



Data Glacier

Your Deep Learning Partner

Bank Marketing Campaign

SURYA CHANDRA KARANAM

PRANEETHA RAJUPALEPU

SELAELO RAMOKGOPA

07/03/2022

Background

- 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).

- Objective:** ABC Bank wants to use ML model to shortlist customers whose chances of buying the product are more so that their marketing channel (tele marketing, SMS/email marketing etc) can focus only on those customers whose chances of buying the product are more.

The analysis will provide information about the following:

- Business and Data Understanding
- Data Acquisition & Preparation
- Modelling
- Deployment

Business Understanding

The plan is to help ABC company to provide a short list of customer that are more likely to buy their product based on their bank details information such as loan.

Marital status, account balance etc.

This goal will be achievable by using a sophisticated machine learning algorithm capable of using a customer record to predict their future action in a blink of an eye to reduce the company's time and resources.

The success criteria:

The business problem would be based on how much maximum number of customers we are able to predict who have subscribed to the product

Data Exploration

Tabular data details:

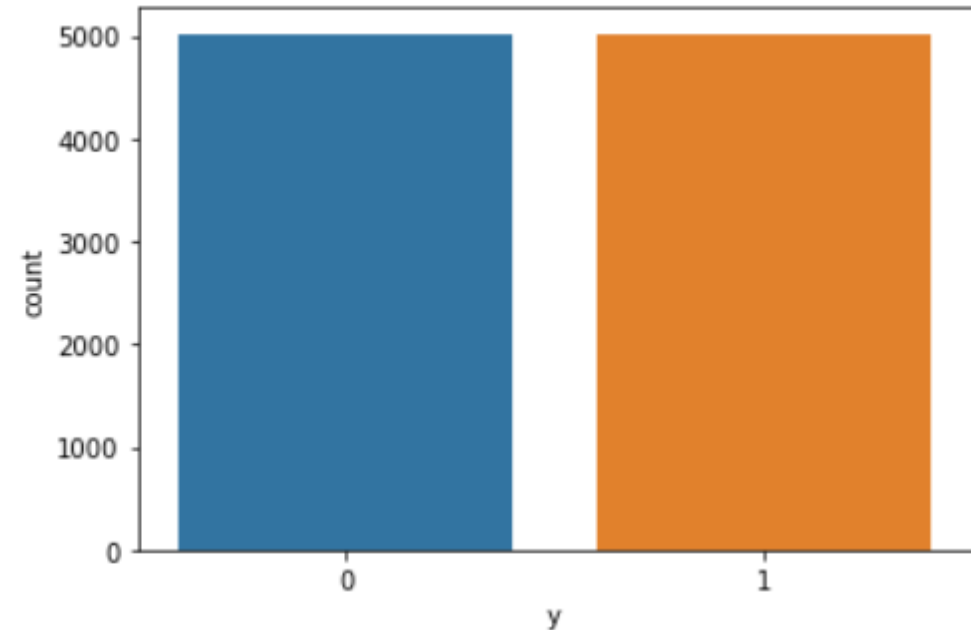
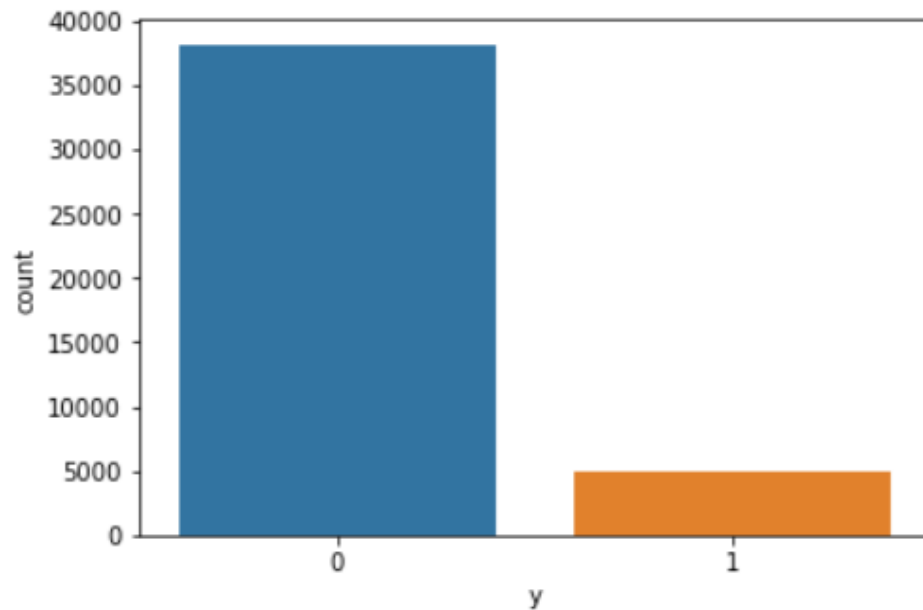
- Total number of observations 41188.
- Total number of files 1.
- Base format of the file .csv Size of the data 4.81 MB.

Approaches:

- Data looks pretty clean and we have unbalanced classes for our target and outliers.
- Categorical imbalance is reduced by eliminating unknown values and also replacing them with mean.
- The dataset is heavily skewed. We looked at a few of the numeric features, and the ones that are skewed will need to be transformed.
- Techniques like IQR score and Weight of Evidence(WOE) and Information Value(IV) are used to deal with outliers and skewness in the data.

