
num_rows = -(-len(columns) // num_cols)  # Calculate required rows

fig, axes = plt.subplots(num_rows, num_cols, figsize=(15, 10))
axes = axes.flatten()  # Flatten the 2D array of axes to 1D for easy
iteration

# Loop through columns and create count plots
for i, col in enumerate(columns):
    sns.countplot(data=data, x=col, ax=axes[i],
edgecolor="black",hue=data["Churn"])
    axes[i].set_title(col, fontsize=12, fontweight="bold")
    axes[i].set_xlabel("")
    axes[i].set_ylabel("Count")
    axes[i].tick_params(axis='x')  # Rotate x labels if needed


# Remove any empty subplots (if the number of columns is not a perfect
multiple of num_cols)
for j in range(i + 1, len(axes)):
    fig.delaxes(axes[j])

# Adjust layout
plt.tight_layout()
plt.show()
```



Customers without OnlineSecurity, OnlineBackup, and TechSupport have a higher churn rate. Fiber optic users churn more than DSL users. PhoneService doesn't show a strong churn impact, but having additional services seems to reduce churn.

```
plt.figure(figsize=(6,4))
ax=sns.countplot(x = "PaymentMethod",data=data,hue="Churn")

ax.bar_label(ax.containers[0])
ax.bar_label(ax.containers[1])
plt.title("Churned Customer By Payment Method")
plt.xticks(rotation=20)

plt.show()
```


Churned Customer By Payment Method
1294
1071
1304
308
1286
258
1290
232
1200
1000
800
600
400
200
0
count
Churn
No
Yes
Electronic check
Mailed check
Bank transfer (automatic)
Credit card (automatic)
PaymentMethod