In [1]:
```
!pip install matplotlib seaborn --quiet
```

```
[notice] A new release of pip is available: 24.0 -> 25.2
[notice] To update, run: C:\Users\jeevan\AppData\Local\Microsoft\WindowsApps\PythonSoftwareFoundation.Python.3.11_qbz5n2kfra8p0
\python.exe -m pip install --upgrade pip
```

In [2]:
```
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.head()
```

Out[2]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	Dev
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 rows × 21 columns

In [3]:
```
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
```

In [4]:
```
df["TotalCharges"] = df["TotalCharges"].replace(" ","0")
df["TotalCharges"] = df["TotalCharges"].astype("float")
```

In [5]:
```
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
```

In [6]:
```
df.isnull().sum().sum()
```

Out[6]:
```
np.int64(0)
```

In [7]:
```
df.describe()
```

Out[7]:

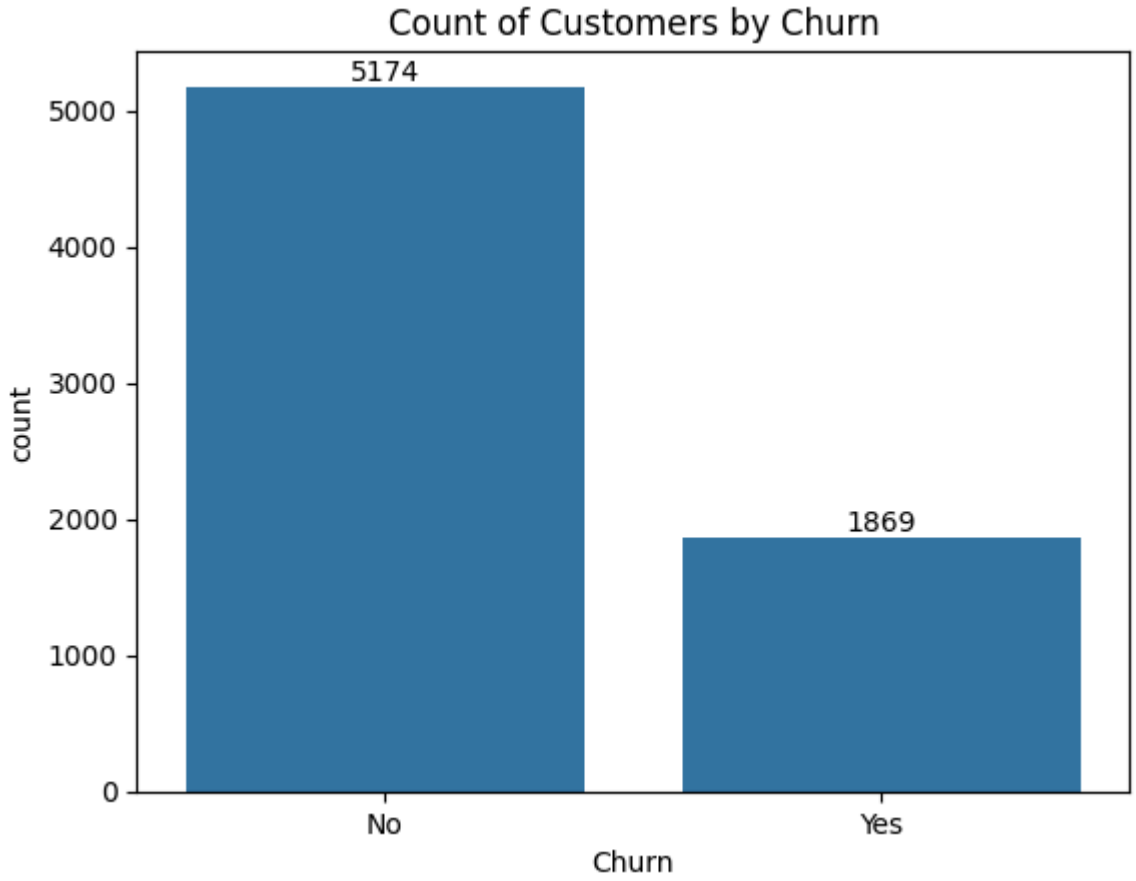
	SeniorCitizen	tenure	MonthlyCharges	TotalCharges
count	7043.000000	7043.000000	7043.000000	7043.000000
mean	0.162147	32.371149	64.761692	2279.734304
std	0.368612	24.559481	30.090047	2266.794470
min	0.000000	0.000000	18.250000	0.000000
25%	0.000000	9.000000	35.500000	398.550000
50%	0.000000	29.000000	70.350000	1394.550000
75%	0.000000	55.000000	89.850000	3786.600000
max	1.000000	72.000000	118.750000	8684.800000

In [8]:
```
df["customerID"].duplicated().sum()
```

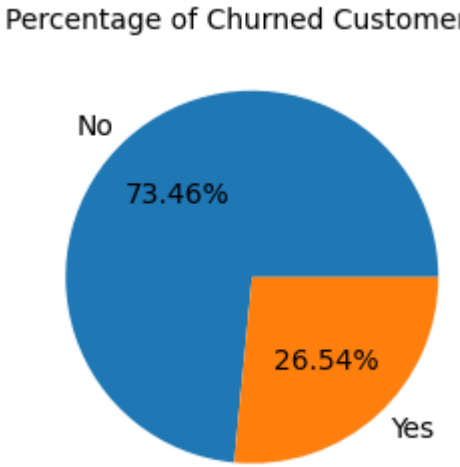
Out[8]:
```
np.int64(0)
```

In [9]:
```
# converted 0 and 1 values of senior citizen to yes/no to make it easier to understand
def conv(value):
    if value == 1:
        return "yes"
    else:
        return "no"

df['SeniorCitizen'] = df["SeniorCitizen"].apply(conv)
```

