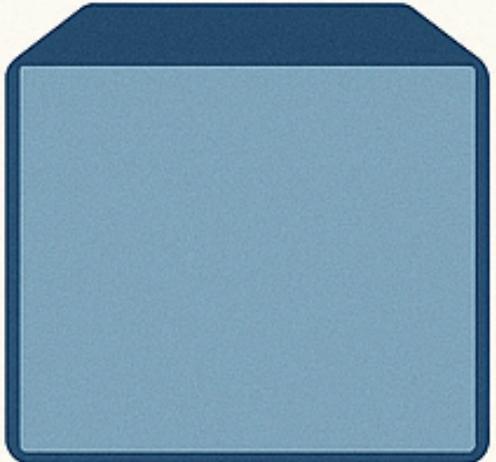


Customer Churn Analysis – Summary



Objective

Identify churn drivers in telecom users to reduce attrition.



Dataset

7,043 customers
26,5% churn rate

Customer Churn Analysis – Key Insights



1. Contract Renewal

54% of churned customers are on month-to-month contracts

2. Phone & Internet Service



- Churn rate is **43%** in Phone+Internet
- Churn rate is **24%** in Internet-only

3. Tech Support

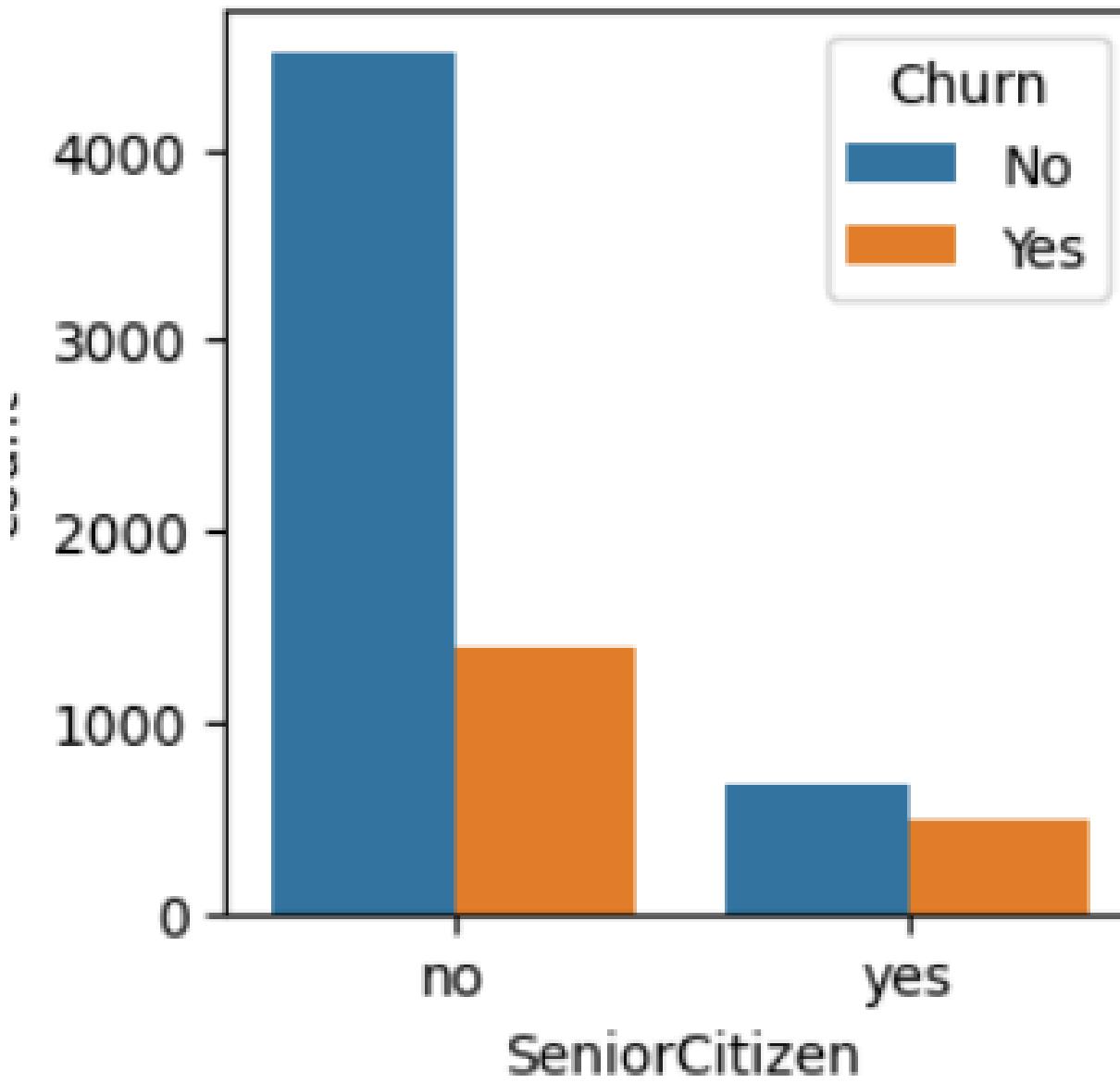


- Customers with tech support have lower churn

Check for Duplicated

```
df.duplicated().sum()  
np.int64(0)  
  
# Checking customerid again as customer id should be unique  
df["customerID"].duplicated().sum()  
  
