

Bank Marketing Campaign



Your Deep Learning Partner

Taimoor Razi
Akshar Chaklashiya
Ogwu Augustine

Date: 15-09-2022

Outline

- Executive Summary
- Problem Statement
- Approach
- Data Overview
- Exploratory Data Analysis (EDA)
- EDA Summary
- Recommendations



Executive Summary

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).

Bank wants to use ML model to shortlist customer 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.



Problem Statement

Data Glacier Objective : Provide actionable insights to help XYZ firm in identifying the right customers for targeting the marketing campaign.

This will save resource and their time (which is directly involved in the cost (resource billing)).

Data Glacier did a 1 month pilot focusing on these tasks:

- Data Intake Report
- EDA Notebook
- ML Model Proposal
- Presentation to ABC's Executive team (Today)

Approach

1 Data Understanding and Extraction

2 Data Cleaning/Modelling

3 **Exploratory Data Analysis**

4 Recommendations

5 ML Model Proposal



UCI ML
Repository



bank-additional-
full.csv



Data Overview

Bank Data

- 21 Features
- 41188 data points
- 5834924 bytes

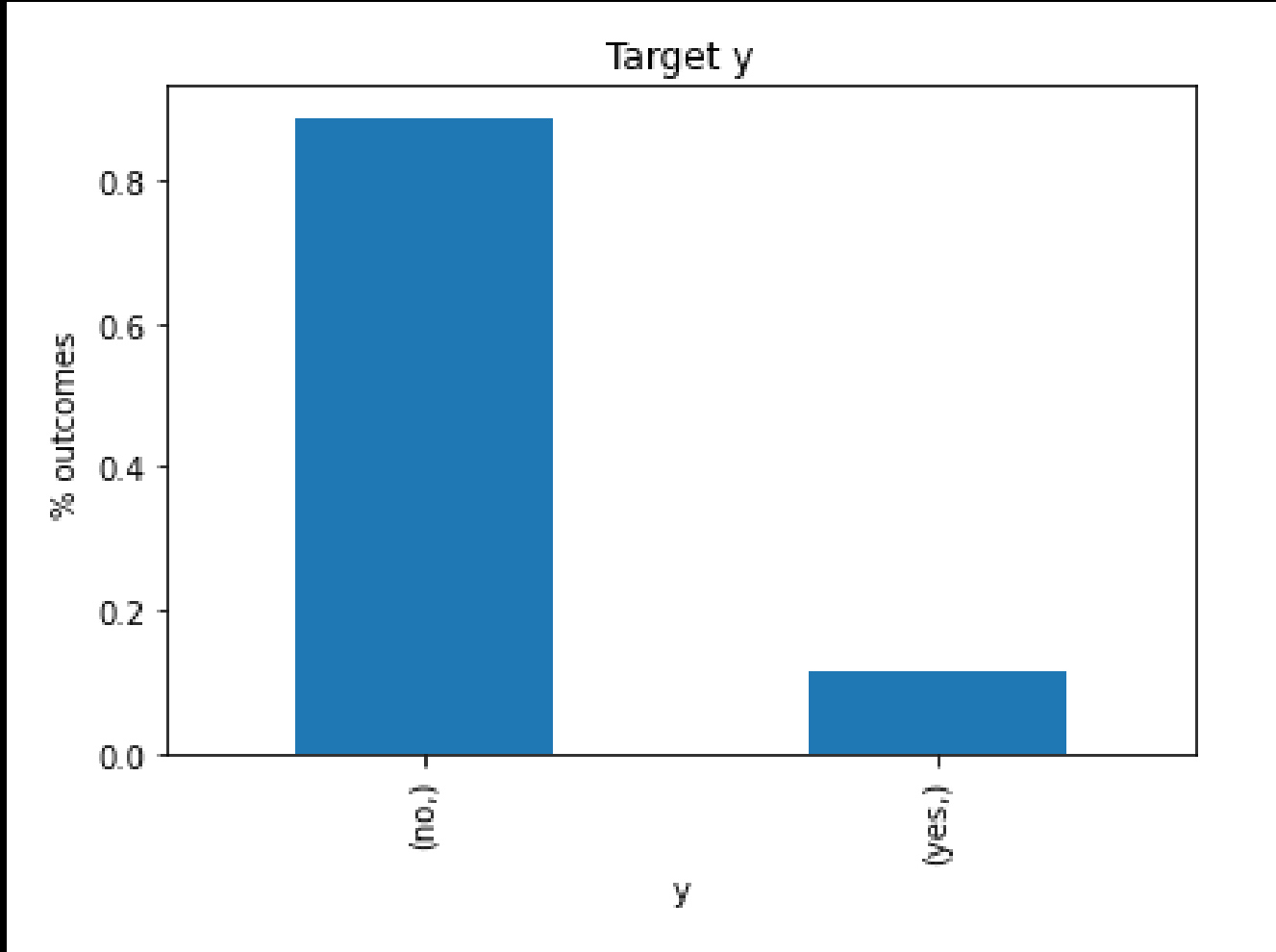


Exploratory Data Analysis (EDA)

The Exploratory Data Analysis (EDA) has been divided into four parts:

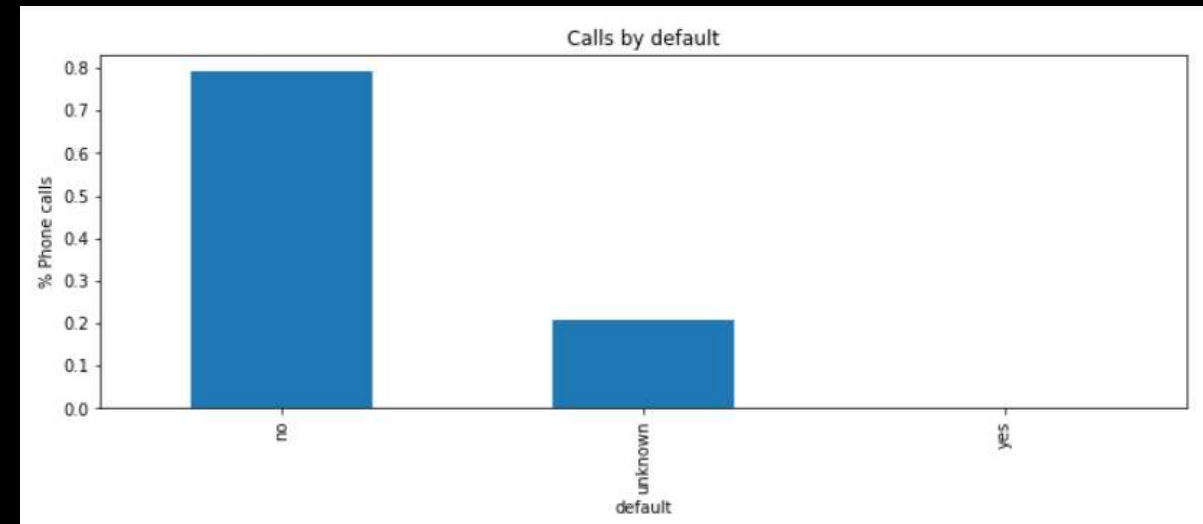
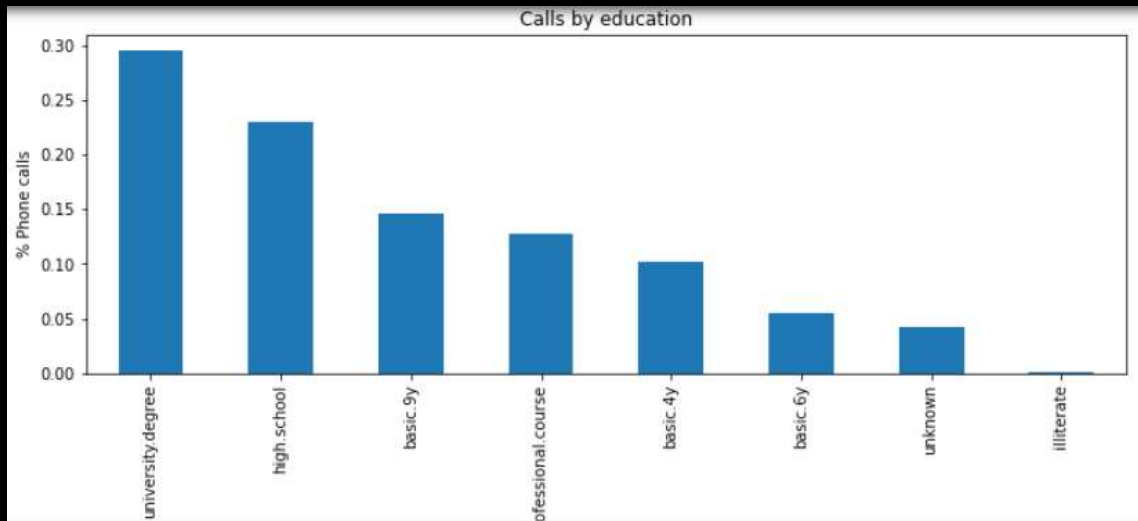
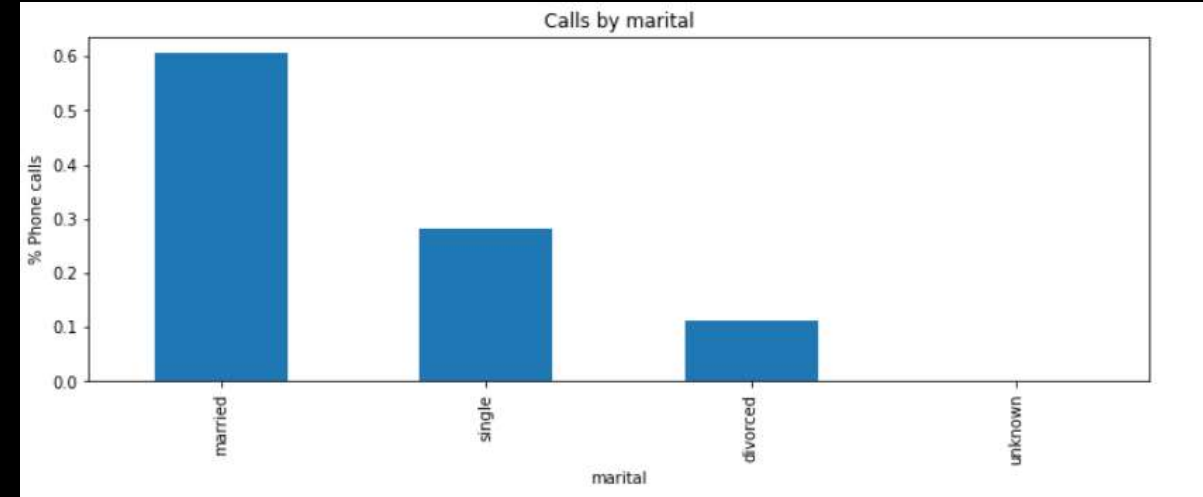
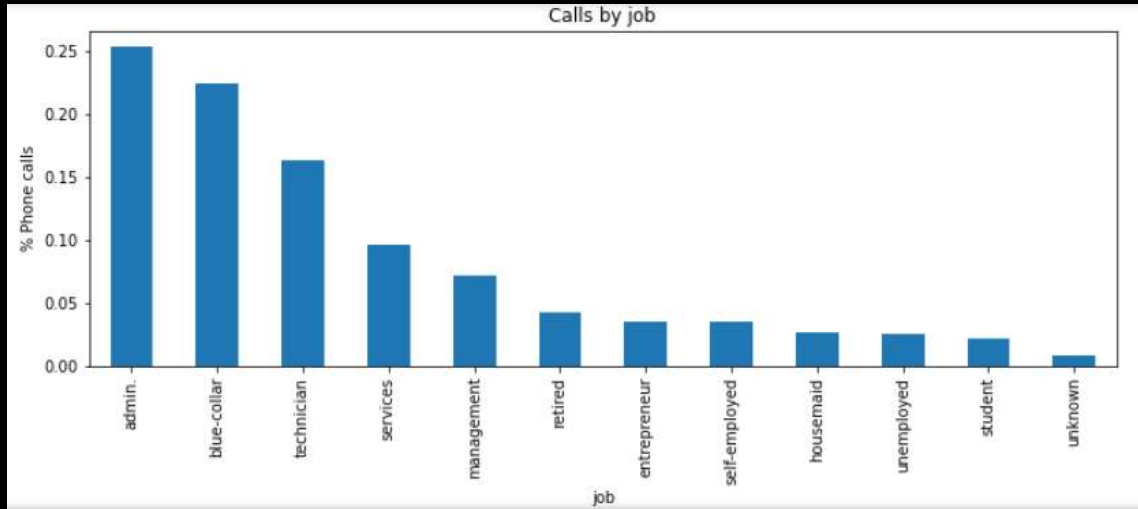
- Market Overview: Categorical and Numerical Variable Analysis
- Target Variable Analysis
- Correlation Analysis
- Time Series Analysis