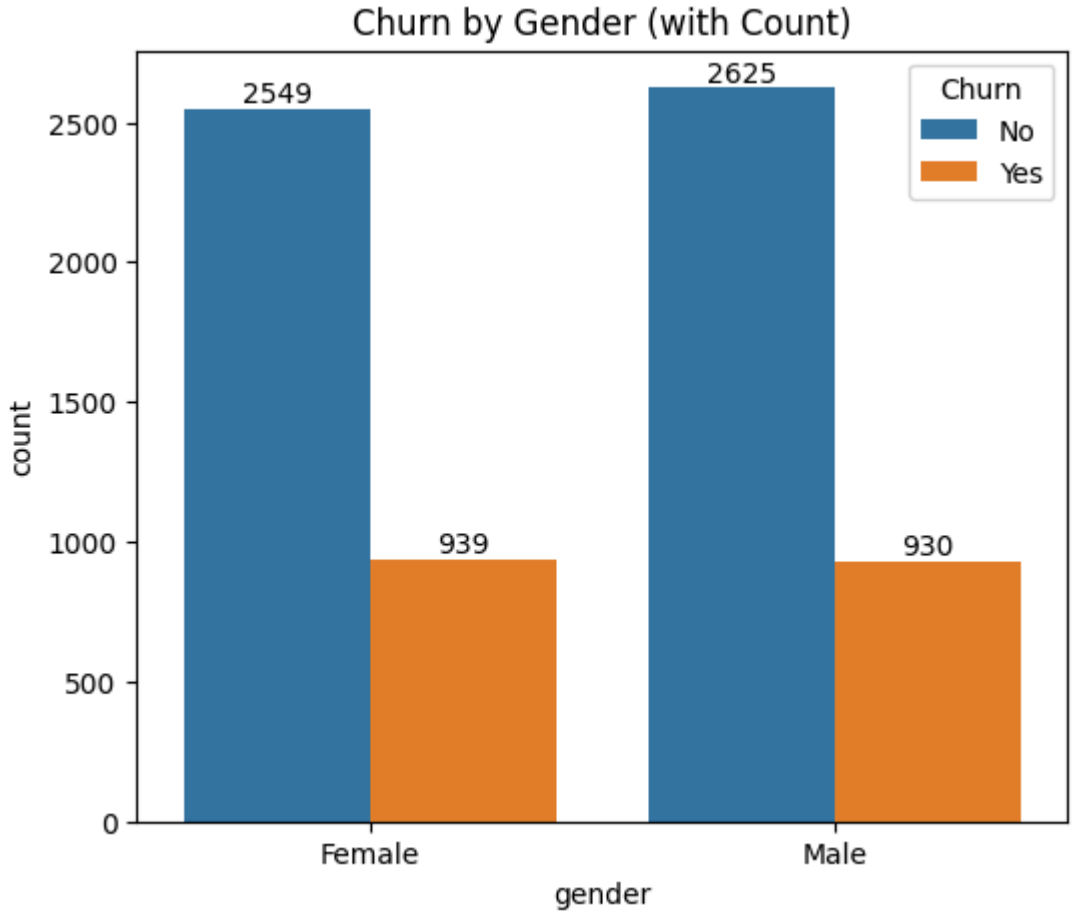
In [10]:
```
# Pie chart showing exact number of customers who churned vs. stayed
ax = sns.countplot(x = 'Churn', data = df)
ax.bar_label(ax.containers[0])
plt.title("Count of Customers by Churn")
plt.show()
```



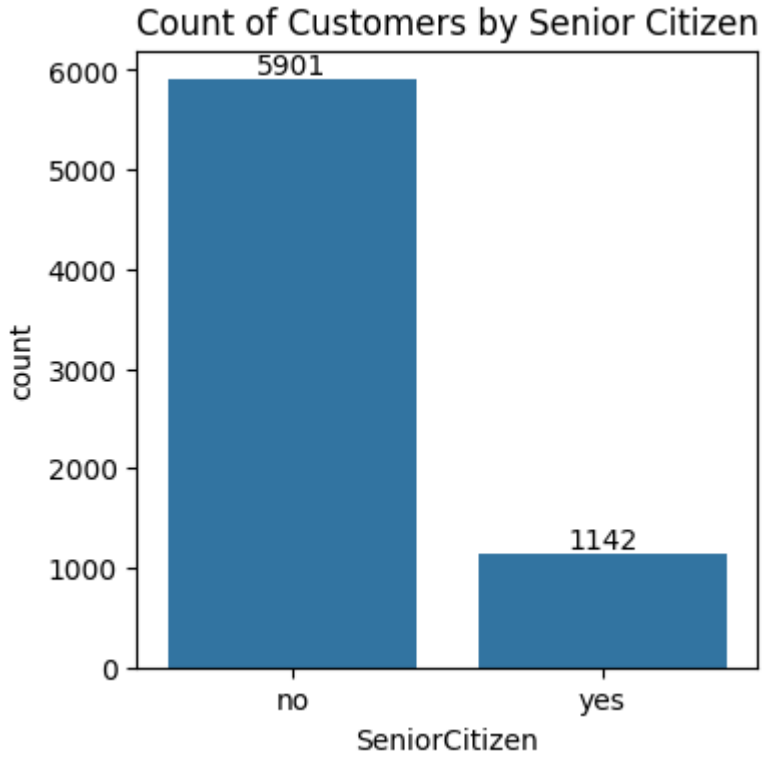
In [11]:
```
#from the given pie chart we can conclude that 26.54% of our customers have churned out.
plt.figure(figsize = (3,4))
gb = df.groupby("Churn").agg({'Churn':"count"})
plt.pie(gb['Churn'], labels = gb.index, autopct = "%1.2f%%")
plt.title("Percentage of Churned Customeres", fontsize = 10)
plt.show()
```



