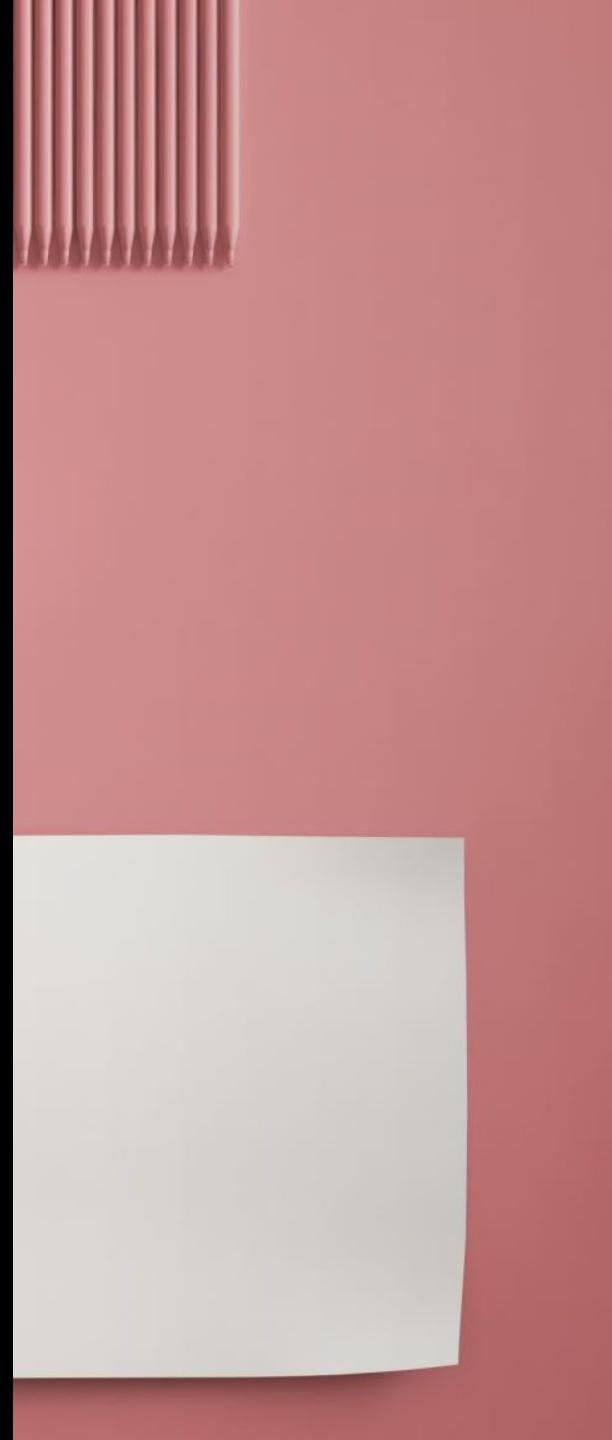# Churning

Analysis

Data Description

- It contains 21 columns and 7043 rows.
- Our Dependent variable is "Churn", which consist of 2 variables, "Yes" and "No".
- There are 11 missing values in dataset.

STEPS

- ■ Handling the NA values.

- ■ Plotting graphs of relevant columns to get useful insights.

- ■ Performing categorical encoding to the columns.