Handling Imbalance Target Data

- Target column → Term Deposit subscription (y) is imbalanced.
- Minority class length: 5021 and Majority class length: 38172
- Imbalance is handled using Under sampling technique.



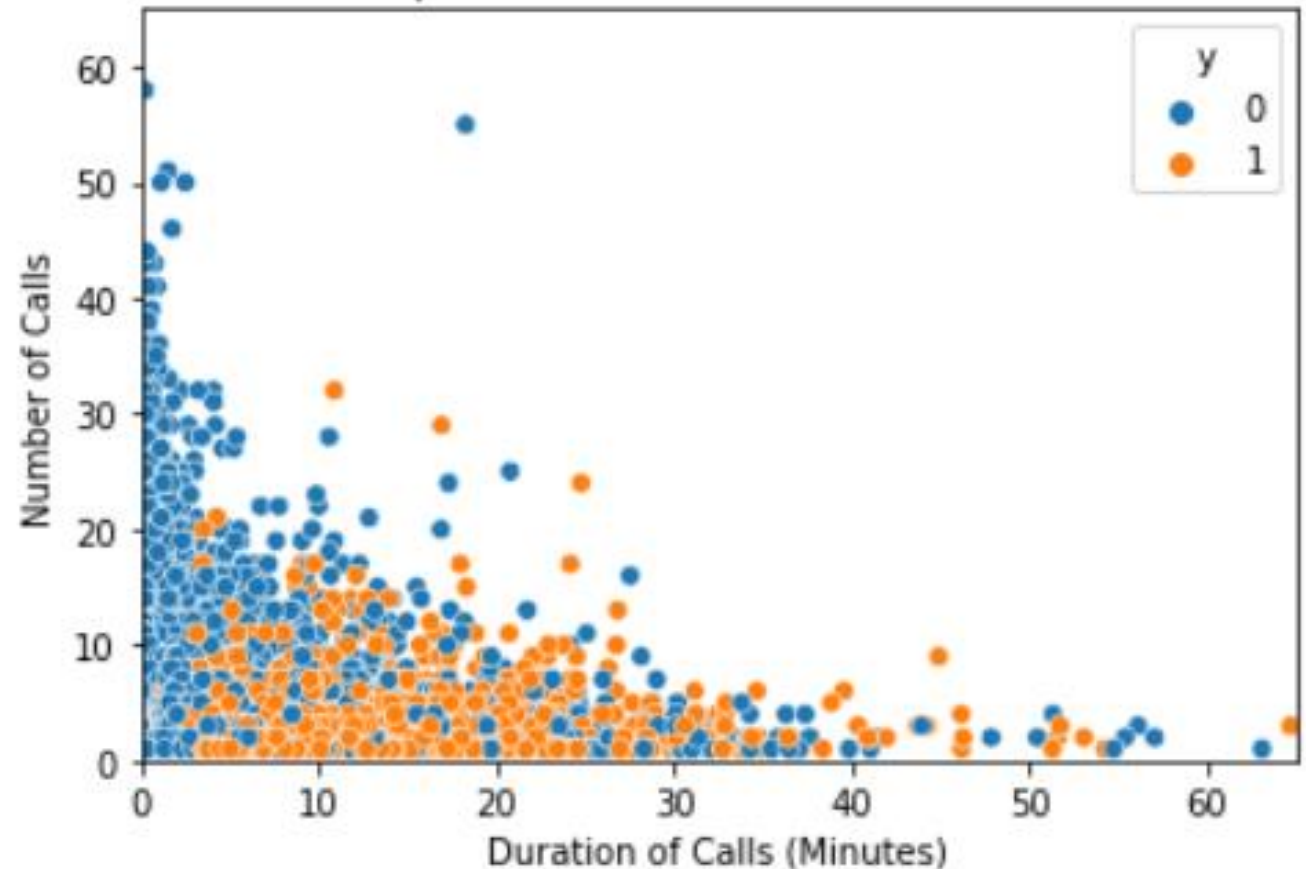
Data Understanding

○ The Relationship between the Number and Duration of Calls:

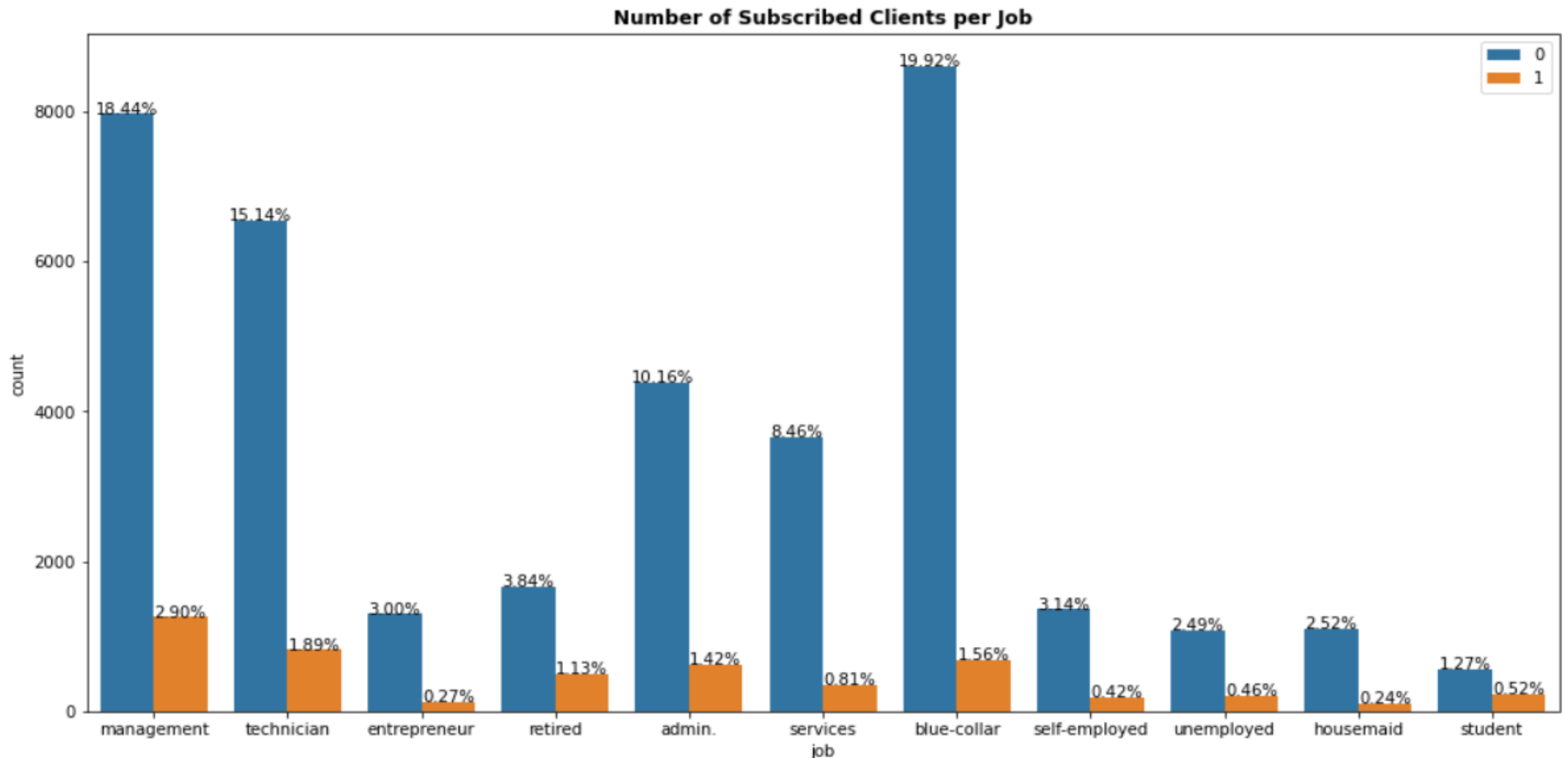
○ According to graph:

Duration and number of calls depends on term subscription:

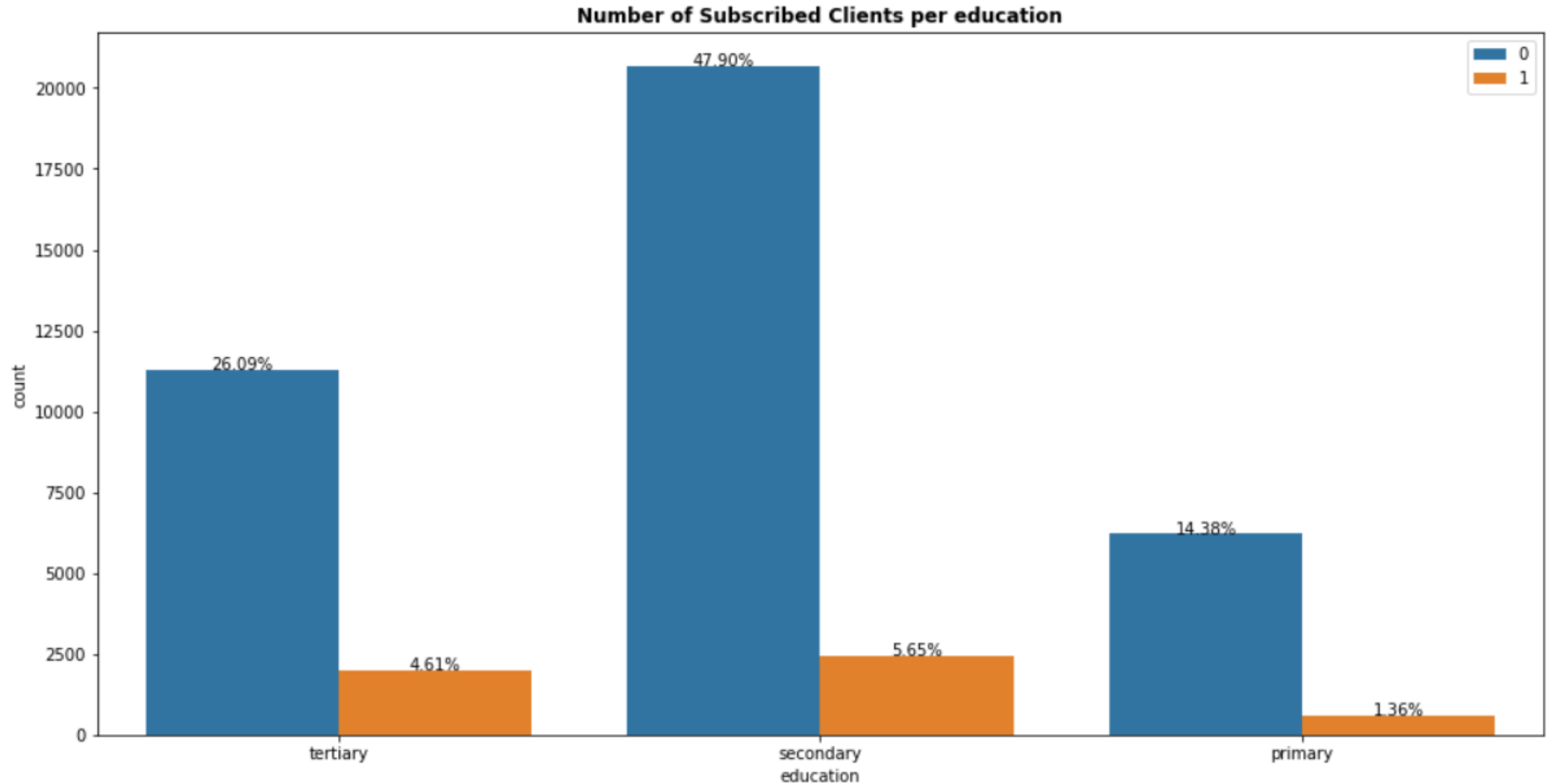
1. Less duration with no subscription
2. 10-20 mins range duration with minimal calls with subscription-based customers.



