



Data Glacier

Your Deep Learning Partner

Exploratory Data Analysis

Project: Bank Marketing Campaign

Name: Supin Hooda

Batch Code: LISUM25

Specialization: Data Science

Agenda

Problem

Statement Dataset

EDA

Recommendations

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 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). This is an application of the organization's marketing data.
- **Objective:** Build a Classification ML model to shortlist customers who are most likely to buy the term deposit product. This would allow the marketing team to target those customers through various channels.

Dataset

<https://archive.ics.uci.edu/ml/datasets/Bank+Marketing>

bank-additional-full.csv: 20 inputs (+1 target variable) and 41118 observations

Assumptions:

- 'Duration' feature is dropped as suggested in the dataset description
- A frequently occurring missing value 'unknown' is considered as another category for the categorical features.
- Duplicate rows were deleted from the dataset.

EDA



Data Glacier

Your Deep Learning Partner

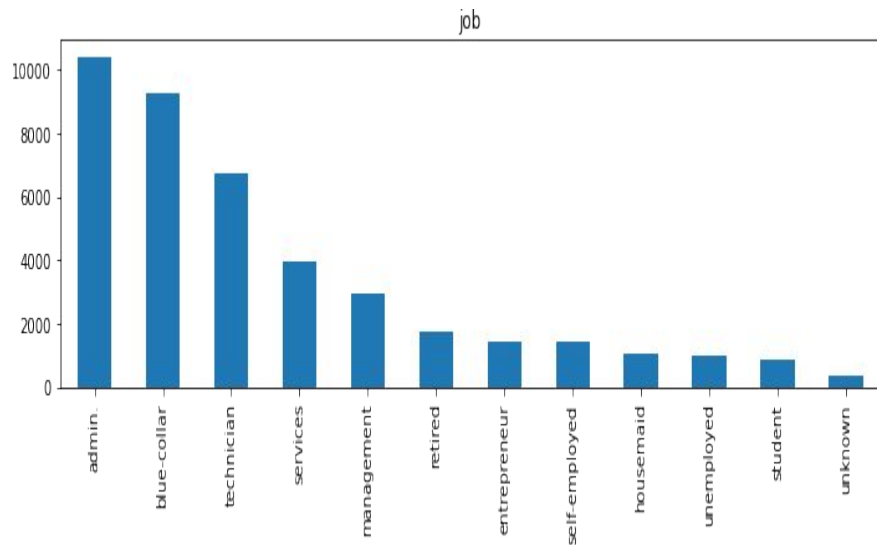
EDA

1. Univariate Analysis



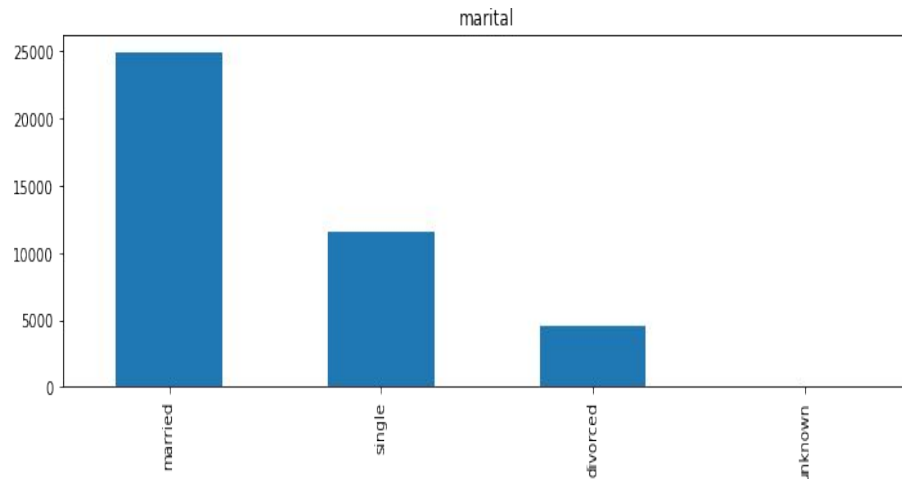
Jobs Distribution

Administrative staff and technicians are the most common job roles.



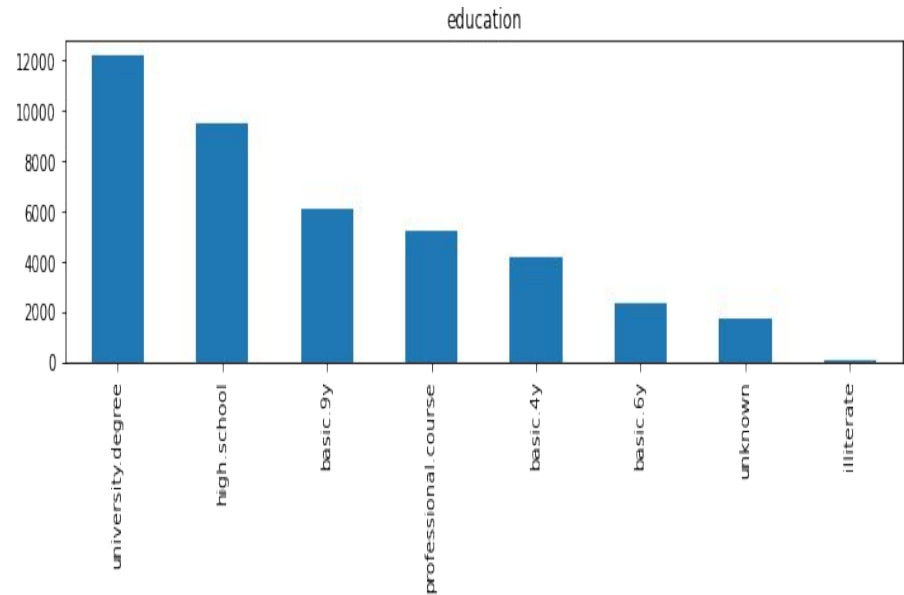
Marital Status and Subscription

Married customers have a higher subscription rate.