Target variable Analysis

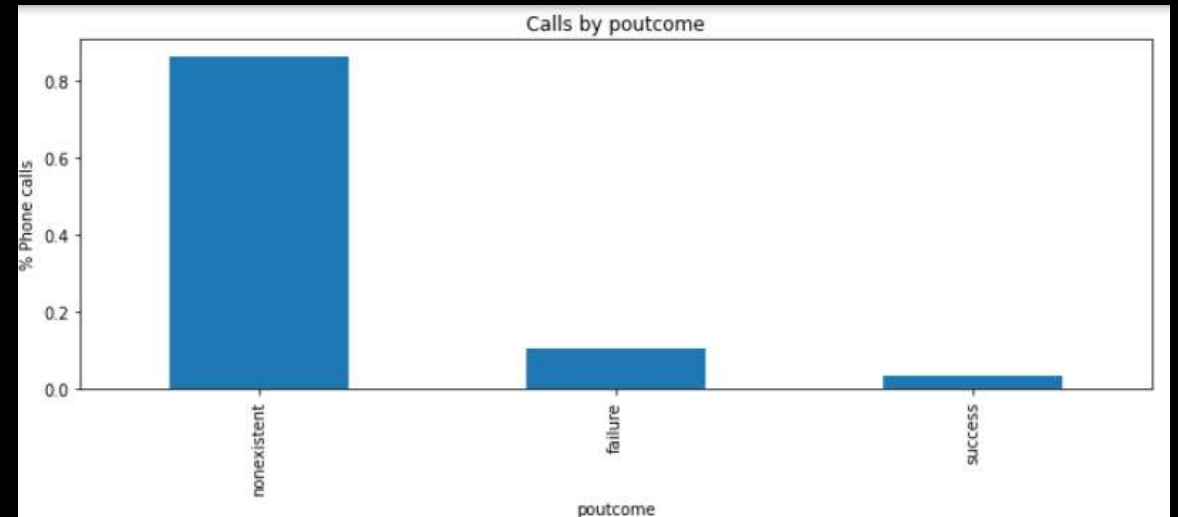
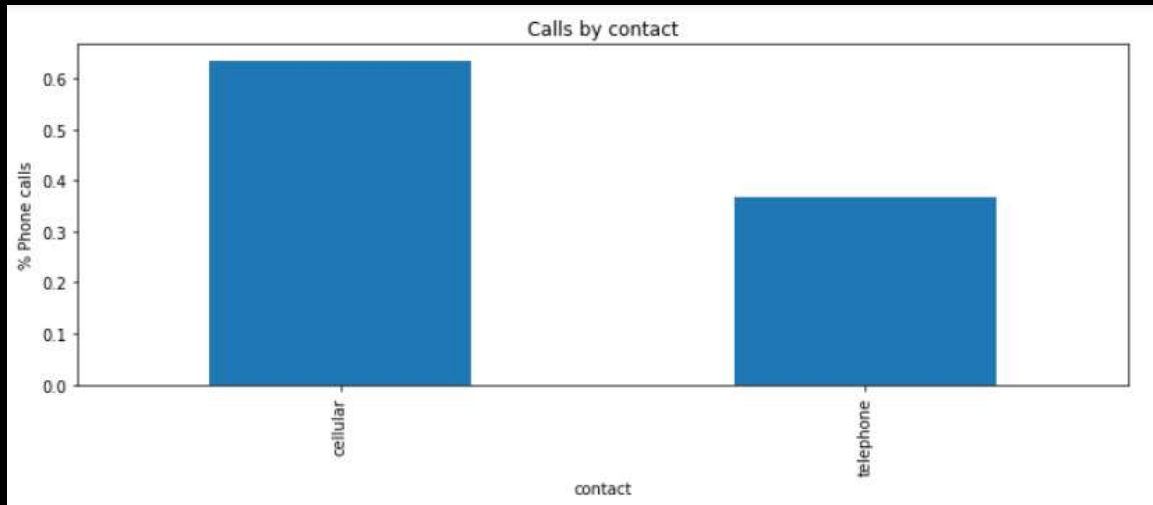
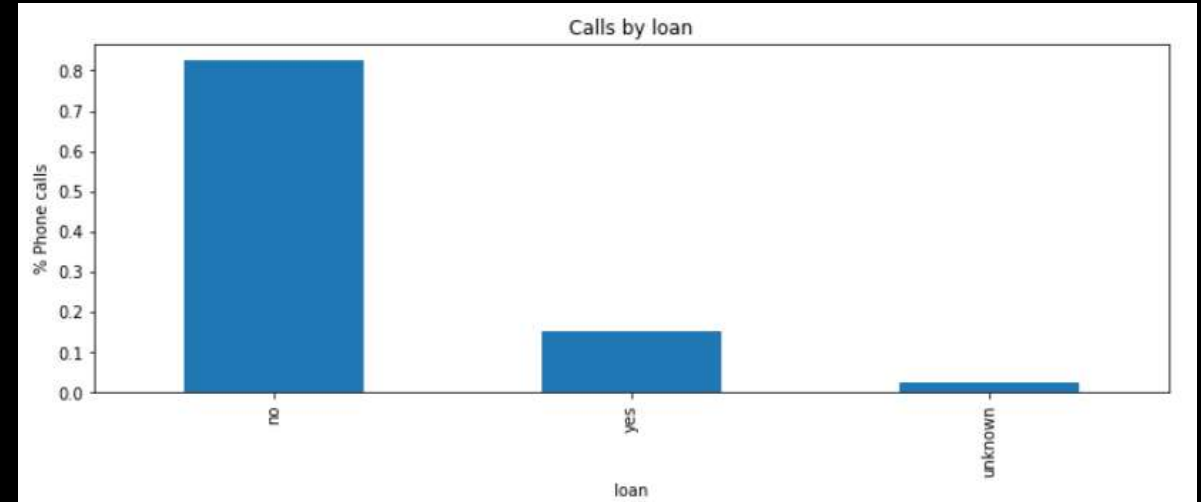
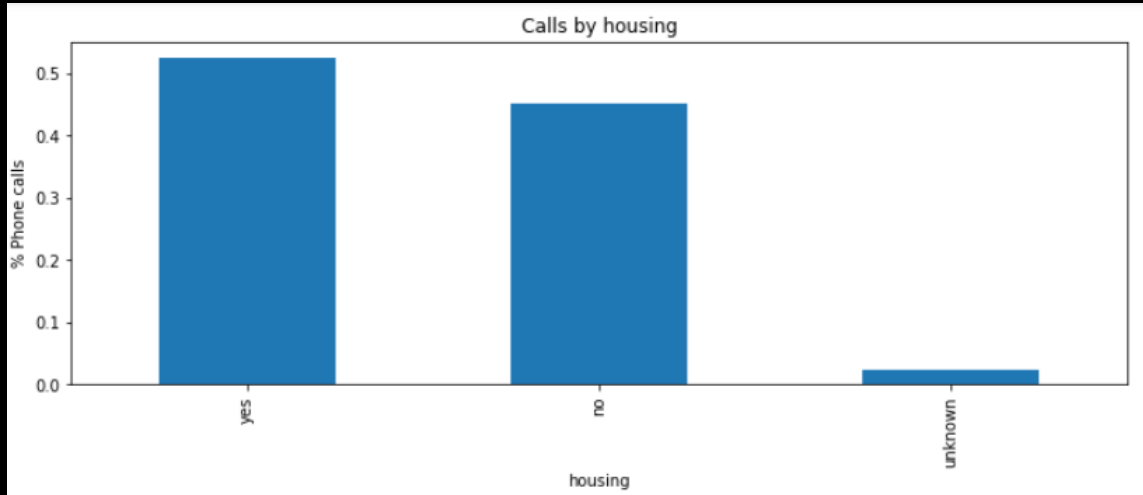


- An imbalance in the target variable.
- An oversampling technique must be applied before model building

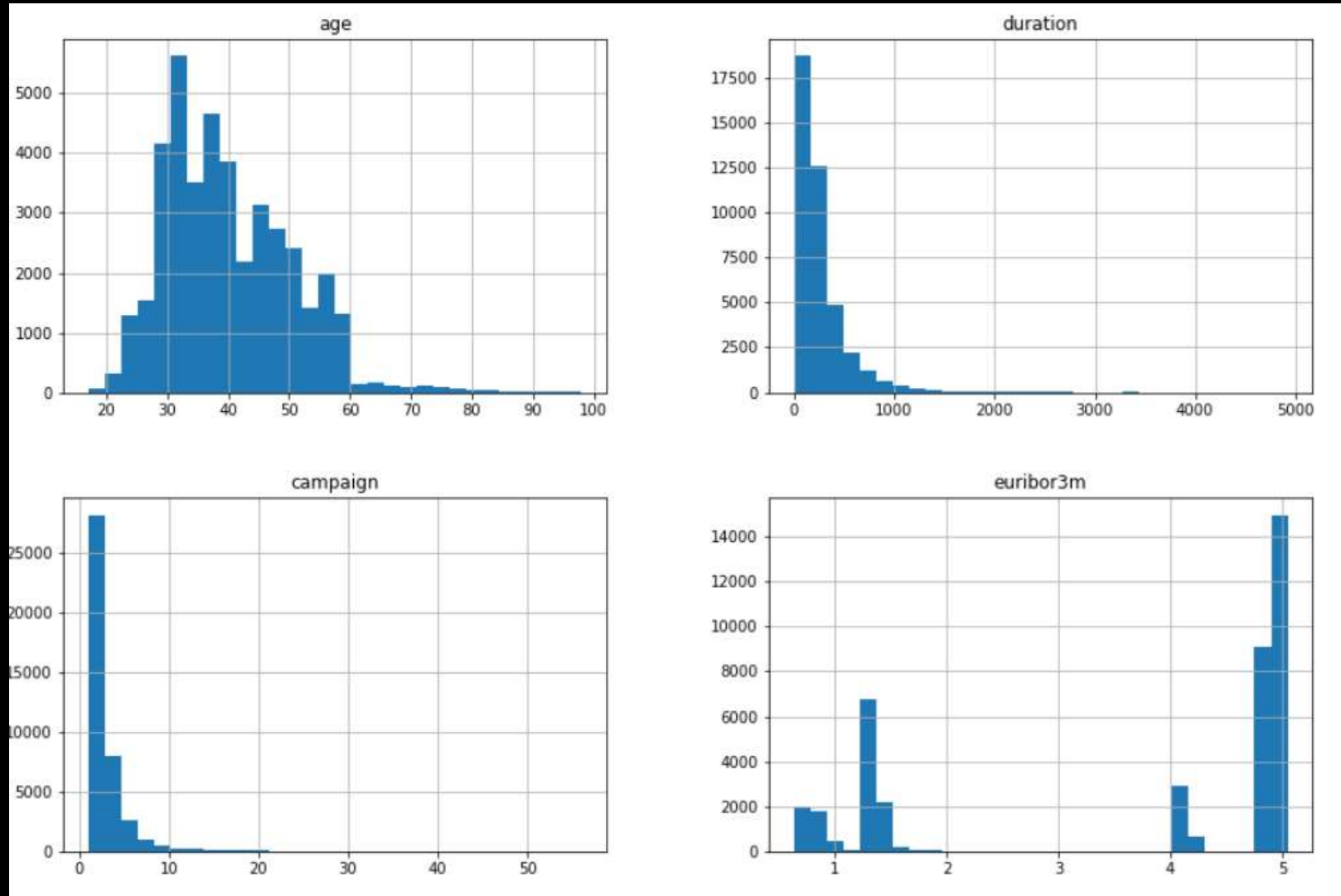
Market Overview – Categorical Variable Analysis



