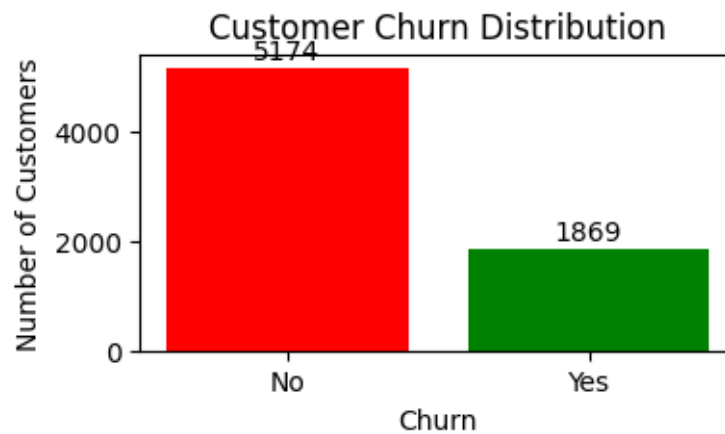# customer-retention-prediction
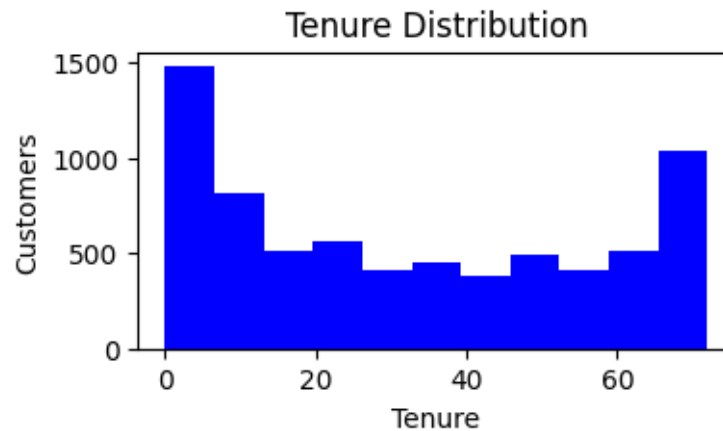
January 6, 2026

```python
[203]: import pandas as pd
       import matplotlib.pyplot as plt
       df = pd.read_excel("Sample.xlsx")
```

```python
[231]: a = df['Churn'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.bar(a.index, a.values, color=['r','g'])
       plt.xlabel("Churn")
       plt.ylabel("Number of Customers")
       plt.title("Customer Churn Distribution")
       plt.bar_label(bars, padding=1)
       plt.show()
```
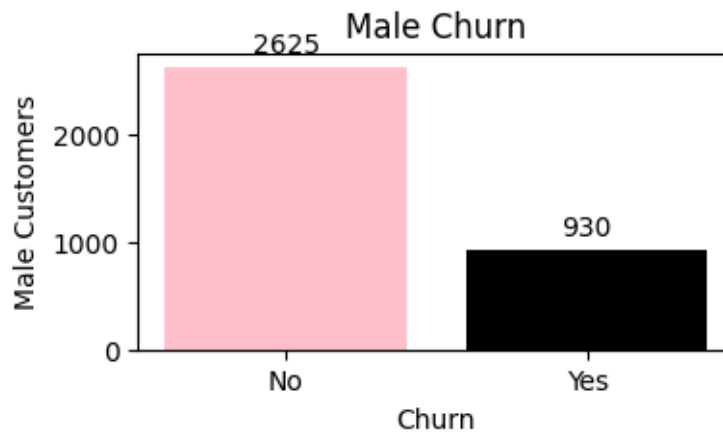


```python
[205]: plt.figure(figsize=(4,2))
       plt.hist(df['tenure'], bins=11, color='b')
       plt.xlabel("Tenure")
       plt.ylabel("Customers")
       plt.title("Tenure Distribution")
       plt.show()
```
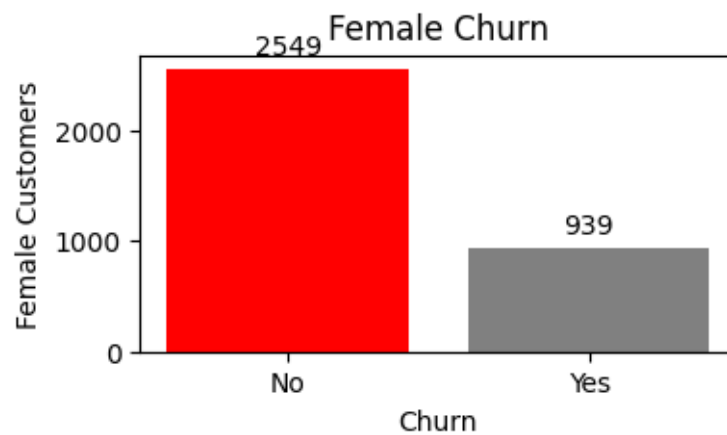
Tenure Distribution


```
[206]: male = df[df['gender'] == 'Male']['Churn'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.bar(male.index, male.values,color=['pink','black'])
       plt.xlabel("Churn")
       plt.ylabel("Male Customers")
       plt.title("Male Churn")
       plt.bar_label(bars, padding=3)
       plt.show()
