CASE STUDY: PROBLEM STATEMENT
Group 11
Reema Yadav and Sakshi Patil

1. **Background Information**

   The data used is related to direct marketing campaigns of Portuguese bank. Marketing campaigns are vital for all the banks to progress. The given data for marketing campaign is based on the phone calls. The aim of this project is to analyze the importance of marketing in the banking sector and significance of phone calls in it.

2. **Data Collection**

   Source of data is https://archive.ics.uci.edu/ml/datasets/Bank+Marketing#

   Attribute information:

   age (numeric)
   job : type of job (categorical)
   marital : marital status (categorical)
   education (categorical)
   default: has credit in default? (categorical)
   housing: has housing loan? (categorical)
   loan: has personal loan? (categorical)
   contact: contact communication type (categorical)
   month: last contact month of year (categorical)
   day_of_week: last contact day of the week (categorical)
   duration: last contact duration, in seconds (numeric).
   campaign: number of contacts performed during this campaign and for this client (numeric)
   pdays: number of days that passed by after the client was last contacted from a previous campaign (numeric)
   previous: number of contacts performed before this campaign and for this client (numeric)
   poutcome: outcome of the previous marketing campaign (categorical)
   emp.var.rate: employment variation rate - quarterly indicator (numeric)
   cons.price.idx: consumer price index - monthly indicator (numeric)
   cons.conf.idx: consumer confidence index - monthly indicator (numeric)
   euribor3m: euribor 3 month rate - daily indicator (numeric)
   nr.employed: number of employees - quarterly indicator (numeric)
   y - has the client subscribed a term deposit? (binary)

3. **The Problem**

   Based on the input of the client, goal is to predict if the consumer will subscribe (yes/no) a term deposit (variable y)

4. **Possible Solution**

   Different Supervised Machine learning techniques can be applied over here like decision tree and Random forest. We will be using different king of regressions on the dataset.