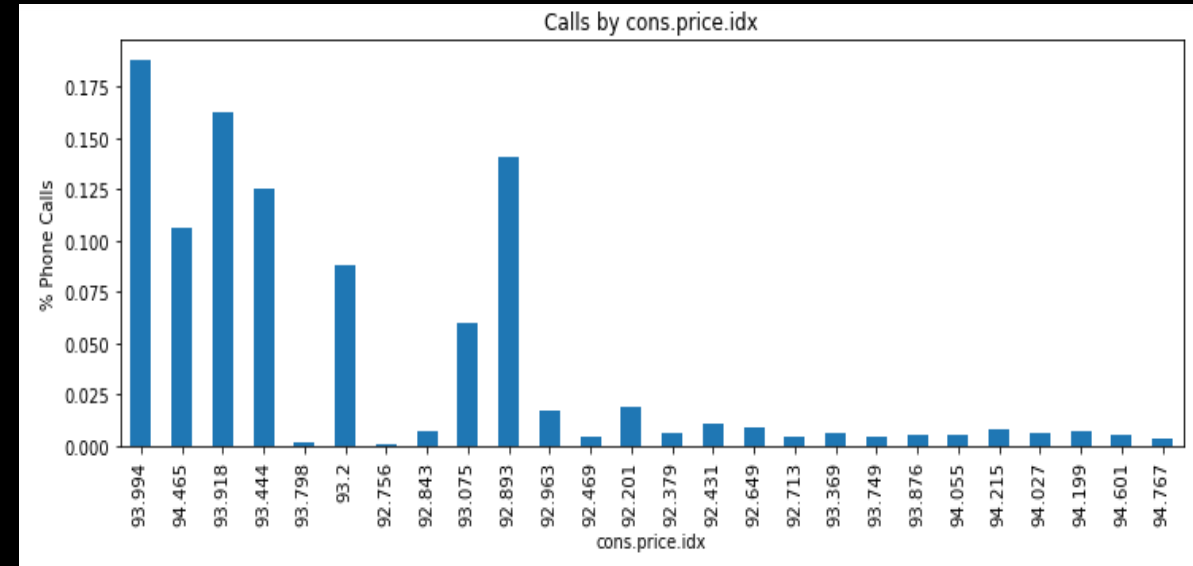
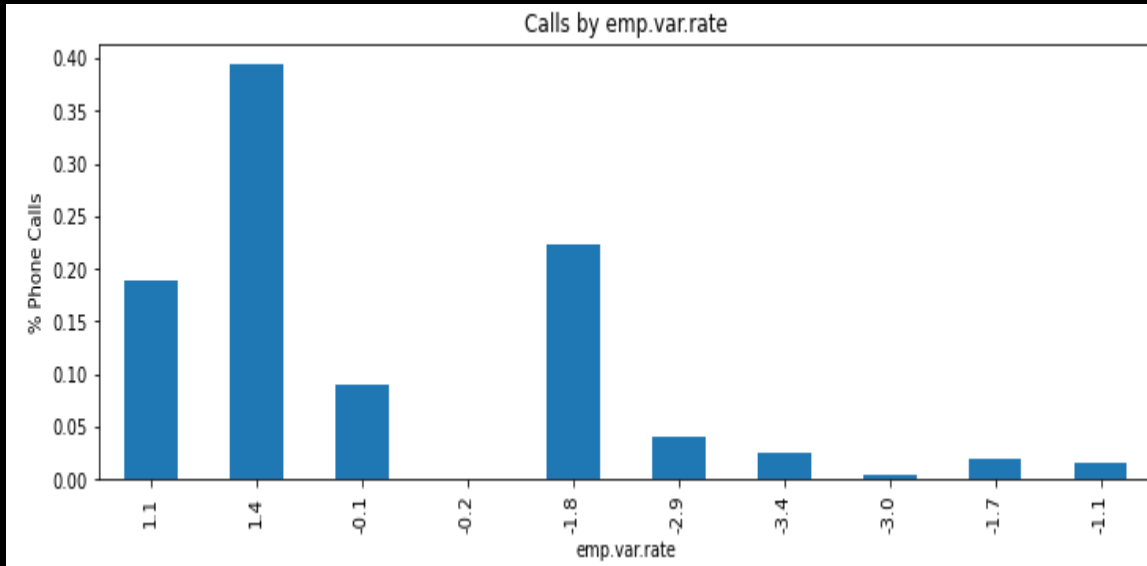
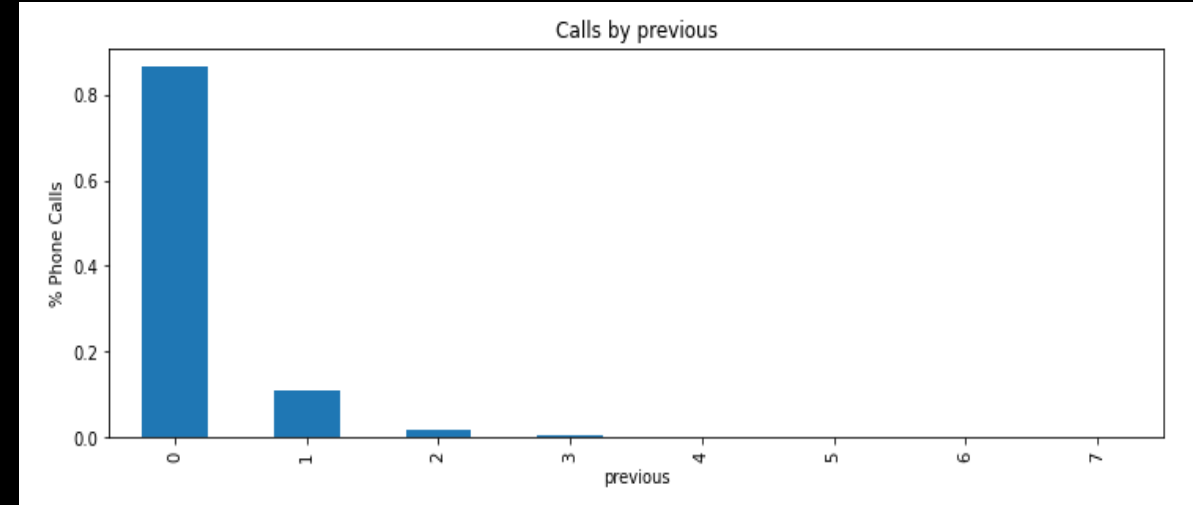
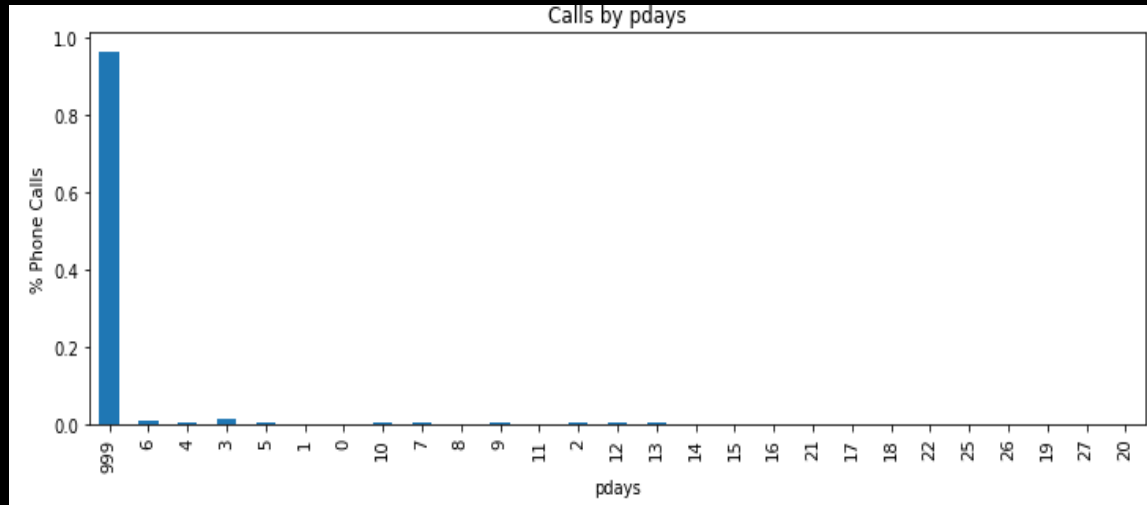
Market Overview – Categorical Variable Analysis



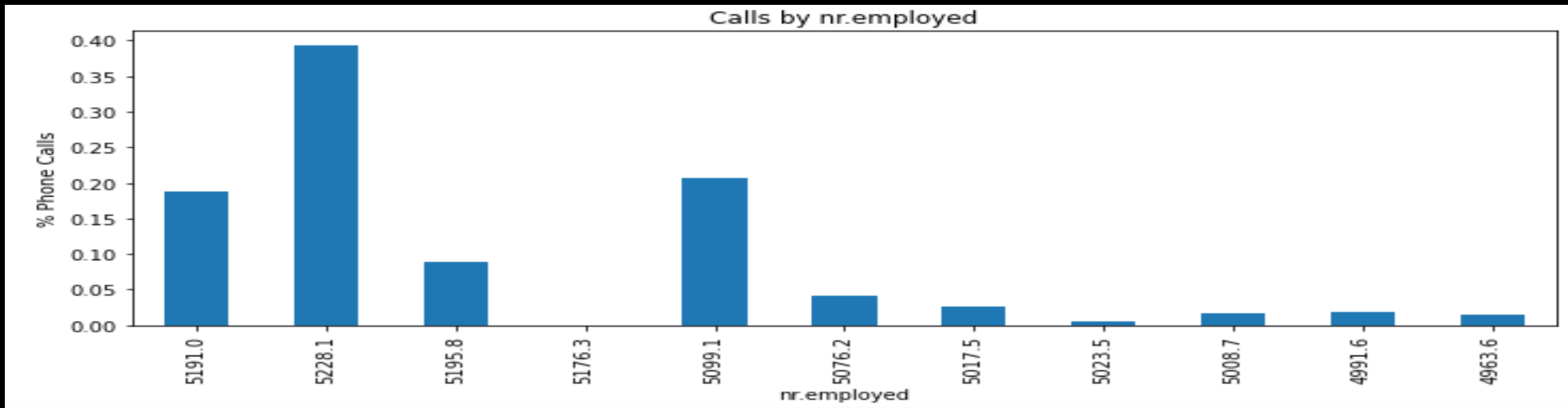
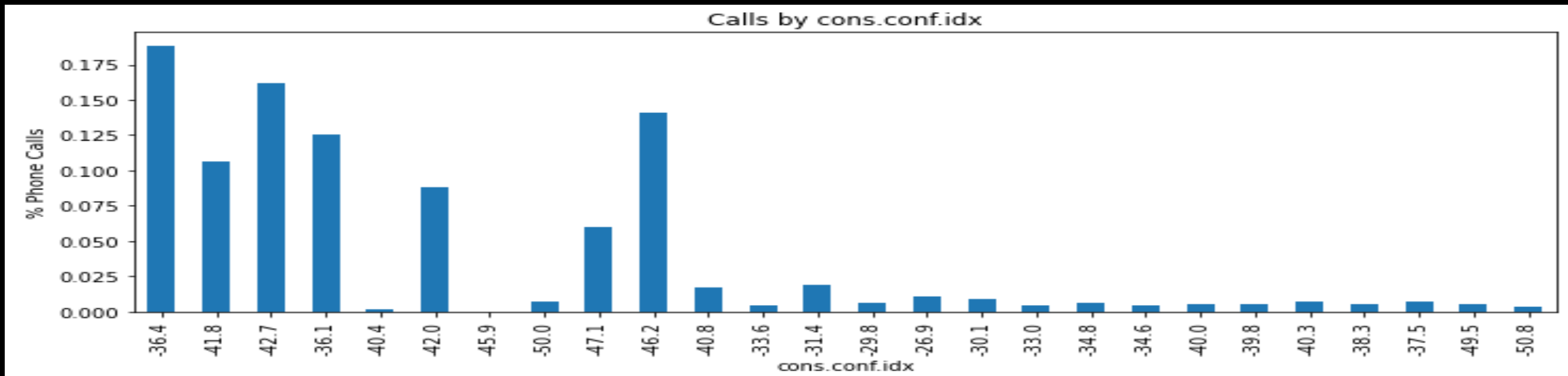
Market Overview – Numerical Variable Analysis



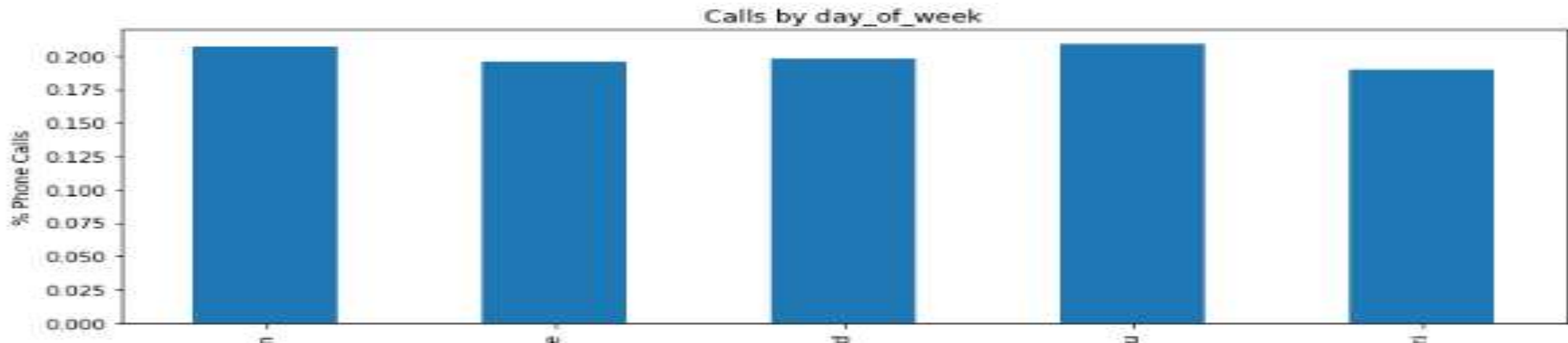
Market Overview – Numerical Variable Analysis