```
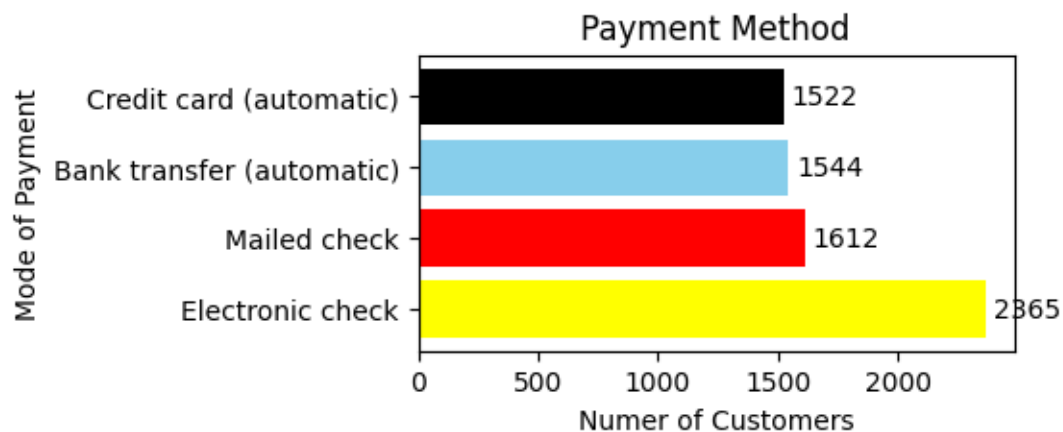
Male Churn


```
[207]: female = df[df['gender'] == 'Female']['Churn'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.bar(female.index, female.values,color=['r','grey'])
       plt.xlabel("Churn")
       plt.ylabel("Female Customers")
       plt.title("Female Churn")
```

```
plt.bar_label(bars, padding=3)
plt.show()
```
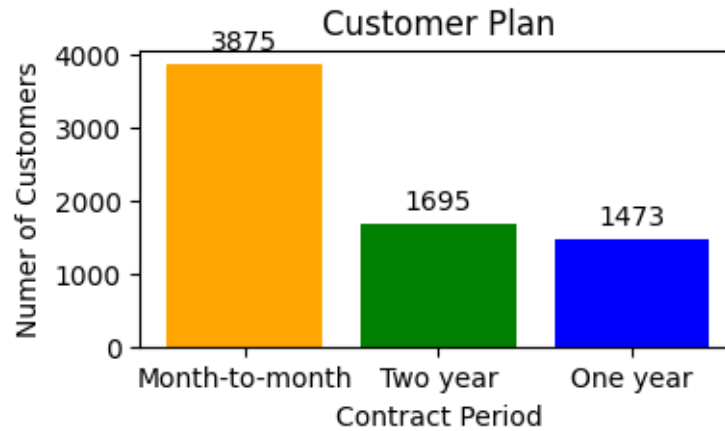


```
[208]:  a=df['PaymentMethod'].value_counts()
        plt.figure(figsize=(4,2))
        bars = plt.barh(a.index,a.values,color=['yellow','r','skyblue','black'])
        plt.xlabel("Numer of Customers")
        plt.ylabel("Mode of Payment")
        plt.title('Payment Method')
        plt.bar_label(bars, padding=3)
        plt.show()
```
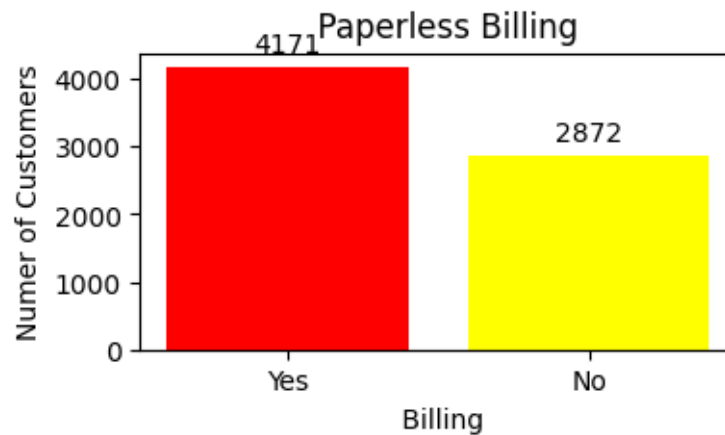


```
[209]:  a=df['Contract'].value_counts()
        plt.figure(figsize=(4,2))
        bars = plt.bar(a.index, a.values, color=['orange','g','b'])
```
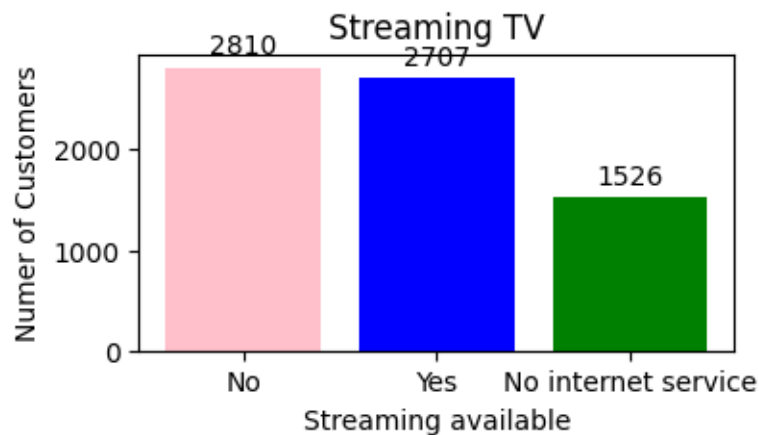
```
plt.xlabel("Contract Period")
plt.ylabel("Numer of Customers")
plt.title('Customer Plan')
plt.bar_label(bars, padding=3)
plt.show()
```
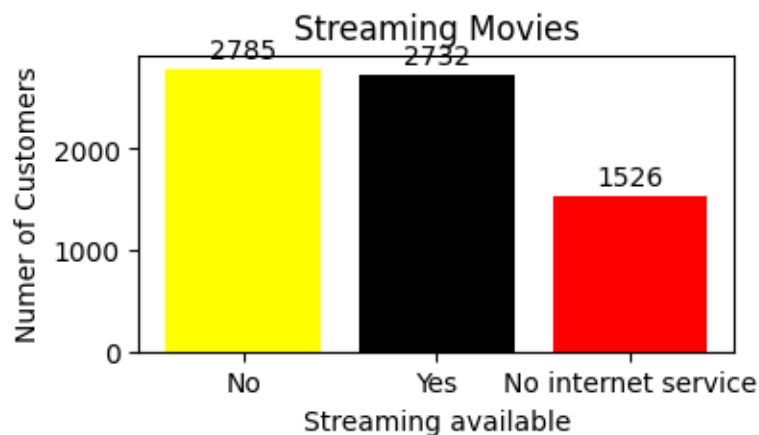


```