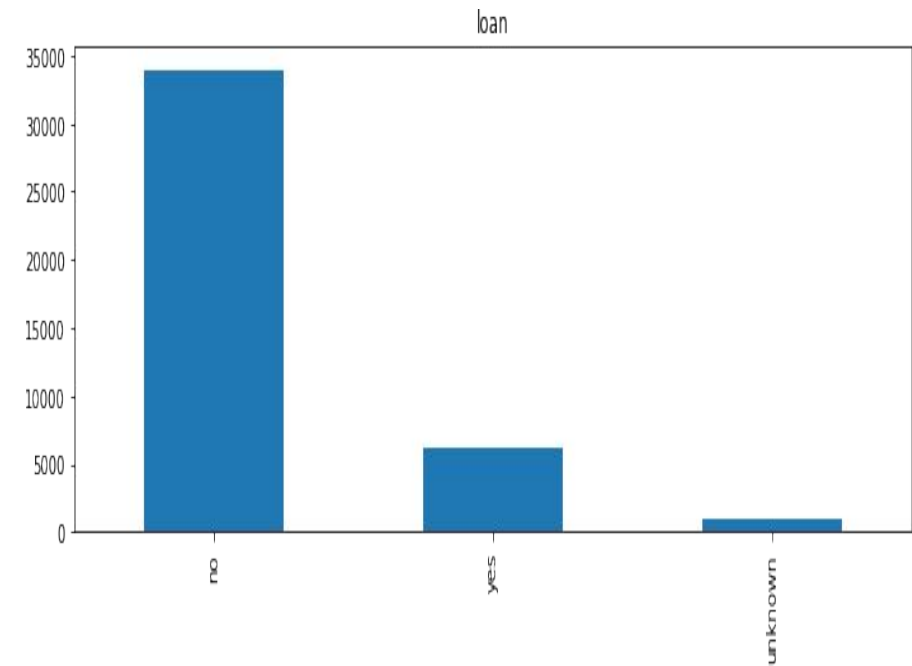
Education Level Distribution

Most customers are university graduates, followed by high school graduates.



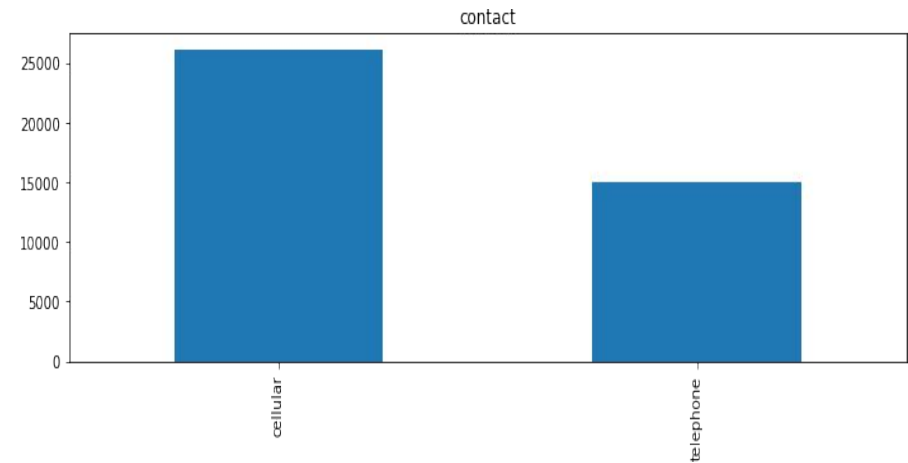
Personal Loan Distribution

Majority of customers do not have a personal loan.



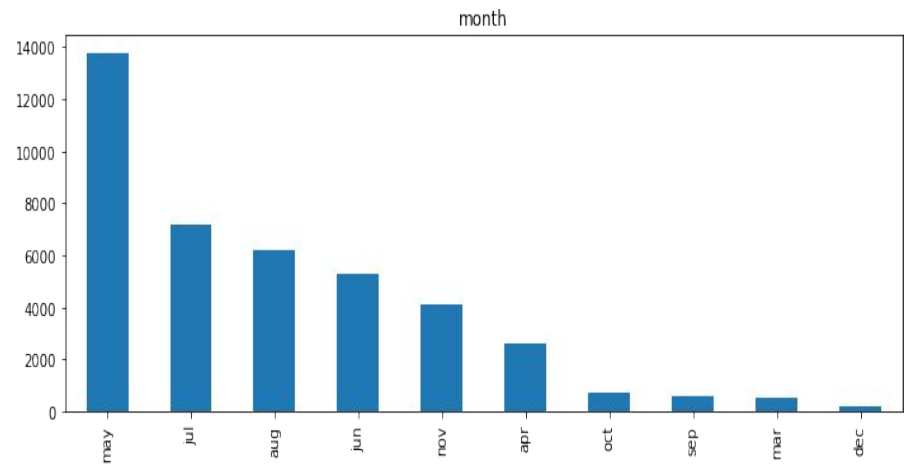
Communication Type

Cellular communication is more prevalent than telephone communication.



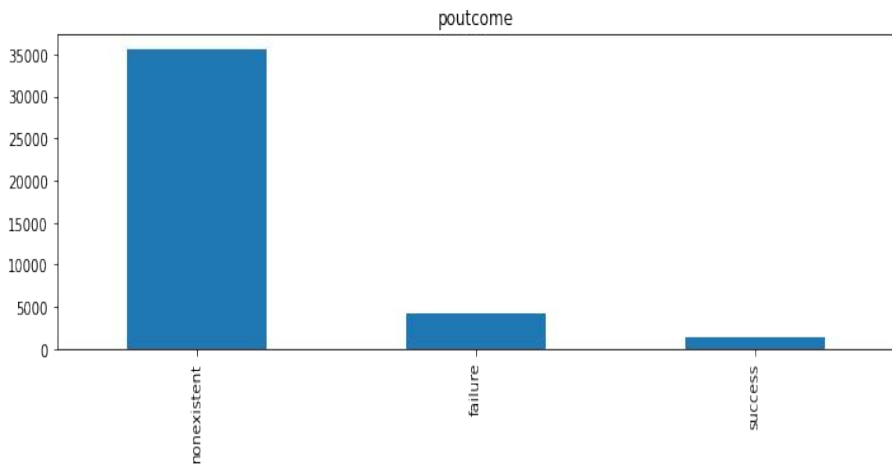
Contact Month Distribution

May has the highest number of contacts.



