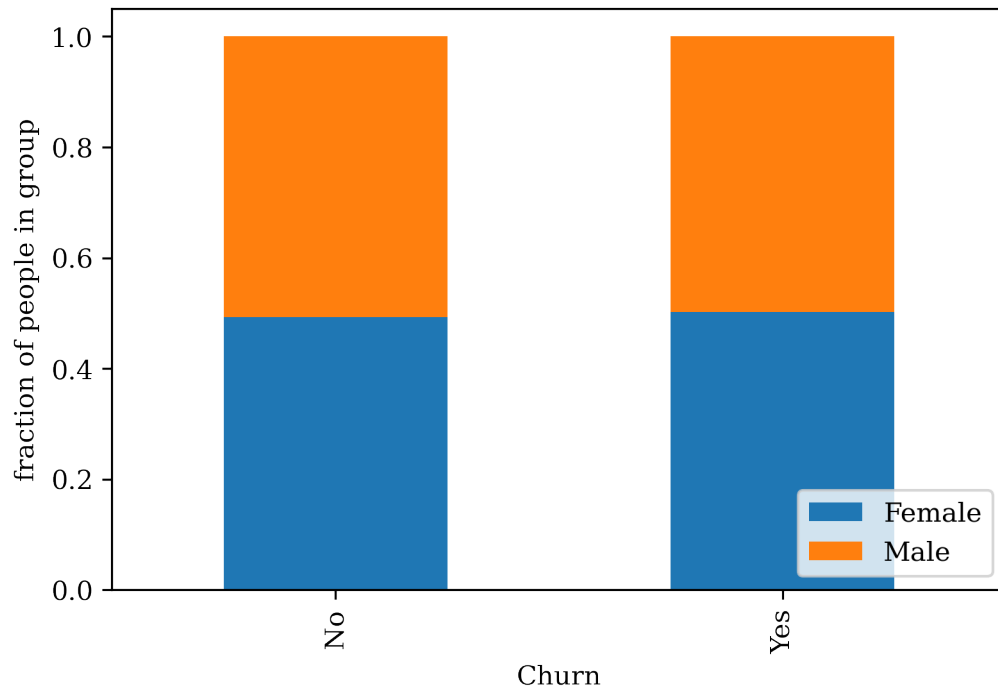
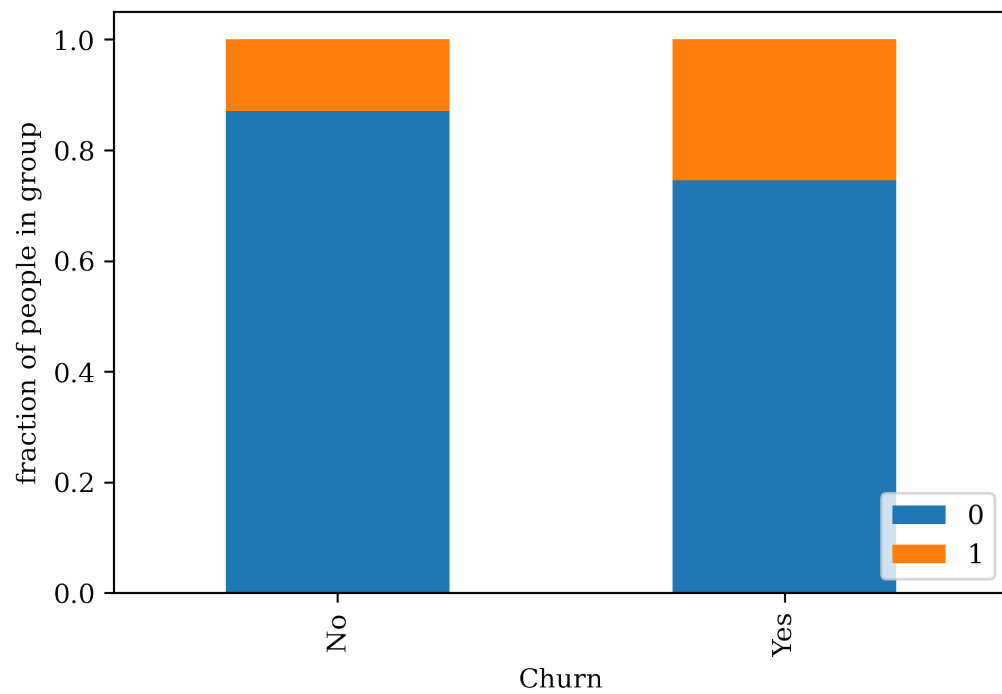


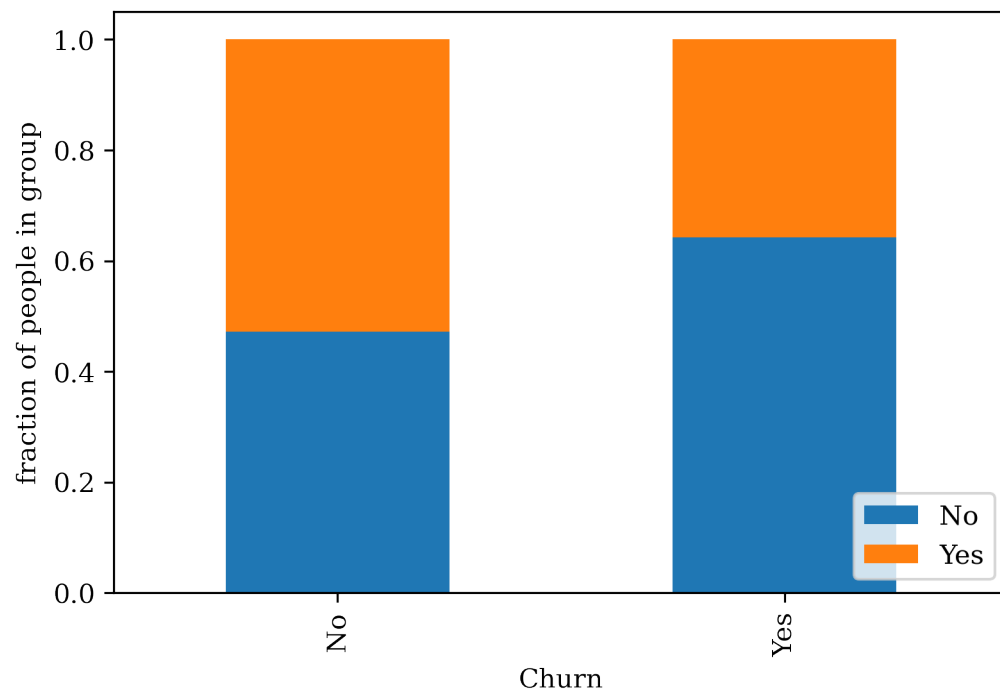
```
plt.title(col + ' Grouped by Churn Category')
plt.show()
print(df[col].describe())
df[col].plot.hist(bins = int(np.sqrt(df.shape[0])))
plt.xlabel(col)
plt.ylabel('Count')
plt.show()
```