[210]: a=df['PaperlessBilling'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.bar(a.index, a.values, color=['r','yellow'])
       plt.xlabel(" Billing")
       plt.ylabel("Numer of Customers")
       plt.title("  Paperless Billing")
       plt.bar_label(bars, padding=3)
       plt.show()
```
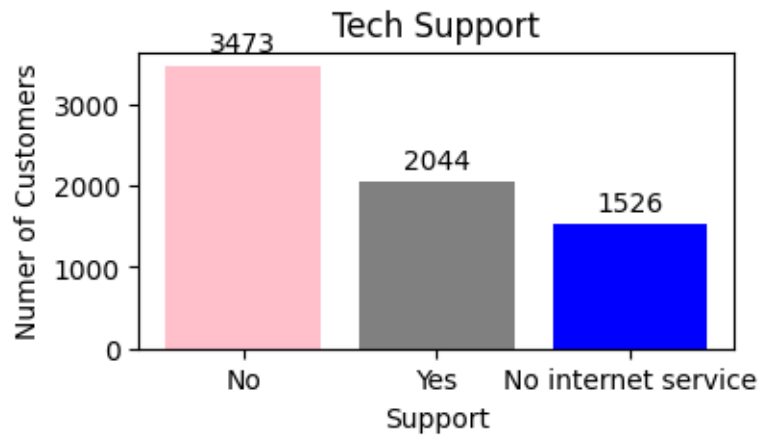
```
[211]: a=df['StreamingTV'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.bar(a.index, a.values, color=['pink','b','g'])
       plt.xlabel("Streaming available")
       plt.ylabel("Numer of Customers")
       plt.title("Streaming TV")
       plt.bar_label(bars, padding=3)
       plt.show()
```
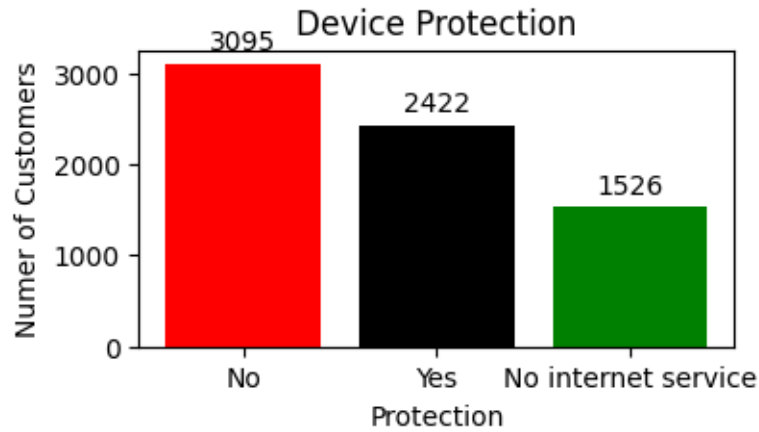


```
[212]: a=df['StreamingMovies'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.bar(a.index, a.values, color=['yellow','black','r'])
       plt.xlabel("Streaming available")
       plt.ylabel("Numer of Customers")
       plt.title("Streaming Movies")
       plt.bar_label(bars, padding=2)
       plt.show()
```
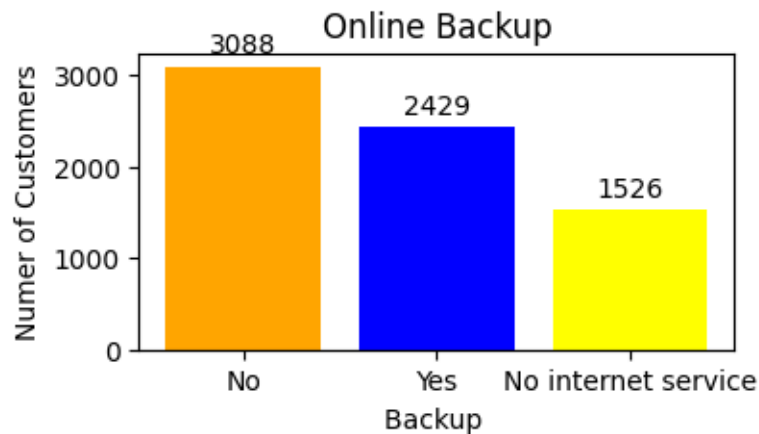
```
[213]: a=df['TechSupport'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.bar(a.index,a.values,color=['pink','grey','b'])