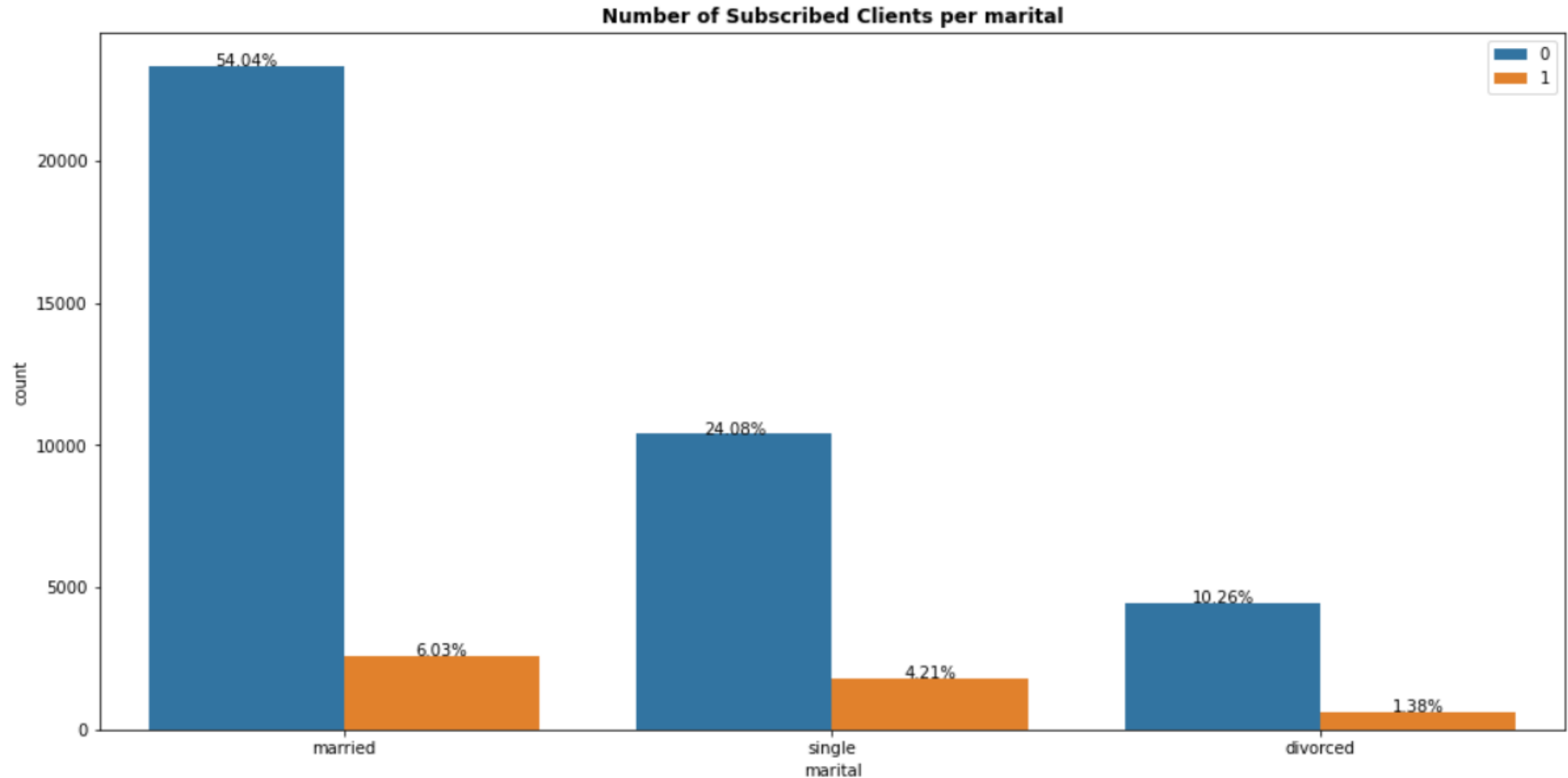
Client subscriptions vs Job



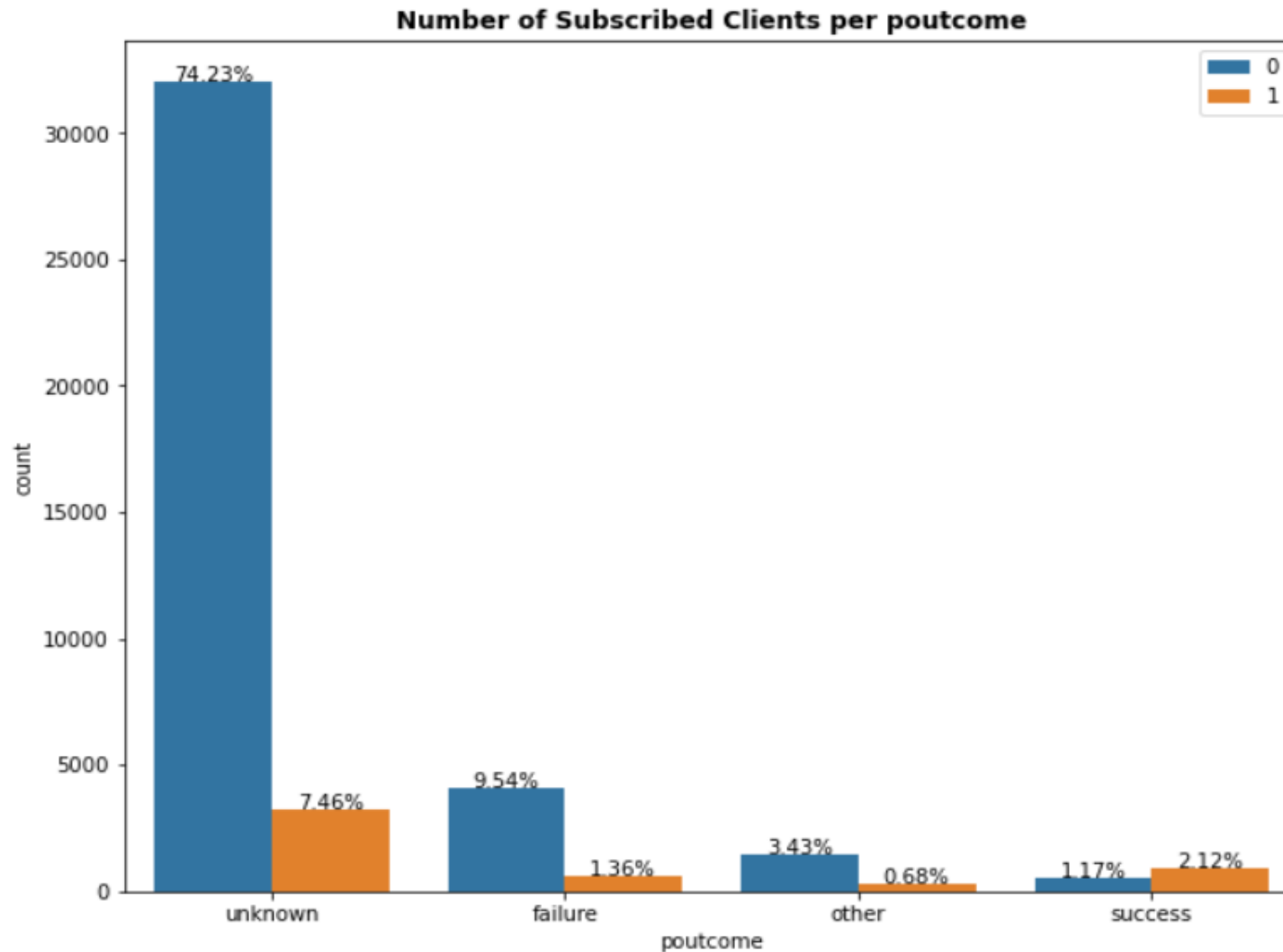
Client subscription vs Education



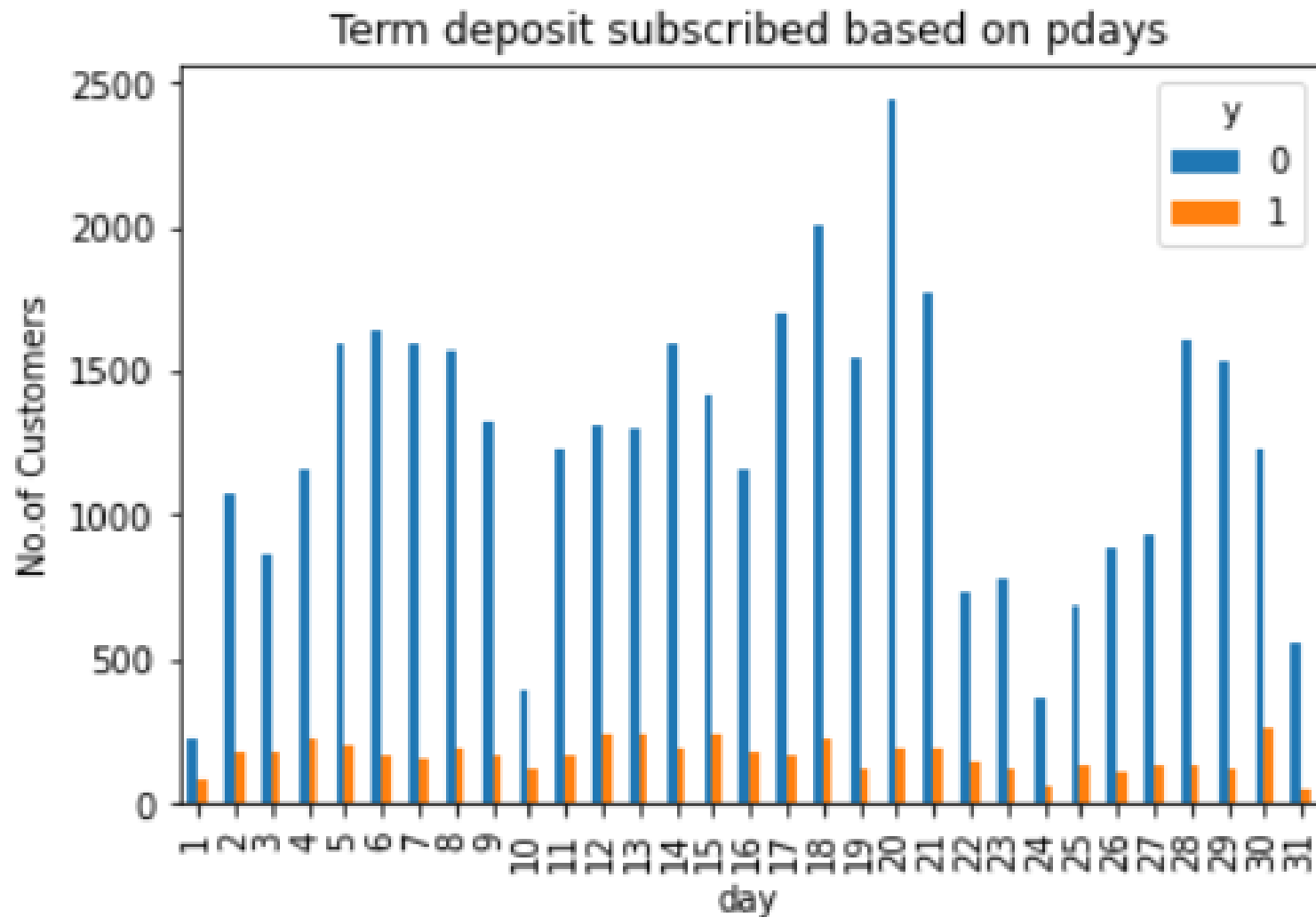
