Analysis made during Data Manipulation, from plots

1. Data cleaning including missing values, outliers and multi-collinearity. Describe your predictive churn model. How did you select variables to be included in the model?
   ⇨ We found missing values and eliminated them from our data set, and observing boxplots we see there are no outliers and upon checking for multi-collinearity we saw that TotalCharges, tenure, PhoneService, MonthlyCharges are highly correlated. Removing TotalCharges multi-collinearity decreased up to certain extent. Based on these deductions we choose our predictors.

2. Demonstrate the performance of the model.
   ⇨ We have built and trained few models and found that accuracy is around 75-80%.

3. What are the key factors that predict customer churn? Do these factors make sense?
   ⇨ Customers with a month-to-month connection have a very high chance to churn if they have subscribed to pay via electronic checks and having Fiber Optical.
   ⇨ People with no online security and no device protection are prone to churn.

4. What offers should be made to which customers to encourage them to remain with company?
   ⇨ As per my assumptions providing good offers to people who are paying on monthly basis can make them to continue the service.
   ⇨ Providing Online Security either at low cost or for free can make some change.
   ⇨ Same goes with Device Protection.
   ⇨ People generally expects many features at low cost.