       plt.xlabel("Support")
       plt.ylabel("Numer of Customers")
       plt.title("Tech Support")
       plt.bar_label(bars, padding=3)
       plt.show()
```



```
[214]: a=df['DeviceProtection'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.bar(a.index,a.values,color=['r','black','g'])
       plt.xlabel("Protection")
       plt.ylabel("Numer of Customers")
       plt.title("Device Protection")
       plt.bar_label(bars, padding=3)
       plt.show()
```
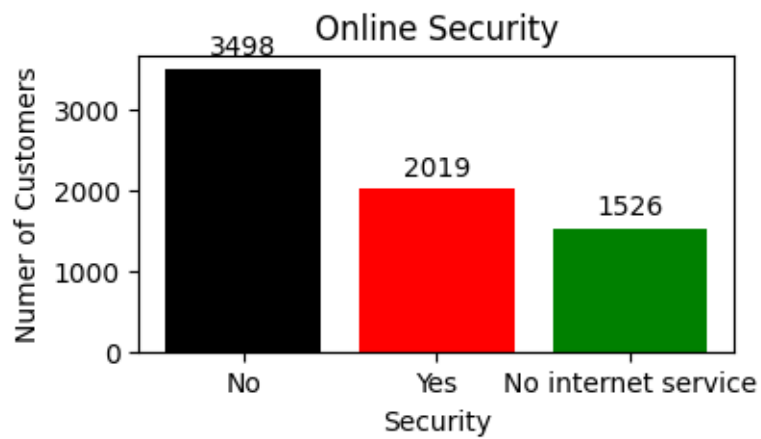
Device Protection


```
[215]: a=df['OnlineBackup'].value_counts()
        plt.figure(figsize=(4,2))
        bars = plt.bar(a.index,a.values,color=['orange','b','yellow'])
        plt.xlabel("Backup ")
        plt.ylabel("Numer of Customers")
        plt.title("Online Backup")
        plt.bar_label(bars, padding=3)
        plt.show()
```

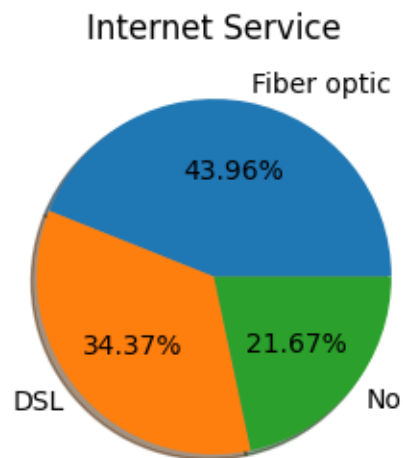## Online Backup


```
[216]: b=df['OnlineSecurity'].value_counts()
        plt.figure(figsize=(4,2))
        bars = plt.bar(b.index,b.values,color=['black','r','g'])
        plt.xlabel("Security")
        plt.ylabel("Numer of Customers")
        plt.title("Online Security")
```