np.int64(0)  
  
def conv(value):  
    if value ==1:  
        return "yes"  
    else:  
        return "no"  
  
df['SeniorCitizen'] = df["SeniorCitizen"].apply(conv)  
  
# Converted 0 and 1 values of senior citizen to yes/No to make it easier to understand  
df.head()  
  
plt.figure(figsize=(3,3))  
plt.title("Churn Rate based on Senior Citizen")  
sns.countplot(x = "SeniorCitizen" , data = df, hue = "Churn")  
plt.show()
```

Churn Rate based on Senior Citizen

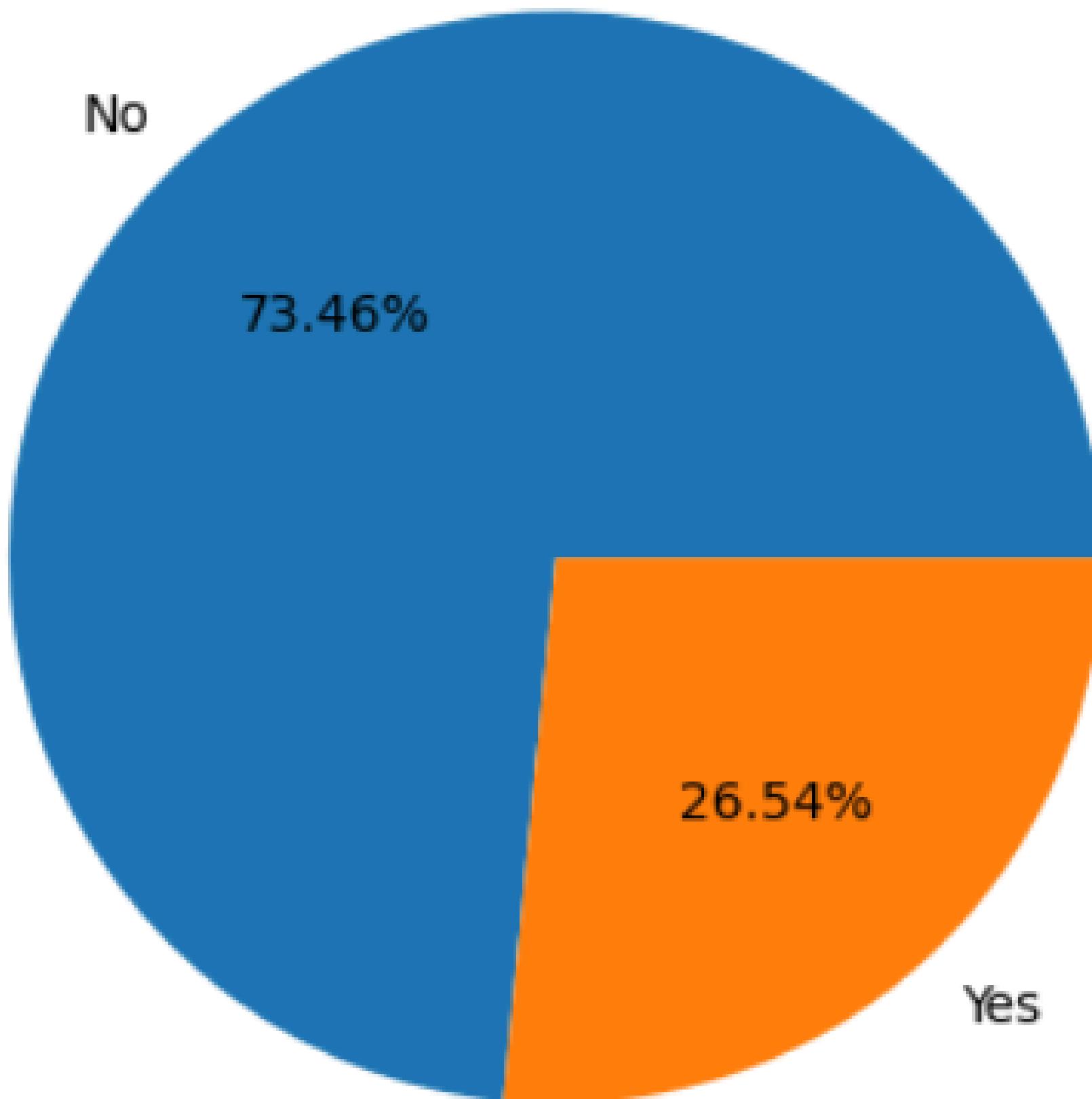


```
gb = df.groupby("Churn").agg({'Churn':"count"})
gb

plt.pie(gb['Churn'], labels= gb.index, autopct = "%1.2f%%")
plt.title("Percentage of Churn Customer")
plt.figure(figsize=(3,4))

plt.show()
```

