Summary

1. We have used logistic regression model to make predictions.
2. EDA indicates following:
    a. The rate of conversion for leads obtained by the business is average 30-35%. We are building the model to improve the conversion rate by predicting which leads are more likely to convert.
    b. We have large amount of data with lot of features recorded. However, a number of features have missing data. Additionally some features have data as 'Select'. It is understood that this is result of no value being selected in an drop down menu.
    c. Highest number of leads are from India, with Mumbai being the most common city.
3. Upon analysis, following is observed:
    a. Biggest factors affecting conversion are Time spent on website and number of visits on the website.
    b. Housewives are most frequent conversions.
    c. Phone calls are most effective method for conversion.