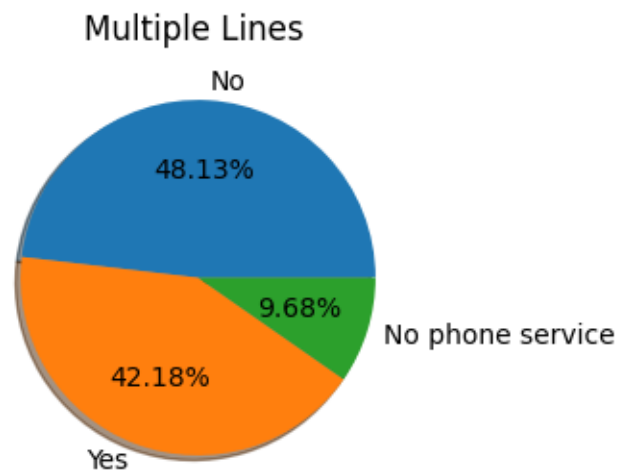
```
plt.bar_label(bars, padding=3)
plt.show()
```


[217]:
```
a=df['InternetService'].value_counts()
b=a.index
plt.figure(figsize=(5,3))
plt.title("Internet Service")
plt.pie(a,labels=b,shadow=True,autopct='%1.2f%%')
plt.show()
```
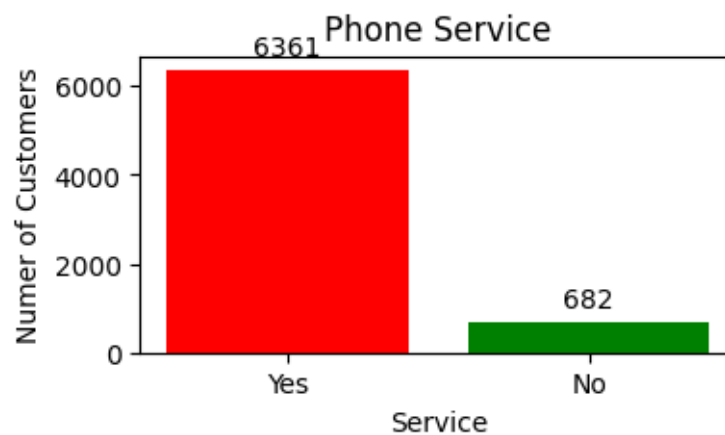

[218]:
```
a=df['MultipleLines'].value_counts()
b=a.index
plt.figure(figsize=(5,3))
plt.title("Multiple Lines ")
```
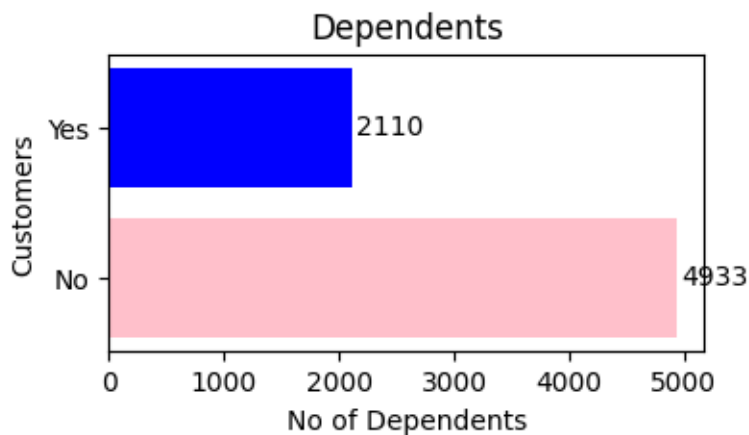
```
plt.pie(a,labels=b,shadow=True,autopct='%1.2f%%')
plt.show()
```

Multiple Lines
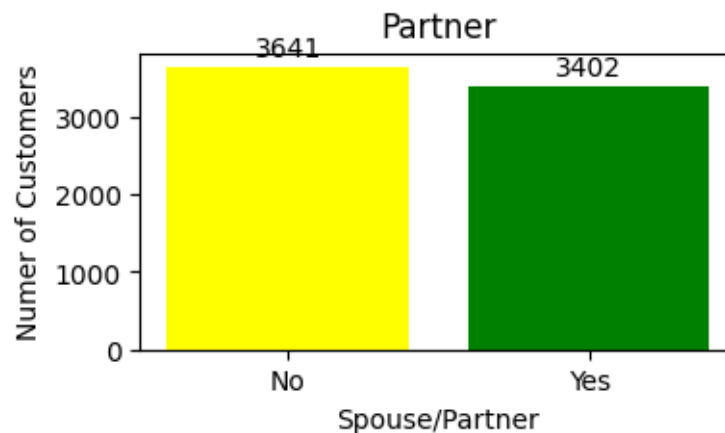


```
[219]: a=df['PhoneService'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.bar(a.index,a.values,color=['r','g'])
       plt.xlabel("Service")
       plt.ylabel("Numer of Customers")
       plt.title("Phone Service")
       plt.bar_label(bars, padding=3)
       plt.show()
```
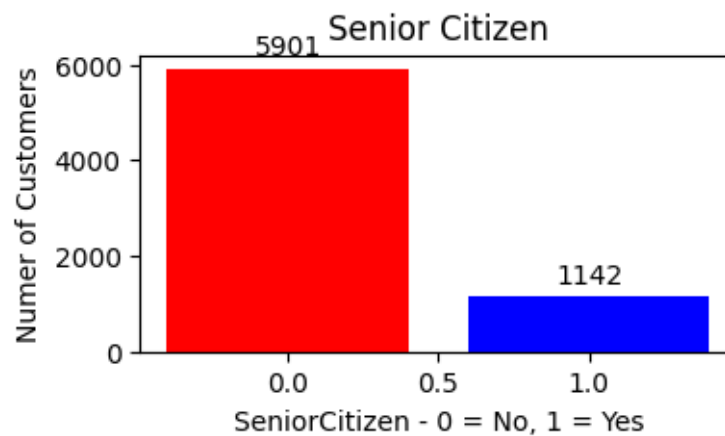
```
[220]: a=df['Dependents'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.barh(a.index,a.values,color=['pink','b'])
       plt.xlabel("No of Dependents")
       plt.ylabel("Customers")
       plt.title("Dependents")
       plt.bar_label(bars, padding=2)