Percentage of Churn Customer



```

# Columns you want to plot (excluding non-categorical or irrelevant ones)
columns_to_plot = [
    'gender', 'SeniorCitizen', 'Partner', 'Dependents',
    'PhoneService', 'MultipleLines', 'InternetService',
    'OnlineSecurity', 'OnlineBackup', 'DeviceProtection',
    'TechSupport', 'StreamingTV', 'StreamingMovies',
    'Contract', 'PaperlessBilling', 'PaymentMethod', 'Churn'
]

# Setup: define subplot grid size
n_cols = 4
n_rows = math.ceil(len(columns_to_plot) / n_cols)

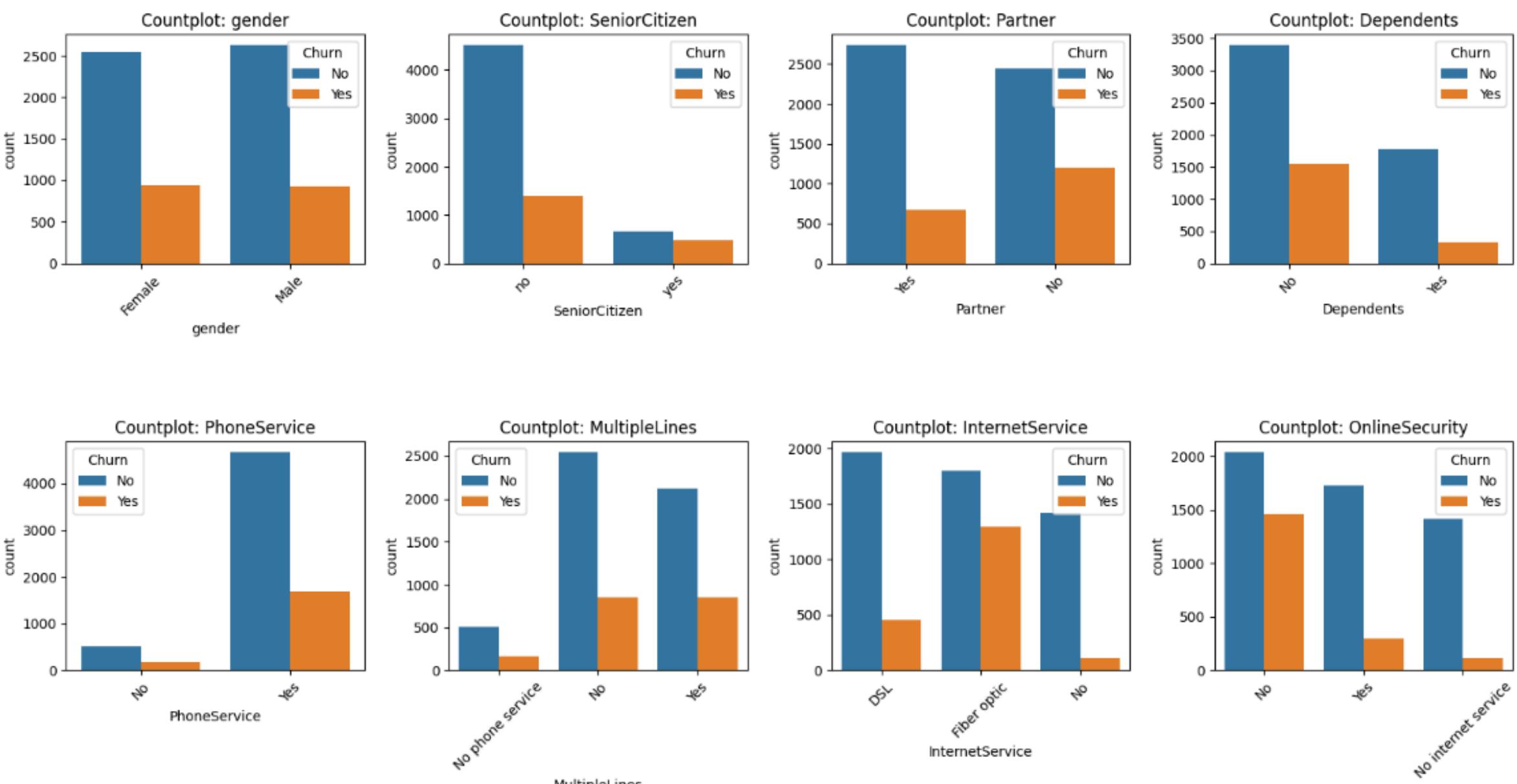
# Create subplots
fig, axes = plt.subplots(n_rows, n_cols, figsize=(n_cols*4, n_rows*4))