Client subscription vs Marital Status



Client subscriptions vs poutcome



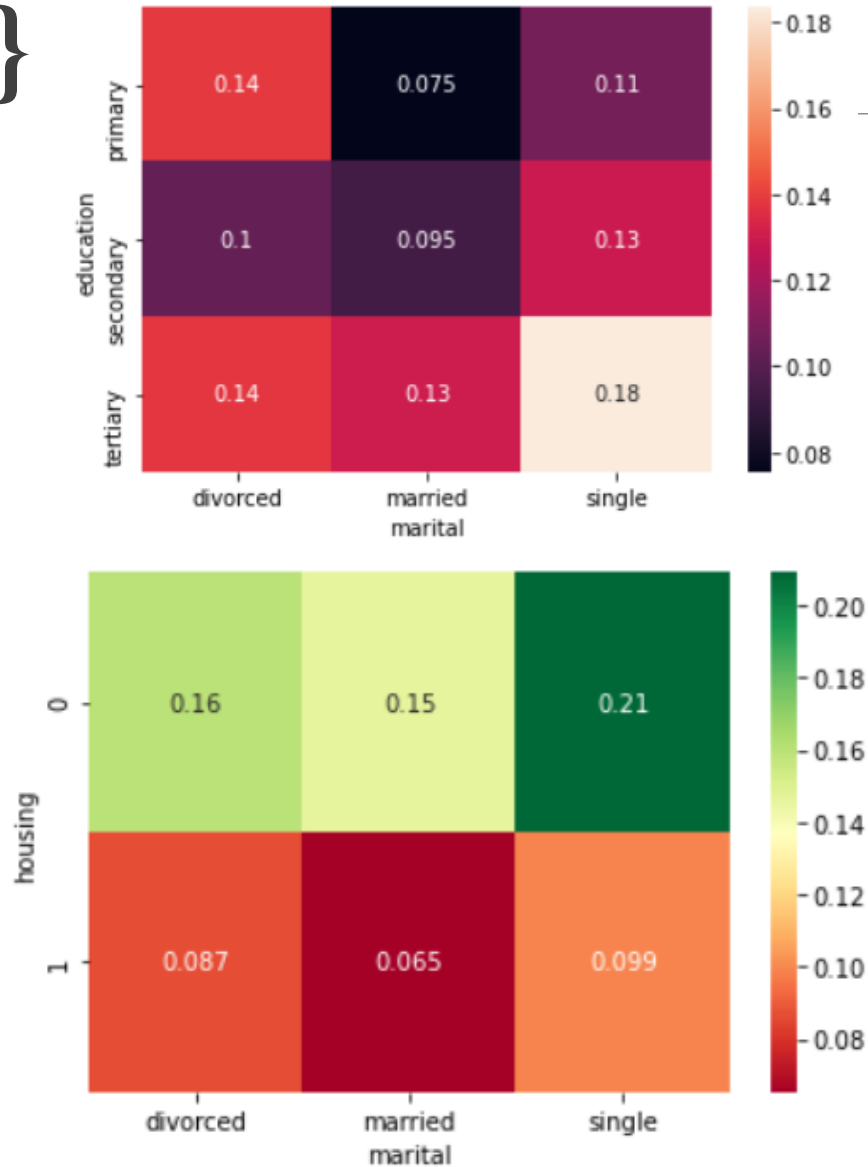
Client subscriptions vs days



