

# Bank Marketing (Campaign)

- I have done final project individually
- Name : Mohammad Aminian
- Email: aminanmuhamad24@gmail.com
- Country : Iran
- College/Company : Tehran Kharazmi University
- Specialization : Data Science

# Problem description

ABC Bank wants to sell its term deposit product to customers and before launching the product they want to develop a model which helps them in understanding whether a particular customer will buy their product or not (based on customer's past interaction with bank or other Financial Institution).

# Data understanding

ABC Bank wants to get clear understanding of its customers and predict customers response

By analyzing their features.find the customers who are more likely to subscribe to term deposits.

# Data understanding

There are 4 Data sets Provided in links:

- bank-additional-full.csv with 41188 rows and 20 columns that is our desired data set to do EDA with.
- bank-additional.csv
- bank-full.csv
- bank.csv

# Data understanding

Data set columns are:

- 1 - age (numeric)
- 2 - job : type of job (categorical: 'admin.', 'blue-collar', 'entrepreneur', 'housemaid', 'management', 'retired', 'self-employed', 'services', 'student', 'technician', 'unemployed', 'unknown')
- 3 - marital : marital status (categorical: 'divorced', 'married', 'single', 'unknown'; note: 'divorced' means divorced or widowed)
- 4 - education (categorical: 'basic.4y', 'basic.6y', 'basic.9y', 'high.school', 'illiterate', 'professional.course', 'university.degree', 'unknown')
- 5 - default: has credit in default? (Categorical: 'no', 'yes', 'unknown')
- 6 - housing: has housing loan? (Categorical: 'no', 'yes', 'unknown')
- 7 - loan: has personal loan? (categorical: 'no', 'yes', 'unknown')
- # related with the last contact of the current campaign:
- 8 - contact: contact communication type (categorical: 'cellular', 'telephone')

# Data understanding

- 9 - month: last contact month of year (categorical: 'jan', 'feb', 'mar', ..., 'nov', 'dec')
- 10 - day\_of\_week: last contact day of the week (categorical: 'mon', 'tue', 'wed', 'thu', 'fri')
- 11 - duration: last contact duration, in seconds (numeric). Important note: this attribute highly affects the output target (e.g., if duration=0 then y='no'). Yet, the duration is not known before a call is performed. Also, after the end of the call y is obviously known. Thus, this input should only be included for benchmark purposes and should be discarded if the intention is to have a realistic predictive model.
- # other attributes:
- 12 - campaign: number of contacts performed during this campaign and for this client (numeric, includes last contact)
- 13 - pdays: number of days that passed by after the client was last contacted from a previous campaign (numeric; 999 means client was not previously contacted)
- 14 - previous: number of contacts performed before this campaign and for this client (numeric)

# Data understanding

- 15 - poutcome: outcome of the previous marketing campaign (categorical: 'failure','nonexistent','success')
- # social and economic context attributes
- 16 - emp.var.rate: employment variation rate - quarterly indicator (numeric)
- 17 - cons.price.idx: consumer price index - monthly indicator (numeric)
- 18 - cons.conf.idx: consumer confidence index - monthly indicator (numeric)
- 19 - euribor3m: euribor 3 month rate - daily indicator (numeric)
- 20 - nr.employed: number of employees - quarterly indicator (numeric)
- 
- Output variable (desired target):
- 21 - y - has the client subscribed a term deposit? (Binary: 'yes','no')

# Data understanding

The data doesn't have any missing values.

Some columns have outliers that we get to dive into them in EDA in next weeks (This is Our method for overcome problems like NA value, outlier).

The data is in CSV format.