# Data Set Information

The data is related to direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be subscribed ('yes') or not ('no') subscribed.

There are two datasets: train.csv with all examples (32950) and 21 inputs including the target feature, ordered by date (from May 2008 to November 2010), very close to the data analyzed in [Moro et al., 2014]

test.csv which is the test data that consists of 8238 observations and 20 features without the target feature

Goal: The classification goal is to predict if the client will subscribe (yes/no) a term deposit (variable y).

The dataset contains train and test data. Features of train data are listed below. And the test data have already been preprocessed.

# Features

| **Feature** | **Feature\_Type** | **Description** |
| --- | --- | --- |
| age | numeric | age of a person |
| job | Categorical, nominal | type of job ('admin.','blue-collar','entrepreneur','housemaid','management','retired','self-employed','services','student','technician','unemployed','unknown') |
| marital | categorical, nominal | marital status ('divorced','married','single','unknown'; note: 'divorced' means divorced or widowed) |
| education | categorical, nominal | ('basic.4y','basic.6y','basic.9y','high.school','illiterate','professional.course','university.degree','unknown') |
| default | categorical, nominal | has credit in default? ('no','yes','unknown') |
| housing | categorical, nominal | has housing loan? ('no','yes','unknown') |
| loan | categorical, nominal | has personal loan? ('no','yes','unknown') |
| contact | categorical, nominal | contact communication type ('cellular','telephone') |
| month | categorical, ordinal | last contact month of year ('jan', 'feb', 'mar', …, 'nov', 'dec') |
| day\_of\_week | categorical, ordinal | last contact day of the week ('mon','tue','wed','thu','fri') |
| duration | numeric | last contact duration, in seconds . Important note: this attribute highly affects the output target (e.g., if duration=0 then y='no') |
| campaign | numeric | number of contacts performed during this campaign and for this client (includes last contact) |
| pdays | numeric | number of days that passed by after the client was last contacted from a previous campaign (999 means client was not previously contacted) |
| previous | numeric | number of contacts performed before this campaign and for this client |
| poutcome | categorical, nominal | outcome of the previous marketing campaign ('failure','nonexistent','success') |

# ****Target variable (desired output):****

| **Feature** | **Feature\_Type** | **Description** |
| --- | --- | --- |
| y | binary | has the client subscribed a term deposit? ('yes','no') |