Exploratory Data Analysis BANK MARKETING - CAMPAIGN

14-June-2022

Background and Agenda

Data Glacier

Your Deep Learning Partner

ABC Bank wants to sell it's term deposit product to customers and before launching the product they want to develop a model which help 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).

The data is related with 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.

There are four datasets:

1) bank-additional-full.csv with all examples (41188) and 20 inputs, ordered by date (from May 2008 to November 2010), very close to the data analyzed in [Moro et al., 2014]

2) bank-additional.csv with 10% of the examples (4119), randomly selected from 1), and 20 inputs.

3) bank-full.csv with all examples and 17 inputs, ordered by date (older version of this dataset with less inputs).

4) bank.csv with 10% of the examples and 17 inputs, randomly selected from 3 (older version of this dataset with less inputs).

The smallest datasets are provided to test more computationally demanding machine learning algorithms (e.g., SVM).

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

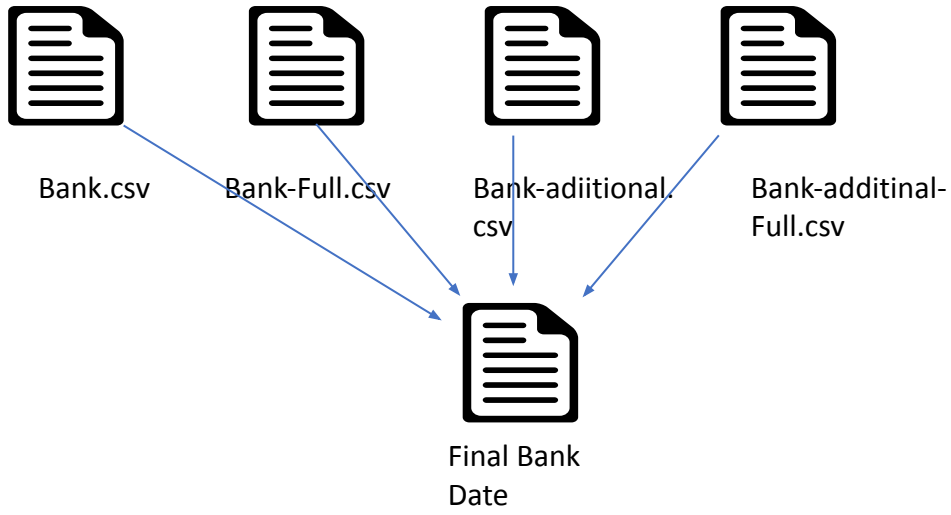
Data Exploration

1) Bank-additional-full.csv with all examples (41188) and 20 inputs, ordered by date (from May 2008 to November 2010), very close to the data analyzed in [Moro et al., 2014]

2) Bank-additional.csv with 10% of the examples (4119), randomly selected from 1, and 20 inputs.

3) Bank-full.csv with all examples and 17 inputs, ordered by date (older version of this dataset with less inputs).

4) Bank.csv with 10% of the examples and 17 inputs, randomly selected from 3 (older version of this dataset with less inputs).



The Bank Data contained four(4) files, namely; Bank, Bank-Full, Bank-additional and the Bank-additional-full documents and all were in the >.csv format.

