







Objective

• Achieve 80% of total responders at the minimum possible cost

Problem Statement

• Predict the probability of response and target most likely respondents in the telemarketing campaign without "duration" variable being considered while building model

Approach

- EDA, data preparation and model building
- Assumed call cost to be 1Rs per min
- A unique ID is given to each prospect for better understanding and analysis of data





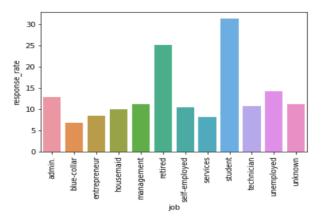
Steps followed

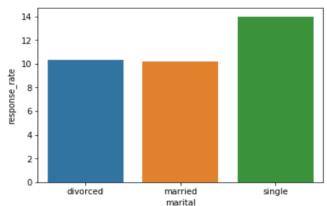
- Model Building –Logistic Regression Model, without using "duration" variable
- Logistic Regression with all variables, PFE and PCA
 - LR with PCA gives the best model
- Identifying the top X% prospect customers that needs to contacted to achieve business objective
- Creating a Lift and Decile Chart
- Identifying the Cost of Acquisition
- Since our objective is to identify True positive rate, we need to mainly focus on "Sensitivity"

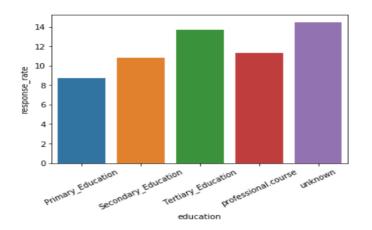


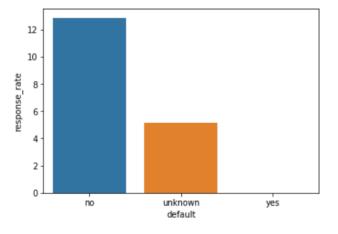


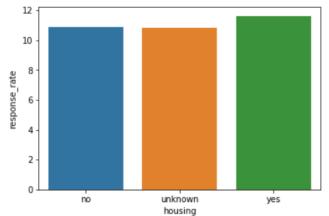
Relationship of different variable with Response

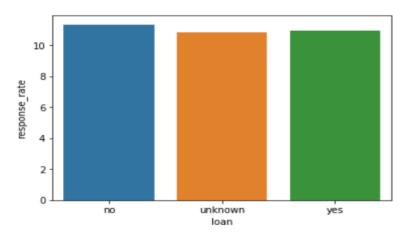








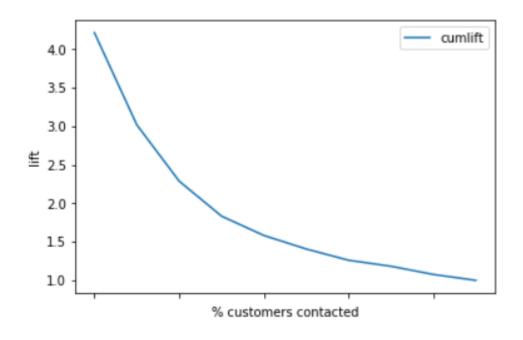


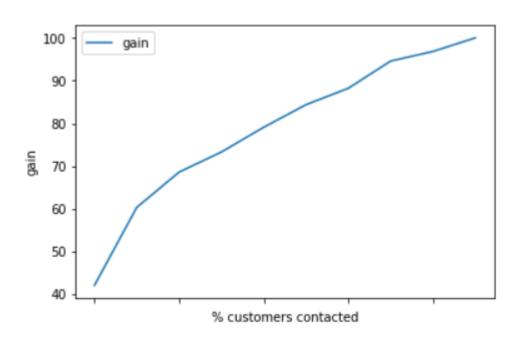






Model Performance – Lift Charts









Cost of Acquisition for 80% response rate

- Cost to be considered = 1*number of contacts made in the current campaign
- Cost= 1* (50% of 41,188) = Rs 20,594/-
- To acquire 80% base we need to contact 50% people





Results

- To achieve our objective of acquiring 80% of total responders at the minimum possible cost; we will need to target 50% people from entire data set.
- Variable to watch out during acquisitions are:

```
# job retired
```

month mar

poutcome success

job student

month_may

cons.price.idx

contact telephone

previous_Nevercontacted

euribor3m

• With model we have achieved 50% efficiency. Instead of calling the entire set of people we just weed need to call 50% people thereby saving money that would otherwise be wasted by calling rest 50% data.