axes = axes.flatten() # Make it easy to iterate

# Create countplots
for i, col in enumerate(columns_to_plot):
    sns.countplot(x=col, data=df, ax=axes[i], hue = "Churn")
    axes[i].set_title(f'Countplot: {col}')
    axes[i].tick_params(axis='x', rotation=45)

# Remove any unused axes
for j in range(len(columns_to_plot), len(axes)):
    fig.delaxes(axes[j])

plt.tight_layout()
plt.show()

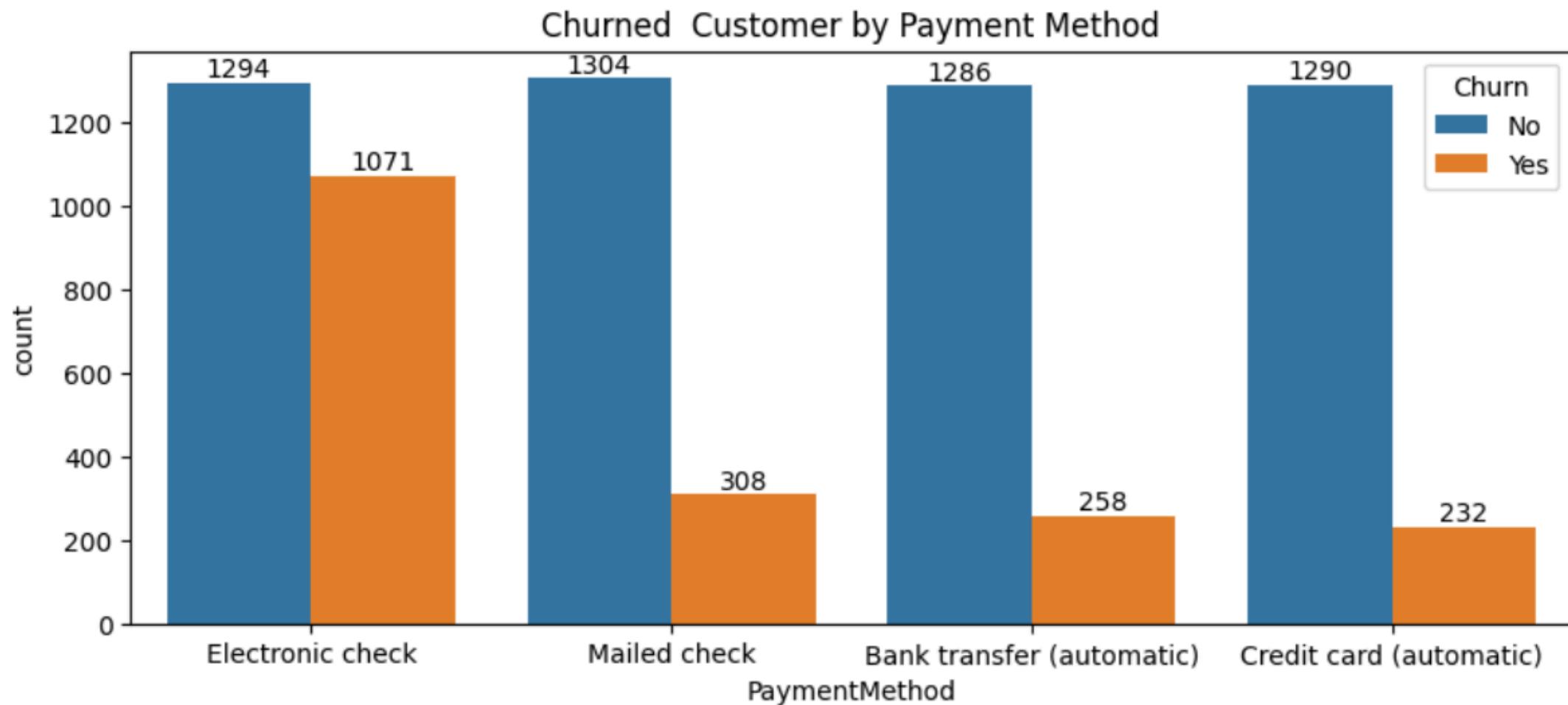
```



```

plt.figure(figsize=(10,4))
ax = sns.countplot(x = "PaymentMethod" , data = df , hue = "Churn")
ax.bar_label(ax.containers[0])
ax.bar_label(ax.containers[1])
plt.title("Churned Customer by Payment Method")