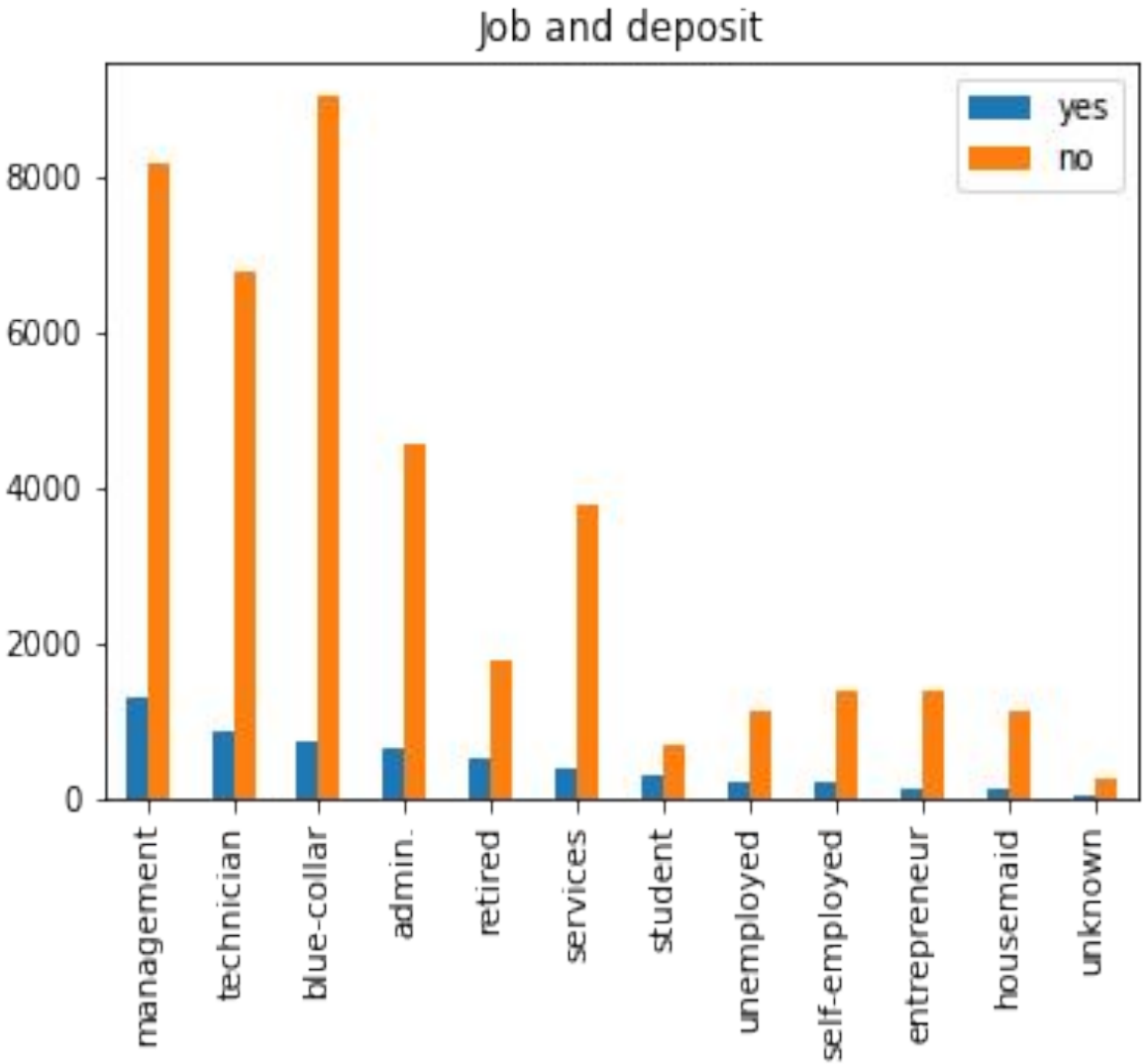
Data Glacier
Your Deep Learning Partner

ANALYSIS OF JOBS AND IMPACT ON DEPOSITS

Data Glacier

Your Deep Learning Partner

JOBS AND DEPOSITS ANALYSIS

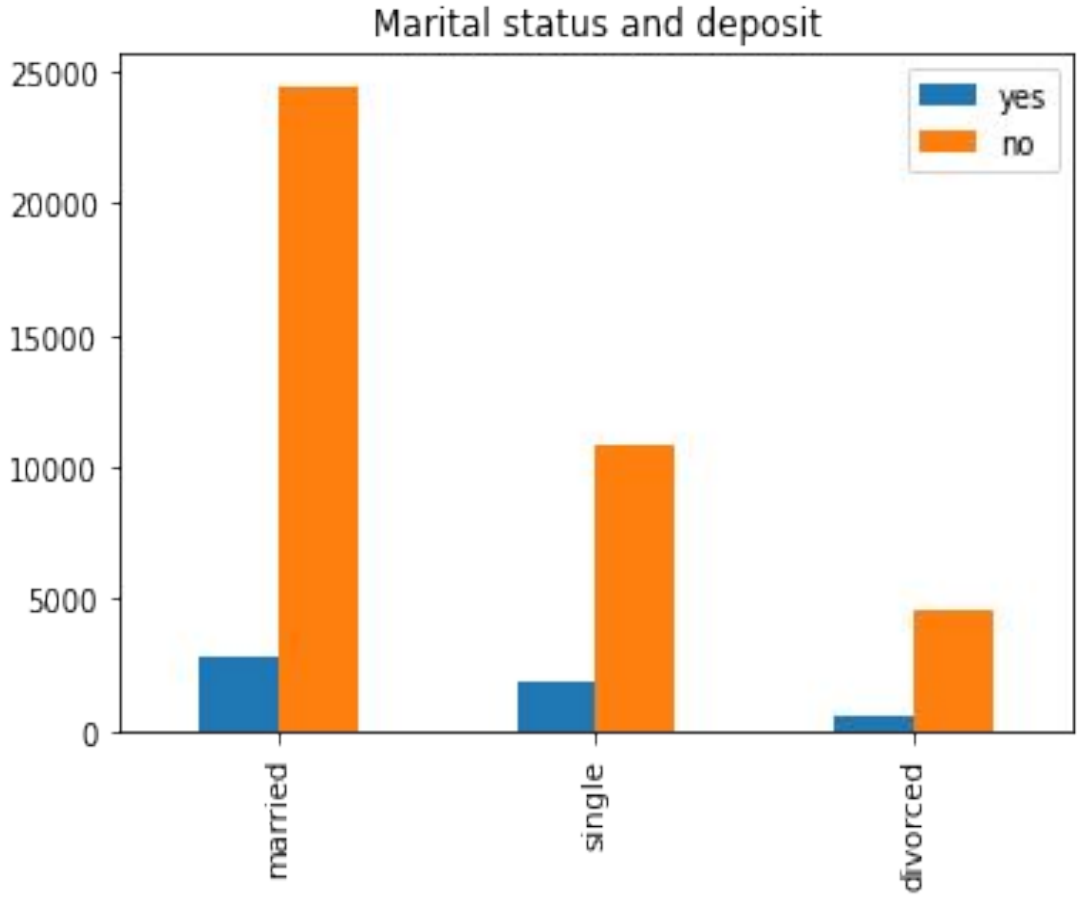


ANALYSIS OF MARITAL STATUS AND IMPACT ON DEPOSITS

Data Glacier

Your Deep Learning Partner