Market Overview – Numerical Variable Analysis



Temporal Variable Analysis



Missing and Unknown Value

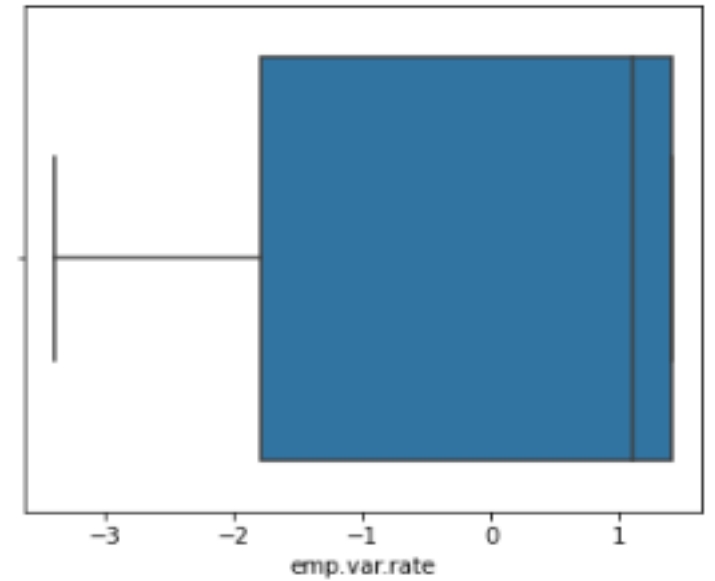
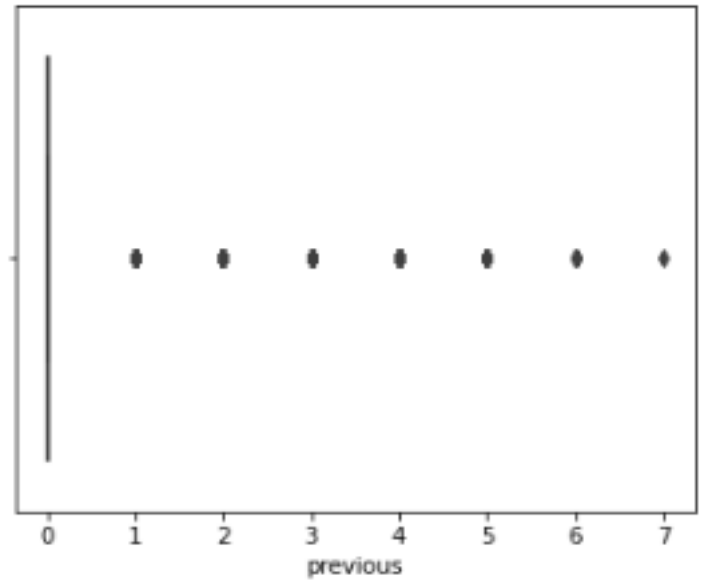
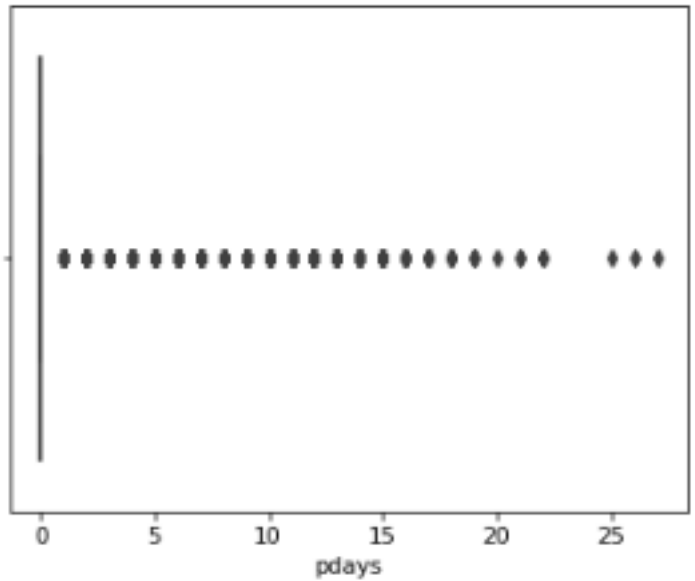
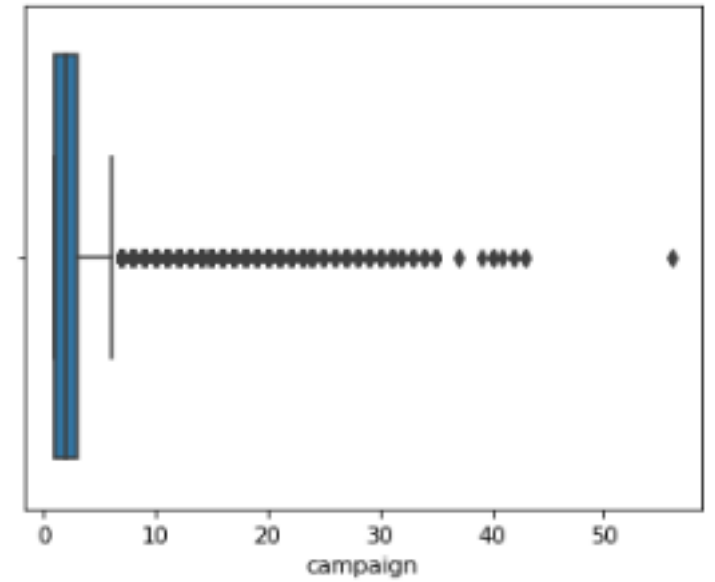
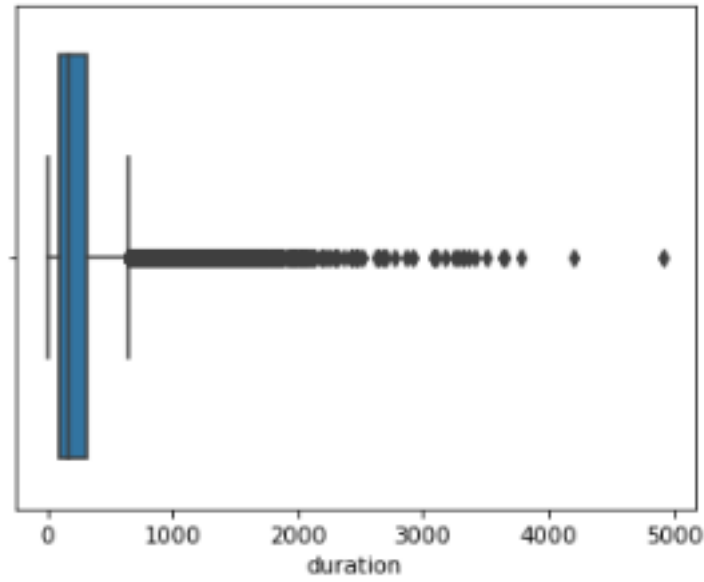
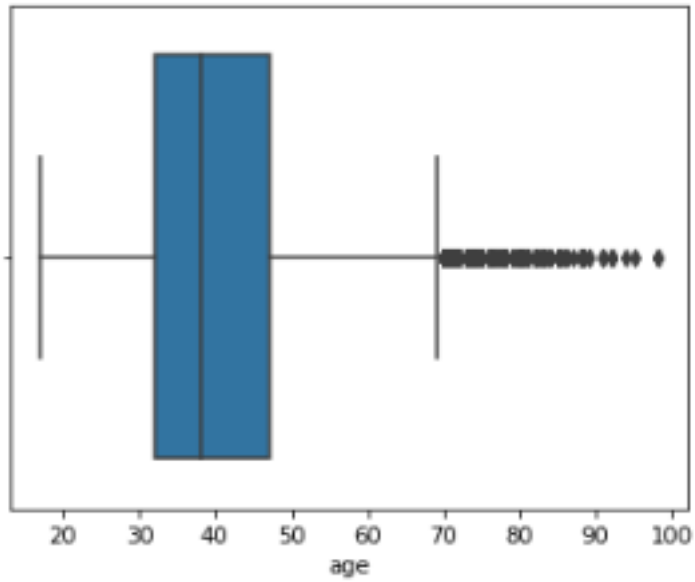
- In our dataset, we don't have any missing value
- For the Unknown value, we found five feature (job, education, default, housing and loan). Row number is relatively small so we dropped those rows.

Handle outlier

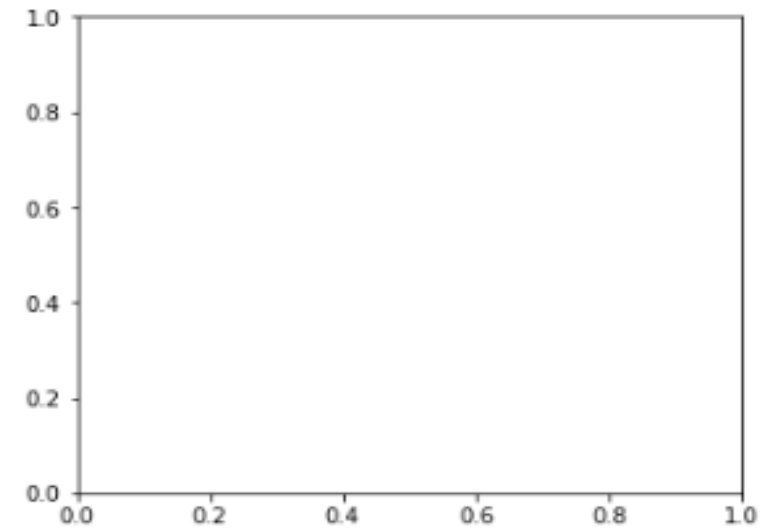
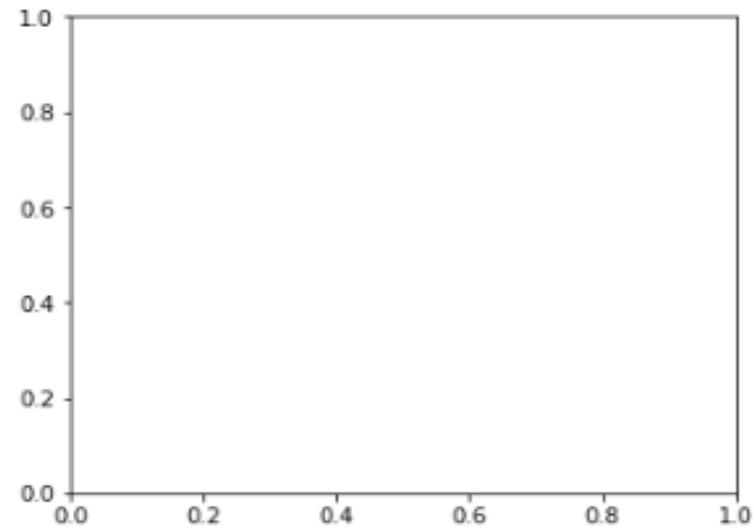
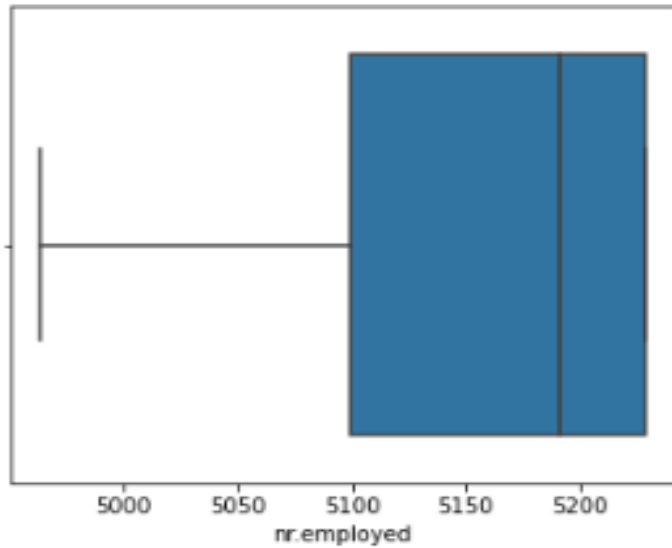
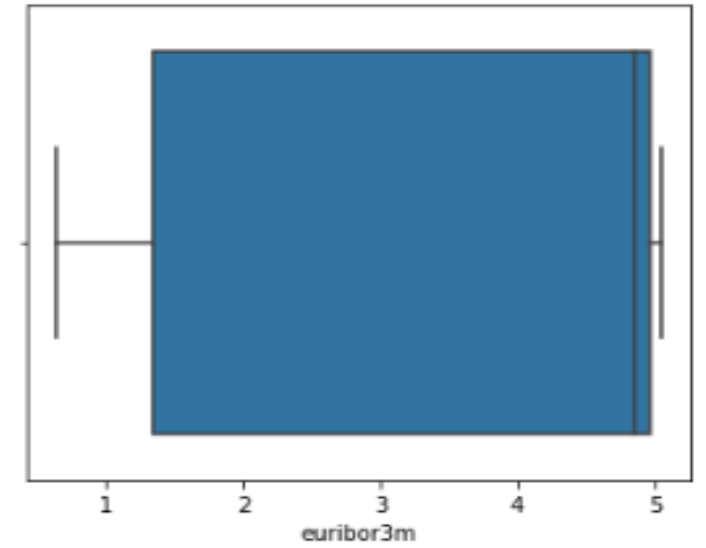
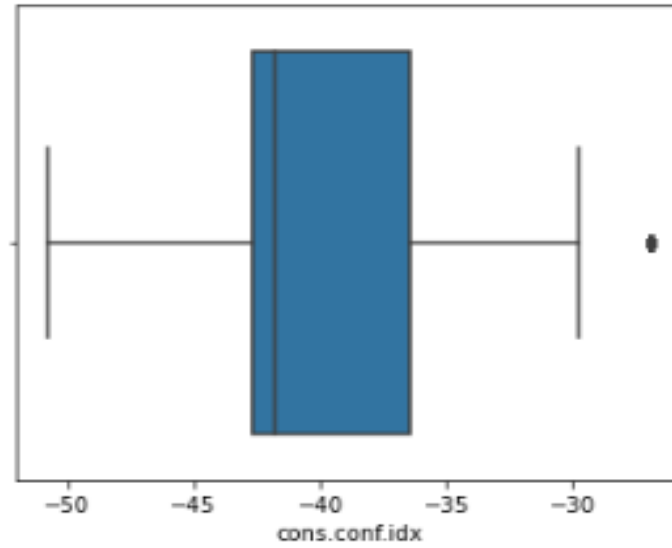
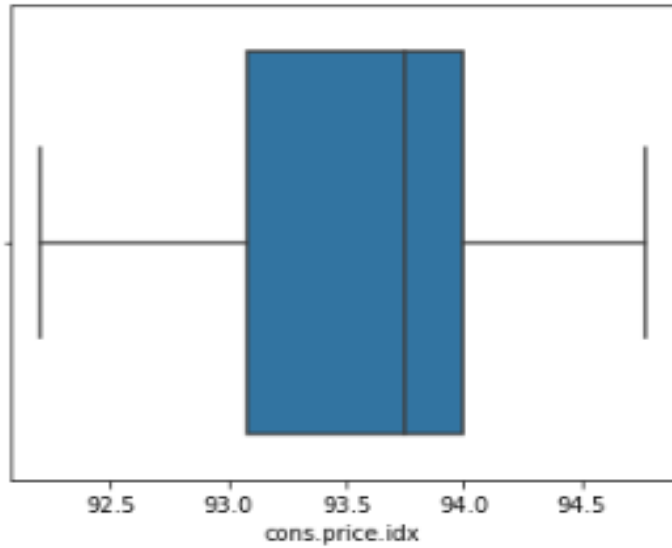
- we are using box plot method to detect outlier.

