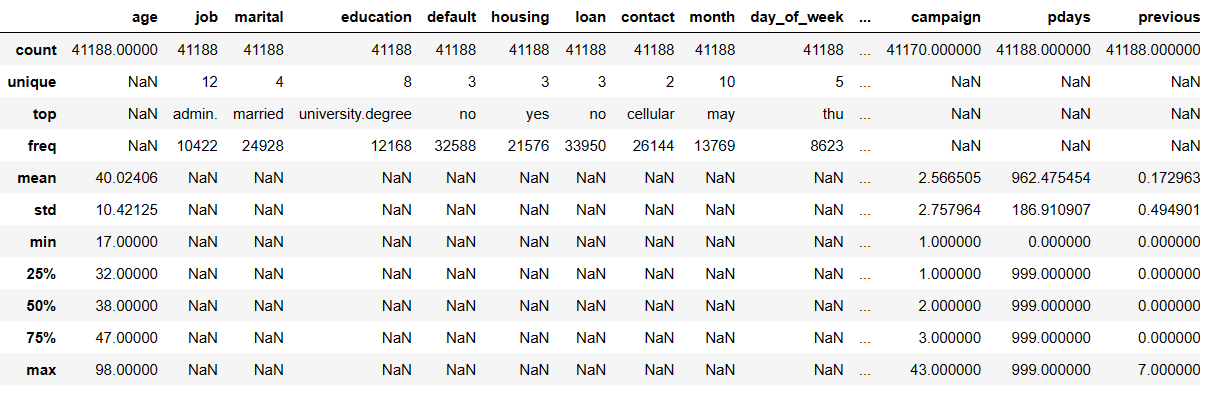
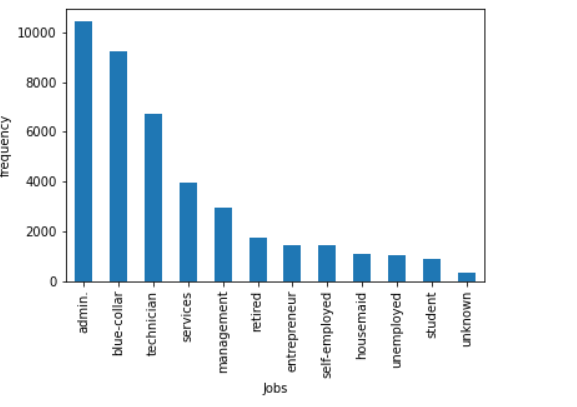
# **Bank Marketing Data Set**

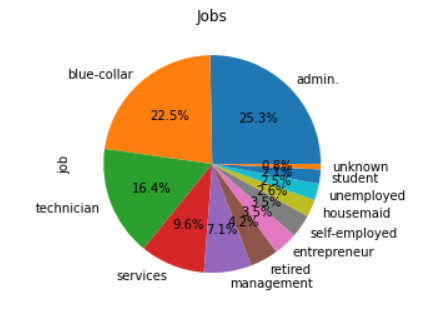
Data source:

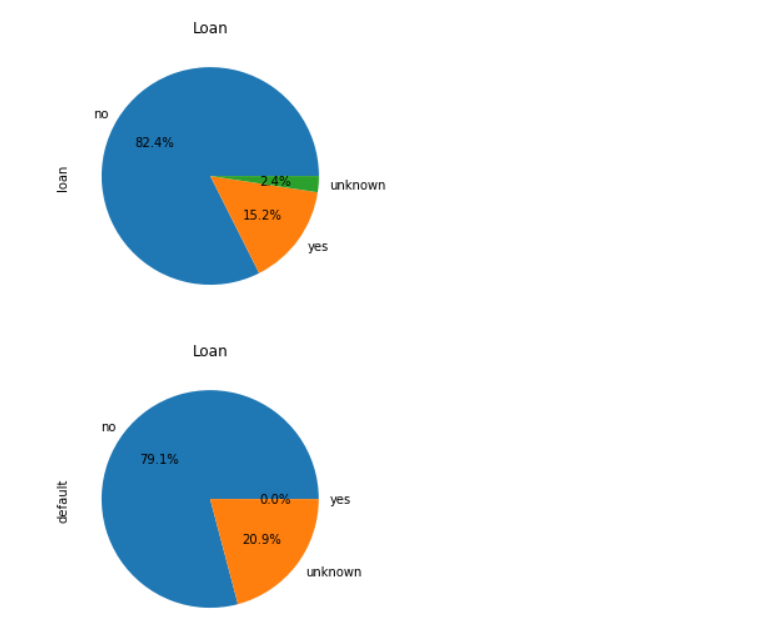
The data is related to direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be ('yes') or not ('no') subscribed. To download the data use this link: <https://archive.ics.uci.edu/ml/datasets/Bank+Marketing>



This data set consists of 21 variables: 20 of them are independent. The variable "y" is a dependent variable. Big amoint of variables consist of character values (categorical variables). The approach to interpreting this variables will be observed later. The data set has quite a big size (6+ MB). Before moving forward, it is necessary to seize the size of the data set







The created graphs show us distribution of categorical variables. What can be noticed:

1. 60% of people contacted are married and have different Job but the most commun ones are; admin, blue-collar , technician and services
2. Variable "Loan" has two levels. Obviously, most of the people contacted 80% does not have personal loan.
3. Most of them has no credit default (79%)

No let's see compare these variables amongs categories (marital, jobs & education)

Git HUB Link