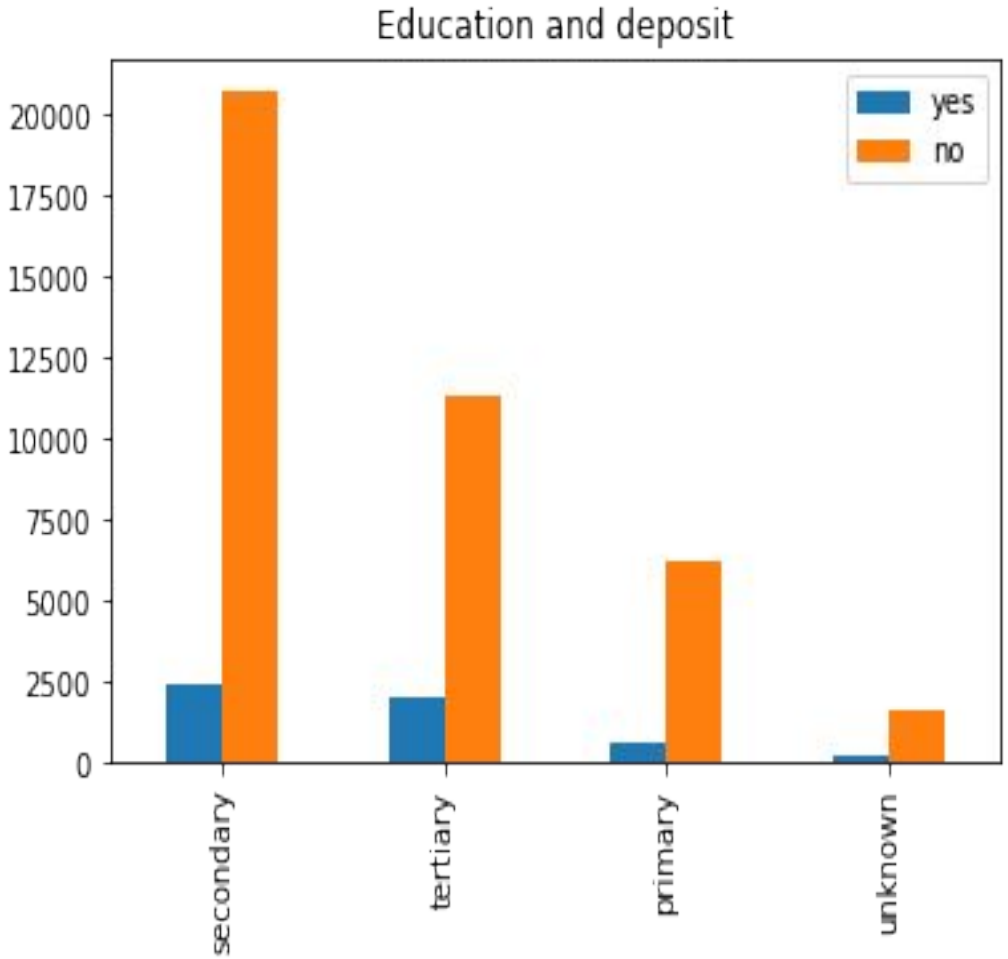
MARITAL STATUS AND DEPOSITS ANALYSIS



ANALYSIS OF EDUCATION AND IMPACT ON DEPOSITS

Data Glacier
Your Deep Learning Partner

EDUCATION AND DEPOSITS ANALYSIS

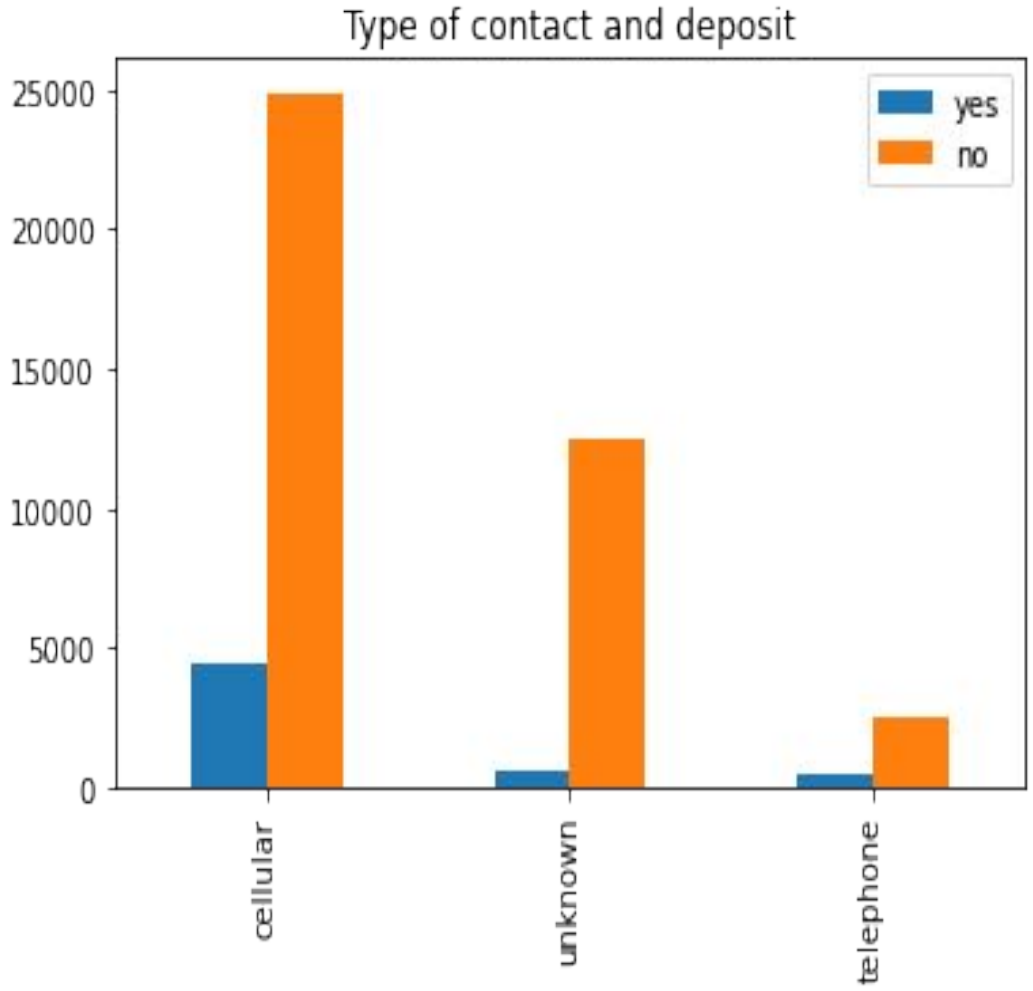


ANALYSIS OF THE METHOD OF CONTACT WITH CUSTOMER AND IMPACT ON DEPOSITS

Data Glacier

Your Deep Learning Partner