In [12]:
```
# Around 28% of females and 26% of males churned — not a big difference, so gender likely doesn't affect churn much
plt.figure(figsize = (6,5))
ax = sns.countplot(x = "gender", data = df, hue = "Churn")
ax.bar_label(ax.containers[0])
ax.bar_label(ax.containers[1])
plt.title("Churn by Gender (with Count)")
plt.show()
```

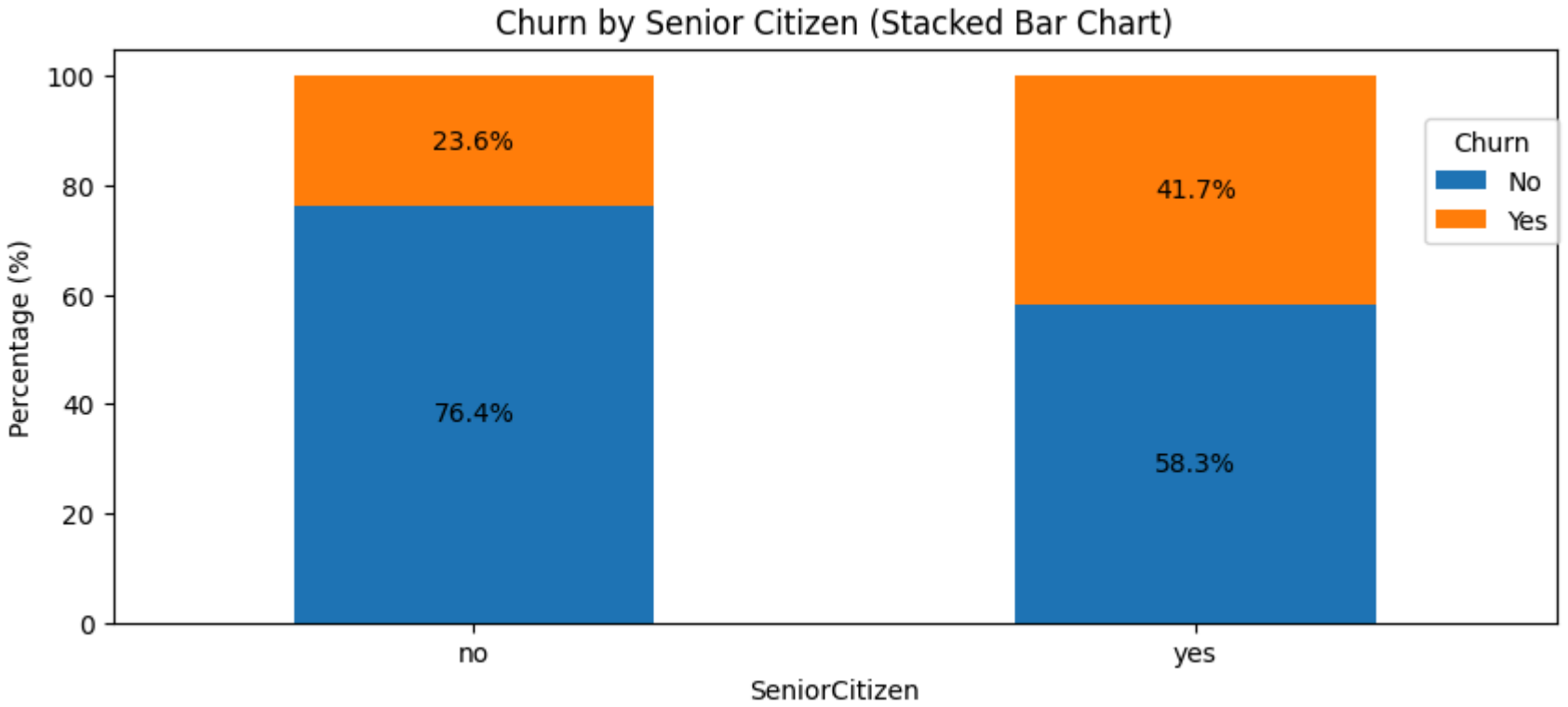


In [13]:
```
plt.figure(figsize = (4,4))
ax = sns.countplot(x = "SeniorCitizen", data = df)
ax.bar_label(ax.containers[0])
plt.title("Count of Customers by Senior Citizen")
plt.show()
```