Correlation Map of Data

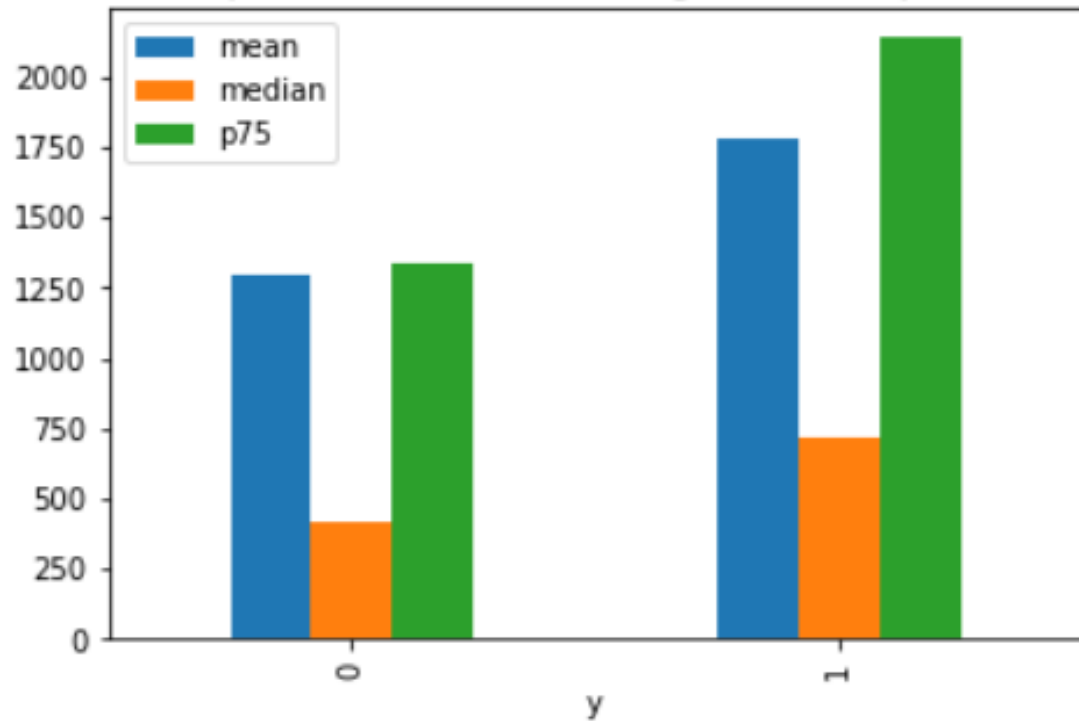


Pivot Tables representation of client subscriptions between marital status and {education, housing, job}