Outlier points = $Q3 + 1.5 \text{ IQR (Upper Quartile)}$
 $Q1 - 1.5 \text{ IQR (Lower Quartile)}$

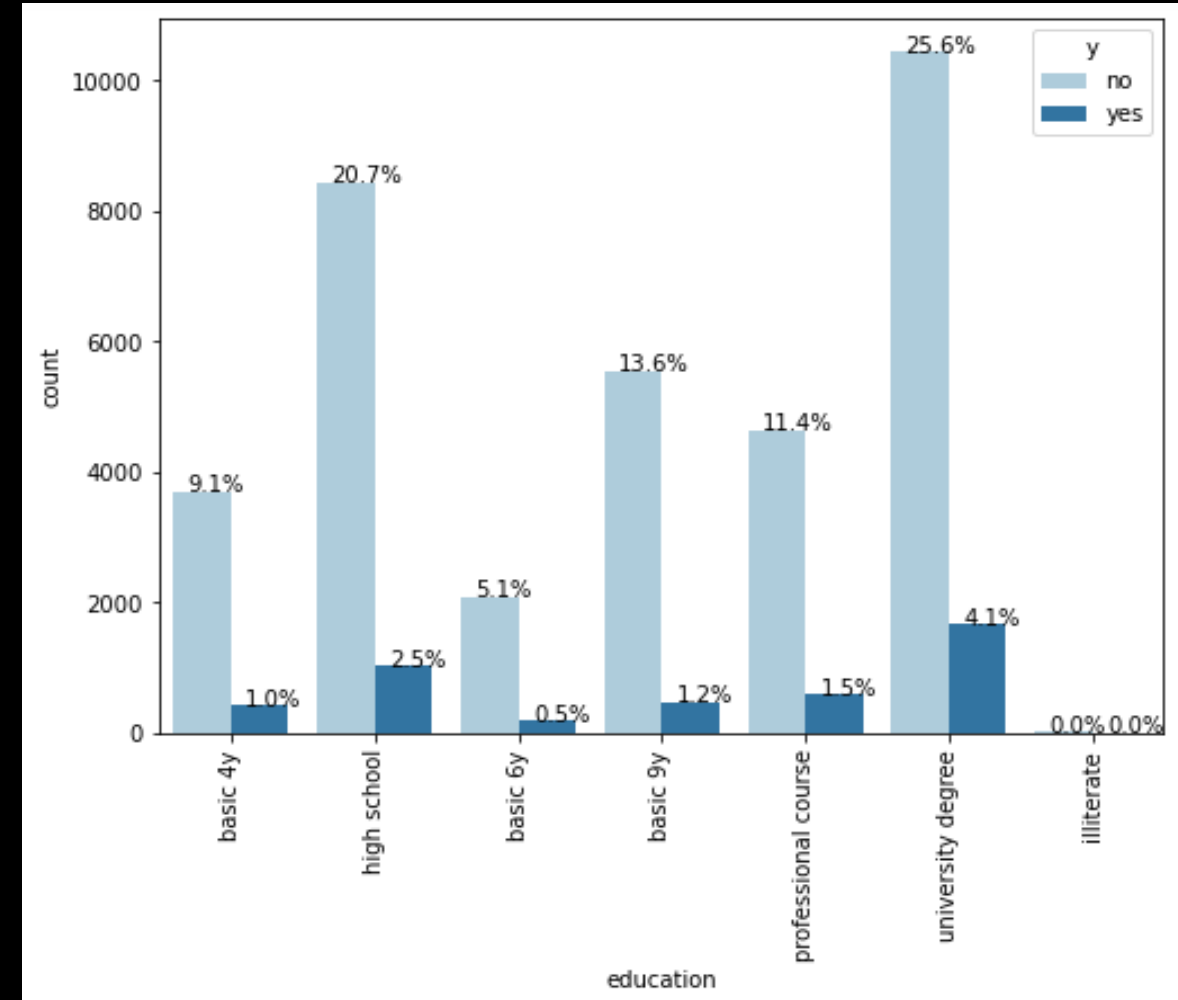
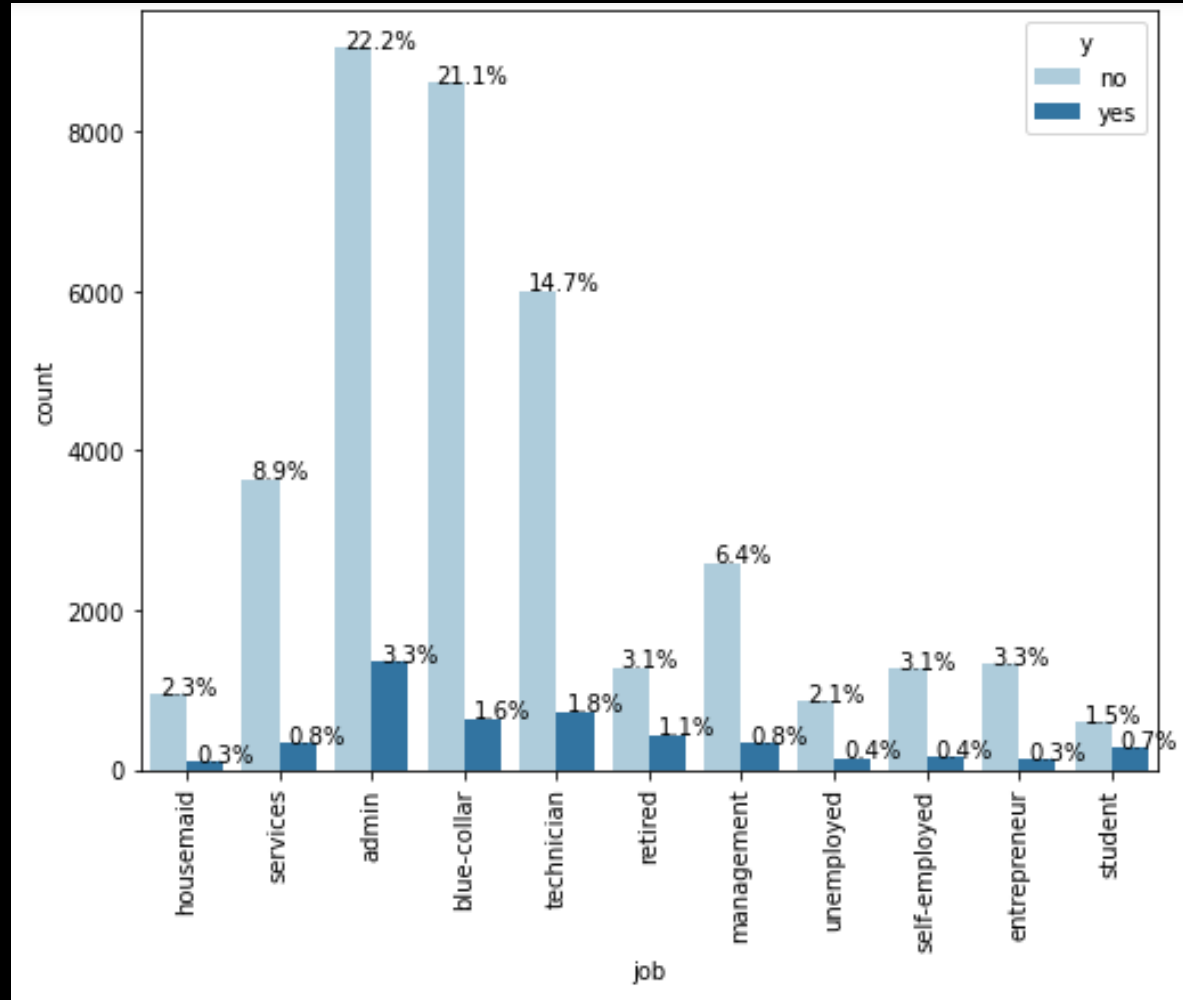
Box plot graph - I



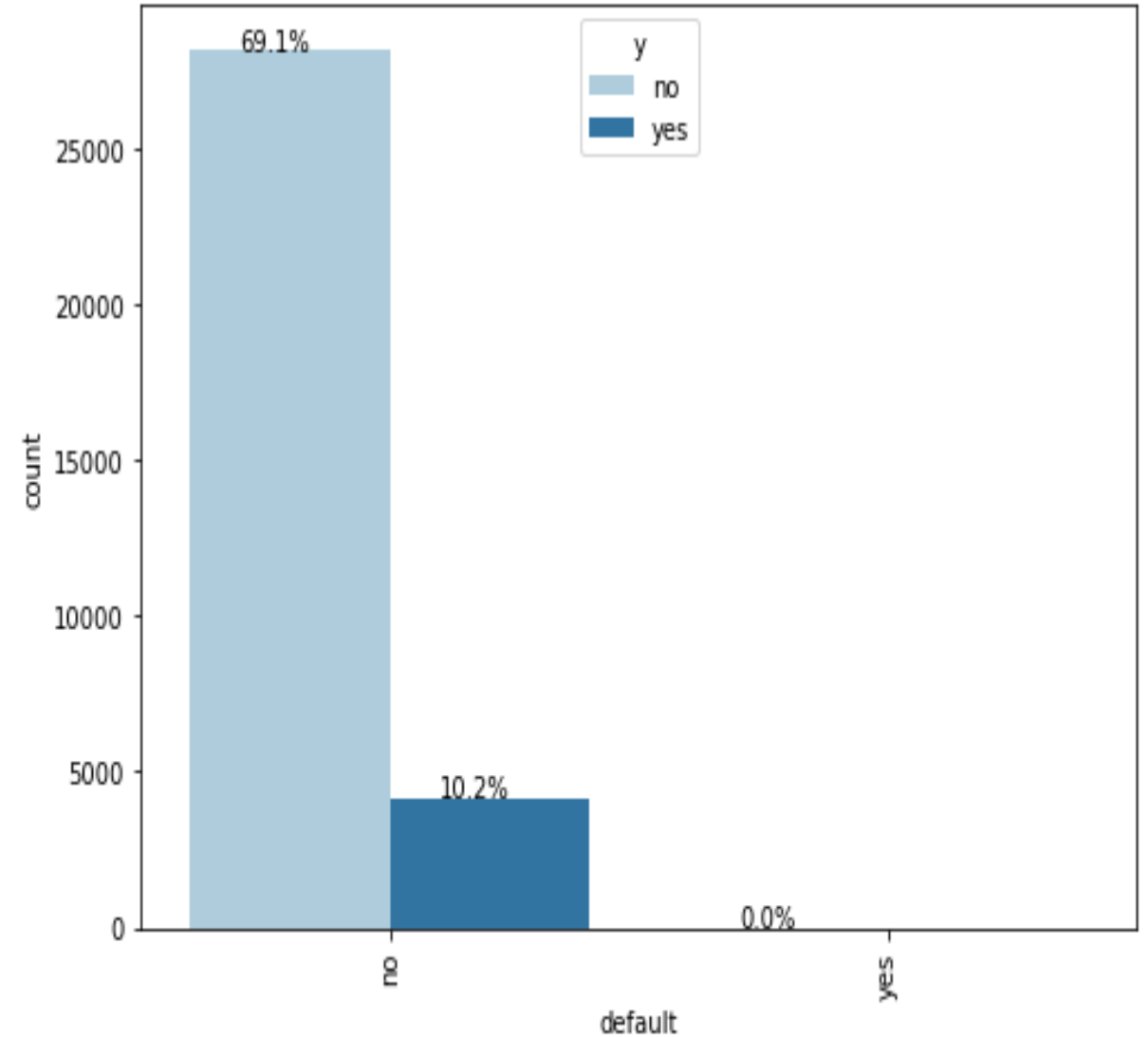
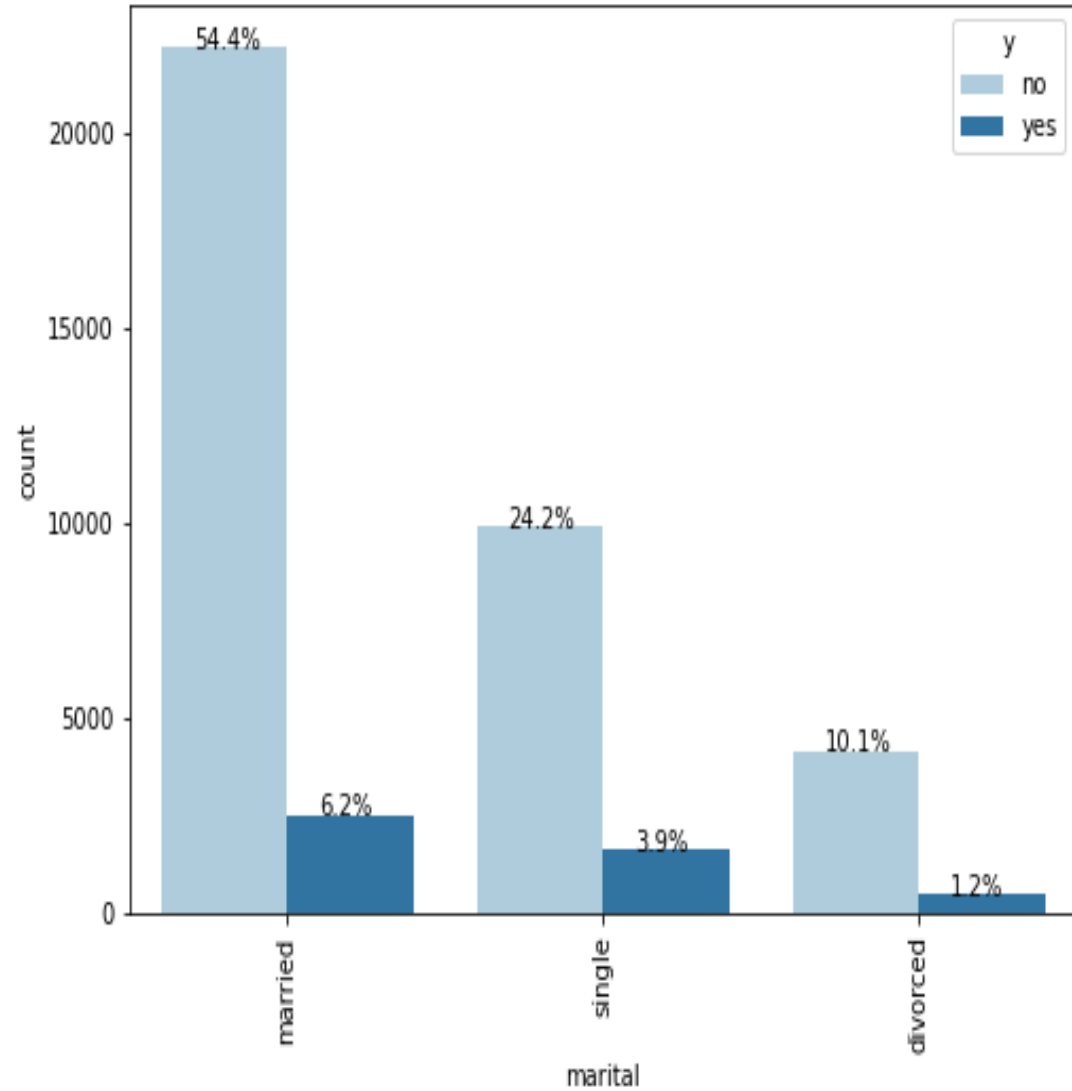
Box plot graph - II