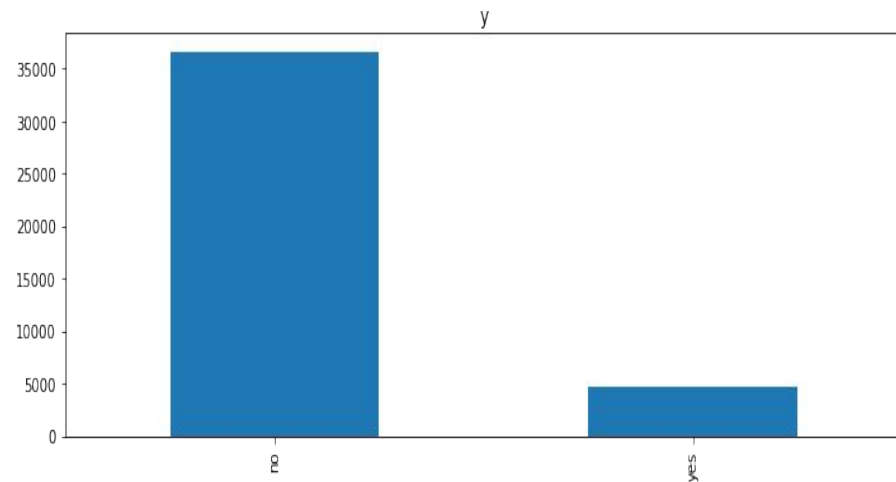
Previous Campaign Results

A significant portion of the dataset does not have information on whether the previous marketing campaign succeeded or failed.



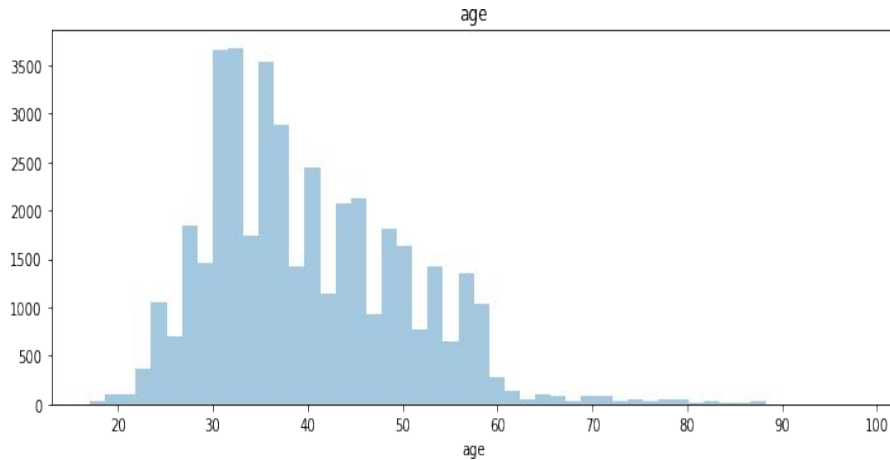
Subscription Rate in Previous Campaign

In the previous campaign, the percentage of people who subscribed to the deposit is less to those who did not.



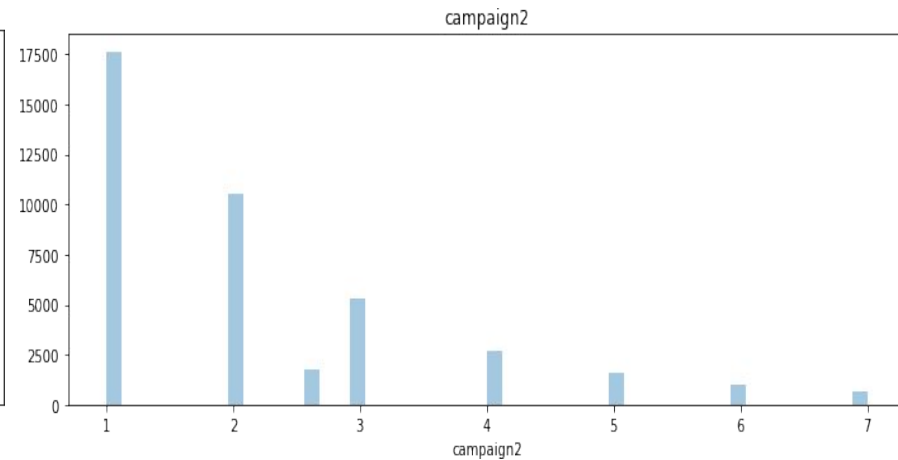
Age Group Distribution

The age group of contacted customers is mostly between 20 to 60.