       plt.show()
```
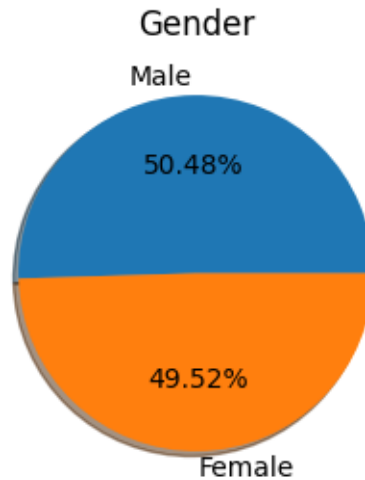


```
[221]: a=df['Partner'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.bar(a.index,a.values,color=['yellow','g'])
       plt.xlabel("Spouse/Partner")
       plt.ylabel("Numer of Customers")
       plt.title("Partner")
       plt.bar_label(bars, padding=2)
       plt.show()
```
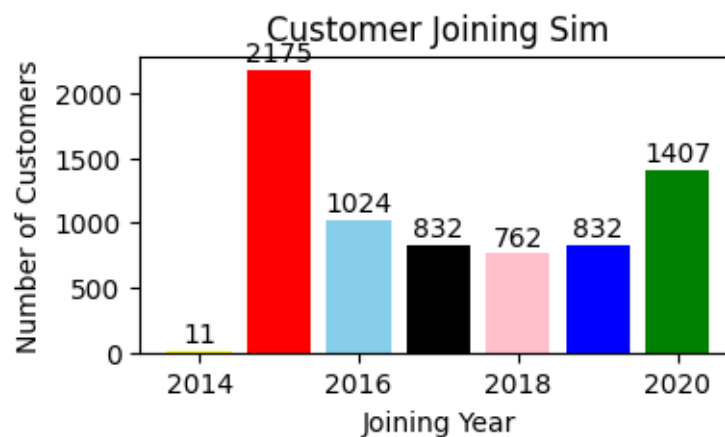
```
[222]: a=df['SeniorCitizen'].value_counts()
       plt.figure(figsize=(4,2))
       bars = plt.bar(a.index,a.values,color=['r','b'])
       plt.xlabel("SeniorCitizen - 0 = No, 1 = Yes")
       plt.ylabel("Numer of Customers")
       plt.title("Senior Citizen")
       plt.bar_label(bars, padding=3)
       plt.show()
```



```
[223]: a=df['gender'].value_counts()
       b=a.index
       plt.figure(figsize=(5,3))
       plt.title("Gender")
       plt.pie(a,labels=b,shadow=True,autopct='%1.2f%%')
       plt.show()
```
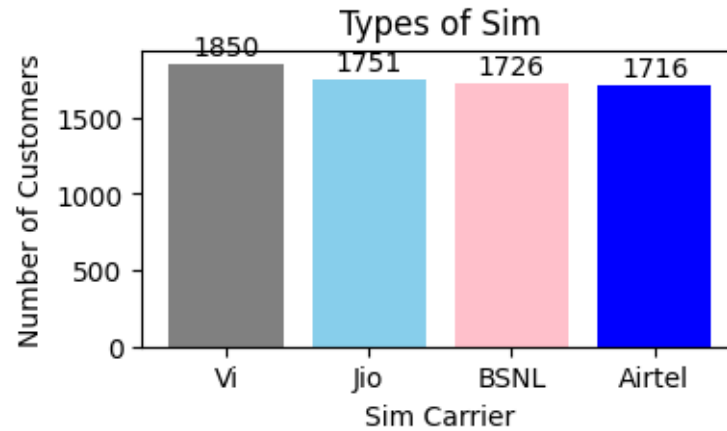
Gender

[224]:
```python
a=df['JoinYear'].value_counts().sort_index()
plt.figure(figsize=(4,2))
bars = plt.bar(a.index,a.
 ↪values,color=['yellow','r','skyblue','black','pink','b','g'])
plt.xlabel('Joining Year')