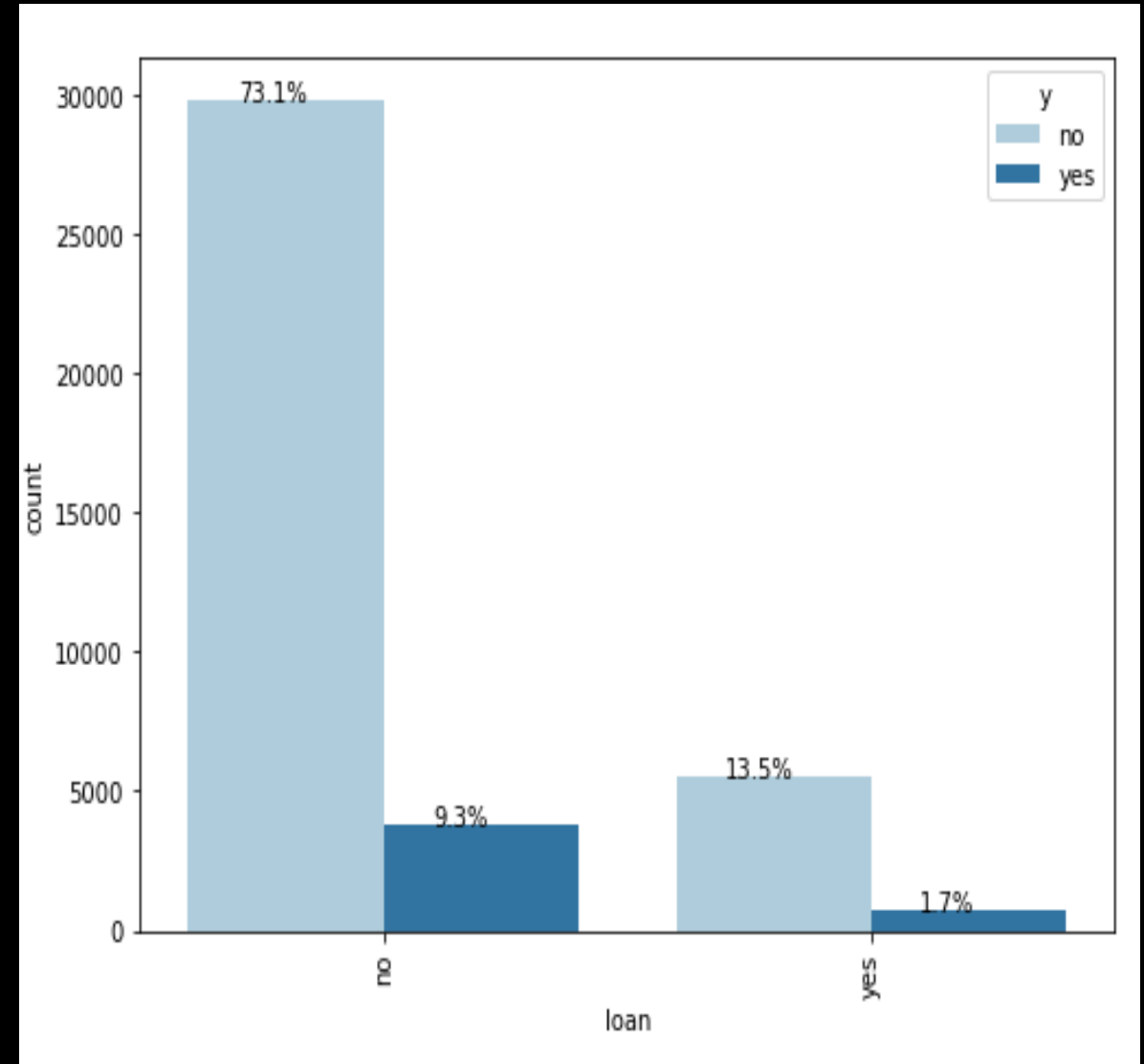
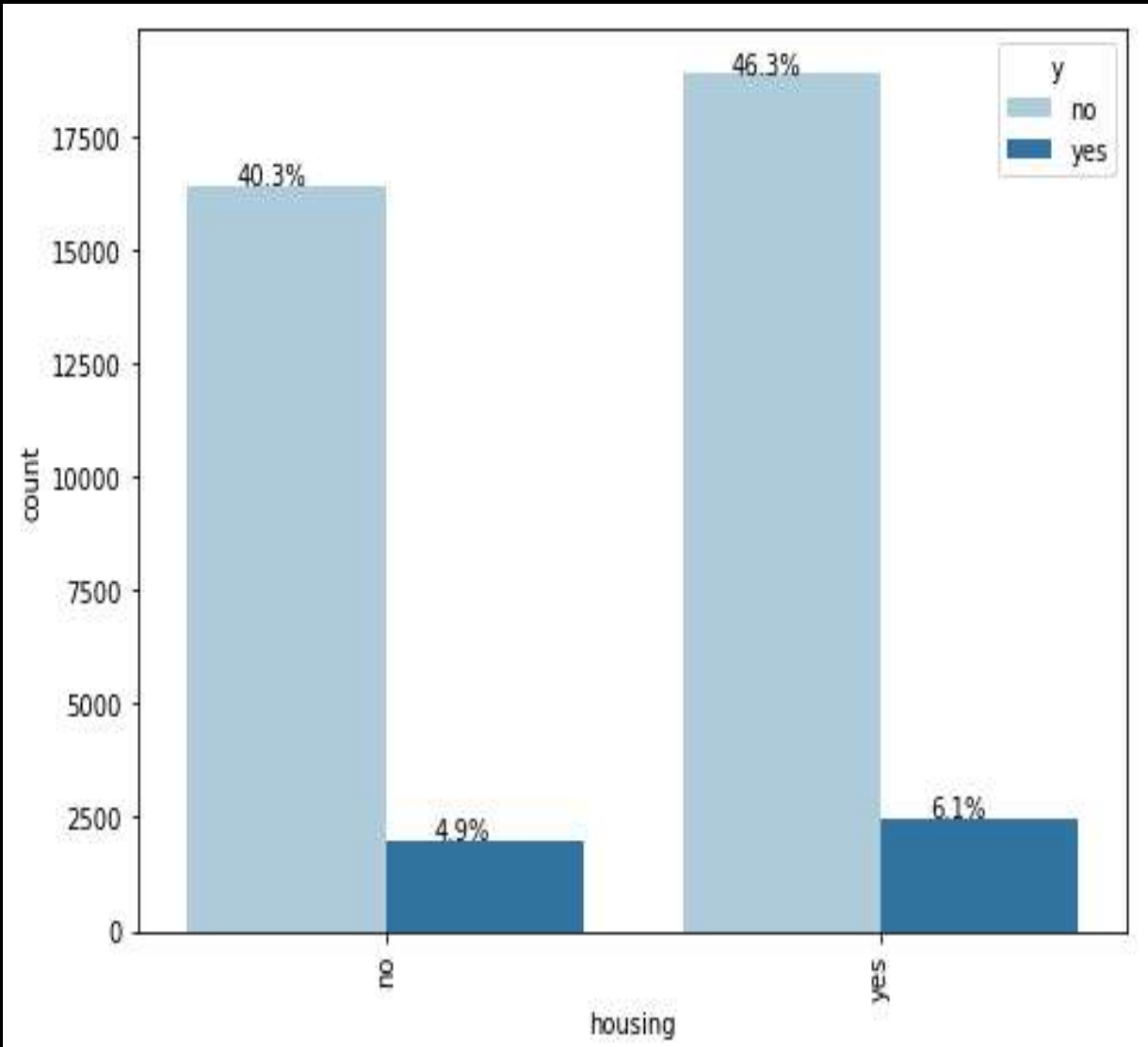
Categorical variable analysis - I