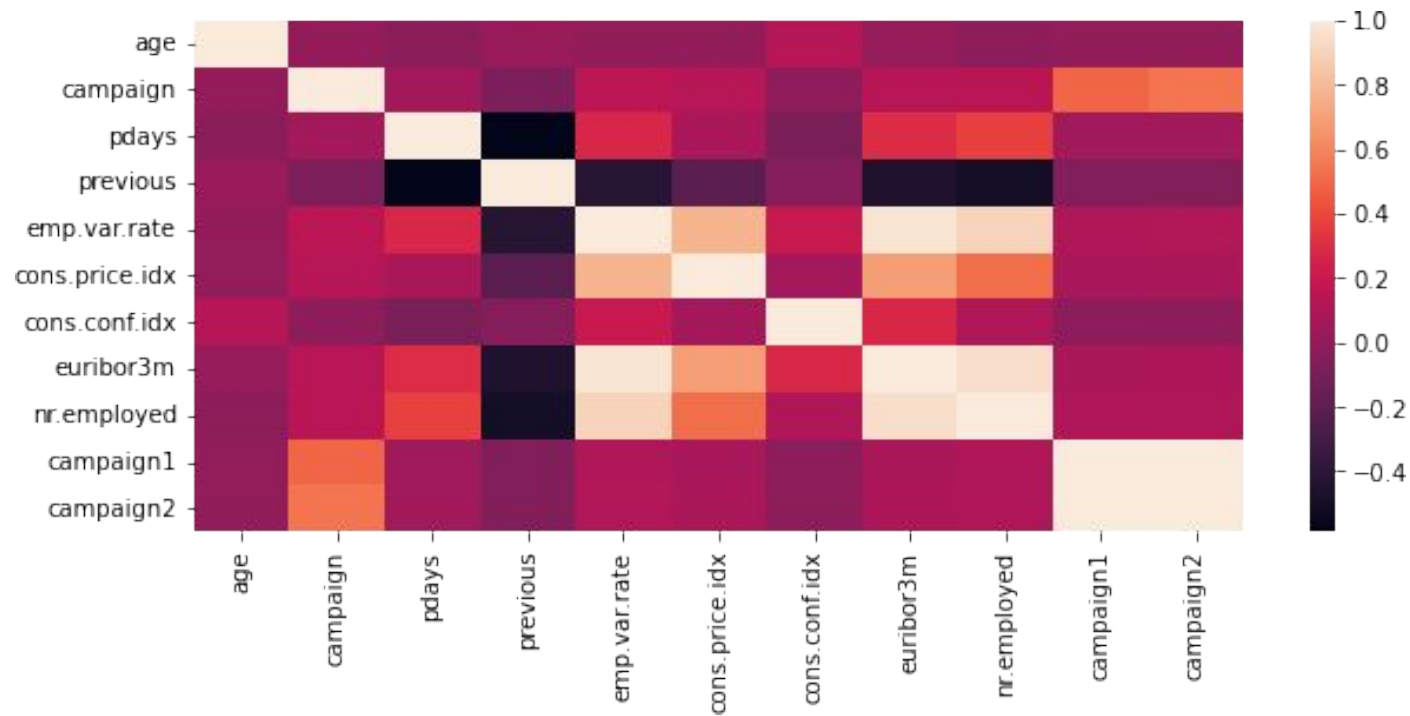


Number of Follow-ups Distribution

The number of follow-ups was relatively low in the previous campaign.