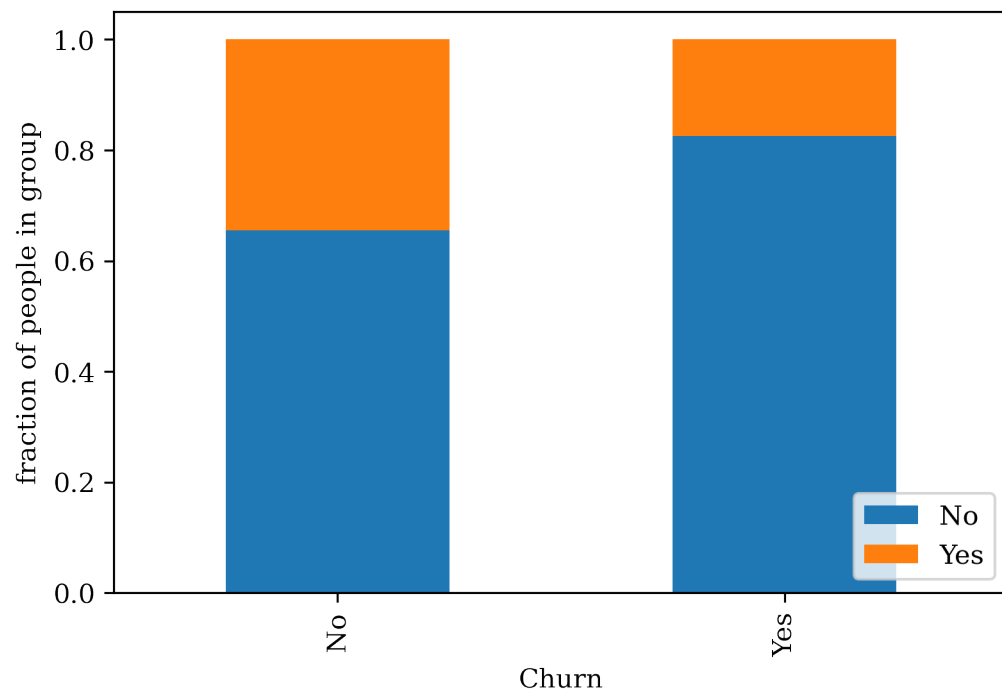
```
Male      3555
Female    3488
Name: gender, dtype: int64
Male      0.504756
Female    0.495244
Name: gender, dtype: float64
```



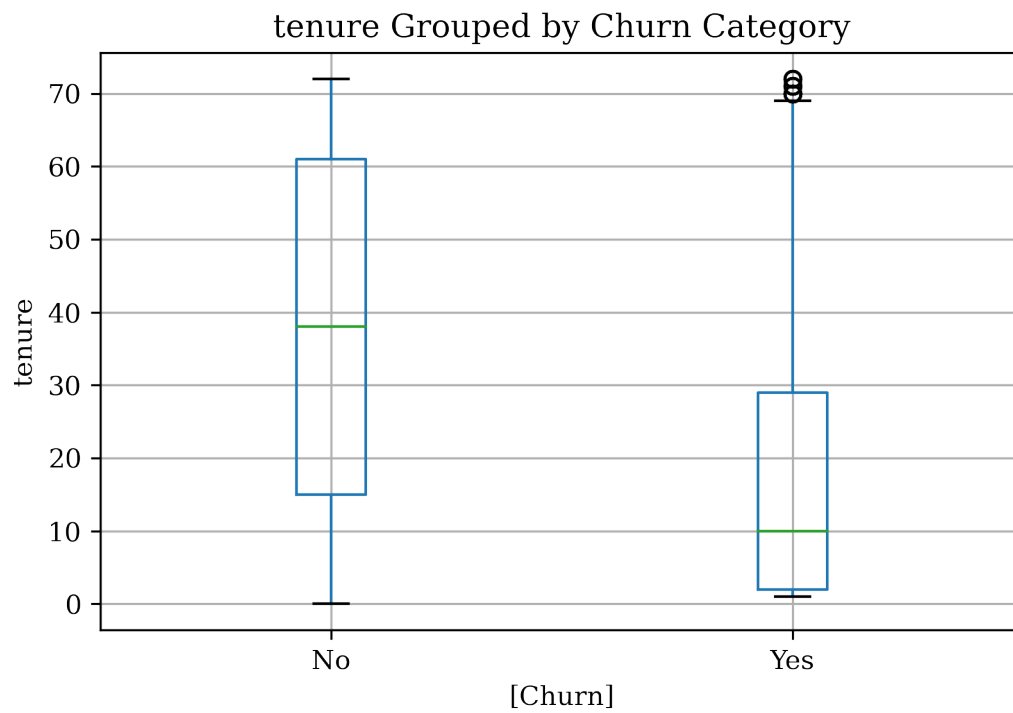
```
0    5901
1    1142
Name: SeniorCitizen, dtype: int64
0    0.837853
1    0.162147
Name: SeniorCitizen, dtype: float64
```