METHOD OF CONTACT AND DEPOSITS ANALYSIS

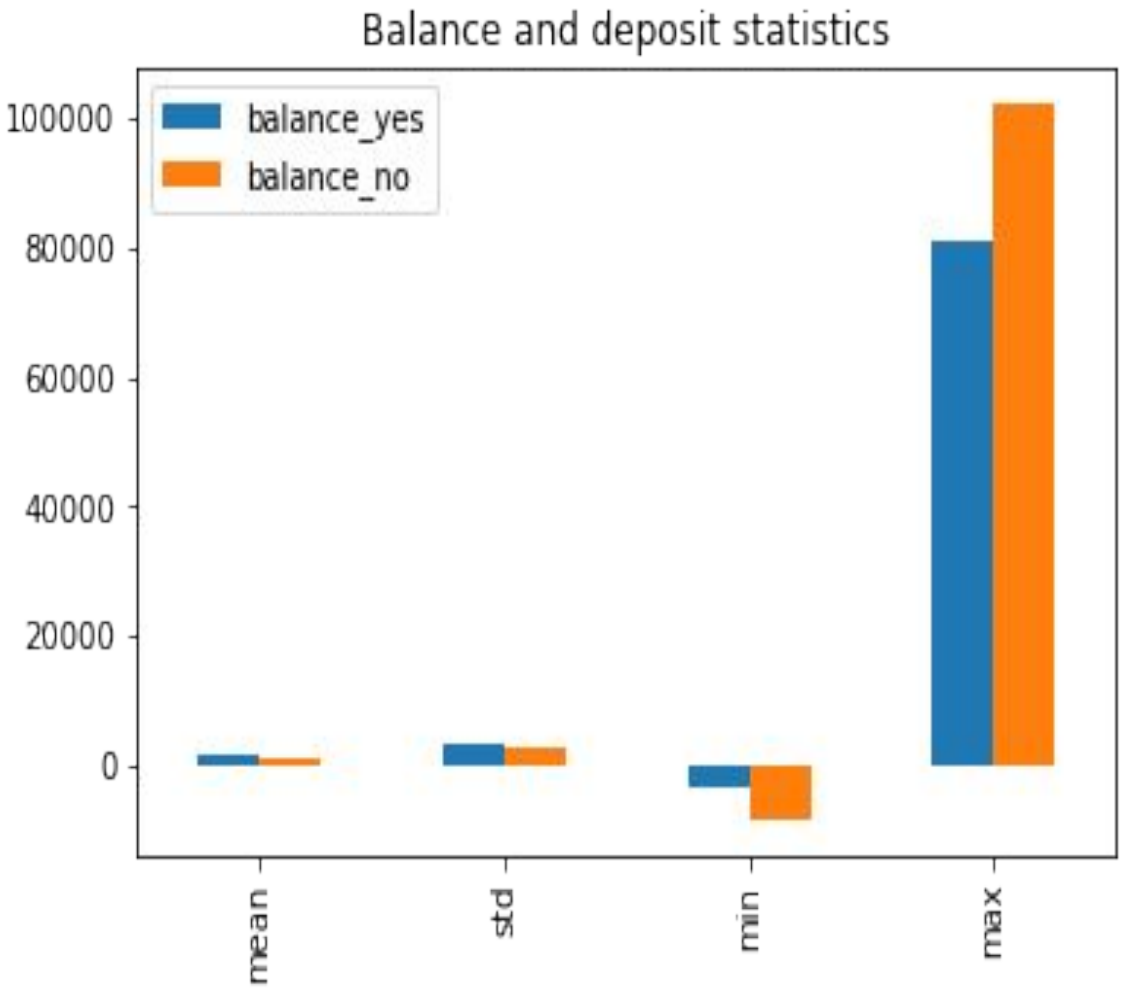


ANALYSIS OF CUSTOMER'S BANK BALANCE AND IMPACT ON DEPOSITS

Data Glacier

Your Deep Learning Partner

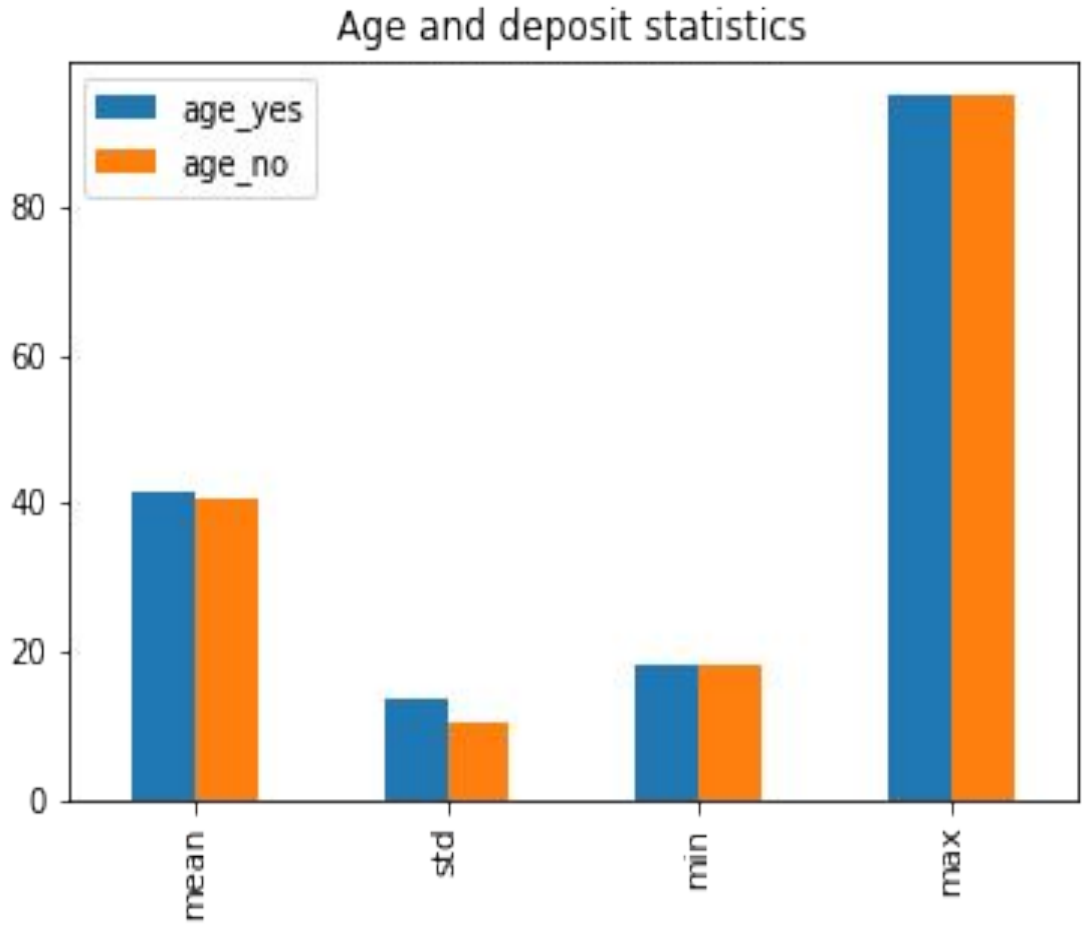
BANK BALANCE AND DEPOSITS ANALYSIS



ANALYSIS OF CUSTOMER'S AGE AND IMPACT ON DEPOSITS

Data Glacier
Your Deep Learning Partner