Correlation map

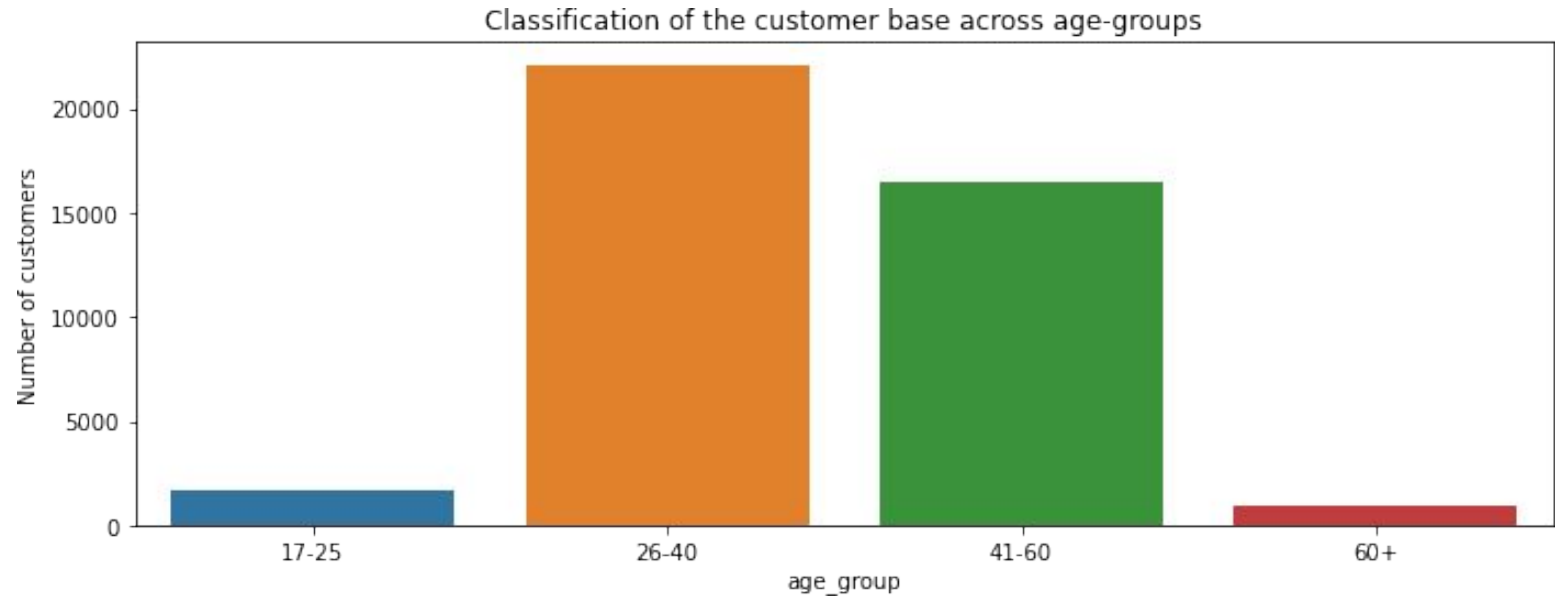


EDA

2. Bivariate Analysis

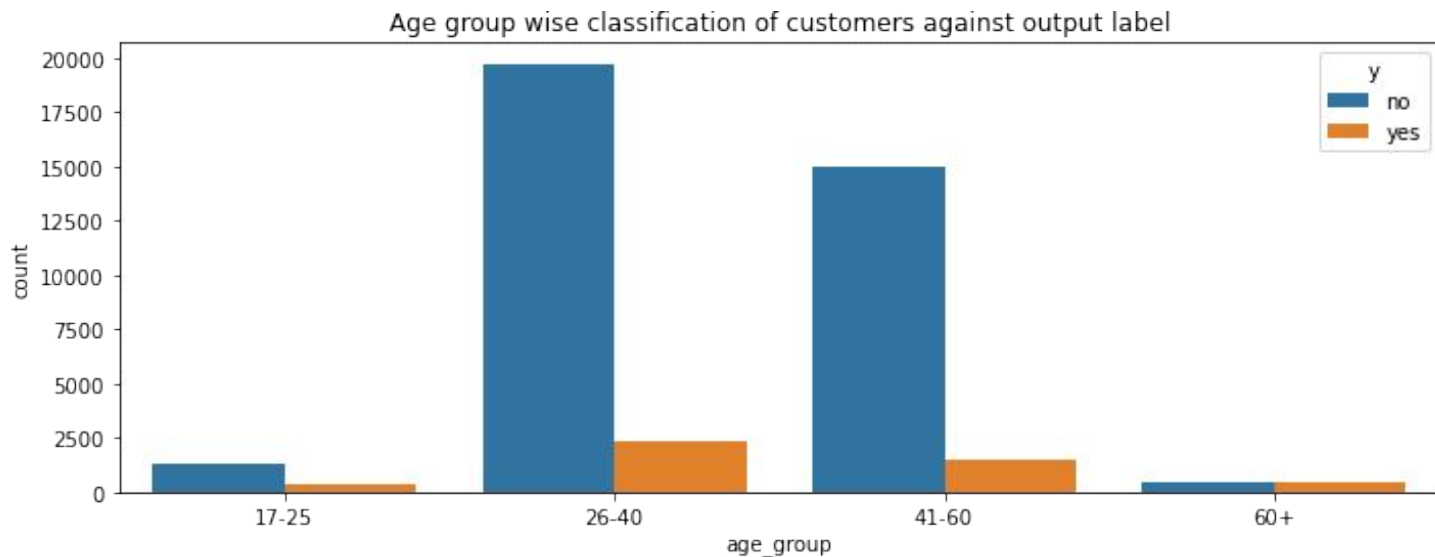
Age-Group Analysis

Most of the customer's belong to the age-groups of 26-40 and 41-60 years



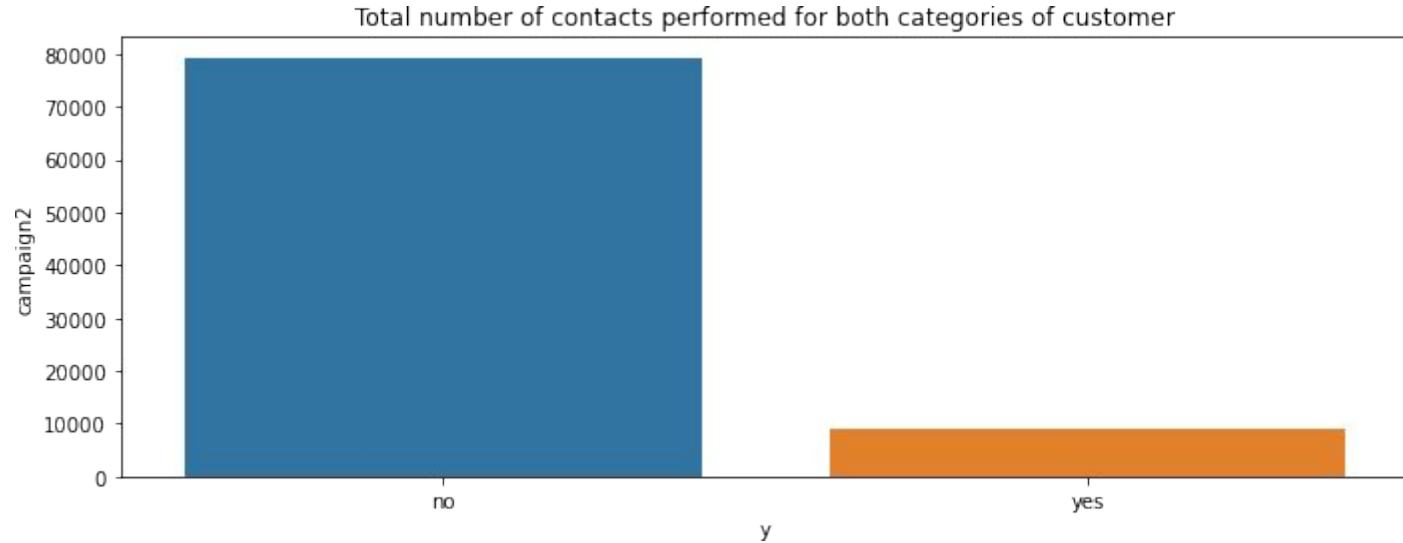
Age-Group Analysis

In the age-groups of 26-40 and 41-60 years, the majority of people did not subscribe to the term deposit plan.



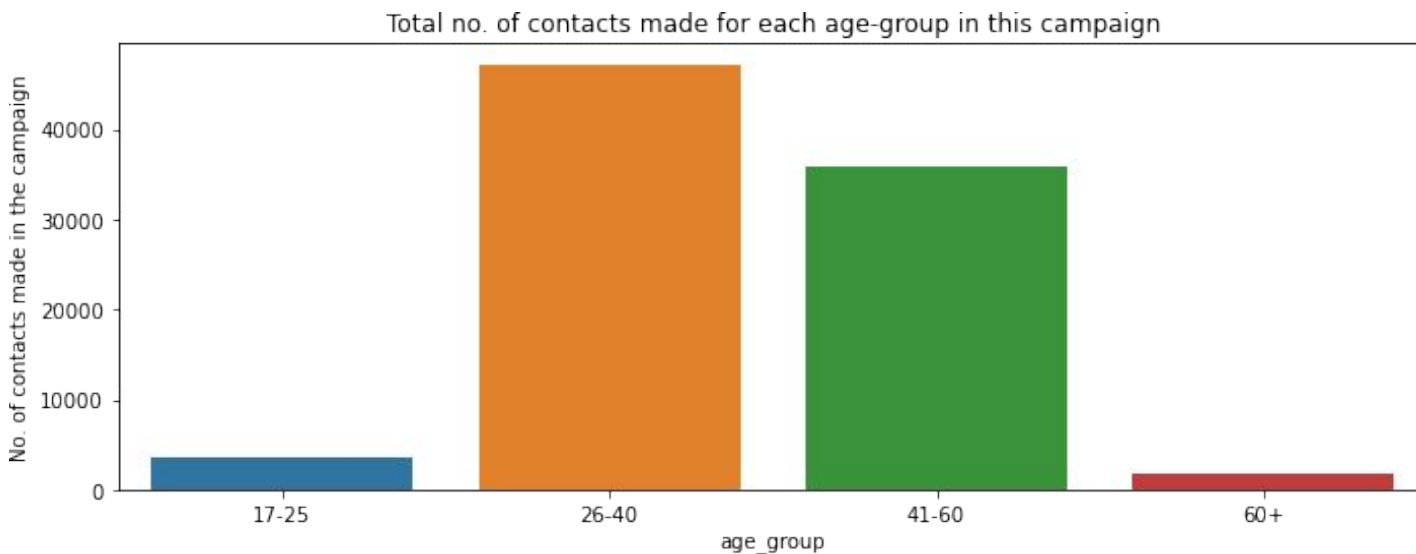
Campaign Contacts vs. Subscription

When a greater number of contacts are made, customers tend not to subscribe to the term deposit plan.