```



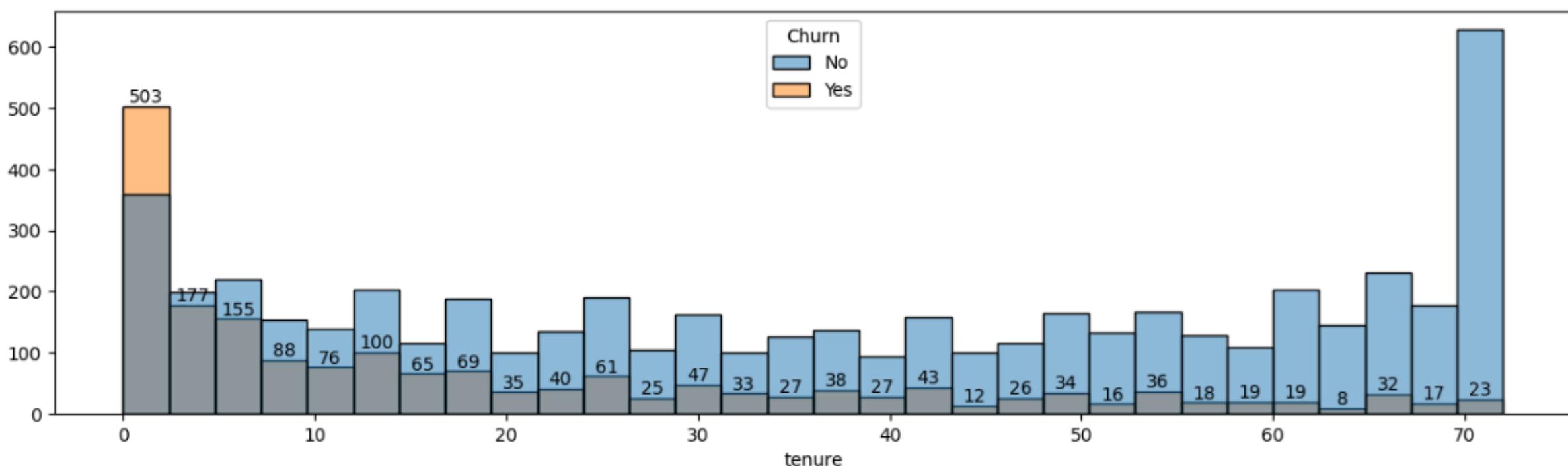
From the Above chart we can see Customers with Electronic Check are more likely to churn and Customer with Automatic