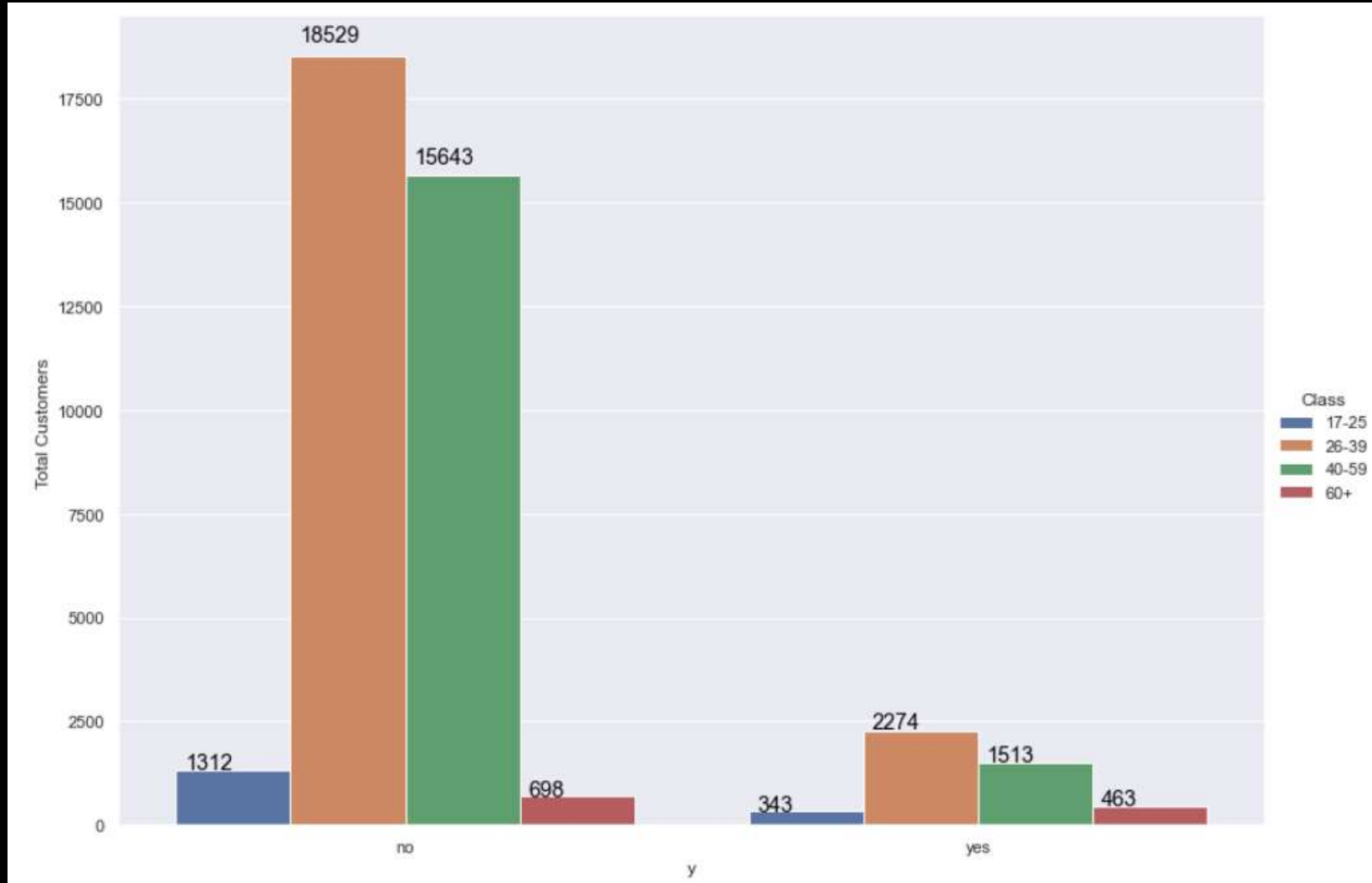
Categorical variable analysis - II



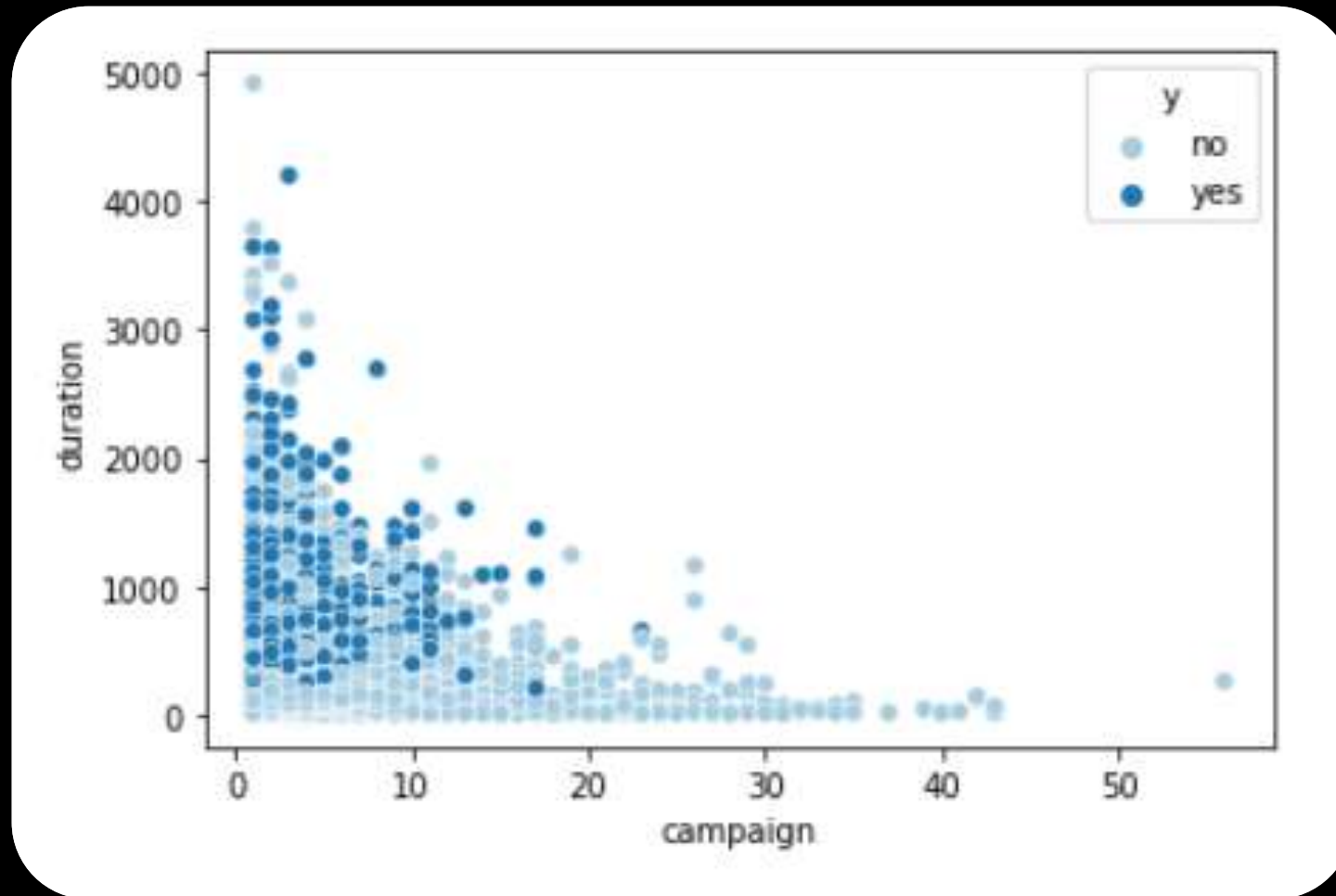
Categorical variable analysis - III



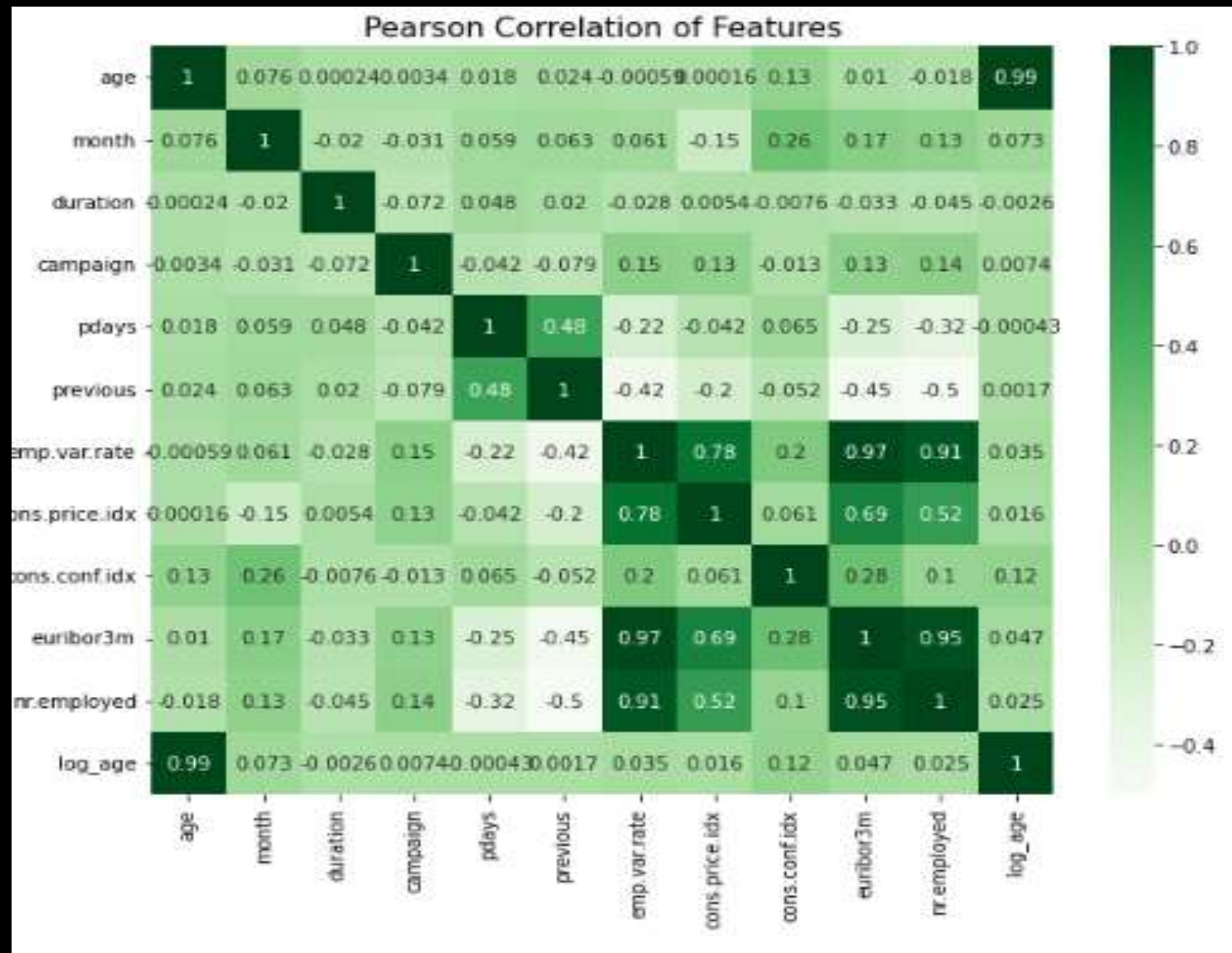
Target Variable Analysis – By Age Groups