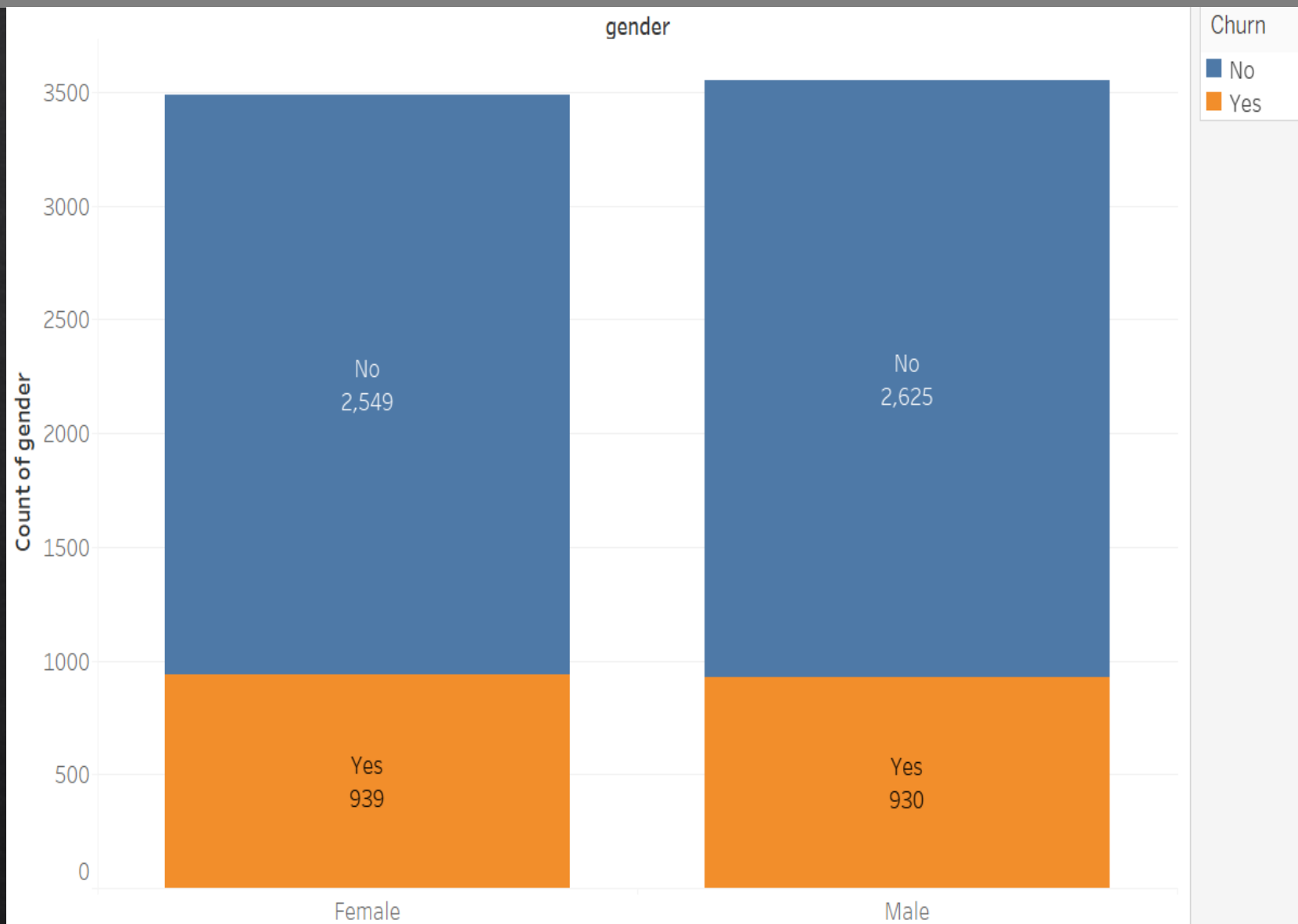- ■ Applying the algorithms of supervised learning.

Categorical Encoding

To make our model efficient.