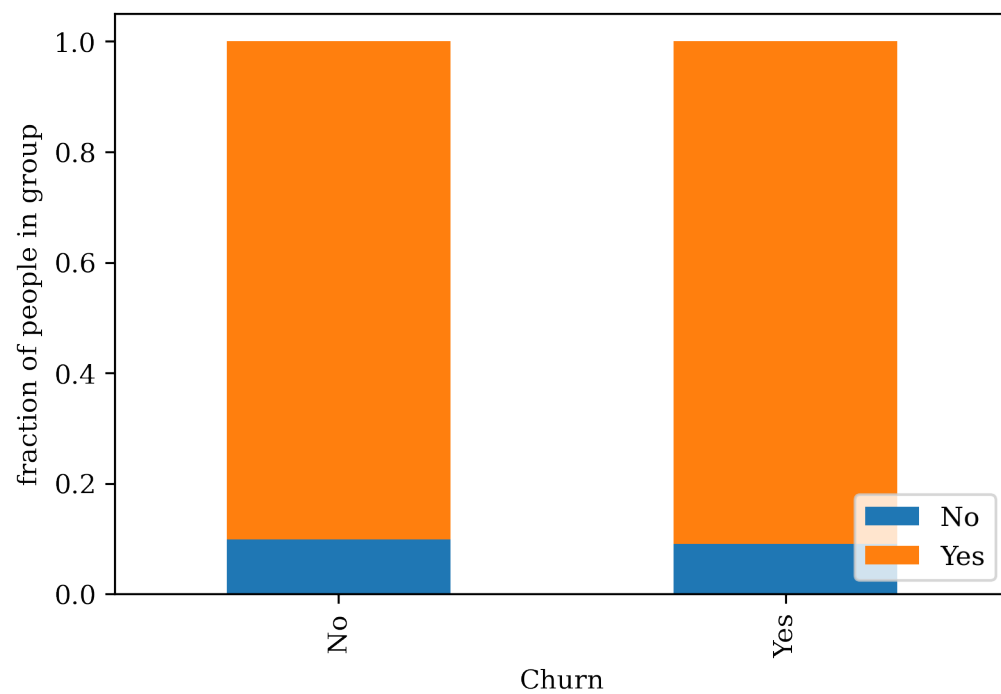
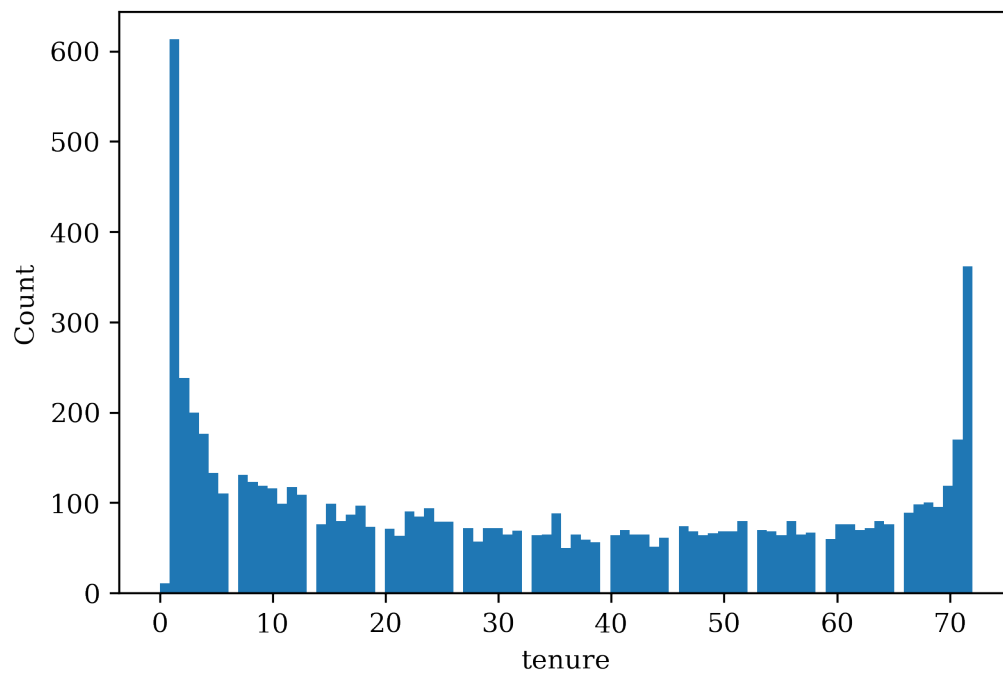
```
No      3641
Yes      3402
Name: Partner, dtype: int64
No      0.516967
Yes      0.483033
Name: Partner, dtype: float64
```



```
No      4933
Yes      2110
Name: Dependents, dtype: int64
No      0.700412
Yes      0.299588
Name: Dependents, dtype: float64
```



```
count    7043.000000
mean      32.371149
std       24.559481
min        0.000000
25%        9.000000
50%       29.000000
75%       55.000000
max       72.000000
Name: tenure, dtype: float64
```

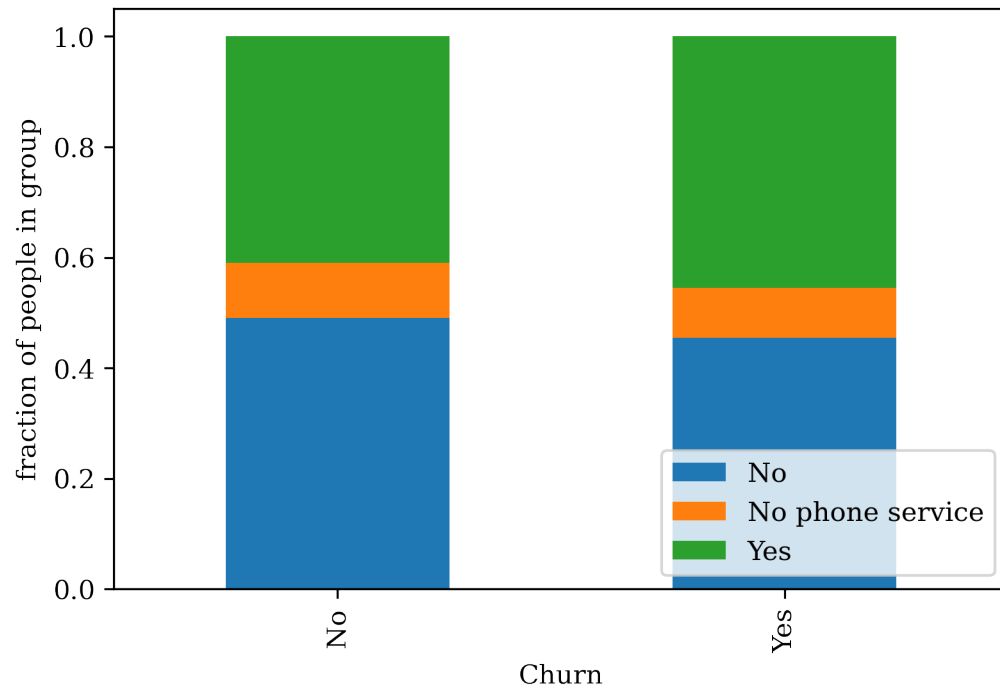


Yes	6361
No	682

```

Name: PhoneService, dtype: int64
Yes    0.903166
No     0.096834
Name: PhoneService, dtype: float64

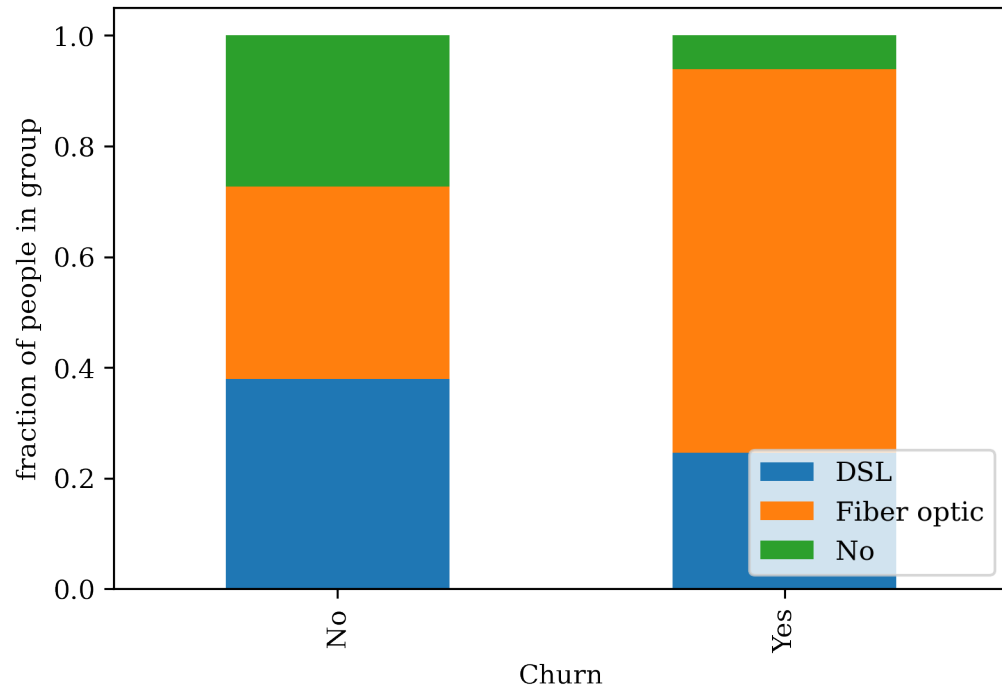
```