Based on tenure

```

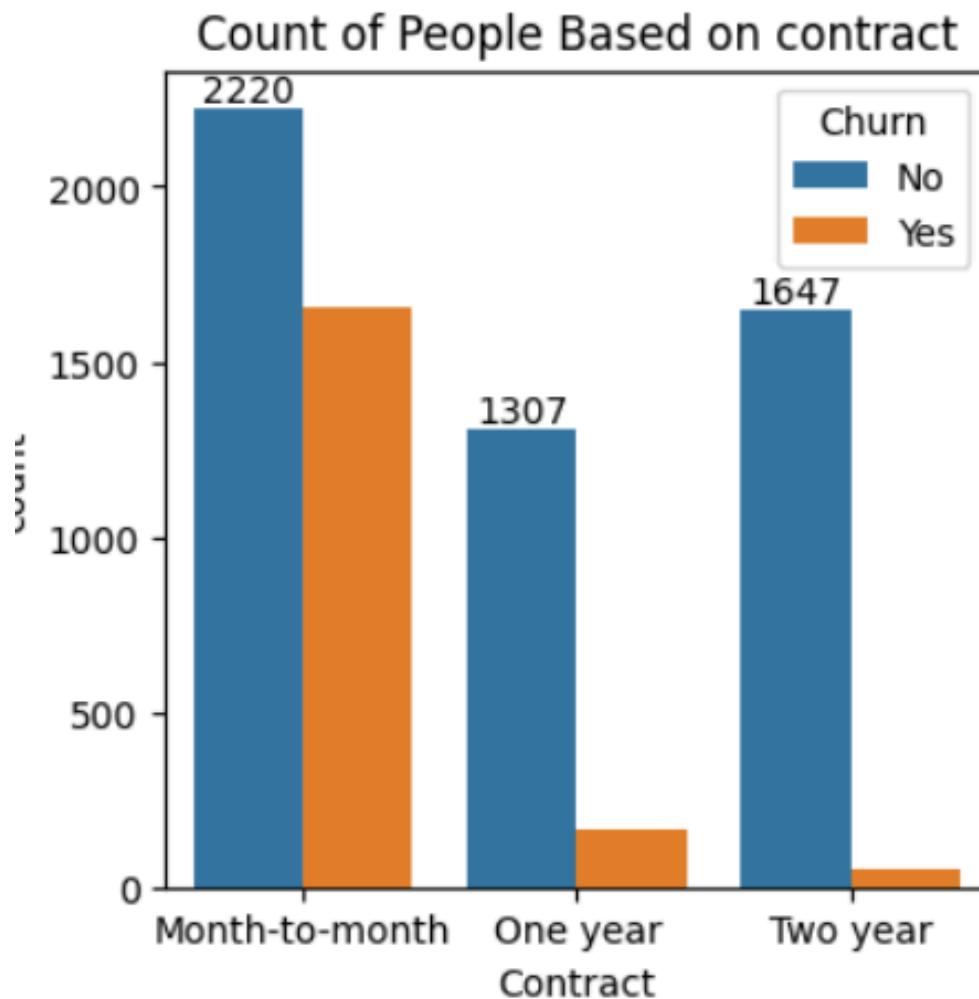
plt.figure(figsize= (15,4))
ax = sns.histplot( x = "tenure" , data = df , bins = 30, hue = "Churn")
ax.bar_label(ax.containers[0])
plt.show()