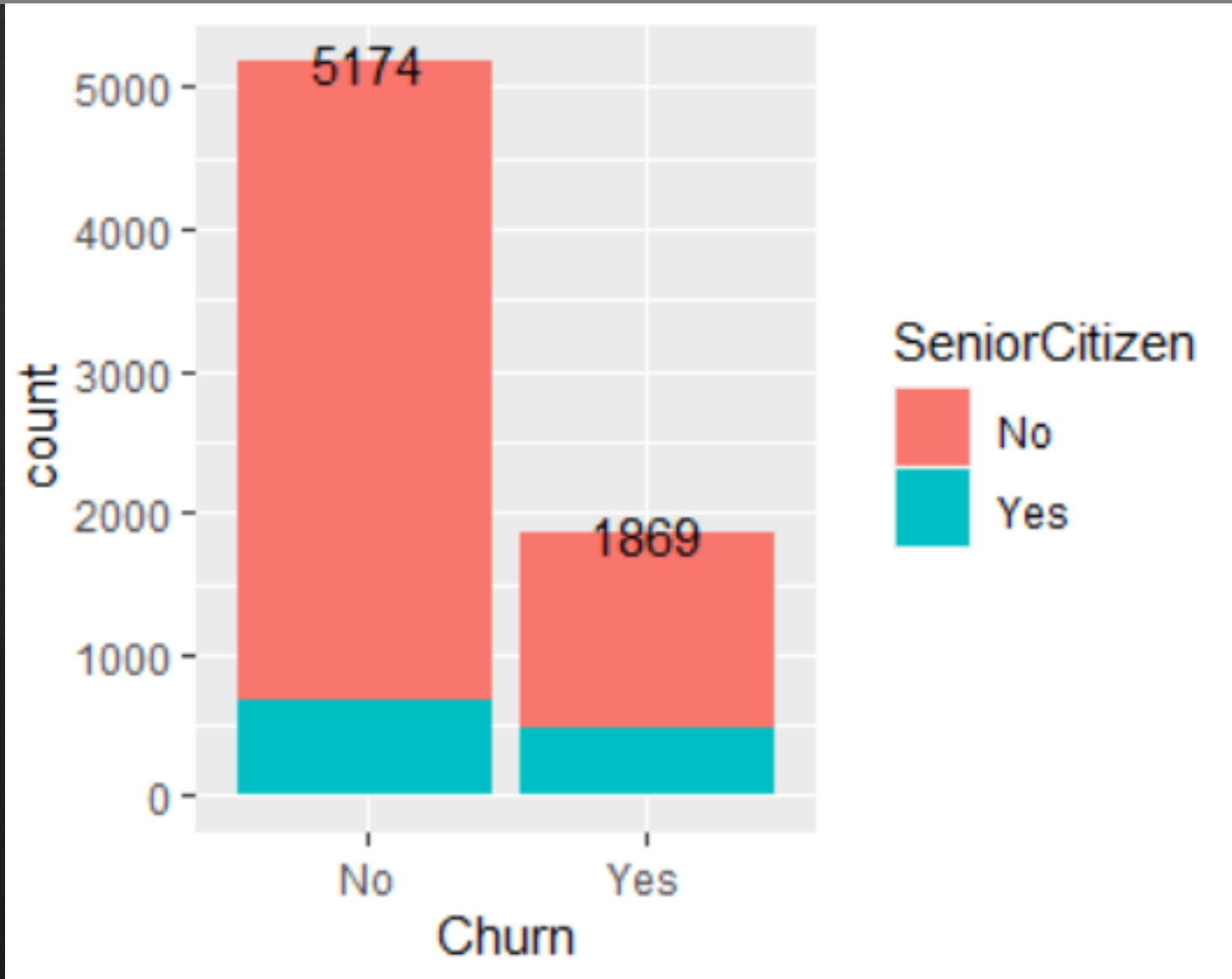◈Replacing category variable as "No internet service" is being used with different names in many columns.
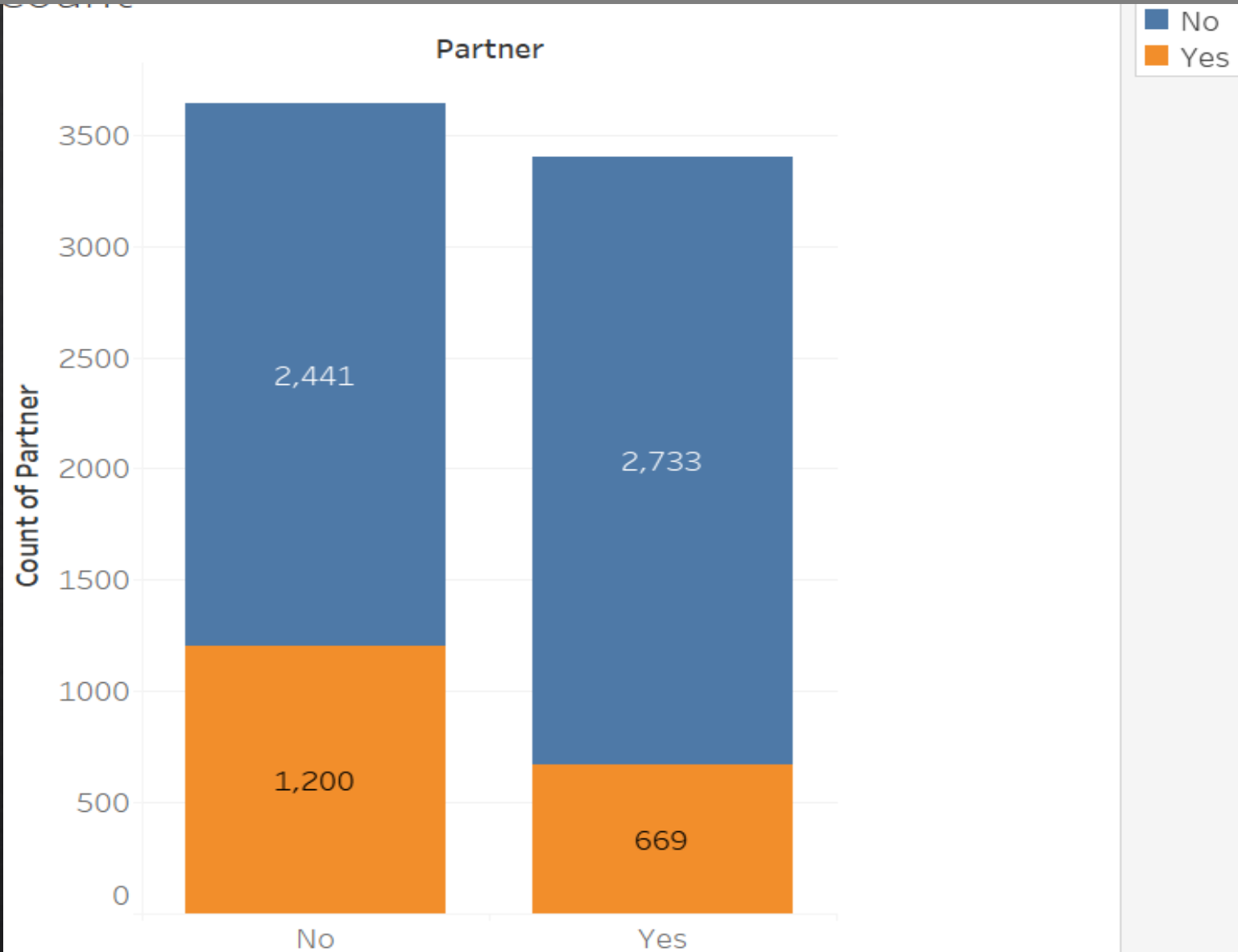