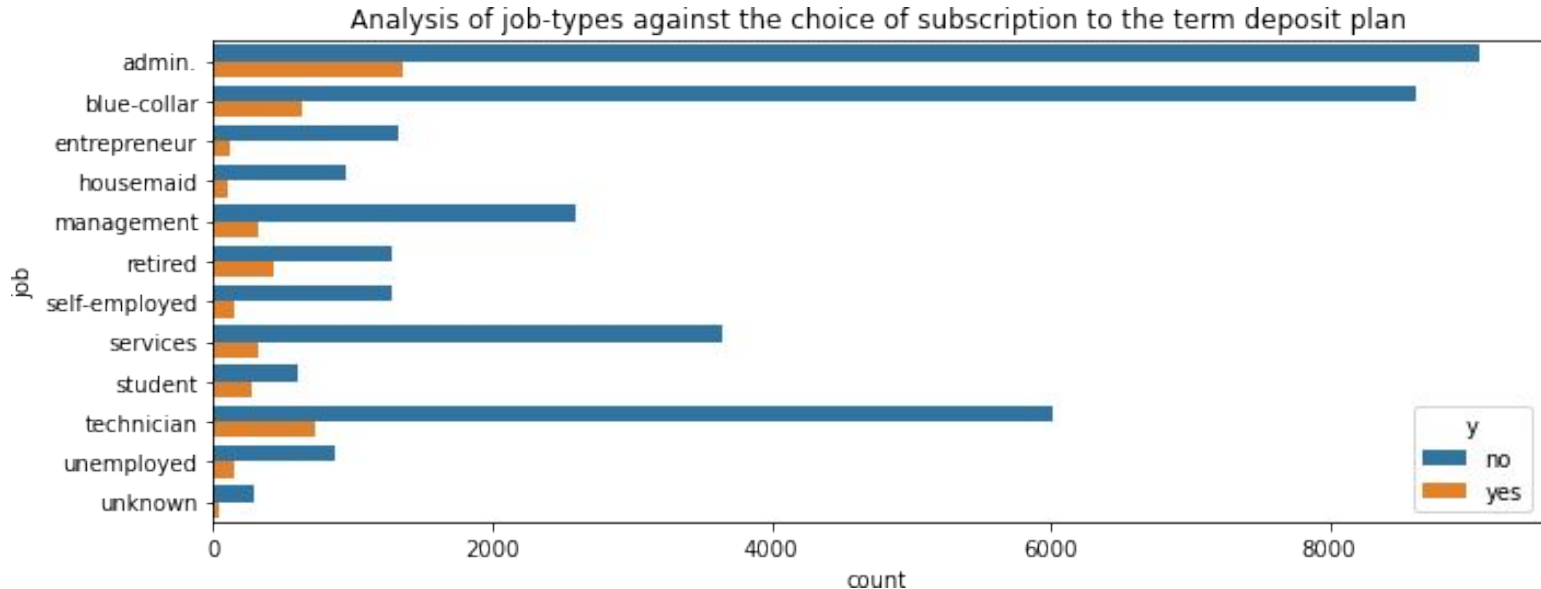
Age-Group and Campaign Contacts

The 26-40 and 41-60 age-groups witness the majority of contacts in this campaign, suggesting these as target groups for the bank.



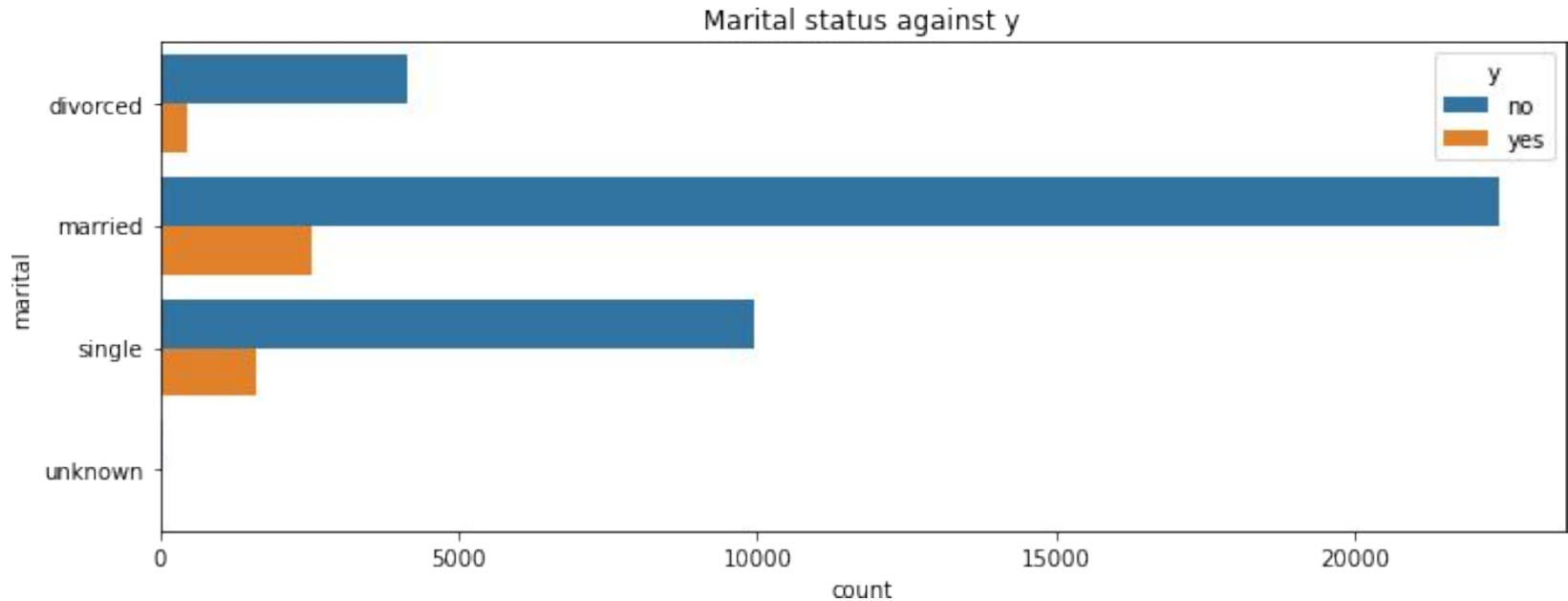
Job Analysis and Subscription

'Admin,' 'blue-collar,' and 'technician' are prominent jobs, with many customers in these roles rejecting the term deposit plan.