```



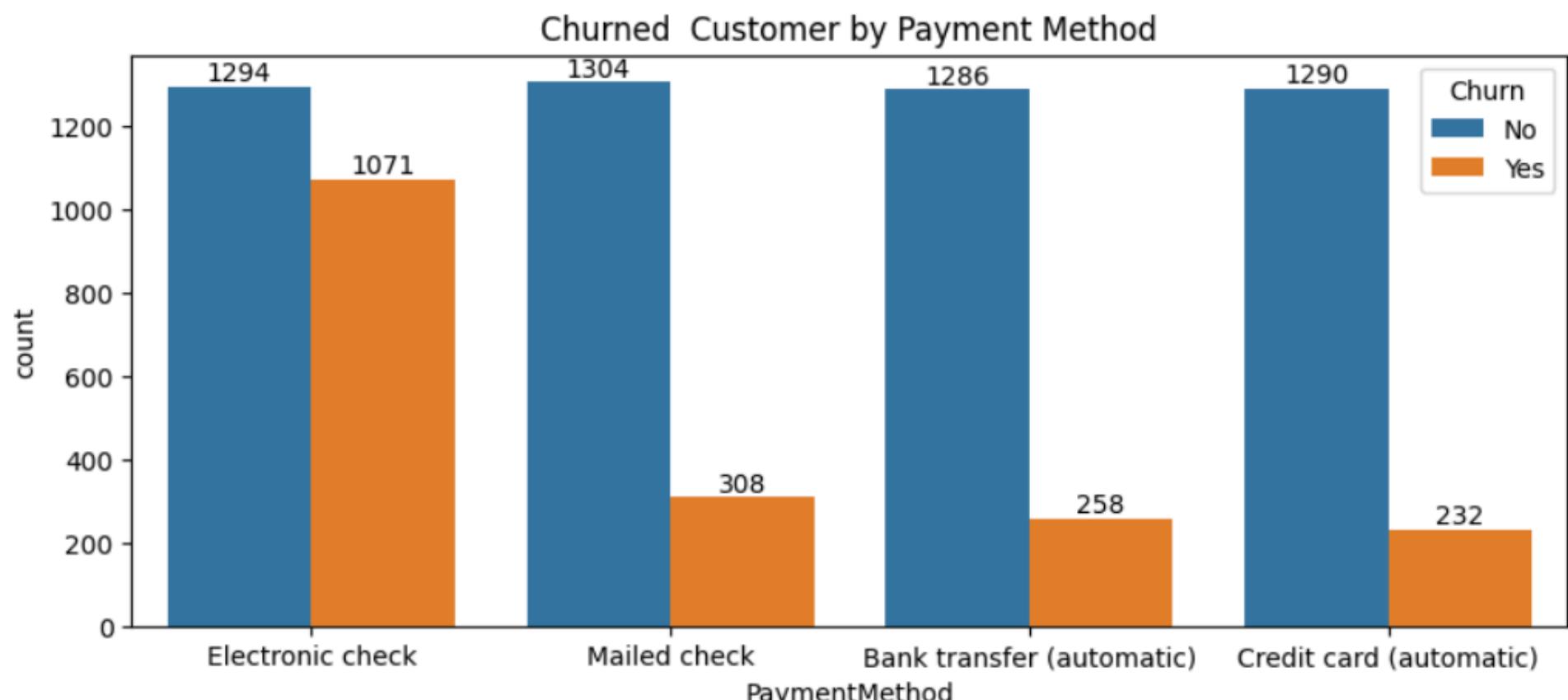
```
# Count of People based on Contract

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



People who have month to month contract are more likely to churn than those who have 1 or 2 years of contract.

```
plt.figure(figsize=(10,4))
ax = sns.countplot(x = "PaymentMethod" , data = df , hue = "Churn")
ax.bar_label(ax.containers[0])
ax.bar_label(ax.containers[1])
plt.title("Churned Customer by Payment Method")
plt.show()
```



From the Above chart we can see Customers with Electronic Check are more likely to churn and Customer with Automatic payment mode are less likely to Churn