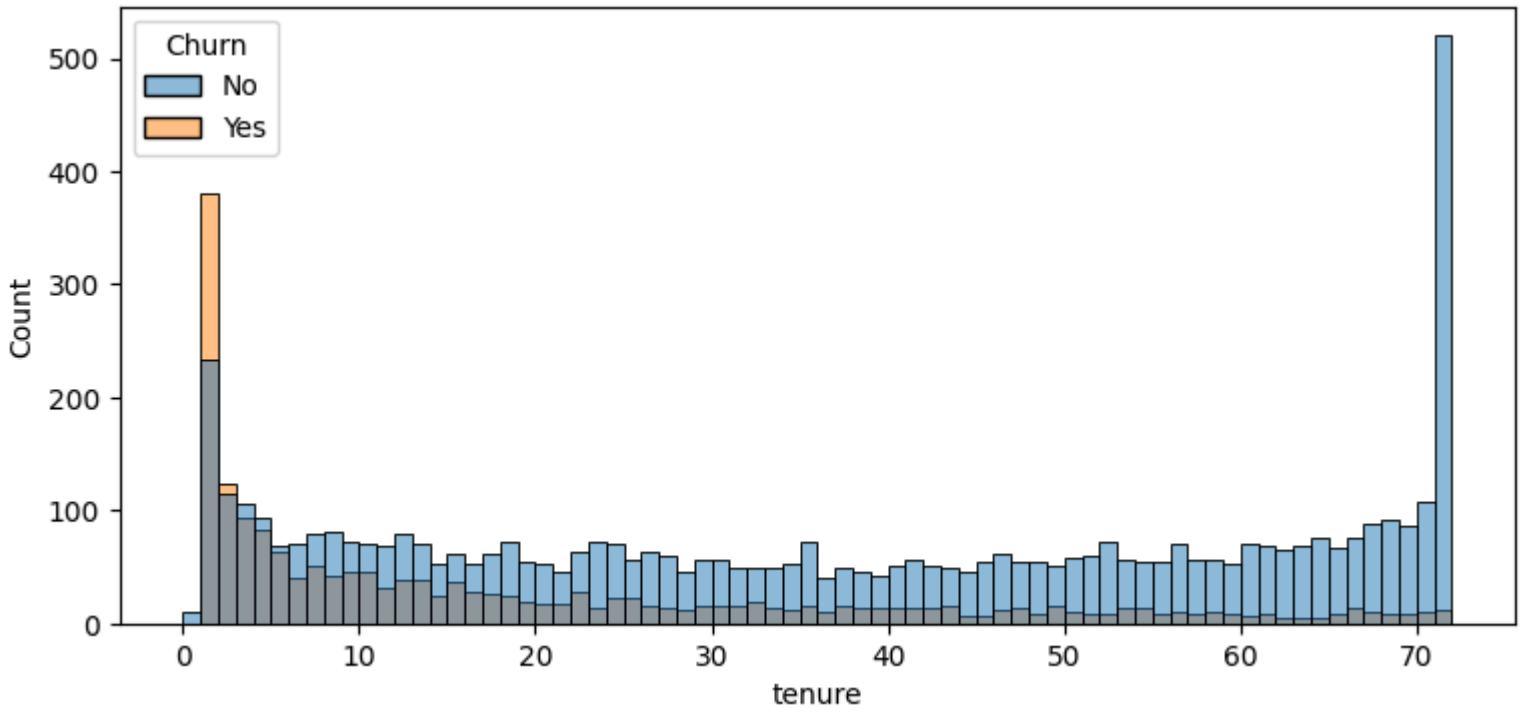
In [14]:
```
total_counts = df.groupby('SeniorCitizen')['Churn'].value_counts(normalize=True).unstack() * 100

# Plot
fig, ax = plt.subplots(figsize=(10, 4))  # Adjust figsize for better visualization

# Plot the bars
total_counts.plot(kind='bar', stacked=True, ax=ax, color=['#1f77b4', '#ff7f0e'])  # Customize colors if desired

# Add percentage labels on the bars
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))  # Customize legend location

plt.show()
```



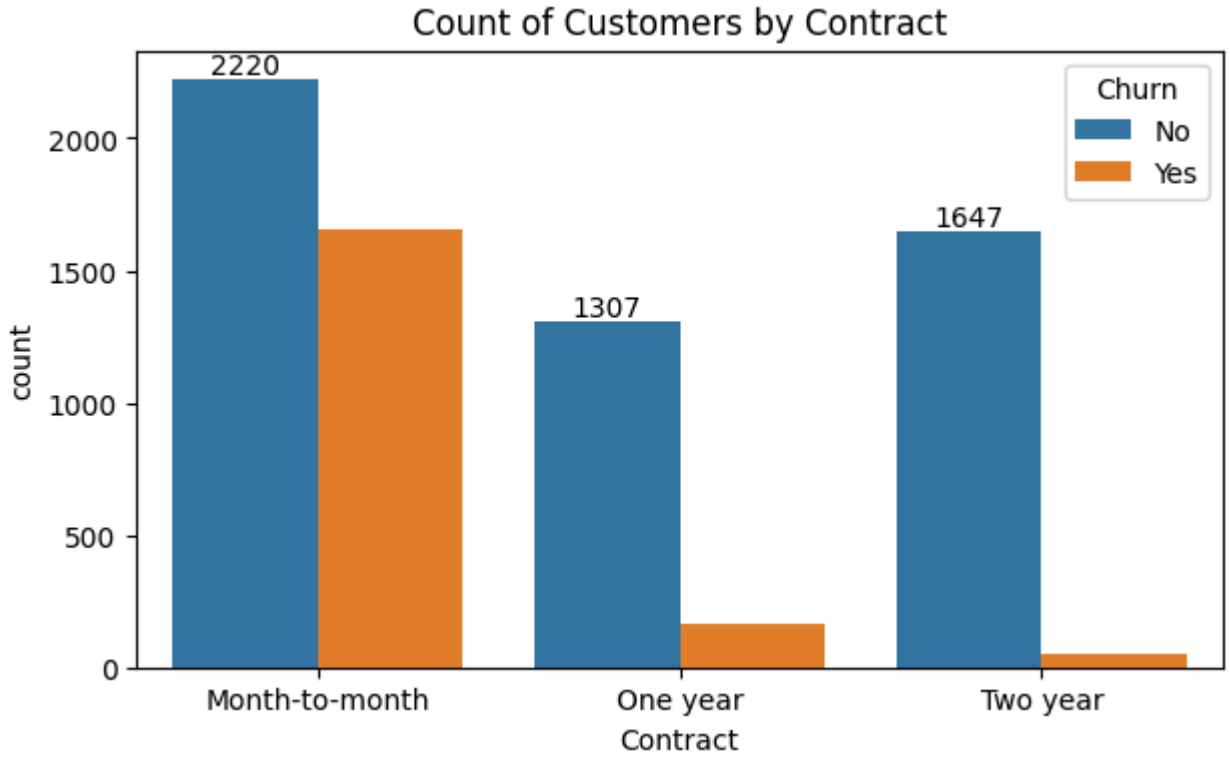
comparative a greater pecentage of people in senior citizen category have churned