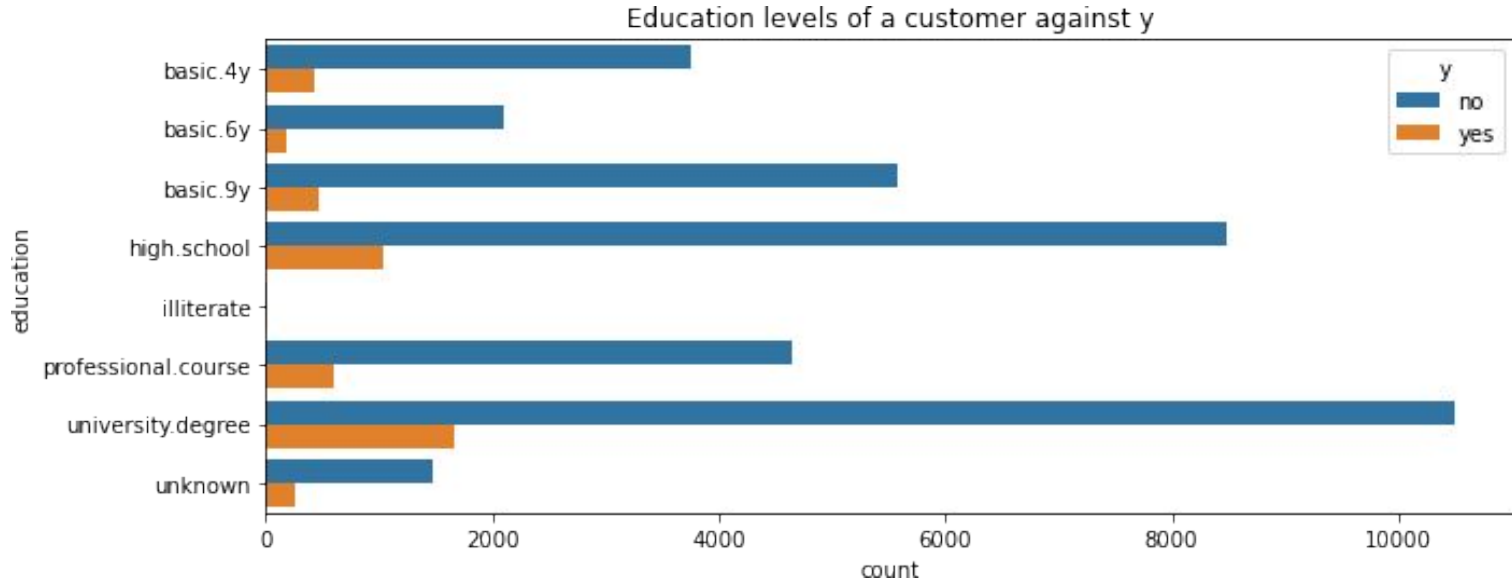
Marital Status Analysis

Married customers are the majority, and comparatively, more of them have taken the term deposit.



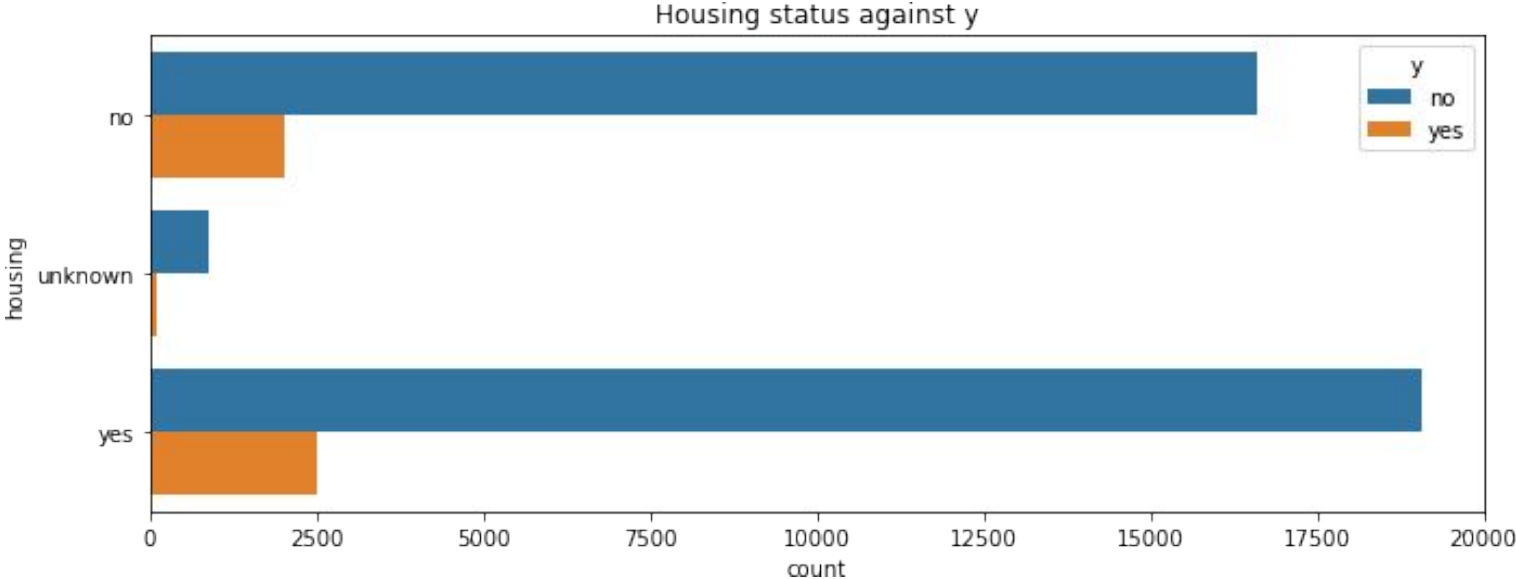
Education Level Analysis

Customers with a university degree are more likely to subscribe to the term deposit.