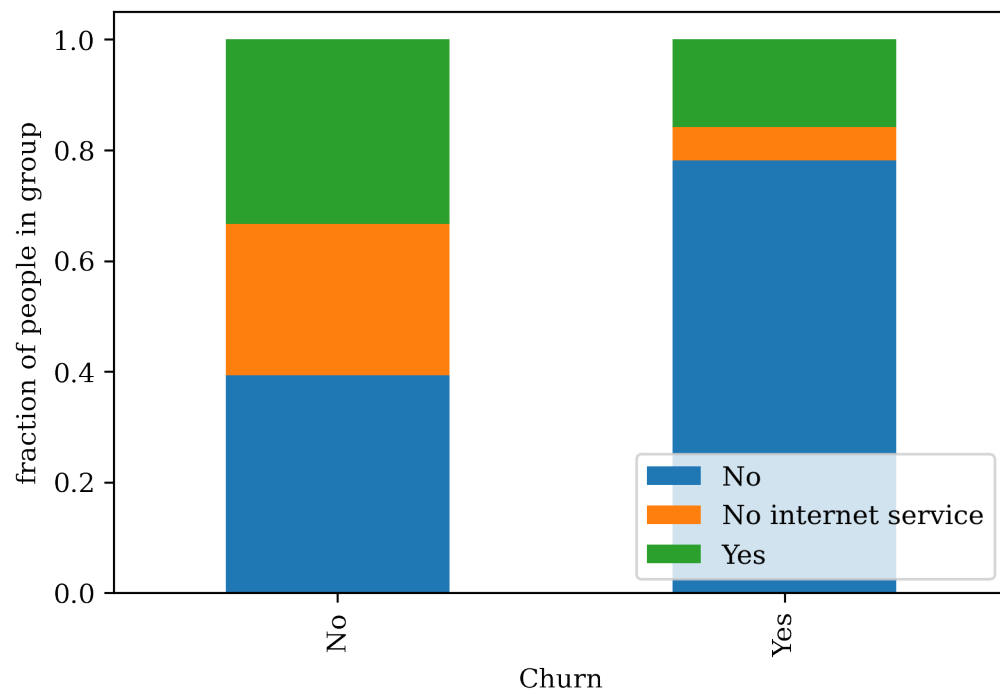
```

No          3390
Yes         2971
No phone service  682
Name: MultipleLines, dtype: int64
No          0.481329
Yes         0.421837
No phone service  0.096834
Name: MultipleLines, dtype: float64