In [15]:
```
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 sevices 1 or 2 months have churned

In [16]:
```
plt.figure(figsize = (7,4))
ax = sns.countplot(x = "Contract", data = df, hue = "Churn")
ax.bar_label(ax.containers[0])
plt.title("Count of Customers by Contract")
plt.show()
```



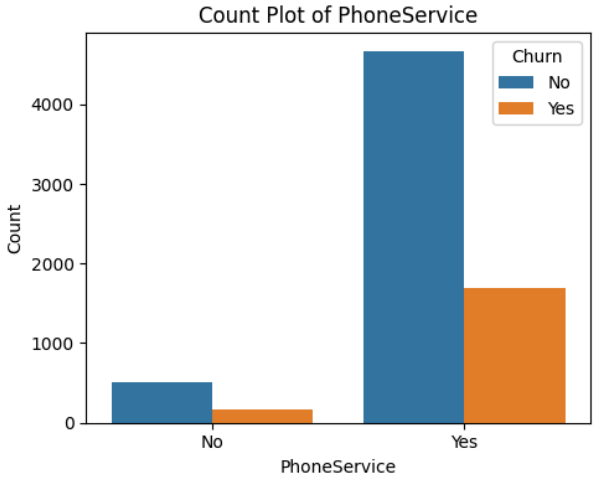
people who have month to month contract are likely to churn then from those who have 1 or 2 years or contract.

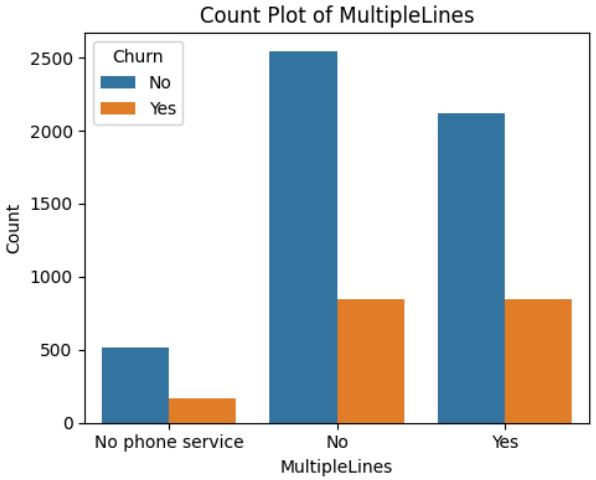
In [17]:
```
df.columns.values
```

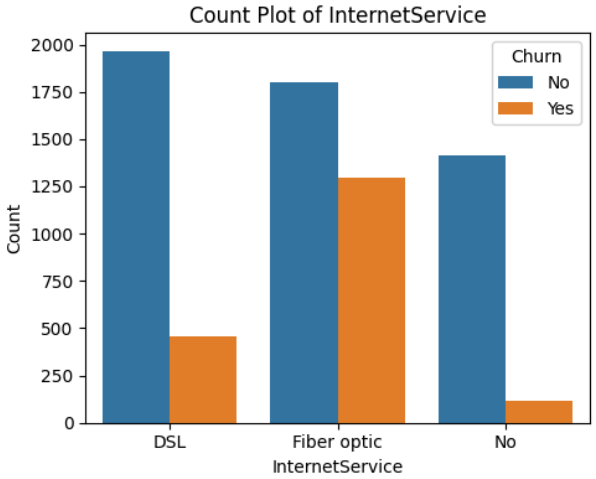
Out[17]:
```
array(['customerID', 'gender', 'SeniorCitizen', 'Partner', 'Dependents',
       'tenure', 'PhoneService', 'MultipleLines', 'InternetService',
       'OnlineSecurity', 'OnlineBackup', 'DeviceProtection',
       'TechSupport', 'StreamingTV', 'StreamingMovies', 'Contract',
       'PaperlessBilling', 'PaymentMethod', 'MonthlyCharges',
       'TotalCharges', 'Churn'], dtype=object)
```

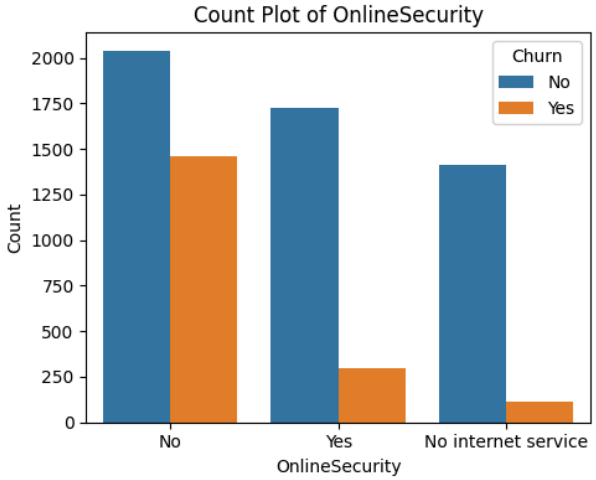
In [18]:
```
columns = ['PhoneService', 'MultipleLines', 'InternetService', 'OnlineSecurity',
           'OnlineBackup', 'DeviceProtection', 'TechSupport', 'StreamingTV', 'StreamingMovies']

