Data Glacier
Your Deep Learning Partner

Exploratory Data Analysis
Bank Marketing (Campaign)

15-Aug-2021

Agenda

Executive Summary

Problem Statement

Approach

EDA Summary

Model Selection.

Data Glacier
Your Deep Learning Partner

Executive Summary

The goal of the project is to develop model which predict if the client will buy term deposit using ML Models to predict clients whose chances of buying the product, the data is related with direct marketing campaigns of a Portuguese banking institution.

Problem Statement

ABC Bank wants to sell its 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)

Approach

-Data Requirements by identifying the necessary data content, formats, and sources for initial data collecting from UCI.

-Data Understanding to check the type of each data and to learn more about the attributes and their names.

-Data Preparation to be sure that data are in the correct format for the machine learning algorithm we desire to use.

-Predictive Modeling that use a training set that is a set of historical data in which the outcomes are already known.

-Model Evaluation using Hold-Out method.

Once the model is evaluated, it is "deployed" and put to the ultimate test.
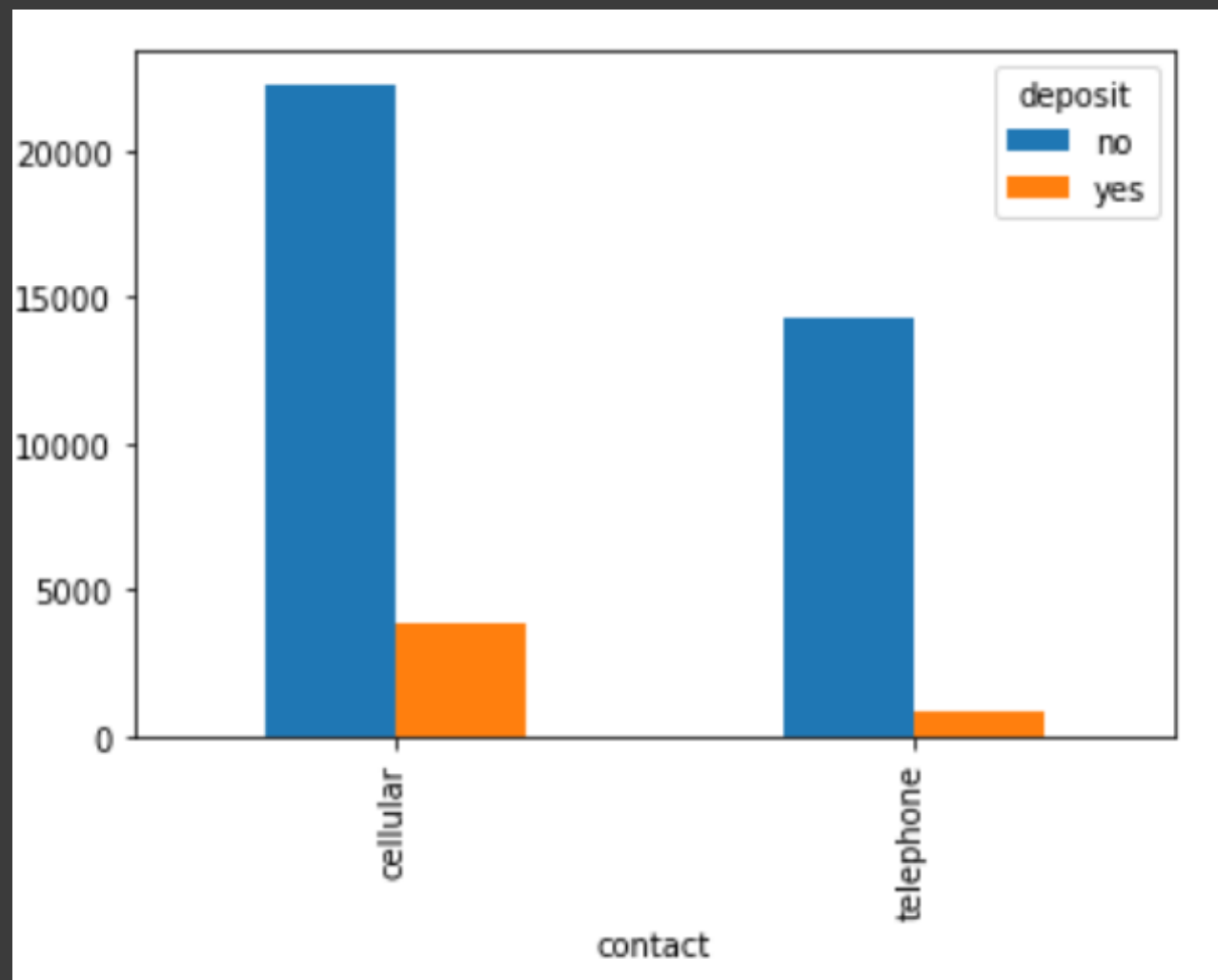
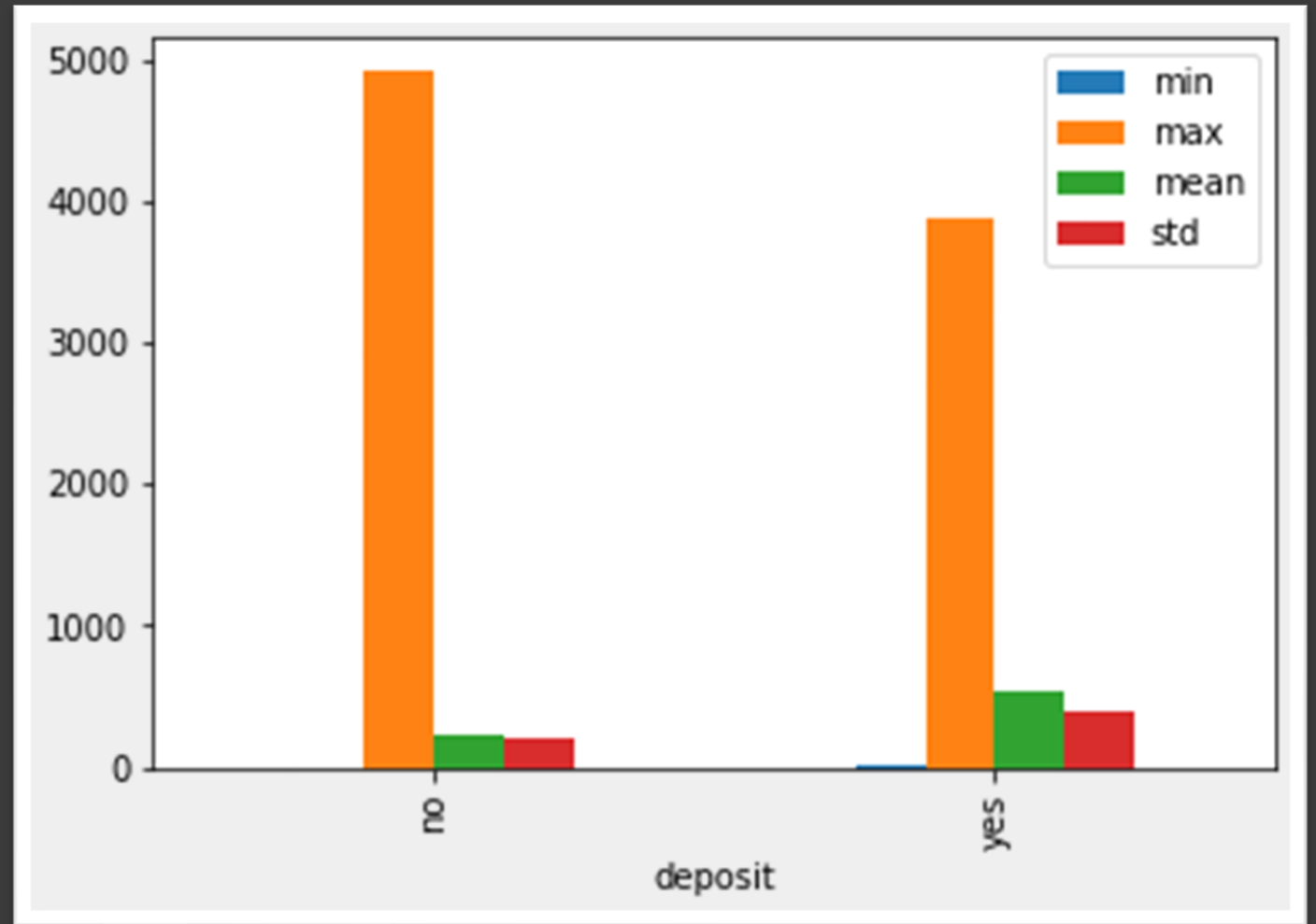# How does customer's job and martial affect the possibility of selling term deposit.
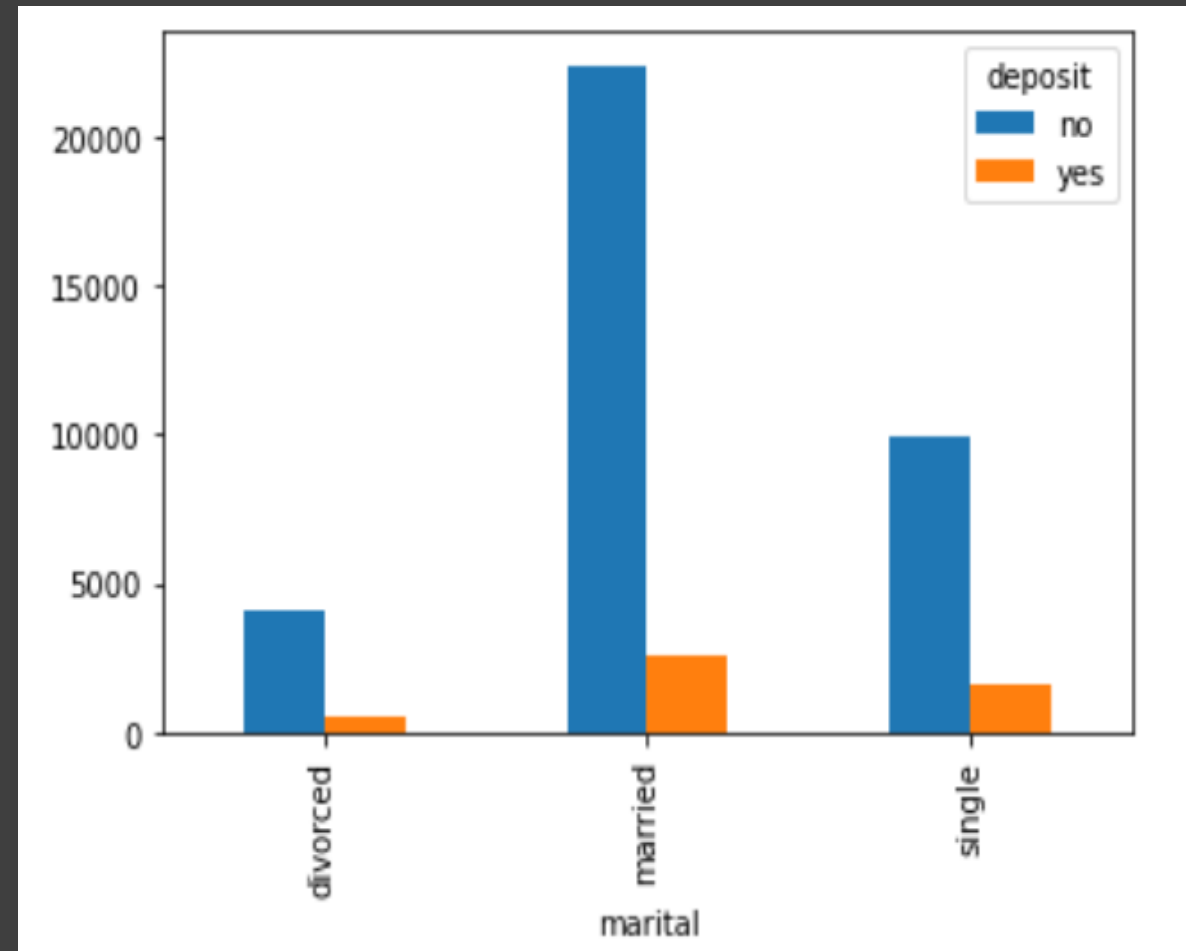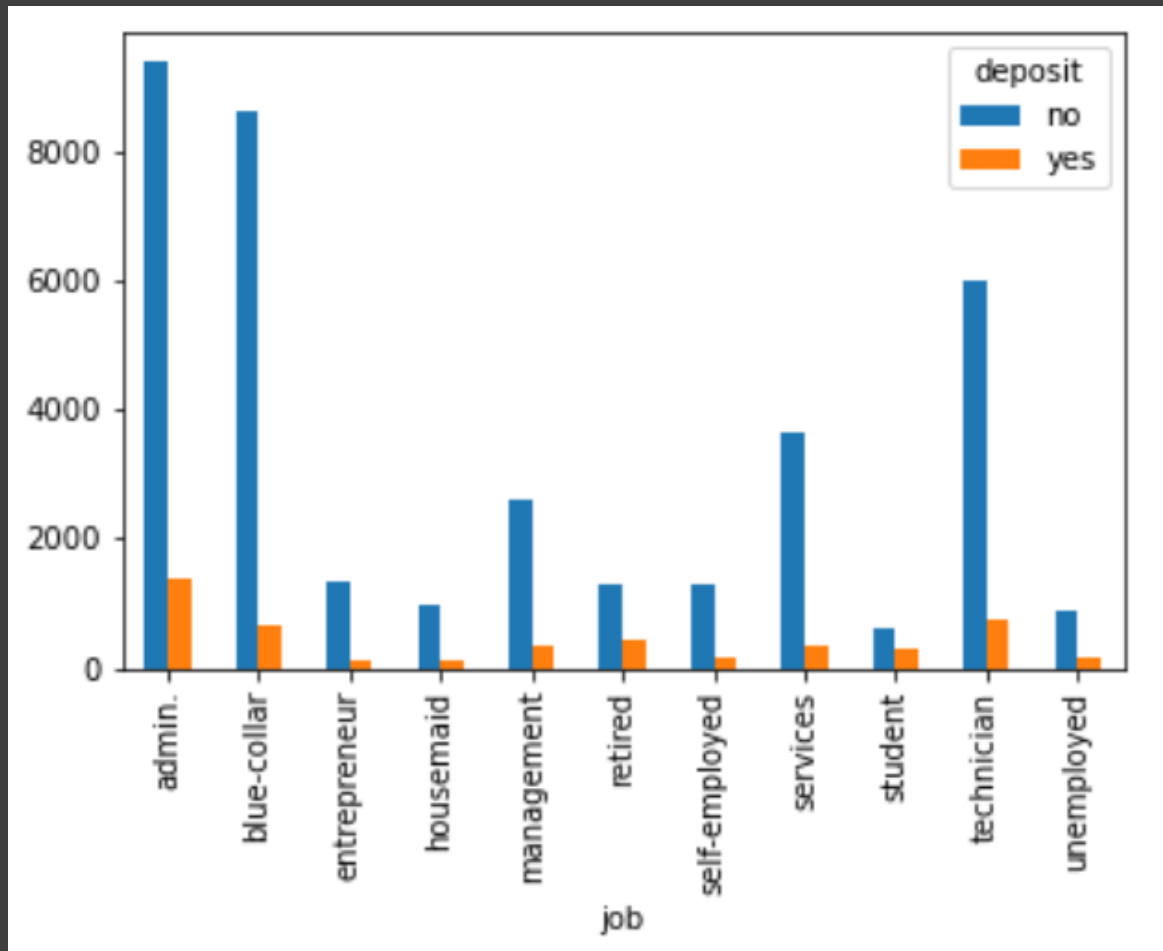