```



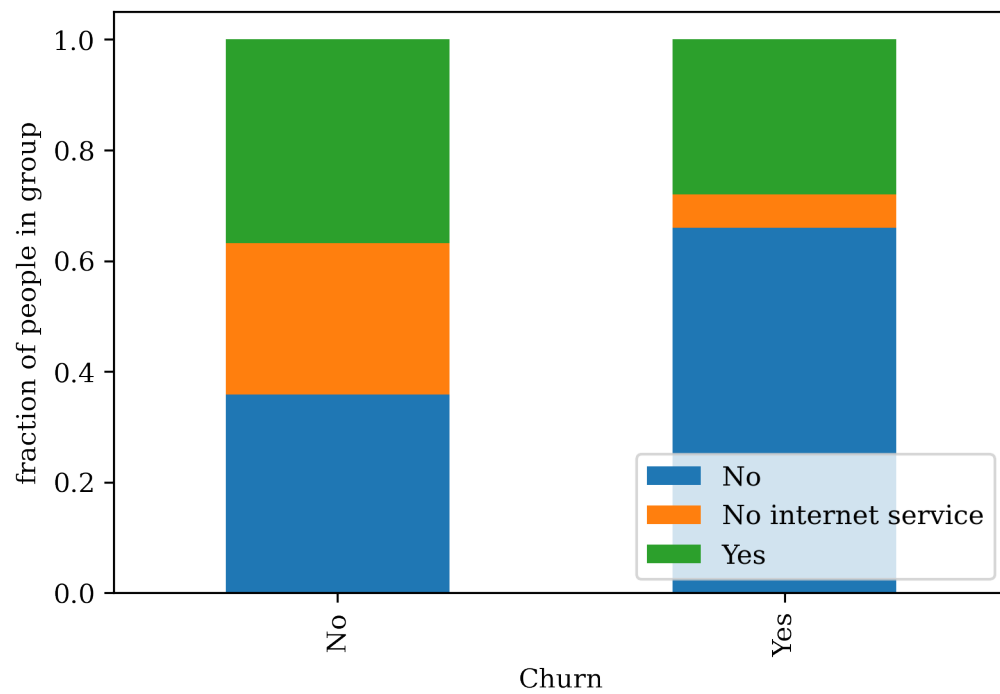
```
Fiber optic    3096
DSL            2421
No             1526
Name: InternetService, dtype: int64
Fiber optic    0.439585
DSL            0.343746
No             0.216669
Name: InternetService, dtype: float64
```

```

No          3498
Yes         2019
No internet service  1526
Name: OnlineSecurity, dtype: int64
No          0.496663
Yes         0.286668
No internet service  0.216669
Name: OnlineSecurity, dtype: float64

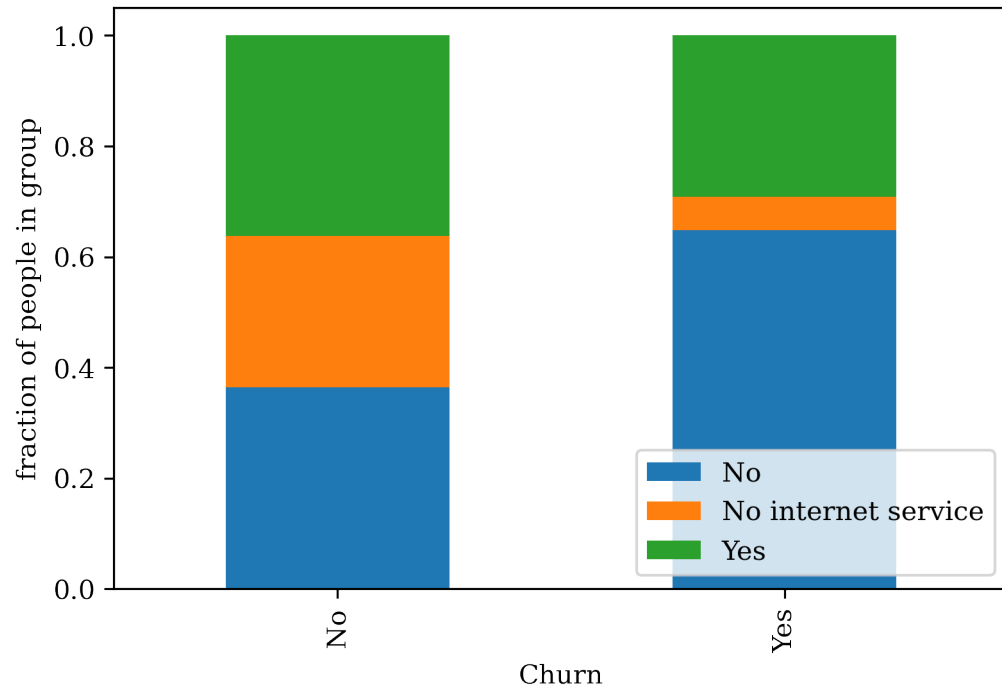
```



```

No          3088
Yes         2429
No internet service  1526
Name: OnlineBackup, dtype: int64
No          0.438450
Yes         0.344881
No internet service  0.216669
Name: OnlineBackup, dtype: float64

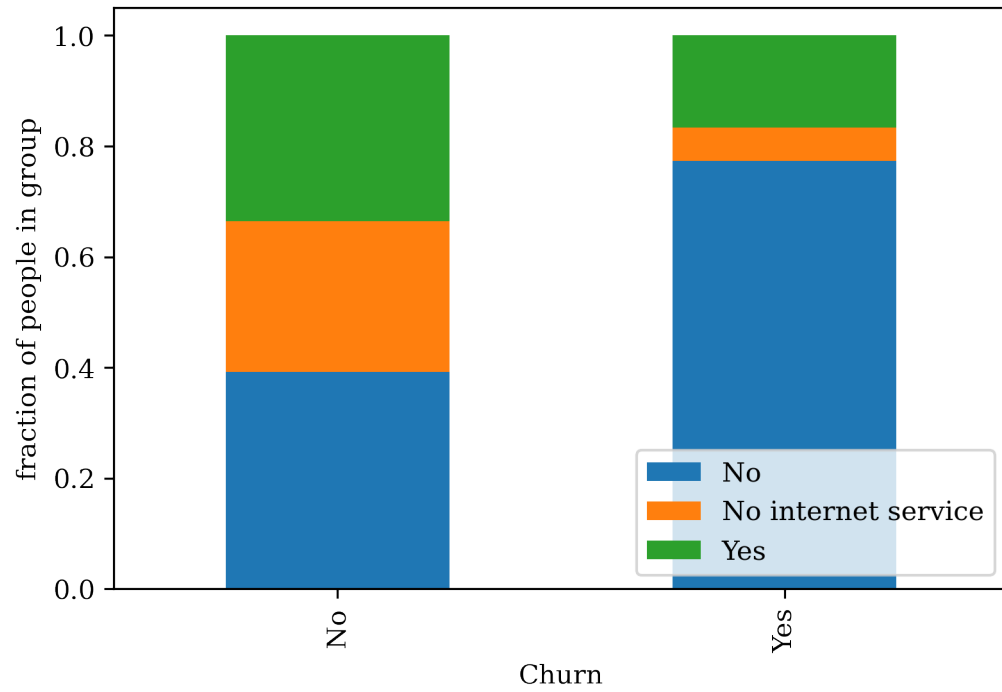
```



```

No          3095
Yes         2422
No internet service  1526
Name: DeviceProtection, dtype: int64
No          0.439443
Yes         0.343888
No internet service  0.216669
Name: DeviceProtection, dtype: float64

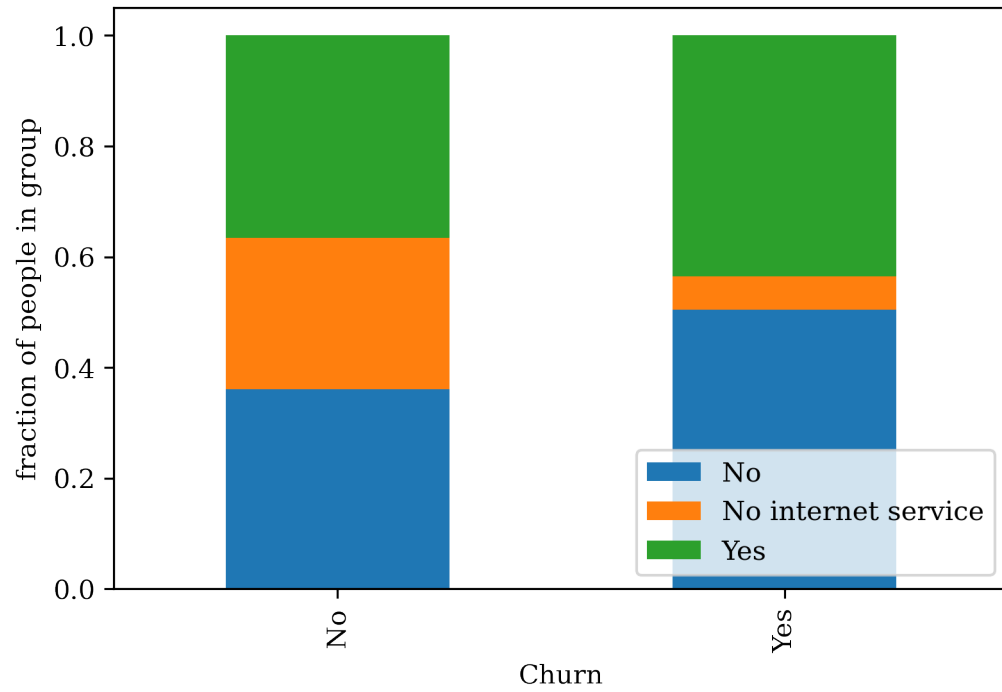
```



```

No          3473
Yes         2044
No internet service  1526
Name: TechSupport, dtype: int64
No          0.493114
Yes         0.290217
No internet service  0.216669
Name: TechSupport, dtype: float64

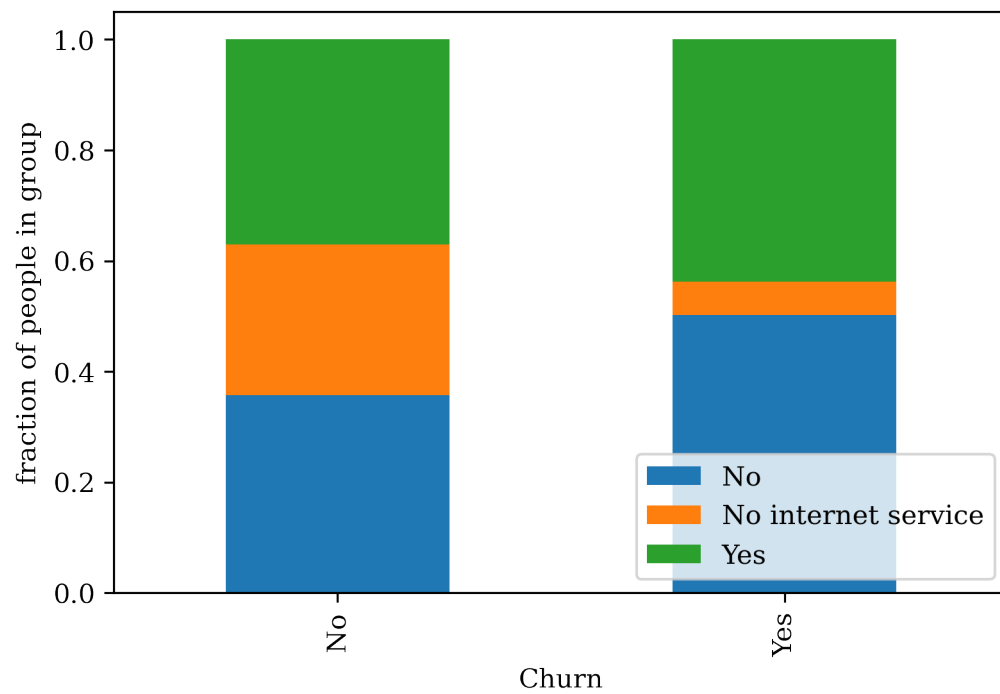
```



```

No                2810
Yes              2707
No internet service  1526
Name: StreamingTV, dtype: int64
No                0.398978
Yes              0.384353
No internet service  0.216669
Name: StreamingTV, dtype: float64

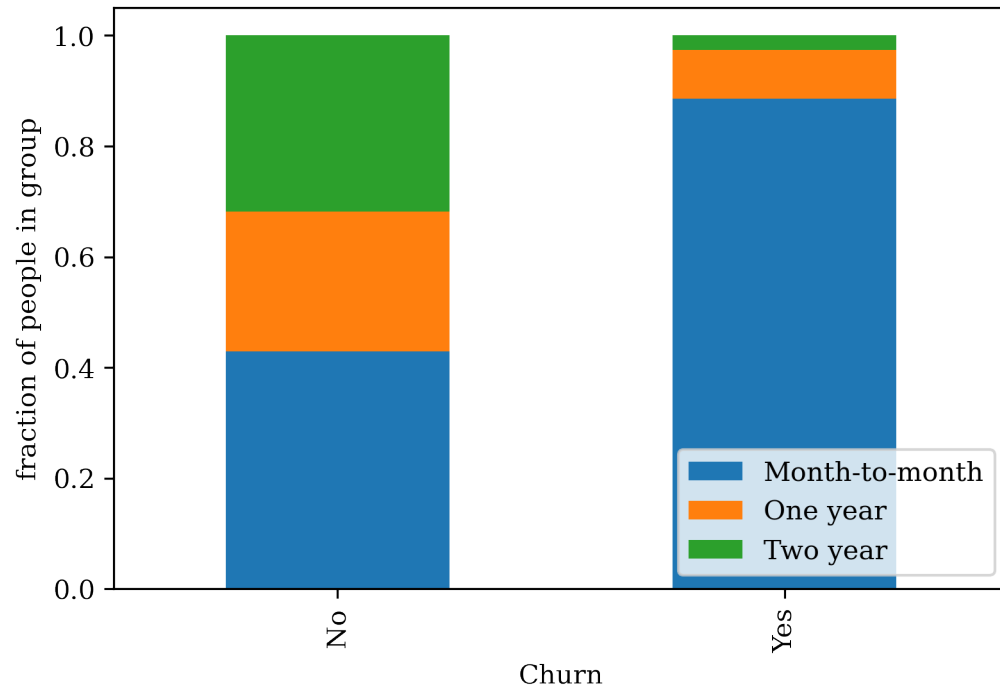
```



```

No                2785
Yes              2732
No internet service  1526
Name: StreamingMovies, dtype: int64
No                0.395428
Yes              0.387903
No internet service  0.216669
Name: StreamingMovies, dtype: float64

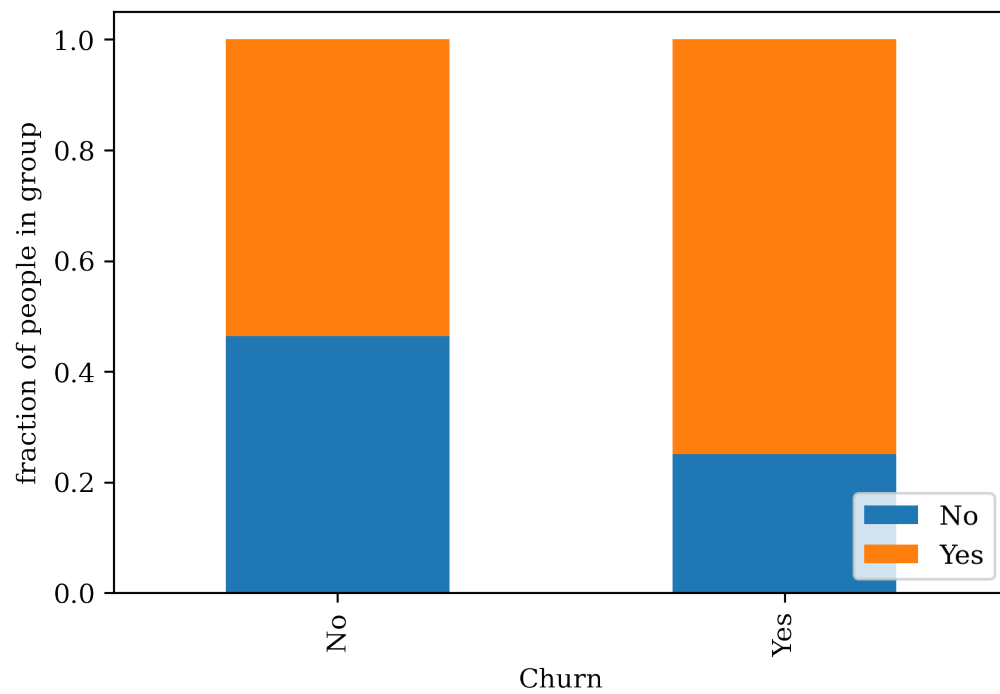
```



