<u>Summary</u>

1. Lead scoring was performed using the logistic regression modelling on the data
   a. Data was cleaned; with removal of columns having more than 40% missing values
   b. Scaling performed using MinMaxScaler
   c. Training/ Test Data split of 70:30
   d. RFE was performed to identify the top 20 variables; with some feature elimination using VIF and p-values
   e. Arrived at Optimal cutoff of 0.4 ; Using the ROC curve & Precision Recall tradeoff
   f. Training Data – model metrics:
      i. Accuracy: 82%
      ii. Precision: 74%
      iii. Recall: 78%
   g. Test Data -model metrics:
      i. Accuracy: 80%
      ii. Precision: 72%
      iii. Recall: 76%
2. Higher Chances of conversion when the lead attributes are:
   - Current Occupation is Working professionals
   - Total time spent on the website
   - Total no of visits
   - Leads who have come from (Lead source or lead origin)
     o Welingak website
     o Lead Add Form
     o Olark Chat
   - Where last activity was a phone conversation with the leads