plt.ylabel('Number of Customers')
plt.title('Customer Joining Sim')
plt.bar_label(bars, padding=1)
plt.show()
```
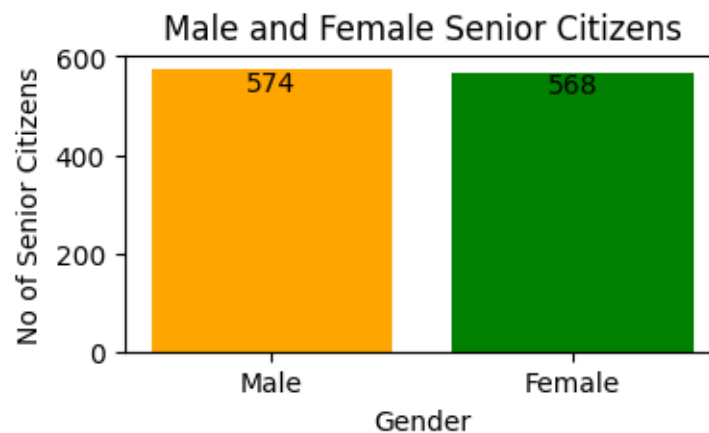

[225]:
```python
a=df['sim'].value_counts()
plt.figure(figsize=(4,2))
bars = plt.bar(a.index,a.values,color=['grey','skyblue','pink','b'])
plt.xlabel('Sim Carrier')
```
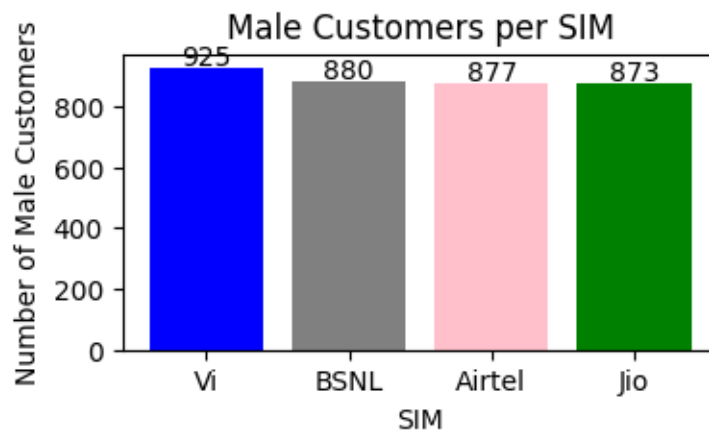
```
plt.ylabel('Number of Customers')
plt.title('Types of Sim')
plt.bar_label(bars, padding=1)
plt.show()
```
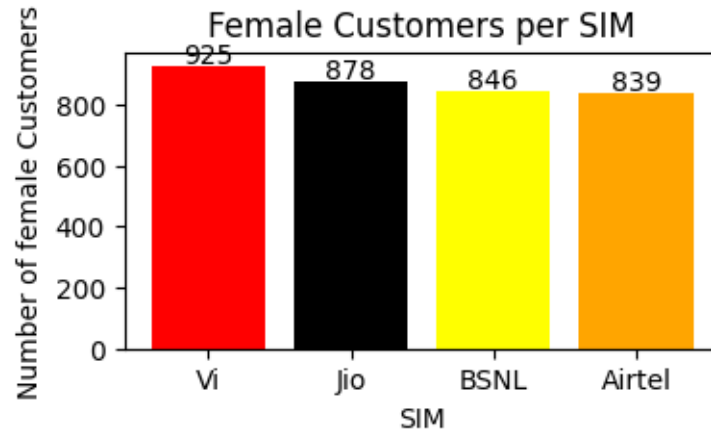
Types of Sim


```
[226]: a = df[df['SeniorCitizen'] == 1]
b = a['gender'].value_counts()
plt.figure(figsize=(4, 2))
bars = plt.bar(b.index, b.values,color=['orange','g'])
plt.xlabel('Gender')
plt.ylabel('No of Senior Citizens')
plt.title('Male and Female Senior Citizens')
plt.bar_label(bars, padding=-10)
plt.show()
```
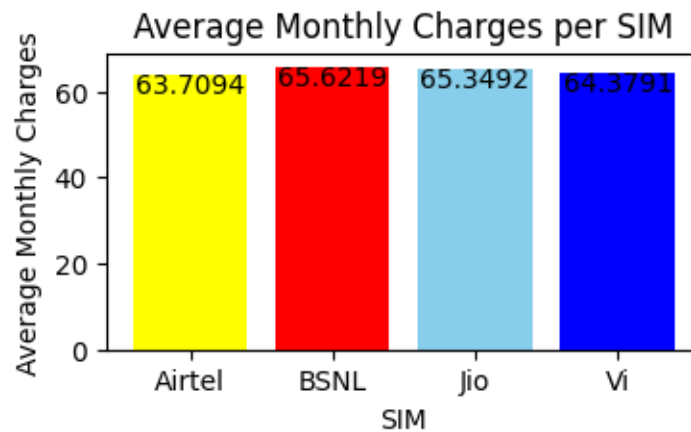
Male and Female Senior Citizens

[227]:
```python
a = df[df['gender'] == 'Male']
b = a['sim'].value_counts()
plt.figure(figsize=(4,2))
bars = plt.bar(b.index, b.values,color=['b','grey','pink','g'])
plt.xlabel("SIM")
plt.ylabel("Number of Male Customers")
plt.title("Male Customers per SIM")
plt.bar_label(bars, padding=-1)