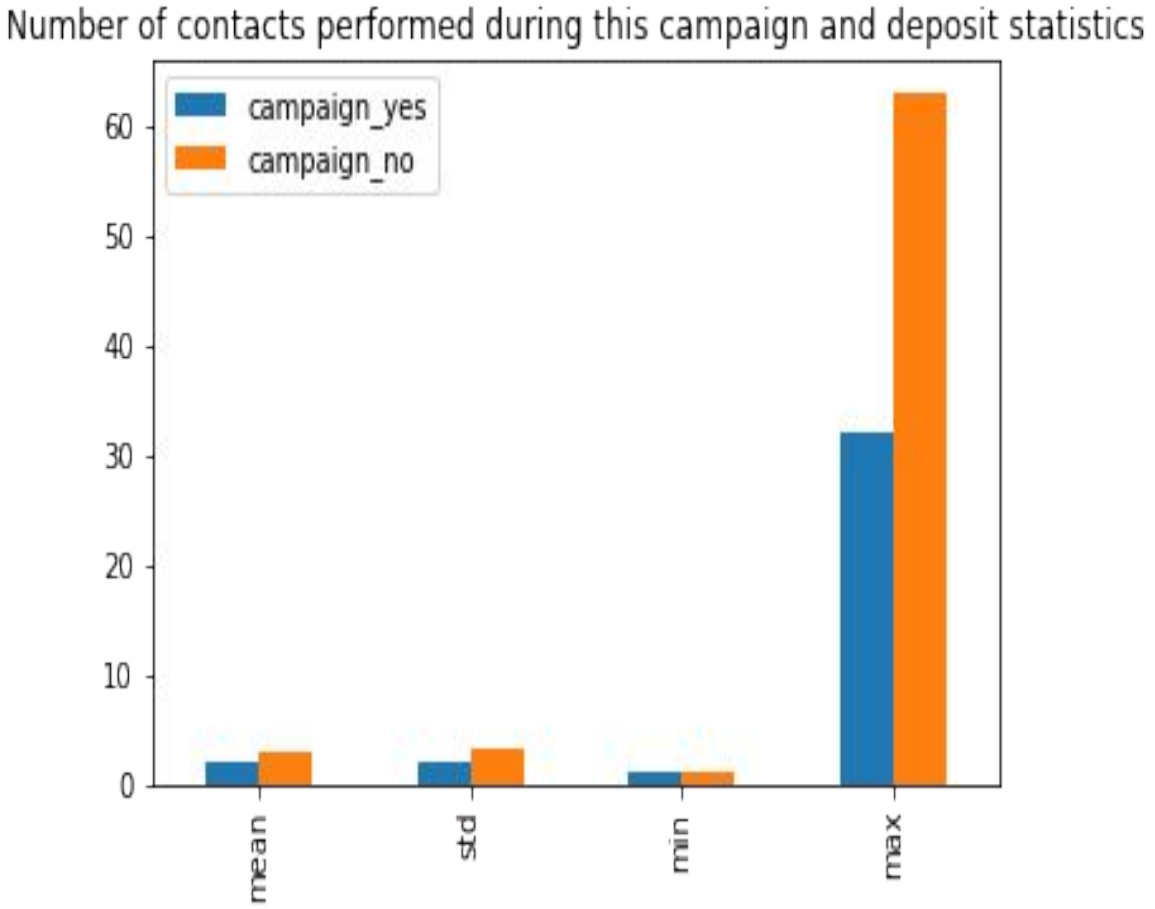
CUSTOMER'S AGE AND DEPOSITS ANALYSIS



ANALYSIS OF THE NUMBER OF TIMES BANK CONTACTED CUSTOMER AND IMPACT ON DEPOSITS

Data Glacier
Your Deep Learning Partner

NUMBER OF CONTACT AND DEPOSITS ANALYSIS



FINAL RECOMMENDATI-ON AND FINDINGS

From Analysis we can tell according to our dataset:

1. Customers with 'blue-collar' and 'services' jobs are less likely to subscribe for term deposit.

2. Married customers are less likely to subscribe for term deposit.

3. Customers with 'cellular' type of contact are less likely to subscribe for term deposit.

4. People who subscribed for term deposit tend to have greater balance and age values.

5. People who subscribed for term deposit tend to have fewer number of contacts during this campaign.

PROPOSED MODELLING TECHNIQUE

THE TEAM IS GOING TO TEST OF THE FOLLOWING MODELS AND PICK WHICH HAS THE BEST ACCURACY SCORE;

1. LOGISTIC REGRESSION
2. RANDOM FOREST CLASSIFIER
3. SUPPORT VECTOR CLASSIFIER (SVC)
4. DECISION TREE CLASSIFIER