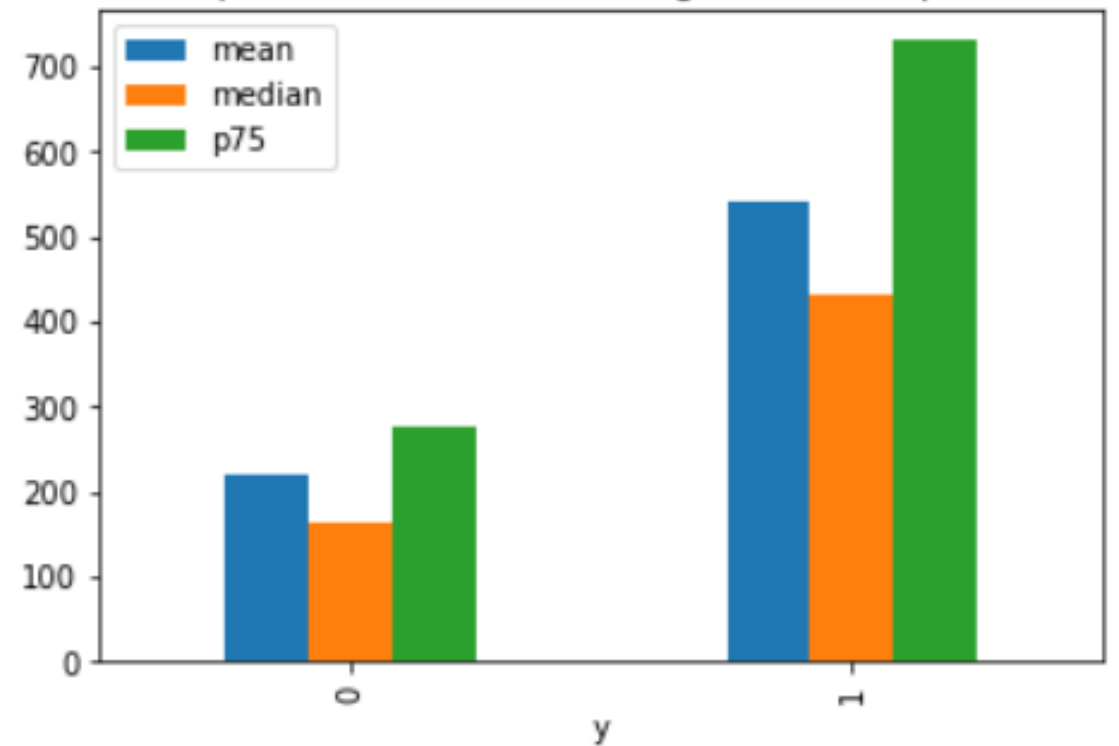


Term deposit estimation according to balance and duration parameters:

term deposit estimation according to balance parameters



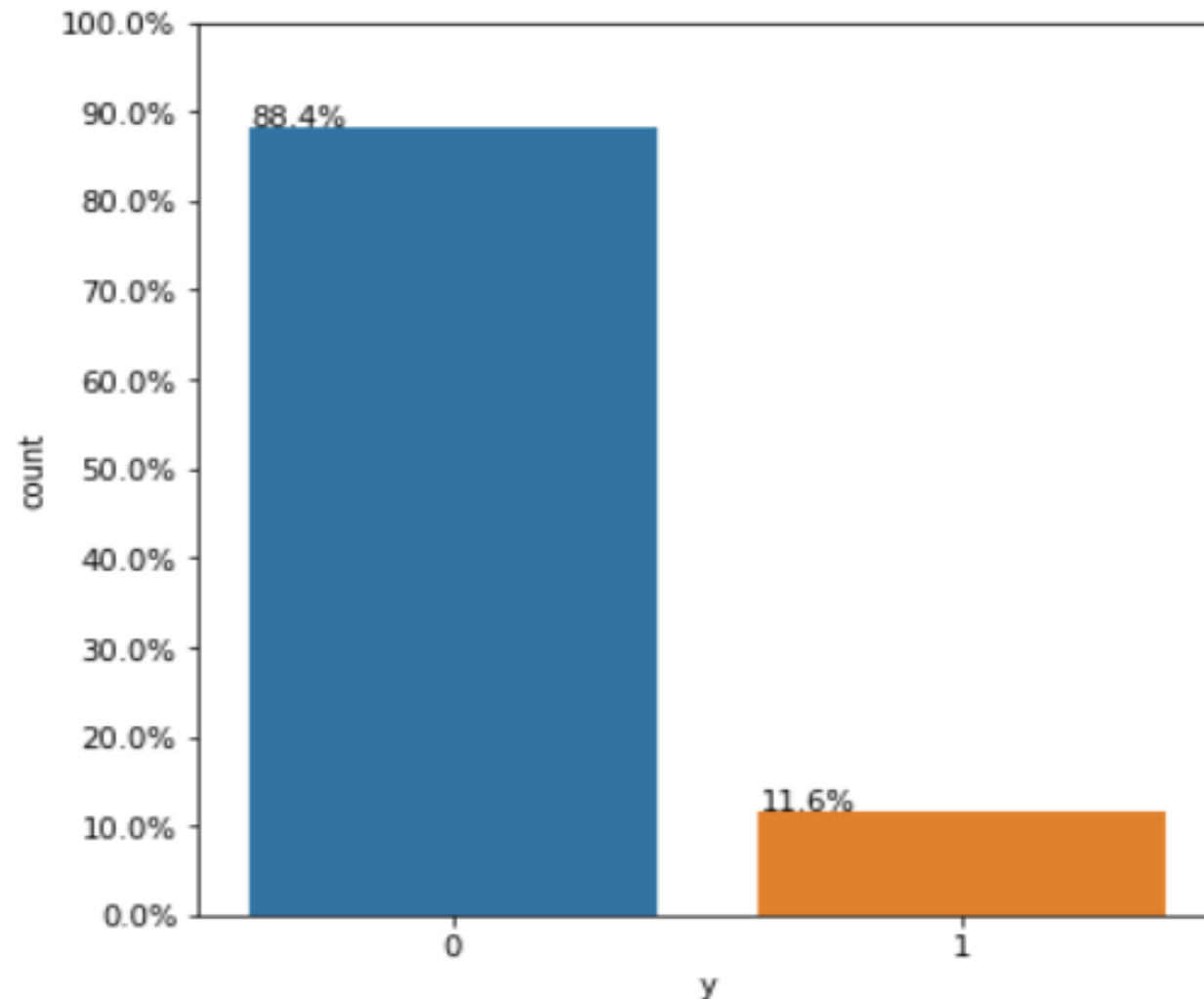
term deposit estimation according to duration parameters



Profit shares by age groups in housing, loan, and education



Client subscriptions total count (in %)



Recommended Models

- Logistic regression.
- Naïve Bayes Classifier.
- Gradient Boosting or Gradient Descent Classifier.

Note: Data analysis and data preparation and data modelling and model building (i.e., the above-mentioned models) model selection depends on performance reporting of business models and business perspectives.