plt.show()
```


[228]:
```python
a = df[df['gender'] == 'Female']
b = a['sim'].value_counts()
plt.figure(figsize=(4,2))
bars = plt.bar(b.index, b.values,color=['r','black','yellow','orange'])
plt.xlabel("SIM")
plt.ylabel("Number of female Customers")
plt.title("Female Customers per SIM")
plt.bar_label(bars, padding=-1)
plt.show()
```

Female Customers per SIM

```
[229]: a = df.groupby('sim')['MonthlyCharges'].mean()
       plt.figure(figsize=(4,2))
       bars = plt.bar(a.index, a.values,color=['yellow','r','skyblue','b'])
       plt.xlabel("SIM")
       plt.ylabel("Average Monthly Charges")
       plt.title("Average Monthly Charges per SIM")
       plt.bar_label(bars,padding=-9)
       plt.show()
```


Average Monthly Charges per SIM

```
[232]: df['Tenure_Quarter'] = (df['tenure'] // 3) + 1
       quarter_counts = df['Tenure_Quarter'].value_counts().sort_index()
       plt.figure(figsize=(5,3))
       plt.bar(quarter_counts.index, quarter_counts.values)
       plt.xlabel("Tenure (Quarters)")
       plt.ylabel("Number of Customers")
```

```
plt.title("Customer Tenure Distribution (Quarterly)")
plt.show()
```


Customer Tenure Distribution (Quarterly)