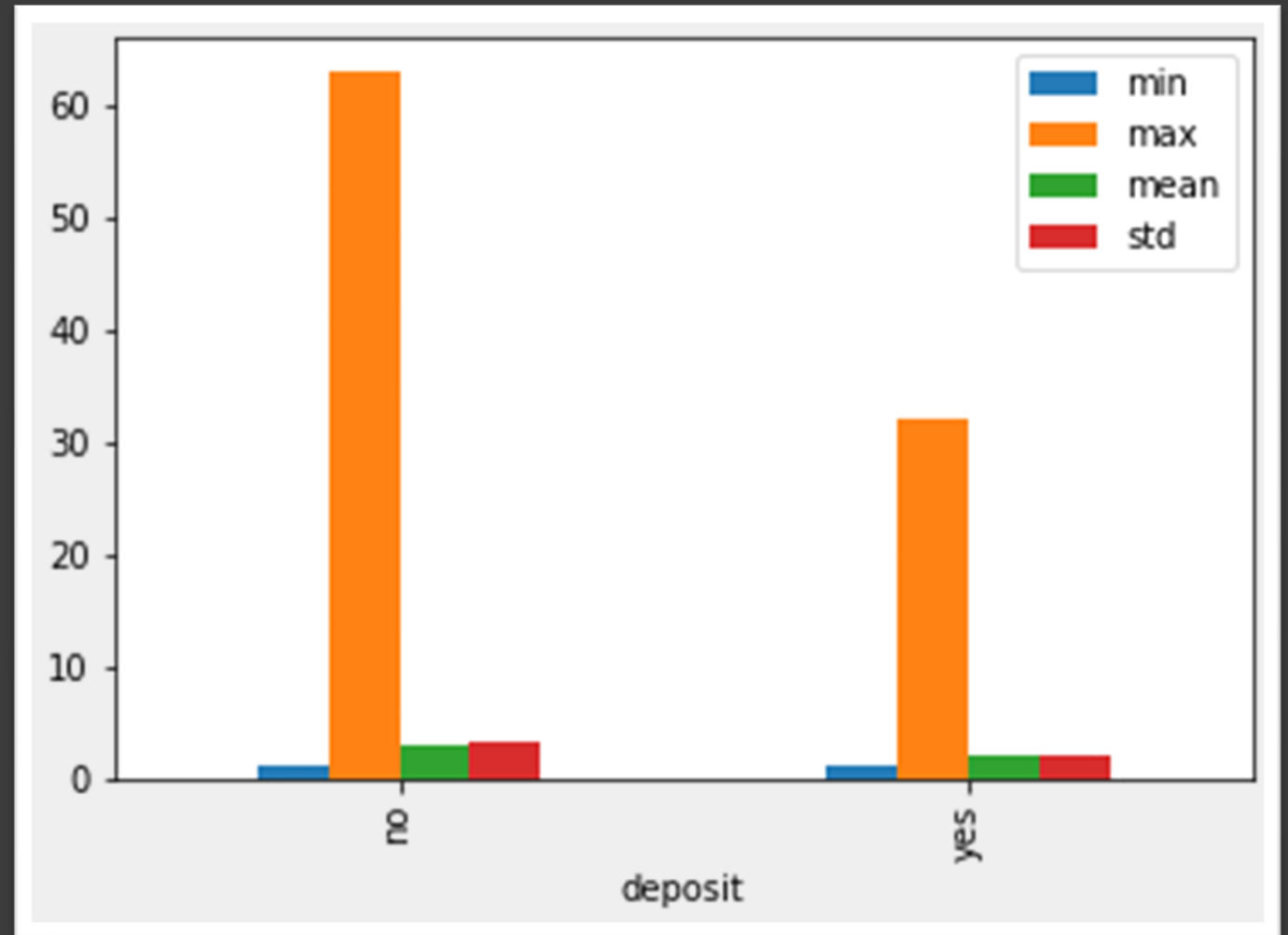How does customer's education affect the possibility of selling term.
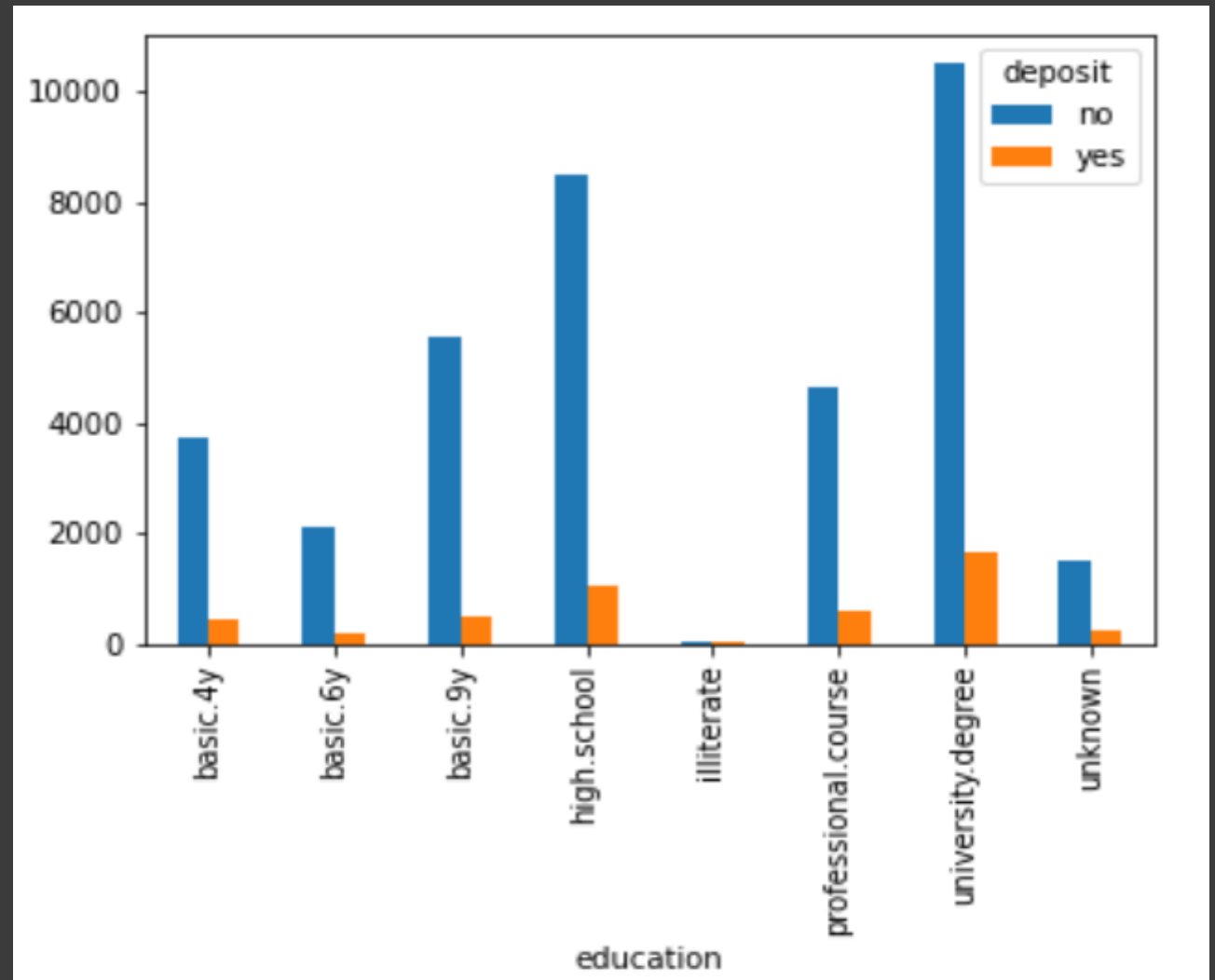
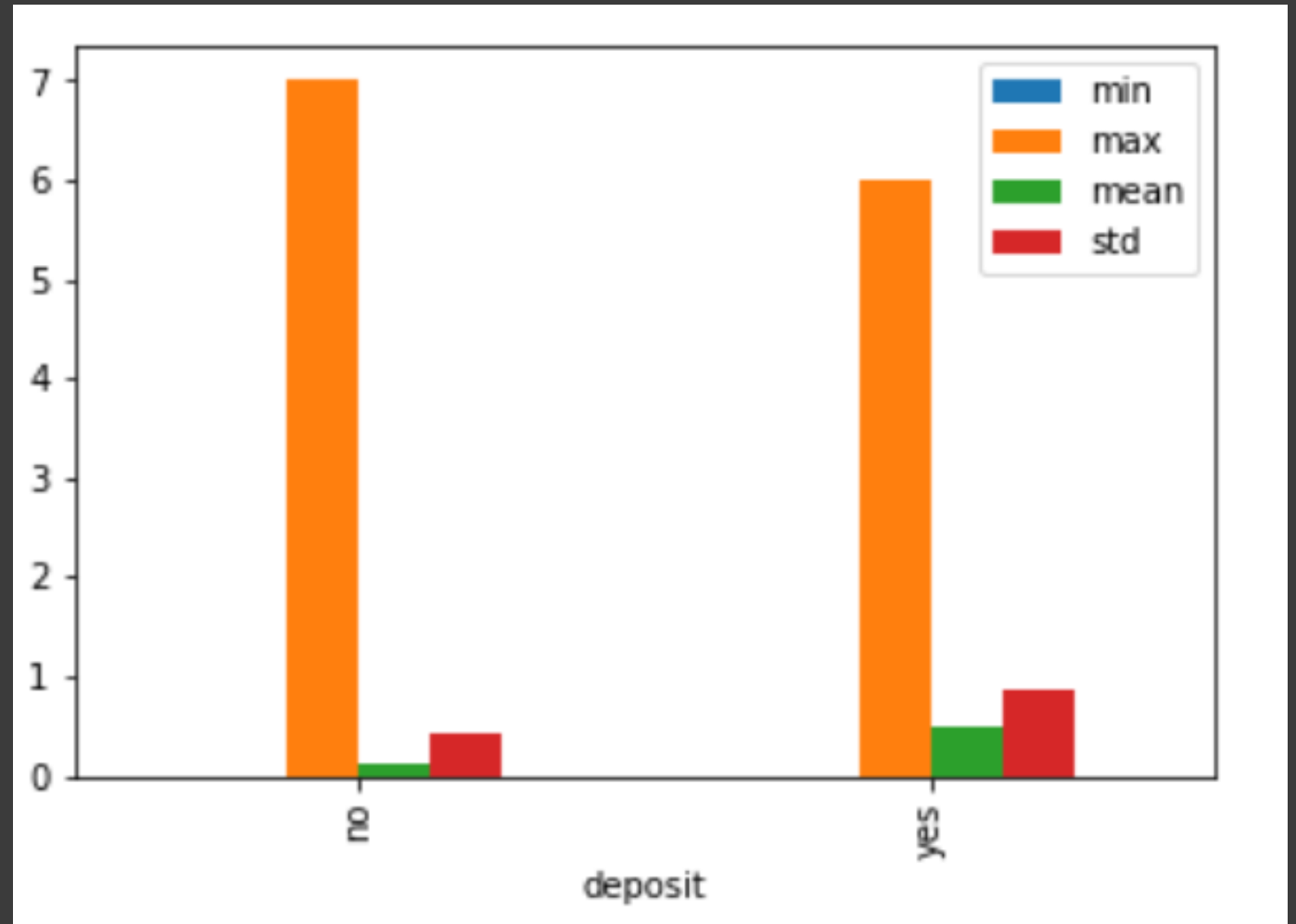Type of connection make difference in selling terms.

