Summary:

1. Observed the data and started analysing the data keeping in mid the problem statement and the business solution needed.
2. The approach was to make a logistic regression model to meet the constraints as per the requirement.
3. The lead conversion success rate was 37.5%. Around 62.5% of the leads are not converted. So in order to achieve a good conversion rate, we started with data sanitization , then did EDA and we started with data preparation for model building.
4. While doing data sanitization we removed columns with ‘Select’ as values , as they are values when nothing is selected.
5. After EDA we removed few columns as there were null values and we kept the cut off as 38 %.
6. Then we did data preparation (created dummies) for the model building .
7. Then started with model building , and depending upon the P value and VIF we removed few columns, once we got the appropriate pvalue (0) and VIF(less than 2) we finalised the model and started with analysing the data with train and test data .
8. We checked for Accuracy, Specificity and sensitivity of the model . Which are in desirable limits, so we concluded that the model is a good fit.
9. We found out the below are important features.
   1. Lead\_Source\_Welingak Website 2.508447
   2. Occupation\_Working Professional 2.333358
   3. Lead\_Origin\_Lead Add Form 2.289549
   4. Last\_Notable\_Activity\_Unreachable 2.034160
   5. Last\_Notable\_Activity\_SMS Sent 1.395211
   6. Time\_on\_Website 1.055896
   7. const 0.149832
   8. Lead\_Source\_Google -1.072325
   9. Occupation\_Other -1.153456
   10. Last\_Activity\_Converted to Lead -1.180340
   11. Lead\_Source\_Organic Search -1.193070
   12. Do\_Not\_Email -1.282749
   13. Lead\_Source\_Direct Traffic -1.443261
   14. Last\_Activity\_Olark Chat Conversation -1.514075
10. And conclude as follows.
    1. if the lead source is form Welingak Website then we should make a call .
    2. Last notable activity is not reachable or sms sent we need to contact them.
    3. Current Occupation is Working professional we need to contact them.
    4. Lead origin is from Lead Add Form , then we need to contact.
    5. if the lead spends more time on the website , then we need to contact them.
    6. if occupation is others, lead source is from google, last activity converted to lead,lead source organic Search then need not contact.
    7. if its do not email need not contact.