# Number of columns for the subplot grid (you can change this)
n_cols = 3
```

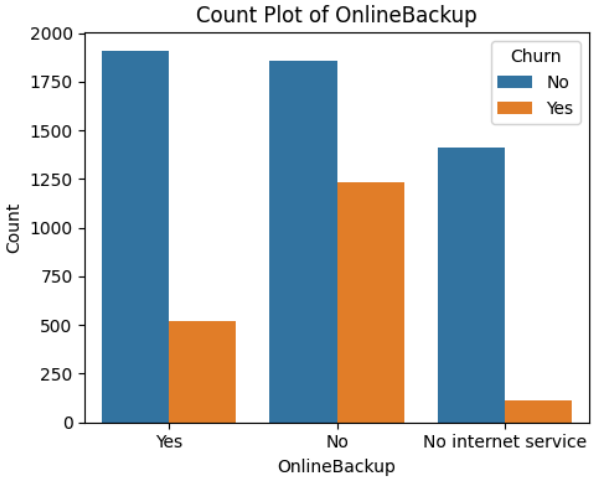
```
n_rows = (len(columns) + n_cols - 1) // n_cols  # Calculate number of rows needed

# Create subplots
fig, axes = plt.subplots(n_rows, n_cols, figsize=(15, n_rows * 4))  # Adjust figsize as needed

# Flatten the axes array for easy iteration (handles both 1D and 2D arrays)
axes = axes.flatten()

# Iterate over columns and plot count plots
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')

# Remove empty subplots (if any)
for j in range(i + 1, len(axes)):
    fig.delaxes(axes[j])

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

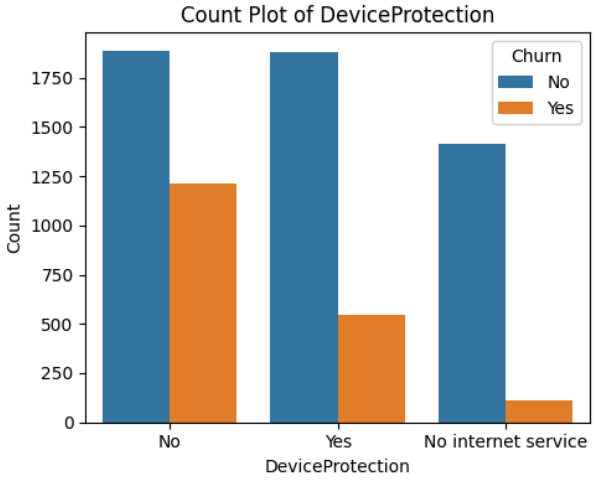

Count Plot of PhoneService
Churn
No
Yes
Count
PhoneService
No
Yes

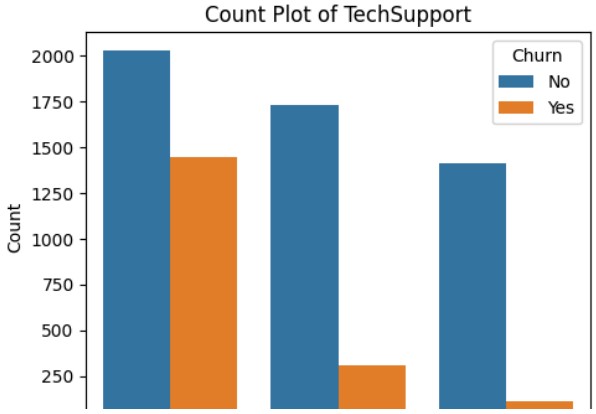

Count Plot of MultipleLines
Churn
No
Yes
Count
MultipleLines
No phone service
No
Yes

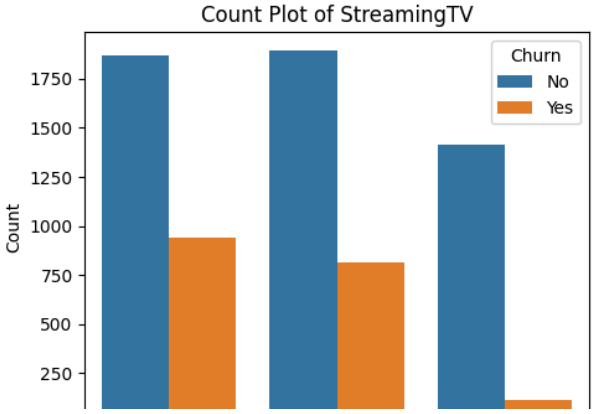

Count Plot of InternetService
Churn
No
Yes
Count
InternetService
DSL
Fiber optic
No