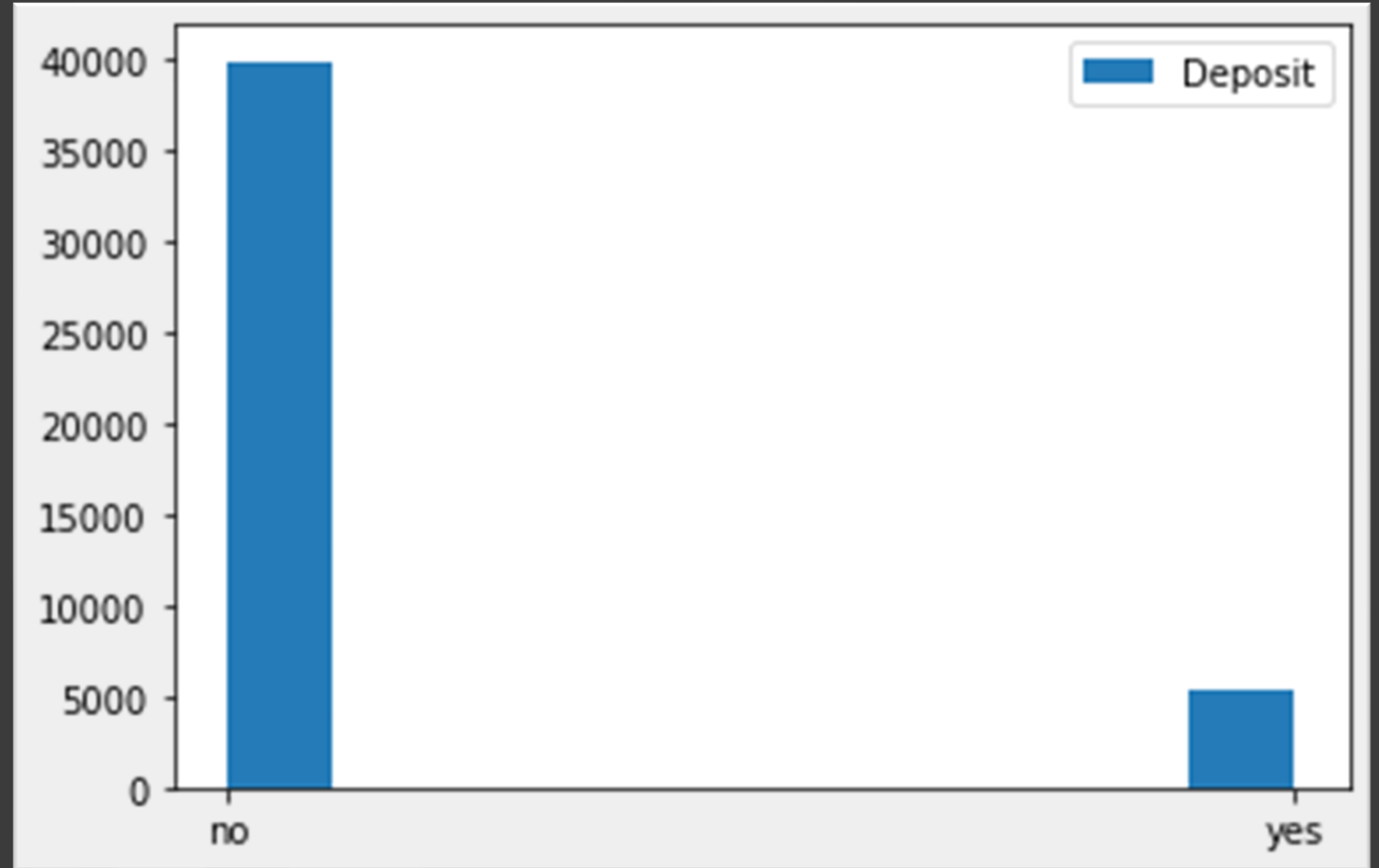Last contact duration and how this attribute highly affects the output target.