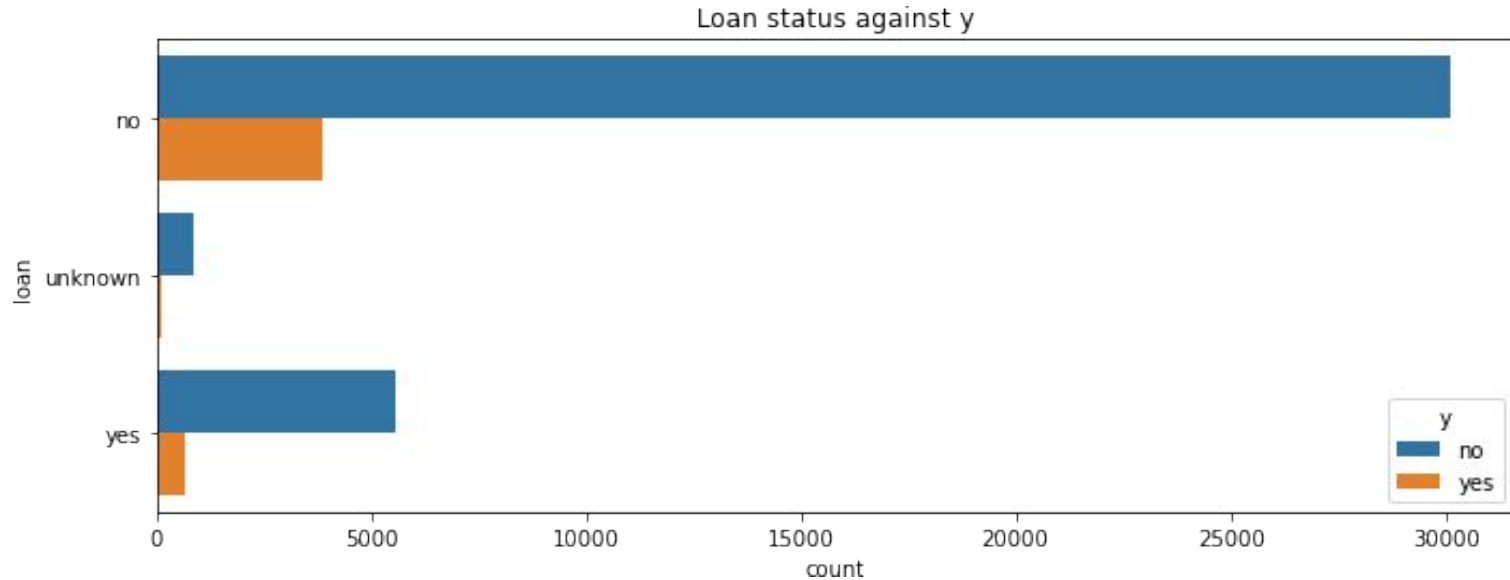
Housing Status Analysis

Customers with a housing loan are more likely to subscribe to the term deposit.



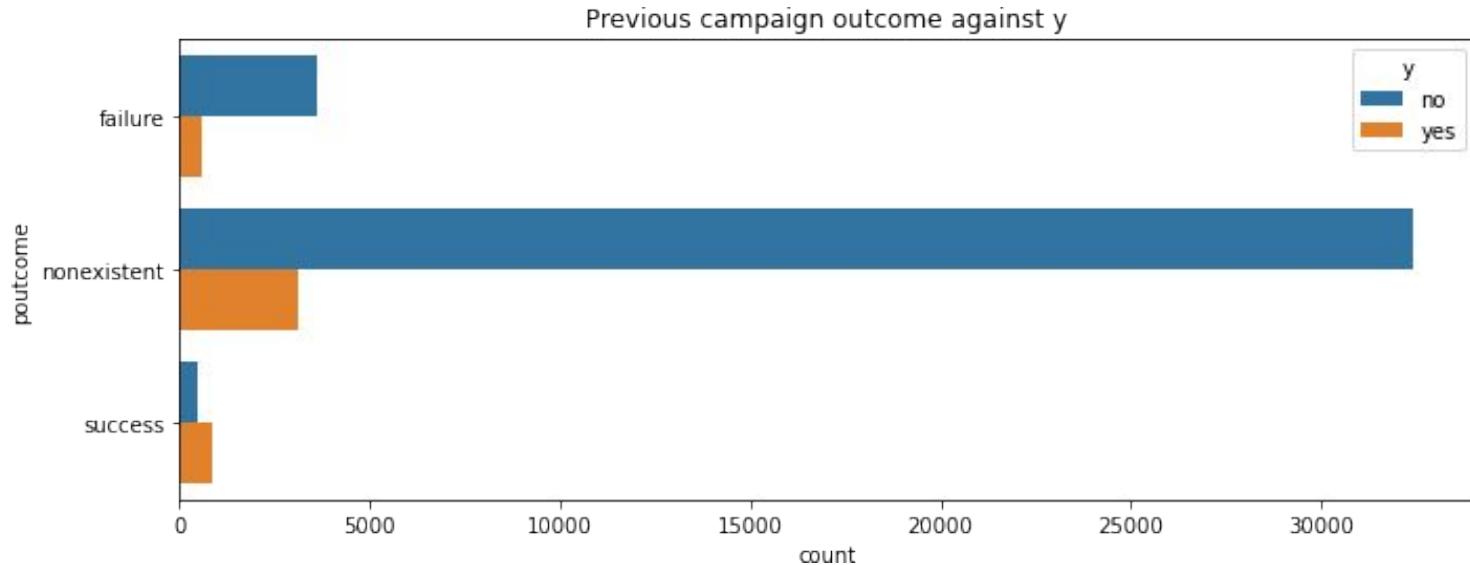
Loan Status Analysis

Customers with a personal loan are comparatively less likely to subscribe to the term deposit.



Previous Campaign Success

The success rate of the previous marketing campaign has resulted in more people subscribing to the term deposit.



Recommendation to improve campaign

Optimal Timing:

May has proven to be the most effective month for customer contact. Consider focusing marketing efforts during this period for maximum impact.

Enhanced Customer Engagement:

Increase the frequency and duration of contacts made to each customer. This extended engagement may lead to improved subscription rates.

Targeted Education Levels:

Place a strategic emphasis on university graduate students and high school degree holders. Tailor marketing strategies to resonate with these educational demographics.

Strategic Age-Group Targeting:

Concentrate marketing efforts on the age-groups of 26-40 and 41-60, as these segments demonstrate a higher proportion among customers. These age-groups present a lucrative target for the marketing team.

Occupational Focus:

Direct marketing efforts towards administrators and technicians. These professional groups show potential for increased subscriptions and warrant additional focus in the campaign strategy.

Recommended Models for Dataset

1. Logistic Regression
2. Naïve Bayes
3. Decision Tree
4. Random Forest
5. Ada Boost

Next Steps: Hyperparameter Tuning and Model Evaluation

- Perform hyperparameter tuning to optimize model performance.
- Evaluate each model's performance using metrics such as accuracy, precision, recall, F1 score, and ROC-AUC.
- Identify the best model and important features through rigorous evaluation.

Thank
You



Data Glacier

Your Cloud Archiving Partner