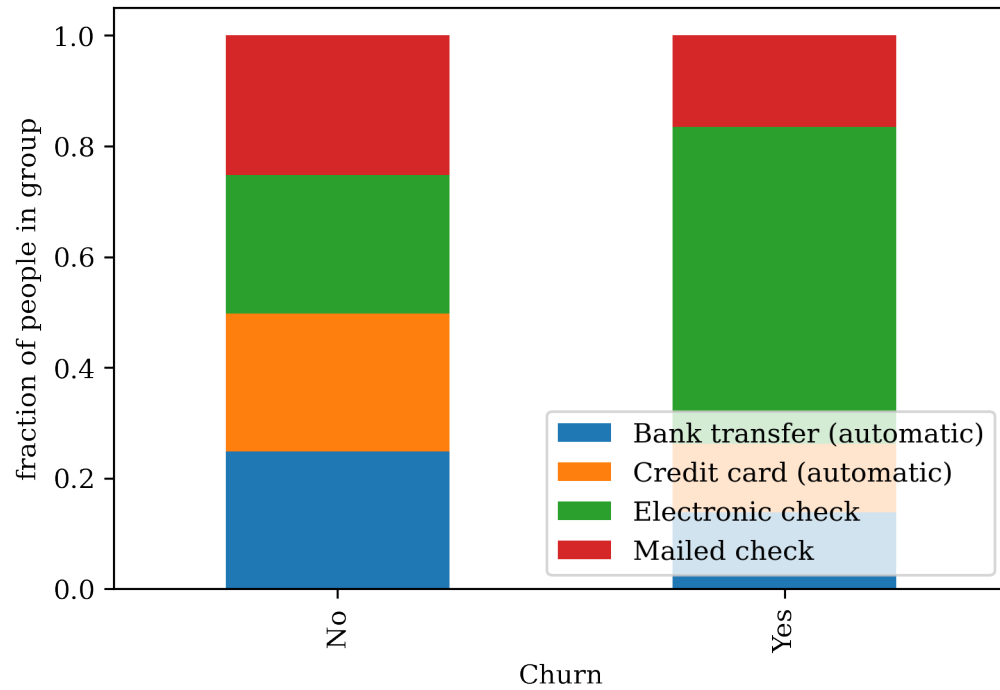
```

Month-to-month    3875
Two year          1695
One year          1473
Name: Contract, dtype: int64
Month-to-month    0.550192
Two year          0.240664
One year          0.209144
Name: Contract, dtype: float64