The relation between number of contacts performed during this campaign for each client and selling of terms.

The relation between number of contacts performed before this campaign for each client and selling of terms.

Deposit in previous campaigns.
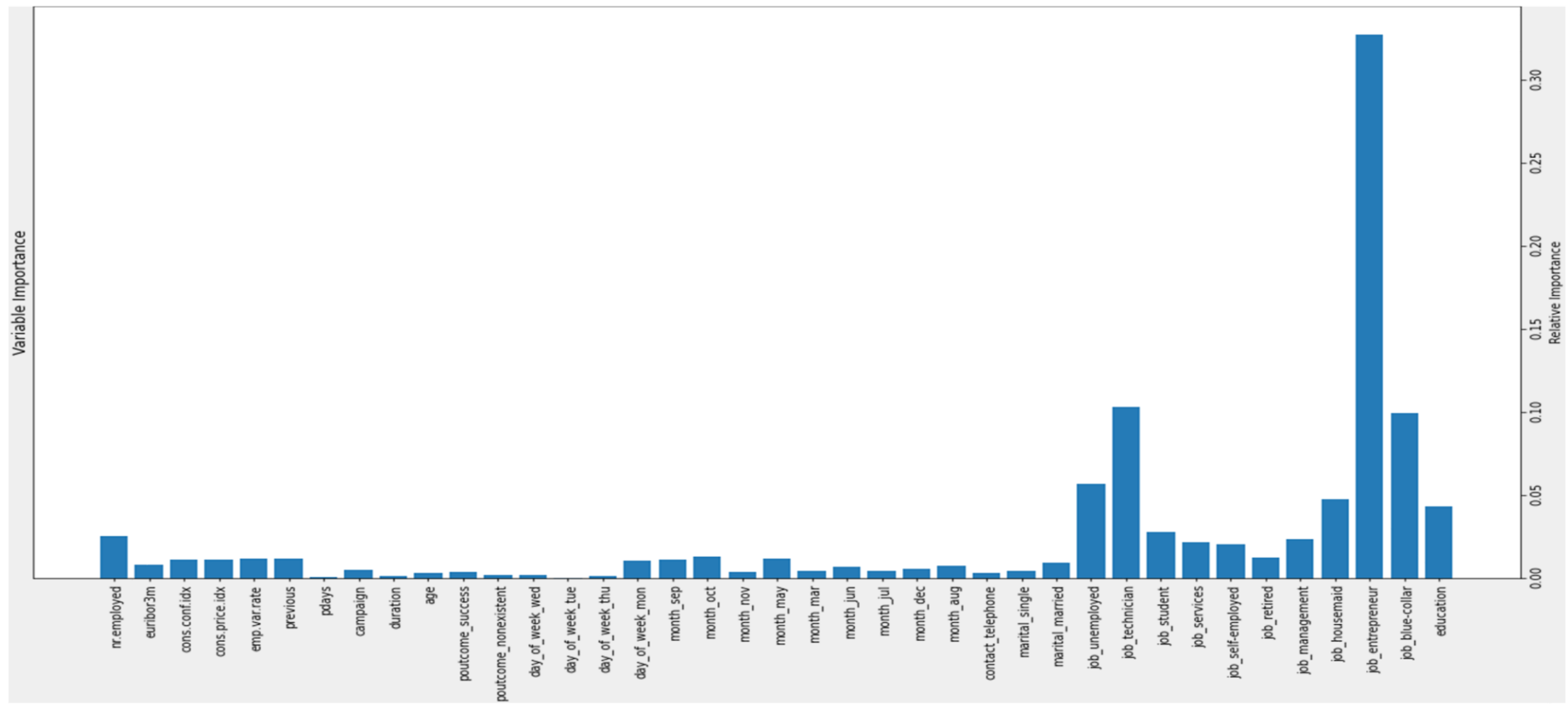Has the clients subscribed a term deposit or not?

In this chart we show the correlation between the numeric variables, in order to analyze the feature selection afterwards.

Relative Importance of features

The most relative importance "job-entrepreneur"

Model Selection.

- Depending on values of y (yes or no) we used as ML models :

Logistic Regression, Bagging Classifier, XGBOOST, Decision Tree.

- After an automatic feature selection, we decided to not apply it since it didn't improve the performance of the model.

- F1 score was used. The models have been trained with 80% of the data and validated in the rest.

Thank You

JSSN Team

Data Glacier

Your Deep Learning Partner