Count of Churn in the dataset
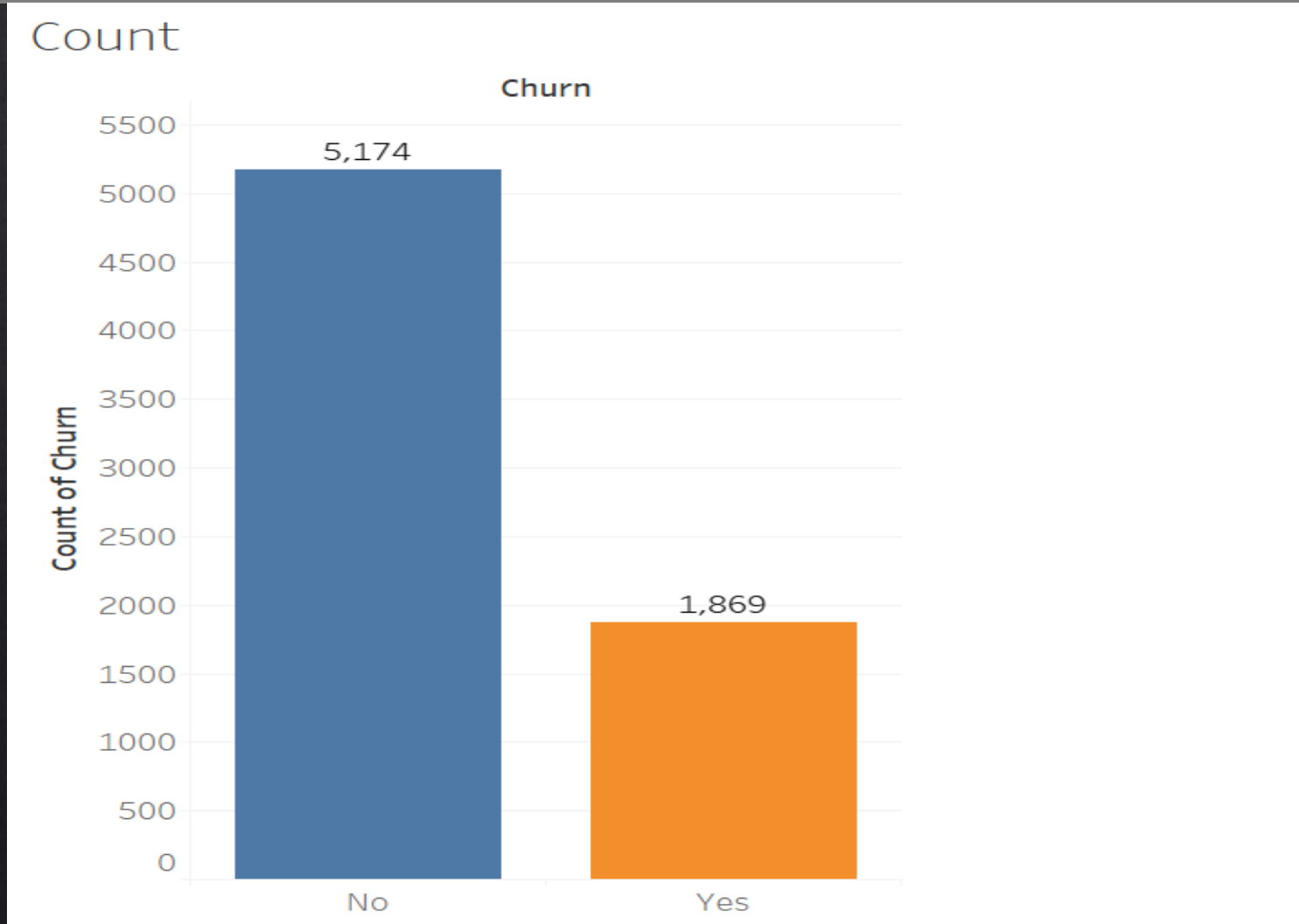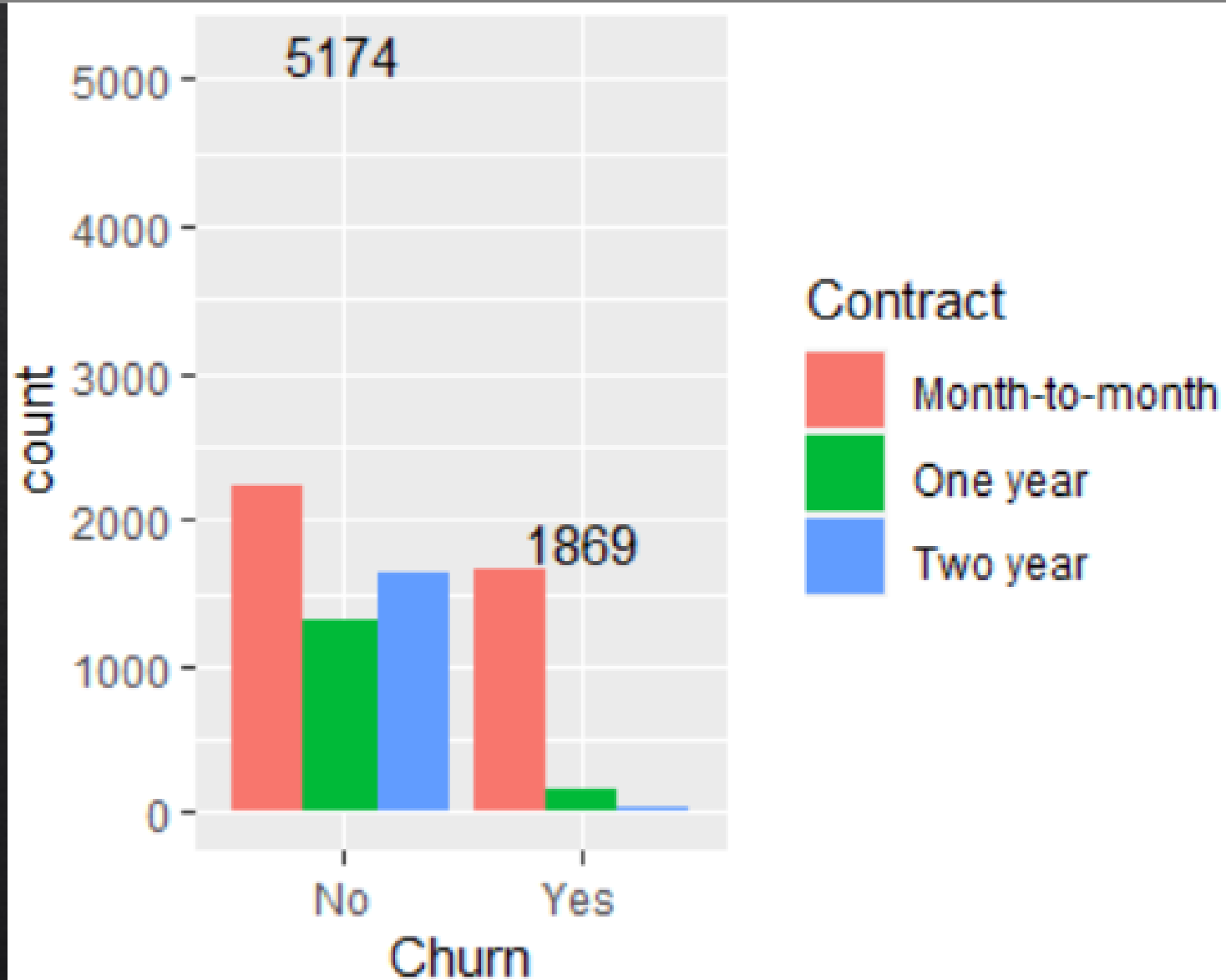
Count of Churn based on gender

Count of Churn based on being a senior citizen

Count of Churn based phone services

Count of Churn based on contract

RANDOMLY SHUFFLING DATA TO AVOID BIASNESS

Sub setting into training and test data set

Training data contains 70% of actual dataset.

Test data contains the 30% of actual dataset.

Applying ALGORITHMS

Using Logistic Regression

Accuracy=79.5%

Using SVM

Accuracy=80%

Using Decision trees

Accuracy=80.9%