```



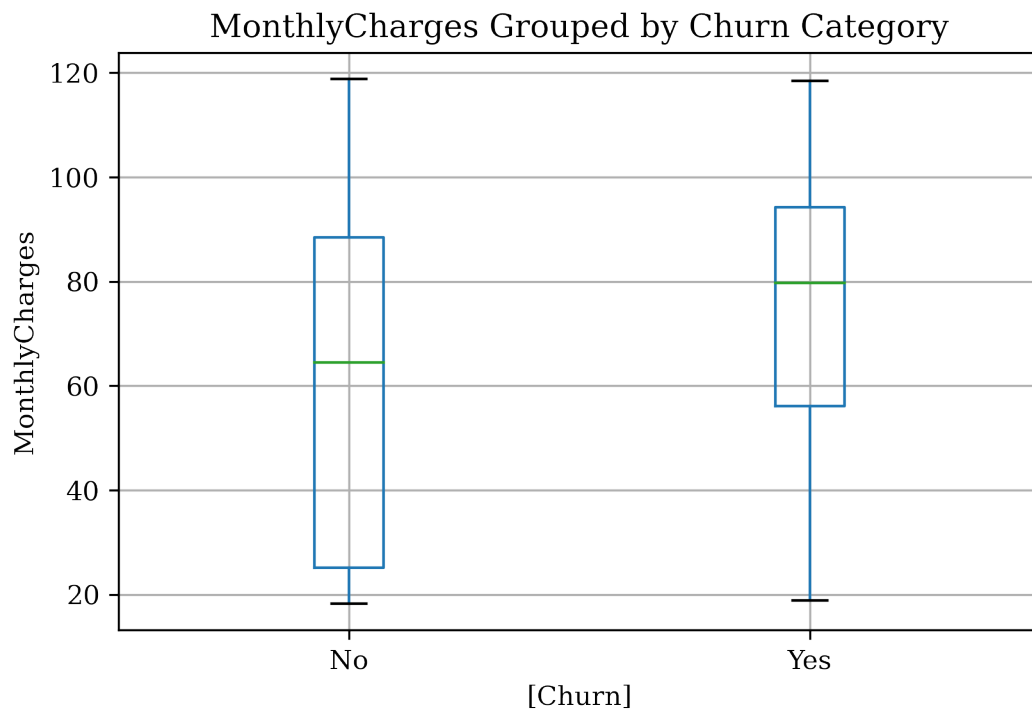
```
Yes    4171
No     2872
Name: PaperlessBilling, dtype: int64
Yes    0.592219
No     0.407781
Name: PaperlessBilling, dtype: float64
```

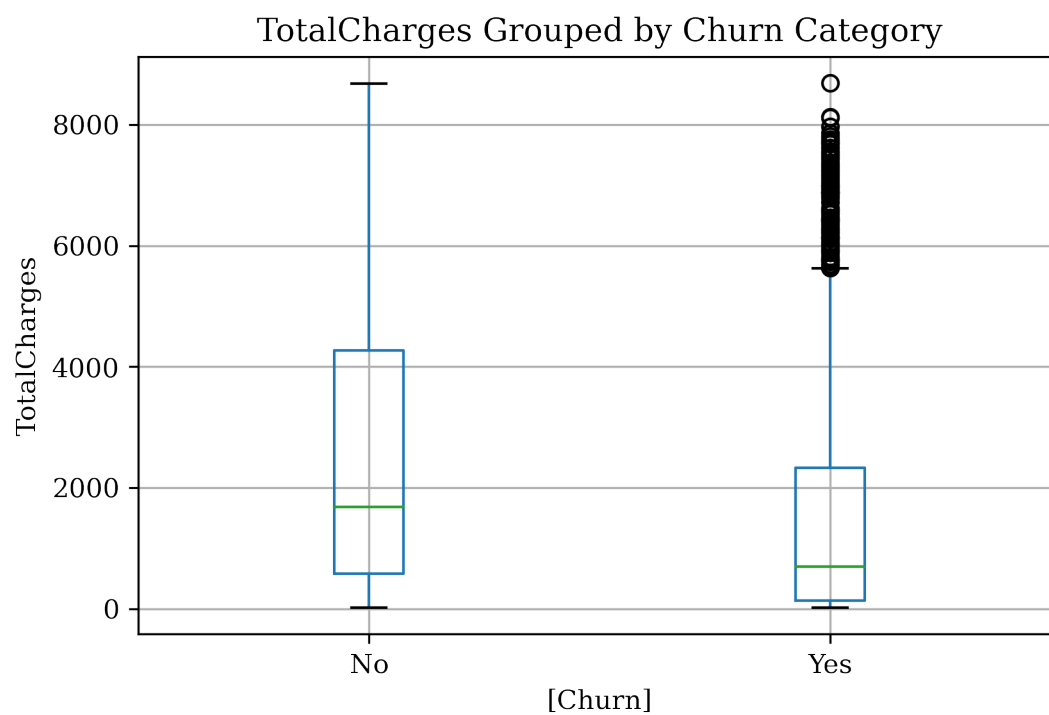
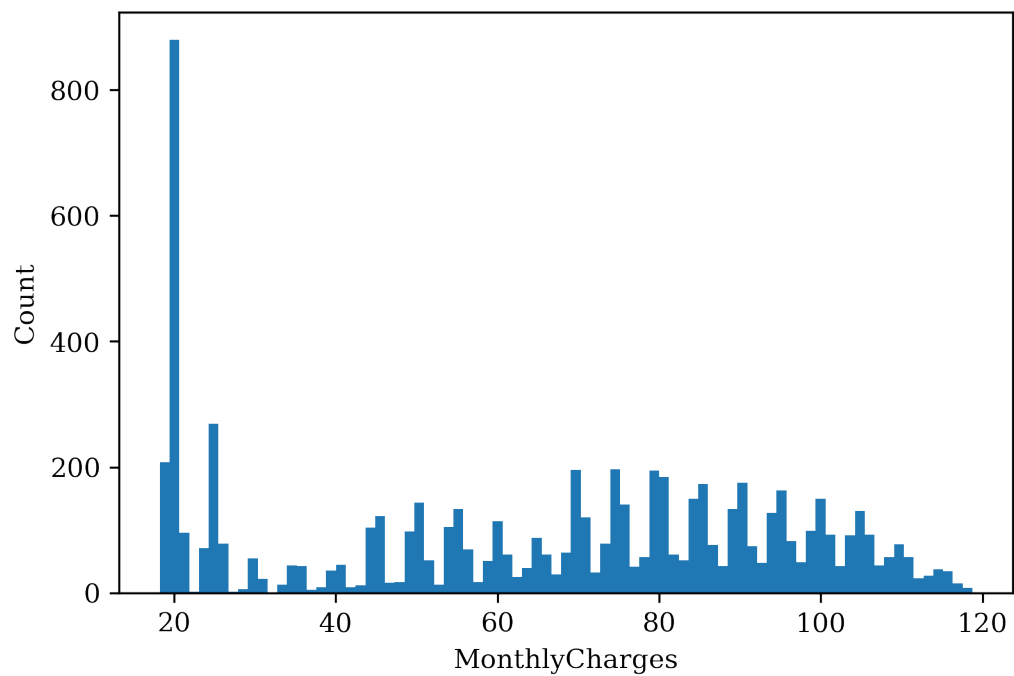
```

Electronic check      2365
Mailed check         1612
Bank transfer (automatic) 1544
Credit card (automatic) 1522
Name: PaymentMethod, dtype: int64
Electronic check      0.335794
Mailed check         0.228880
Bank transfer (automatic) 0.219225
Credit card (automatic) 0.216101
Name: PaymentMethod, dtype: float64