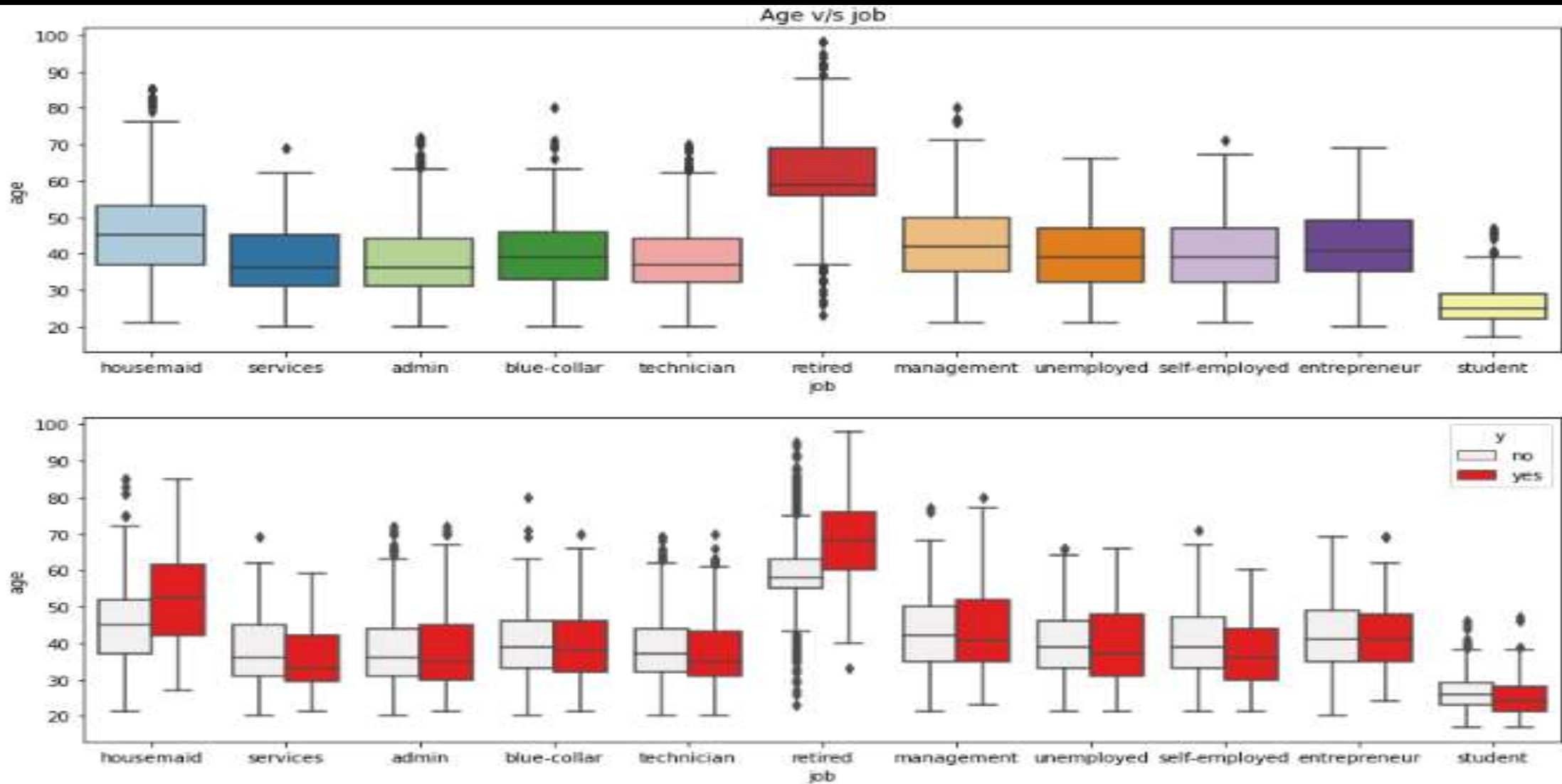
Correlation Analysis



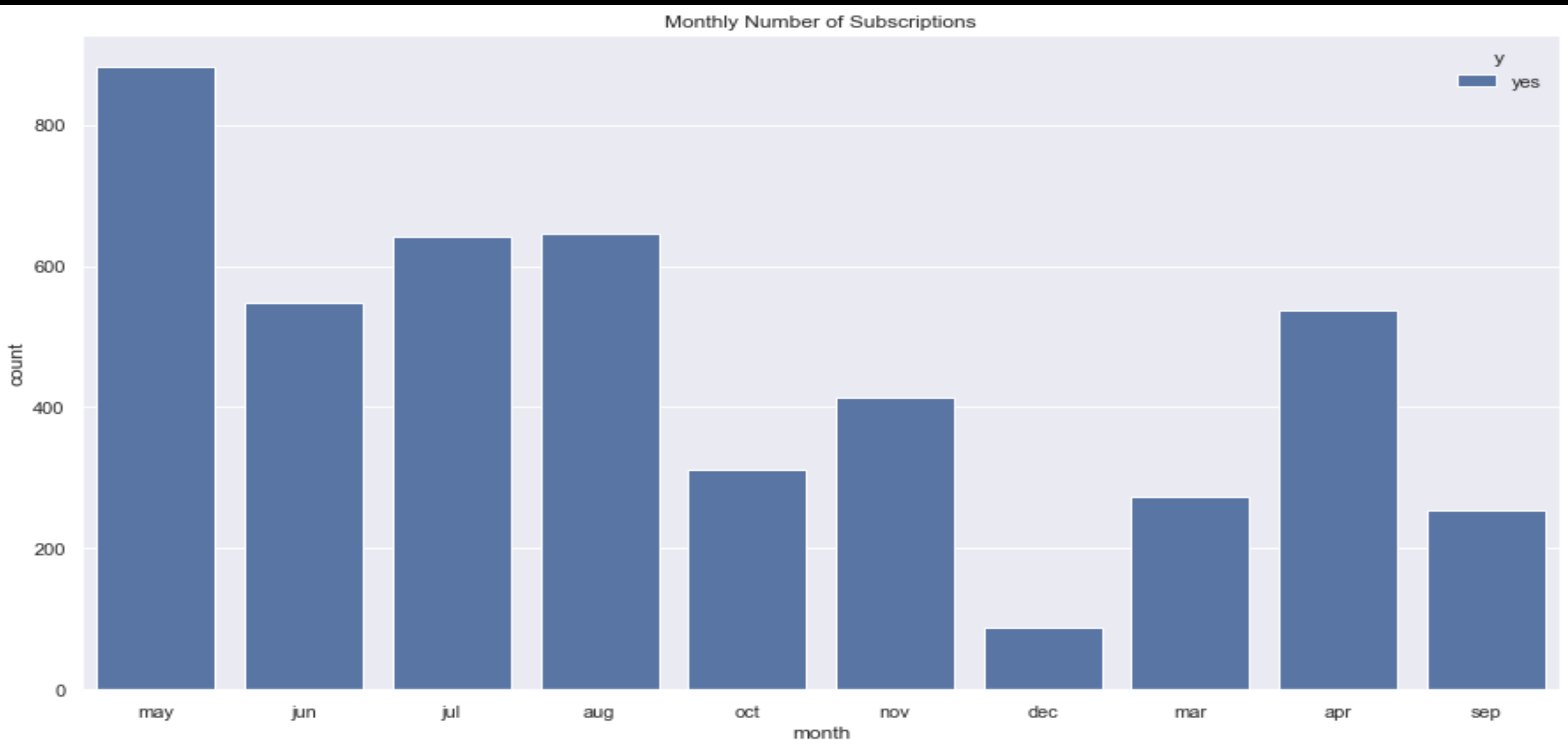
Correlation Analysis



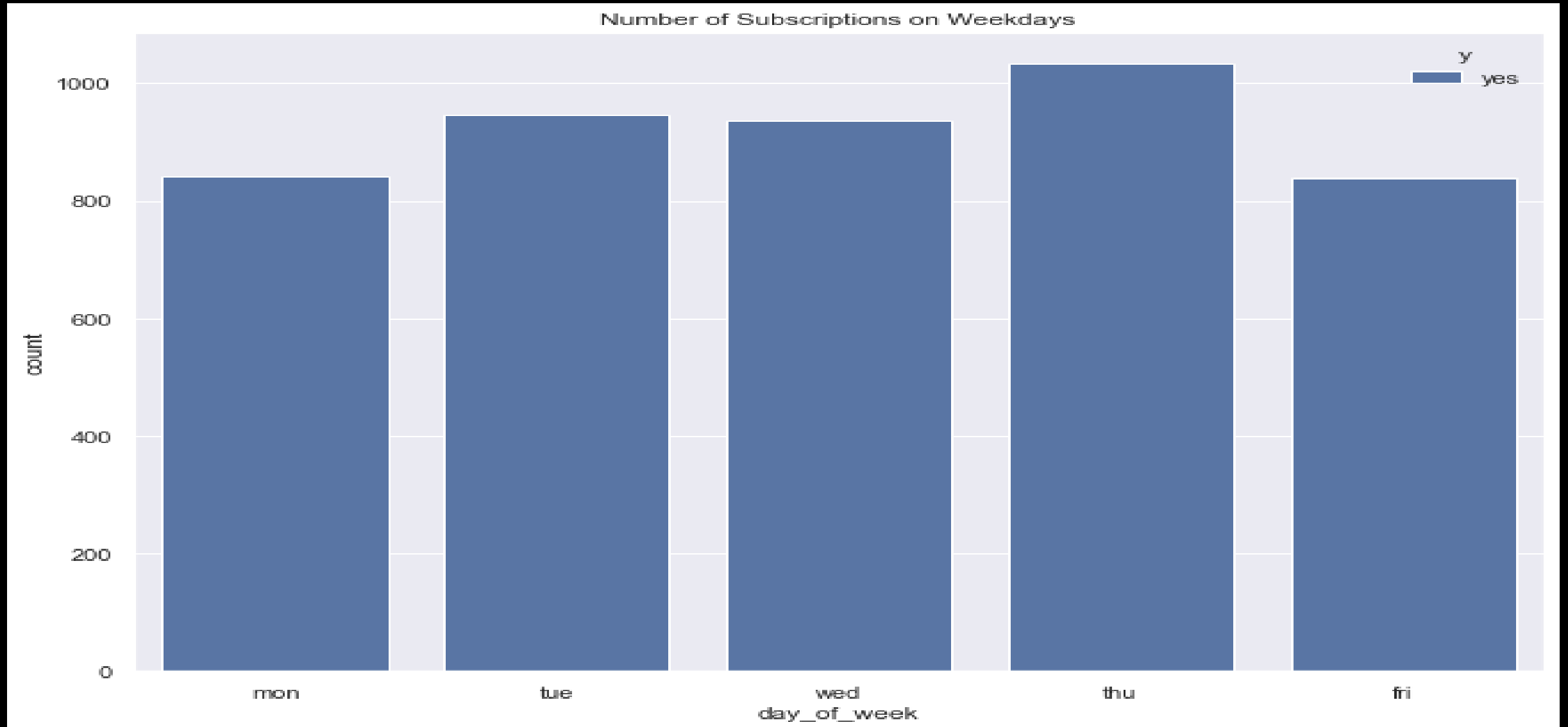
Age vs Job Analysis



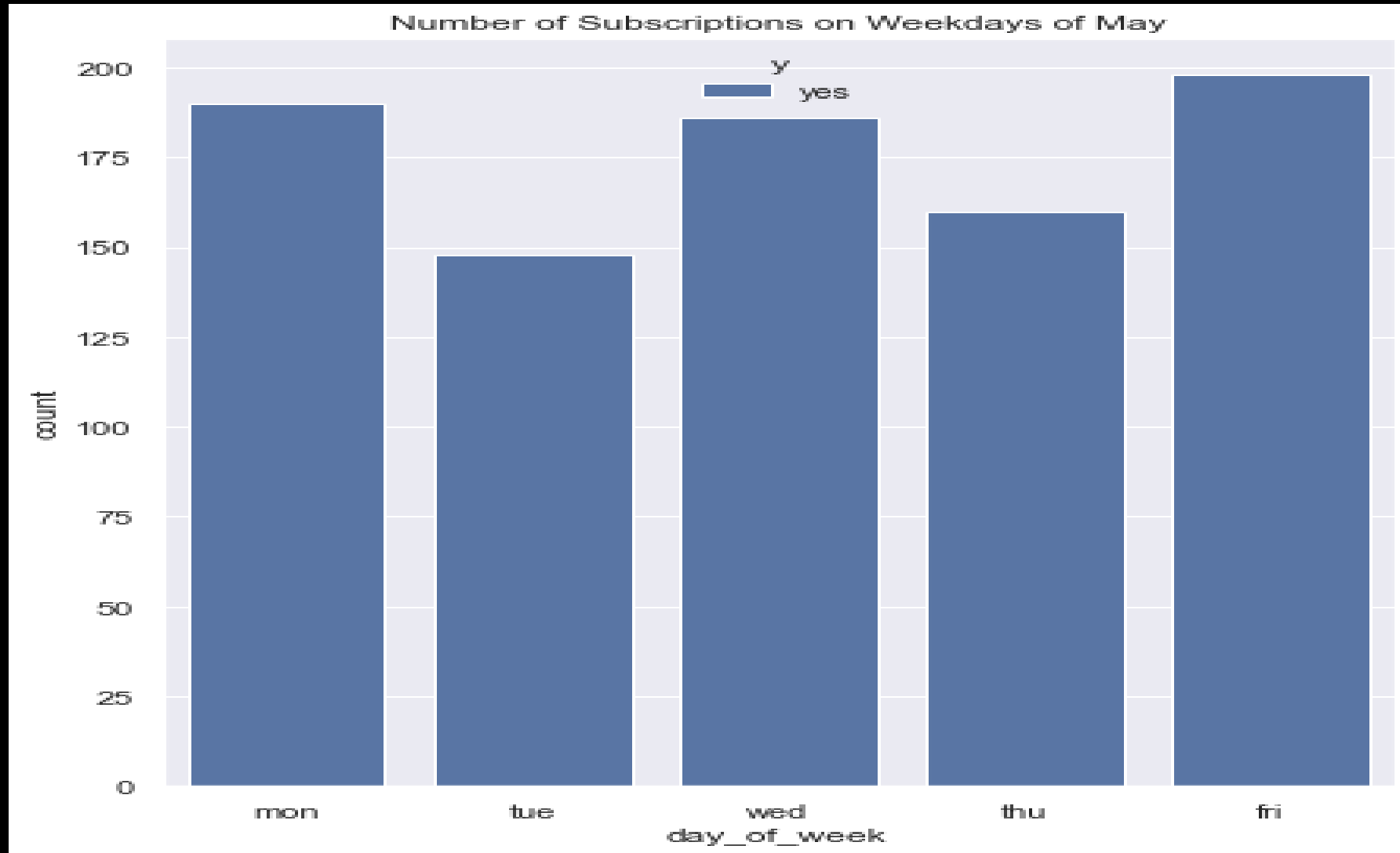
Time Series Analysis – Monthly Demand



Time Series Analysis – Week Days Demand



Time Series Analysis –Week-Day Demand of May



EDA Summary - Insights

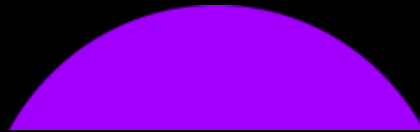
May

Month with most subscribers



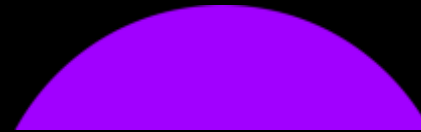
26-39

Age Group with most transactions



Thursday

Day of the week with most subscribers



Cellular

Best Mode of Contact



EDA Summary - Insights

May

Month with most subscribers



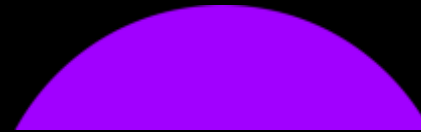
26-39

Age Group with most transactions



Thursday

Day of the week with most subscribers



Cellular

Best Mode of Contact





Recommendations

1. The outcome of a previous marketing campaign is a key feature for determining user behavior. If the previous campaign is successful then the user is more likely to subscribe than reject it. But if it is a failure then also some portion of the people still subscribe to the new one.
2. The people who are contacted through cellular have a higher conversion rate than telephone. Campaign should focuss more on cellular calls.
3. The people who are in default should not be contacted.
4. Retired and students people should be approached more as they show high conversion rate. However unemployed and housemaids should be contacted less as they show the lowest conversion rate.
5. Once the customer is contacted 25 times they should not be contacted further.



Recommendations

6. The older age group 60+ is more vulnerable to marketing campaign as they have a high conversion rate to subscription and therefore should be targeted more.
7. The more the customers are contacted the more they will subscribe. The efficiency of the salesmen should be supervised.
8. Mondays seem to have a lower conversion rate. The people should be contacted more on other days especially Thursdays.
9. There should be an increase in the contacts during March, September and October which shows promising conversion rate.
10. Thursdays of April seems to have an unusual large subscriptions. The reason behind this should be further investigated.



Model Proposal

- Logistic Regression
- Neural Network
- Random Forest

Thank You!



Data Glacier

Your Deep Learning Partner