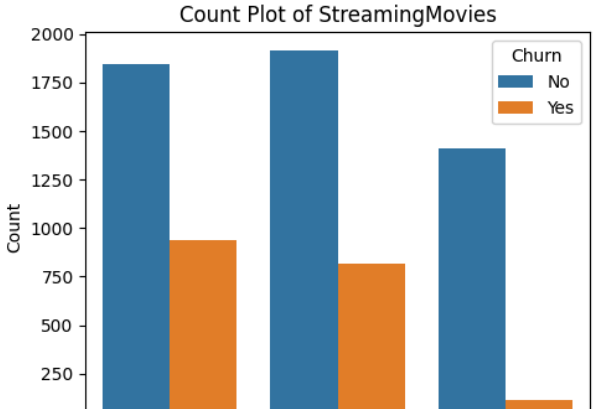

Count Plot of OnlineSecurity
Churn
No
Yes
Count
OnlineSecurity
No
Yes
No internet service

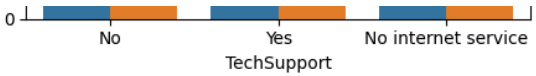

Count Plot of OnlineBackup
Churn
No
Yes
Count
OnlineBackup
Yes
No
No internet service

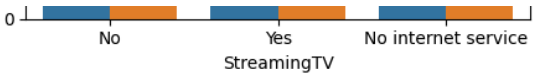

Count Plot of DeviceProtection
Churn
No
Yes
Count
DeviceProtection
No
Yes
No internet service

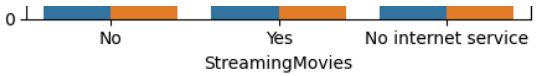

Count Plot of TechSupport
Churn
No
Yes
Count


Count Plot of StreamingTV
Churn
No
Yes
Count


Count Plot of StreamingMovies
Churn
No
Yes
Count







PhoneService

- Most customers have phone service.
- Customers with phone service show higher churn than those without.

MultipleLines

- Customers with multiple lines tend to churn more.
- Those with no phone service churn the least.

InternetService

- Fiber optic users have the highest churn rate.
- DSL users churn less.
- No internet service = very low churn.

OnlineSecurity

- Customers without online security churn much more.
- Online security helps reduce churn.

OnlineBackup

- No online backup → higher churn.
- Backup users are more loyal.

DeviceProtection

- Device protection lowers churn.
- Lack of it increases the risk of churn.

TechSupport

- Lack of tech support = high churn.
- Tech support users rarely churn.

StreamingTV

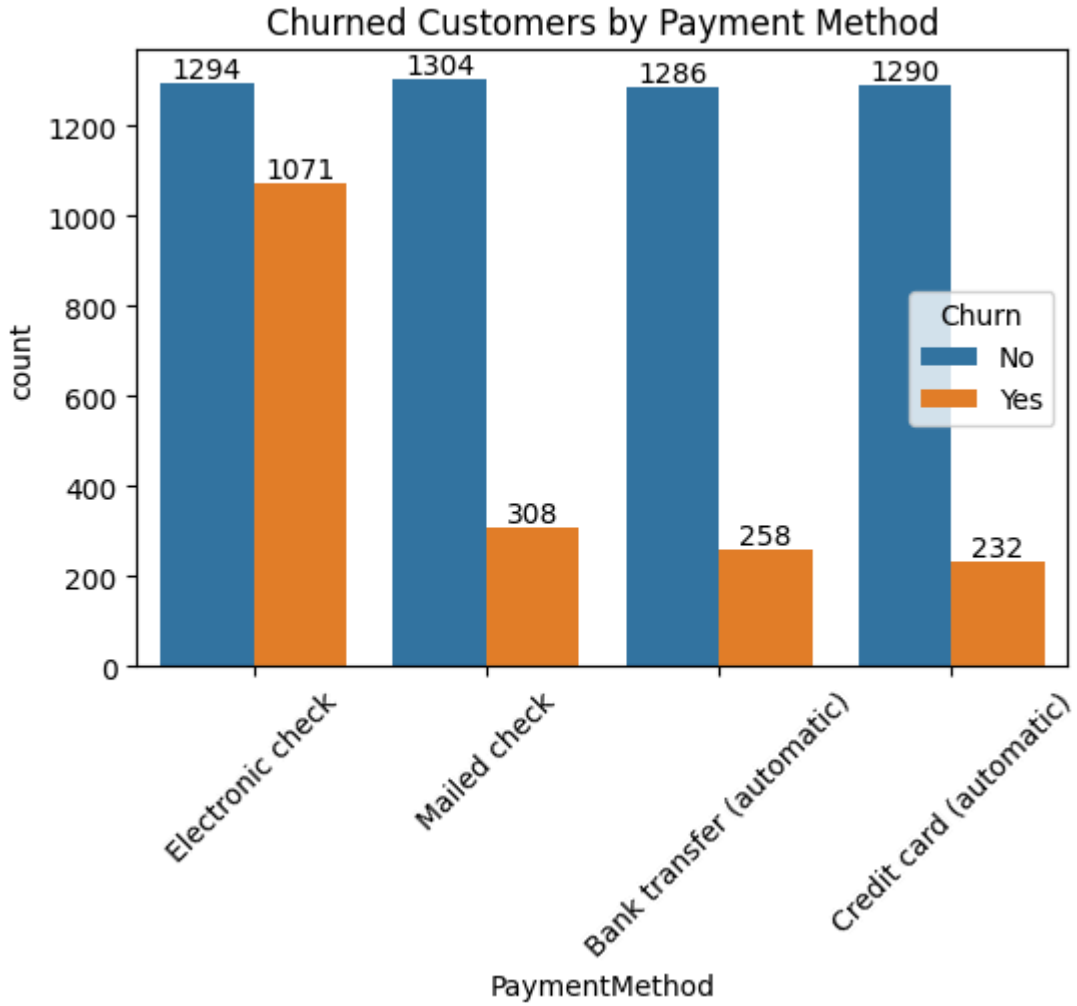
- Slightly higher churn in customers without streaming TV.
- Not a strong churn factor.