```

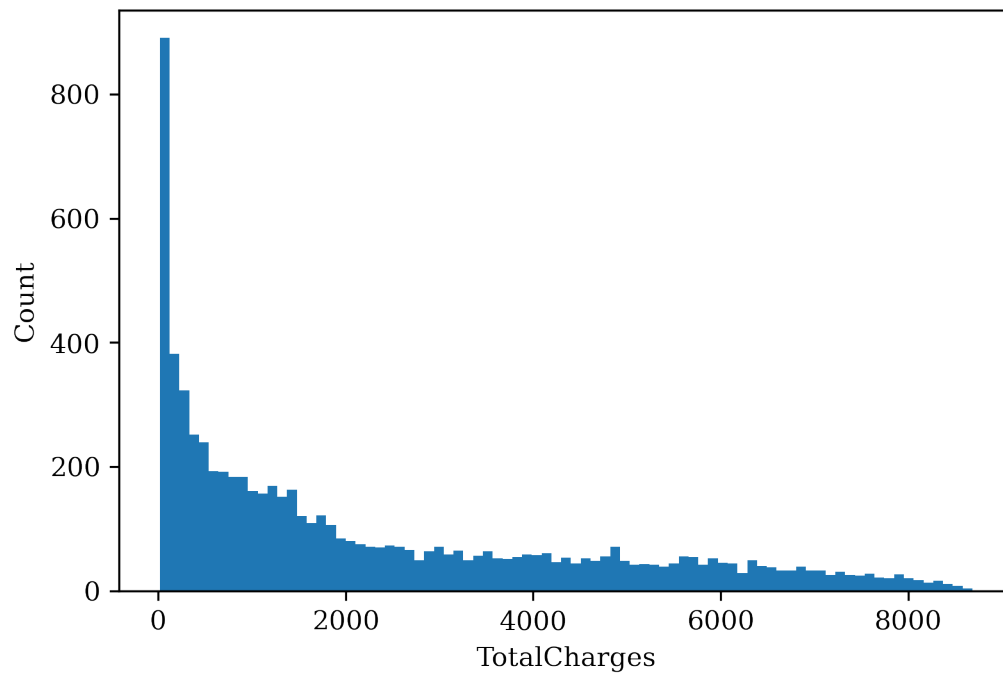


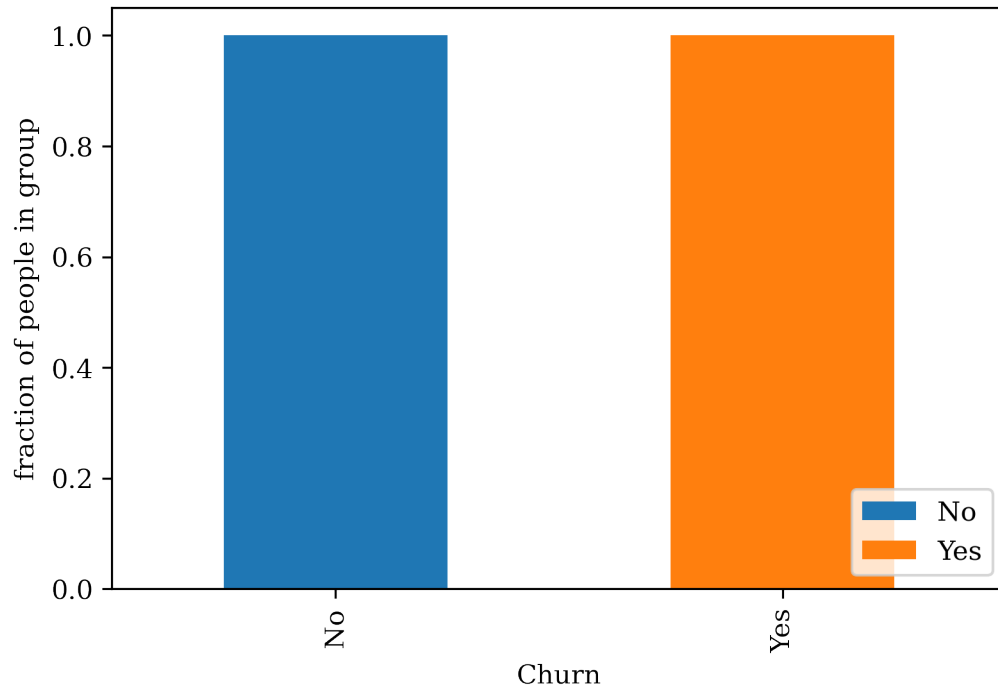
```
count    7043.000000
mean      64.761692
std       30.090047
min       18.250000
25%       35.500000
50%       70.350000
75%       89.850000
max       118.750000
Name: MonthlyCharges, dtype: float64
```



```
count    7032.000000
mean     2283.300441
```

```
std      2266.771362
min       18.800000
25%      401.450000
50%     1397.475000
75%     3794.737500
max     8684.800000
Name: TotalCharges, dtype: float64
```





```
No      5174
Yes      1869
Name: Churn, dtype: int64
No      0.73463
Yes      0.26537
Name: Churn, dtype: float64
```

```
[215]: # Only three continuous variables for scatter matrix

df_continuous = df[['tenure', 'MonthlyCharges', 'TotalCharges']]

pd.plotting.scatter_matrix(df_continuous, figsize=(9, 9),
    ↪marker='o', hist_kws={'bins': 50},
                                s=30, alpha=.1)

plt.show()
```