StreamingMovies

- Similar to StreamingTV — no service = slightly higher churn.

The majority of customers who do not churn tend to have services like PhoneService, InternetService (particularly DSL), and OnlineSecurity enabled. For services like OnlineBackup, TechSupport, and StreamingTV, churn rates are noticeably higher when these services are not used or are unavailable.

In [19]:
```
plt.figure(figsize = (6,4))
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 a payment method.

In [20]:
```
pip install nbconvert
```

```
Requirement already satisfied: nbconvert in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2k
fra8p0\localcache\local-packages\python311\site-packages (7.16.6)
Requirement already satisfied: beautifulsoup4 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qb
z5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (4.13.4)
Requirement already satisfied: bleach!=5.0.0 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz
5n2kfra8p0\localcache\local-packages\python311\site-packages (from bleach[css]!=5.0.0->nbconvert) (6.2.0)
Requirement already satisfied: defusedxml in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2
kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (0.7.1)
Requirement already satisfied: jinja2>=3.0 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n
2kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (3.1.6)
Requirement already satisfied: jupyter-core>=4.7 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11
_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (5.8.1)
Requirement already satisfied: jupyterlab-pygments in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.
11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (0.3.0)
Requirement already satisfied: markupsafe>=2.0 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_q
bz5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (3.0.2)
Requirement already satisfied: mistune<4,>=2.0.3 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11
_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (3.1.3)
Requirement already satisfied: nbclient>=0.5.0 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_q
bz5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (0.10.2)
Requirement already satisfied: nbformat>=5.7 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz
5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (5.10.4)
Requirement already satisfied: packaging in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2k
fra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (25.0)
Requirement already satisfied: pandocfilters>=1.4.1 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.
3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (1.5.1)
Requirement already satisfied: pygments>=2.4.1 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_q
bz5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (2.19.2)
Requirement already satisfied: traitlets>=5.1 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qb
z5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbconvert) (5.14.3)
Requirement already satisfied: webencodings in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5
n2kfra8p0\localcache\local-packages\python311\site-packages (from bleach!=5.0.0->bleach[css]!=5.0.0->nbconvert) (0.5.1)
Requirement already satisfied: tinycss2<1.5,>=1.1.0 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.
3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from bleach[css]!=5.0.0->nbconvert) (1.4.0)
Requirement already satisfied: platformdirs>=2.5 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11
_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from jupyter-core>=4.7->nbconvert) (4.3.8)
Requirement already satisfied: pywin32>=300 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5
n2kfra8p0\localcache\local-packages\python311\site-packages (from jupyter-core>=4.7->nbconvert) (311)
Requirement already satisfied: jupyter-client>=6.1.12 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.pytho
n.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbclient>=0.5.0->nbconvert) (8.6.3)
Requirement already satisfied: fastjsonschema>=2.15 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.
```

```
3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbformat>=5.7->nbconvert) (2.21.1)
Requirement already satisfied: jsonschema>=2.6 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from nbformat>=5.7->nbconvert) (4.25.0)
Requirement already satisfied: soupsieve>1.2 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from beautifulsoup4->nbconvert) (2.7)
Requirement already satisfied: typing-extensions>=4.0.0 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from beautifulsoup4->nbconvert) (4.14.1)
Requirement already satisfied: attrs>=22.2.0 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from jsonschema>=2.6->nbformat>=5.7->nbconvert) (25.3.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from jsonschema>=2.6->nbformat>=5.7->nbconvert) (2025.4.1)
Requirement already satisfied: referencing>=0.28.4 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from jsonschema>=2.6->nbformat>=5.7->nbconvert) (0.36.2)
Requirement already satisfied: rpds-py>=0.7.1 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from jsonschema>=2.6->nbformat>=5.7->nbconvert) (0.26.0)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from jupyter-client>=6.1.12->nbclient>=0.5.0->nbconvert) (2.9.0.post0)
Requirement already satisfied: pyzmq>=23.0 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from jupyter-client>=6.1.12->nbclient>=0.5.0->nbconvert) (27.0.0)
Requirement already satisfied: tornado>=6.2 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from jupyter-client>=6.1.12->nbclient>=0.5.0->nbconvert) (6.5.1)
Requirement already satisfied: six>=1.5 in c:\users\jeevan\appdata\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\python311\site-packages (from python-dateutil>=2.8.2->jupyter-client>=6.1.12->nbclient>=0.5.0->nbconvert) (1.17.0)
Note: you may need to restart the kernel to use updated packages.
[notice] A new release of pip is available: 24.0 -> 25.2
[notice] To update, run: C:\Users\jeevan\AppData\Local\Microsoft\WindowsApps\PythonSoftwareFoundation.Python.3.11_qbz5n2kfra8p0\python.exe -m pip install --upgrade pip
```

In [ ]:

In [ ]: