



**Data Glacier**

Your Deep Learning Partner

# Exploratory Data Analysis

## Bank Marketing Campaign

---

Nuri Öztürk

Chitra S Chaudhari

16-March-2023

# Agenda

- Executive Summary
- Data Understanding
- EDA
- EDA Summary
- Recommendations

# Executive Summary

## Client:

ABC Bank wants to sell its term deposit product to customers. Before launching the product, they want to predict which clients are most likely to subscribe to a term deposit. In this way, the bank wants to save time and money by running the marketing campaign more effectively and successfully.

## Problem description:

- Create an ML model for the bank that will shortlist customers who are more likely to subscribe to the term deposit product.
- This will allow the marketing team to target those customers more efficiently.



# EDA

The analysis has been divided into three parts:

- Data Understanding
- Univariate and Bivariate Analysis
- Model Recommendations

# Data Understanding

- Datasets:
  - bank\_additional\_full : 21 features and 41119 observations
  - bank\_additional: 21 features and 4119 observations
  - bank\_full: 18 features and 45211 observations
  - bank: 18 features and 4521 observations
- Assumptions:
  - Timeline of Observations: May 2008 to November 2010
  - There are several missing values in some categorical attributes, all coded with the “Unknown” label.
  - Duplicate rows are dropped from dataset

# Data Exploration

Proportion of people who agreed to a term deposit (positive class) compared to the people who did not is 11.3%. That tells us that data seems unbalanced

Clients  
Subscribed to  
Term Deposit

Y

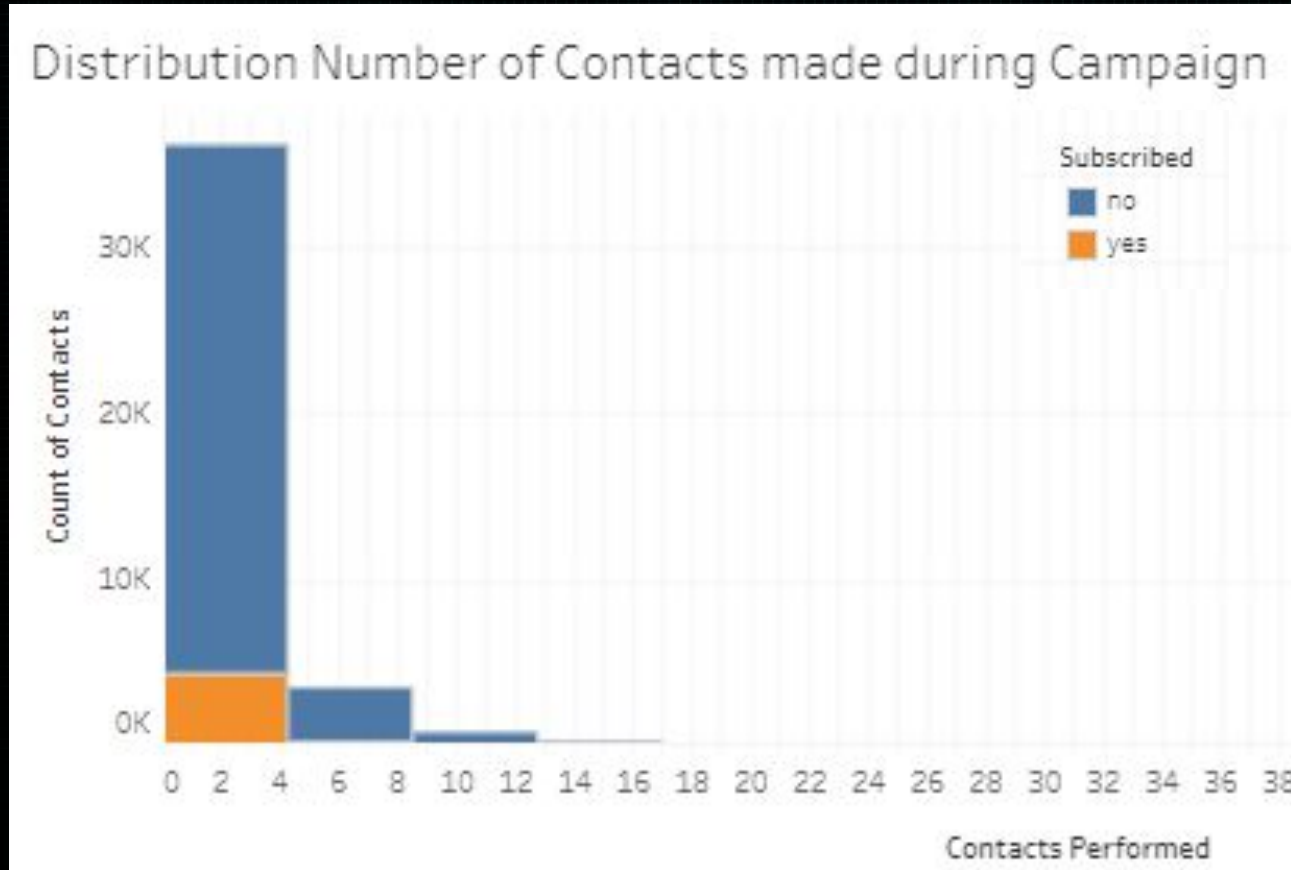
no

36,548

yes

4,540

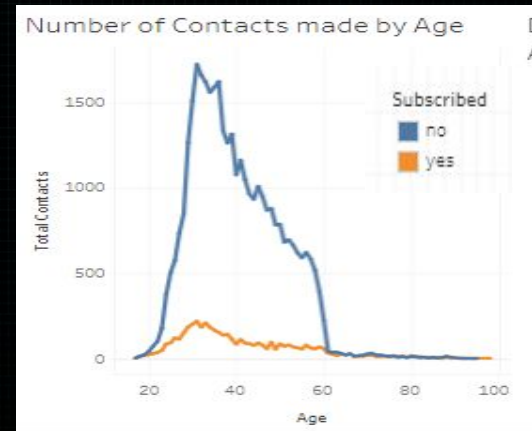
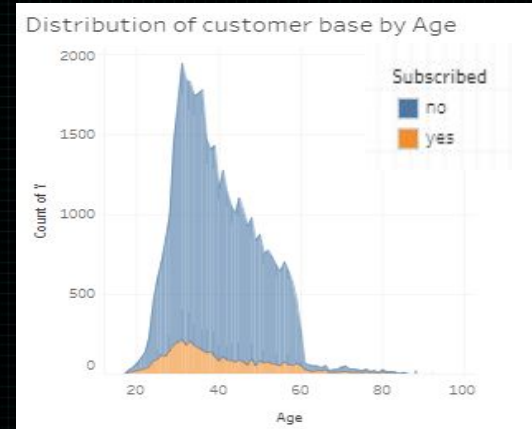
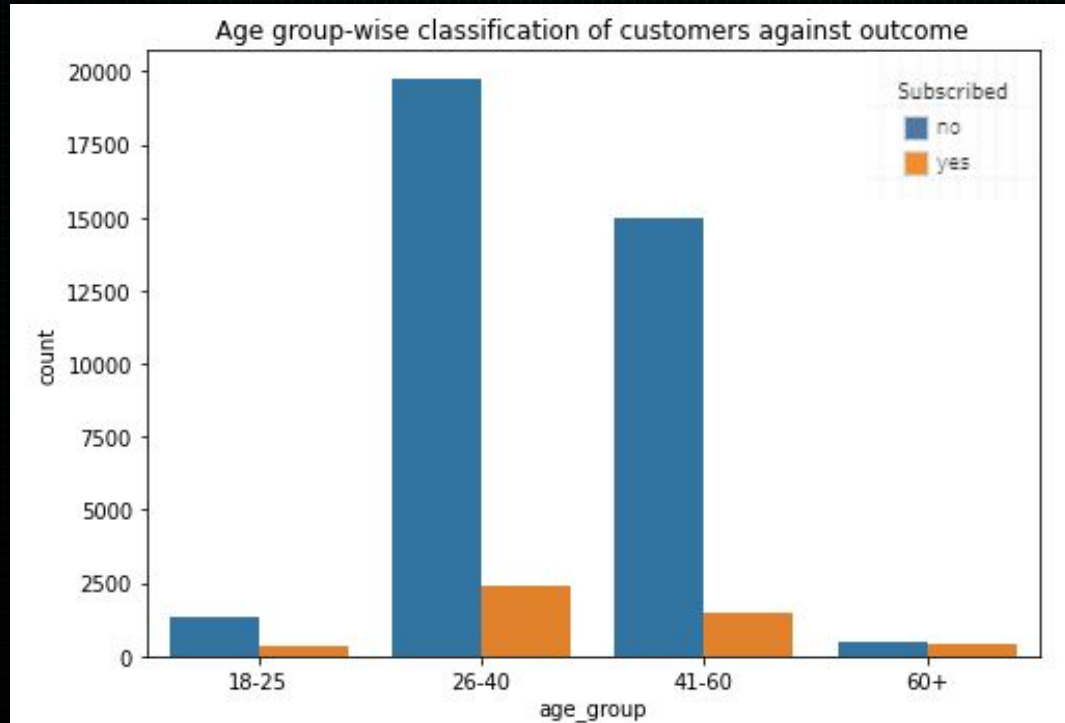
# Data Exploration



As number of contacts performed increased less customers opted for term deposit



# Customer Base by Age

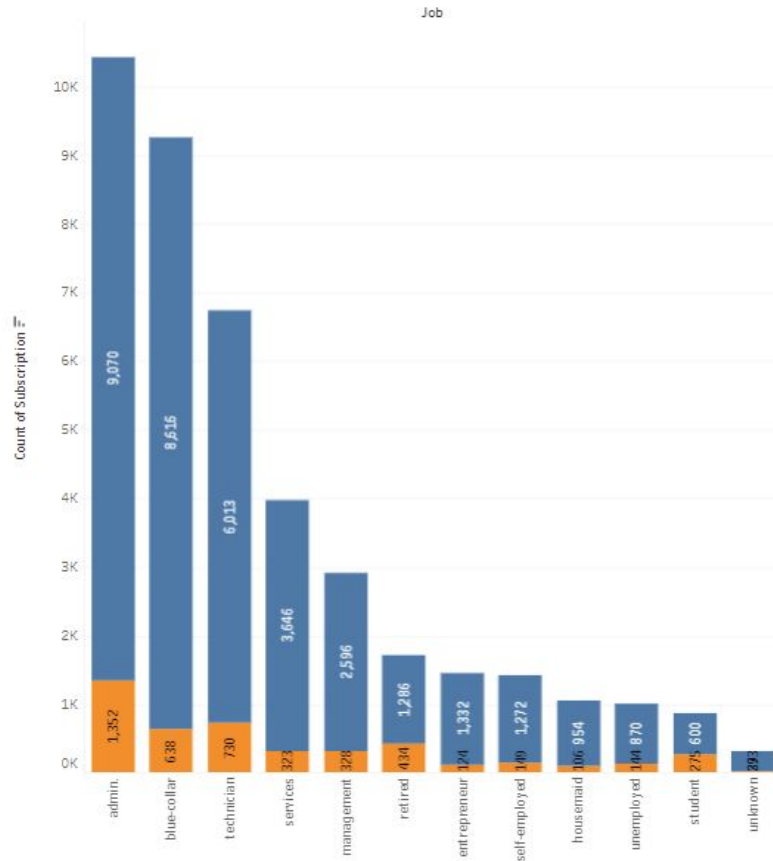


The 26-60 Age group occupy a portion of 97% of the customer base and more contacts are made to them over 50000

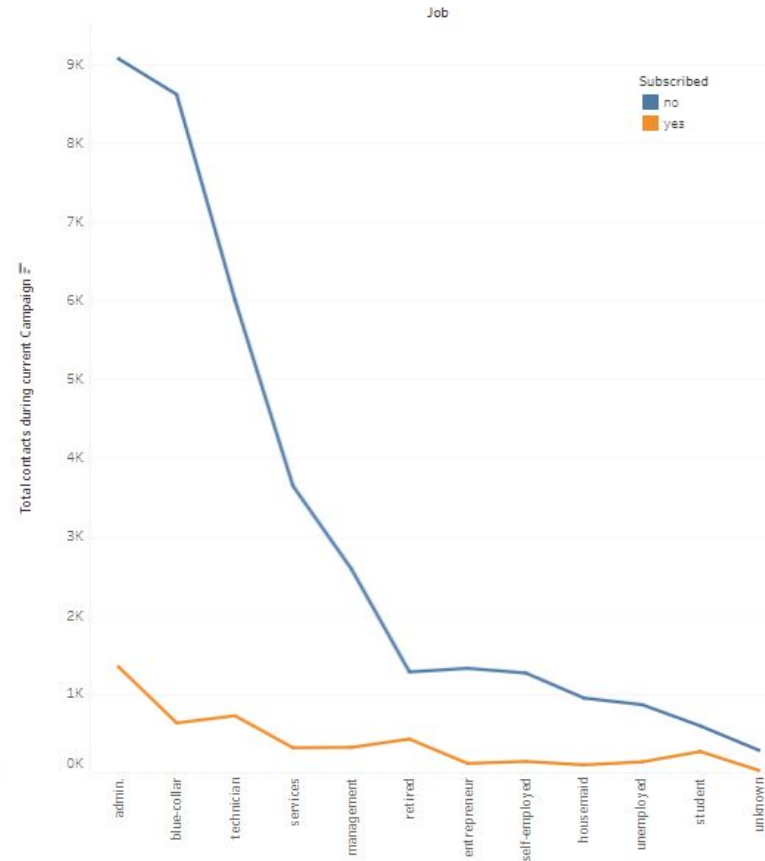


# Customer Base by Job Analysis

Distribution of customer base by Job Categories



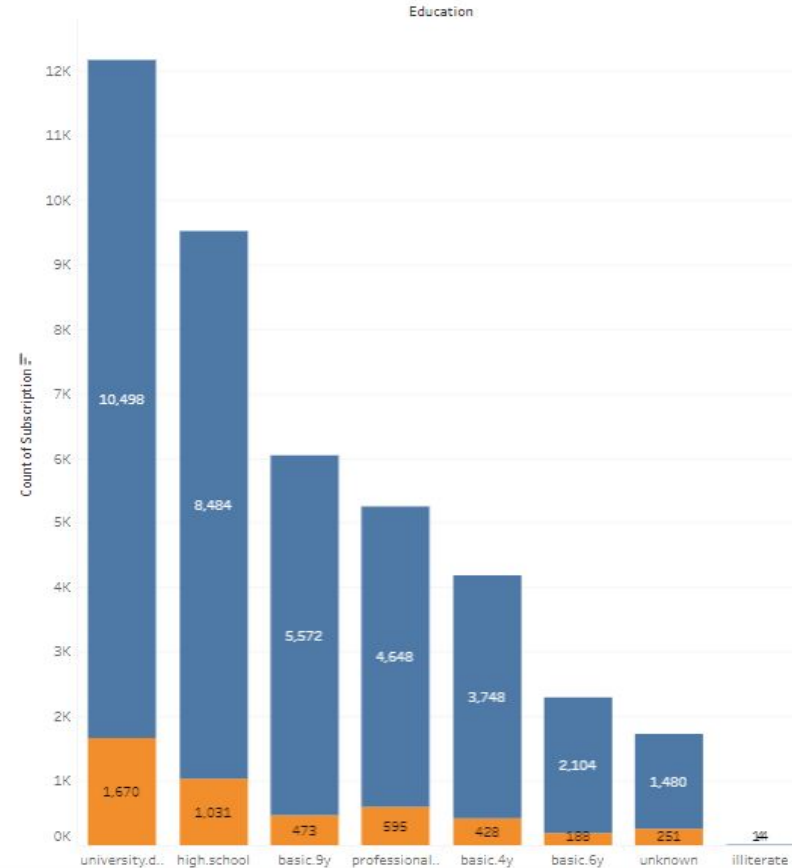
Number of Contacts made by Job



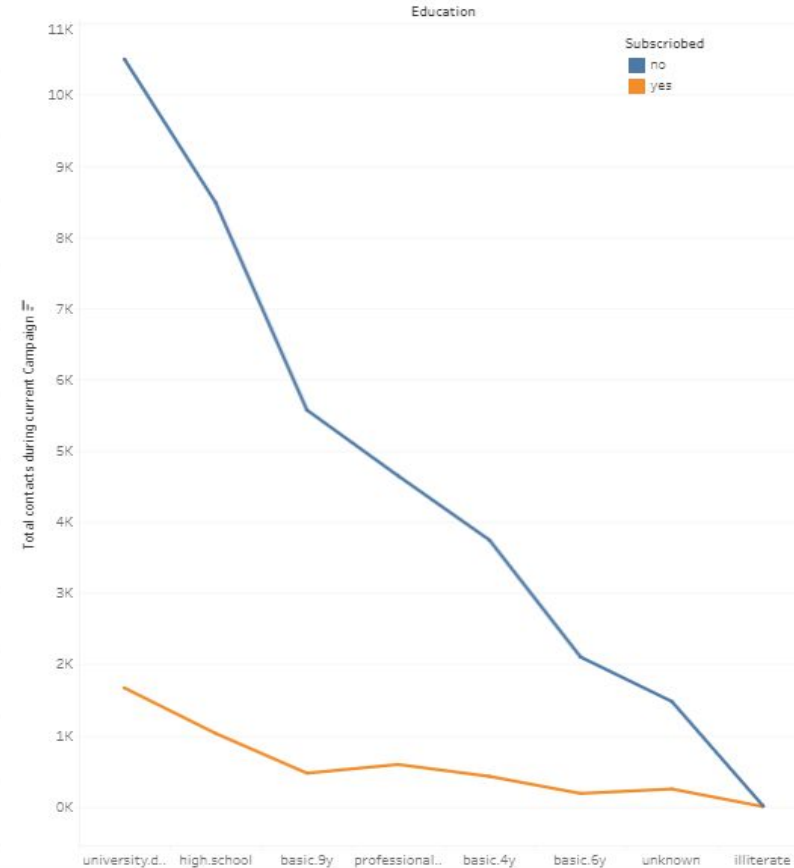
- In the customer base more than 63% are in admin, blue-collar, technician.
- Whereas services and management are around 41%.
- But only 13.5% of the customers are accepted to Term Deposit plan.

# Customer Base by Education Analysis

Distribution of customer base by Education

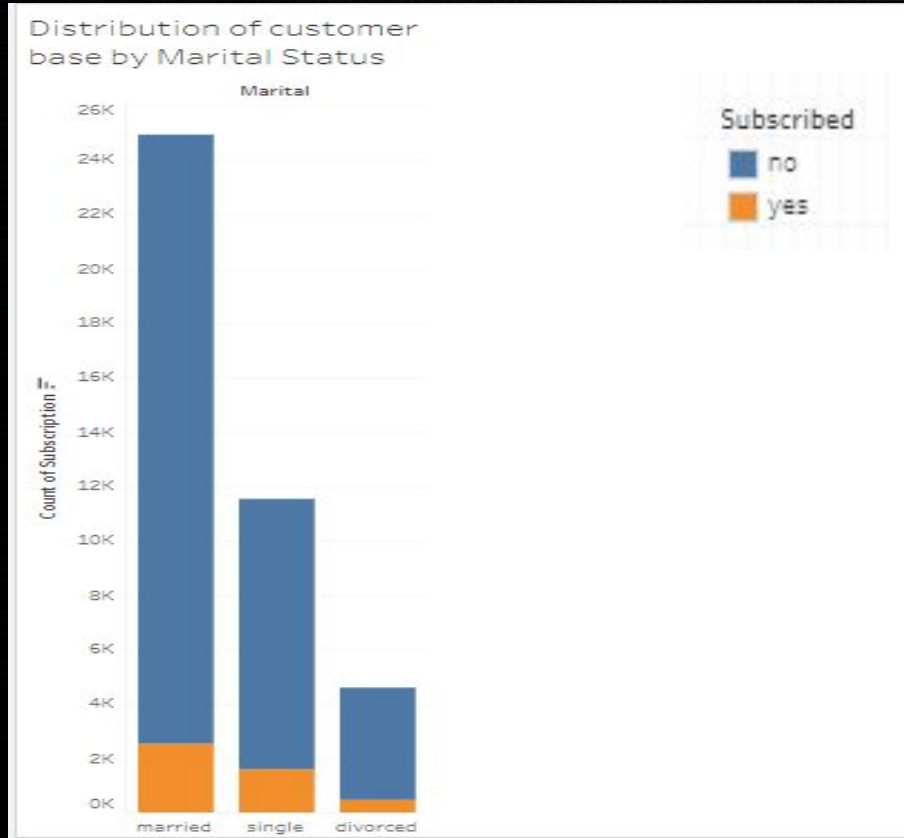


Number of Contacts made by Education level



- Over 10000 customers with an education level of university degree show rejection to term deposit plan
- Also there are close to 1700 customers with unknown education level.

# Customer base analysis by Marital Status

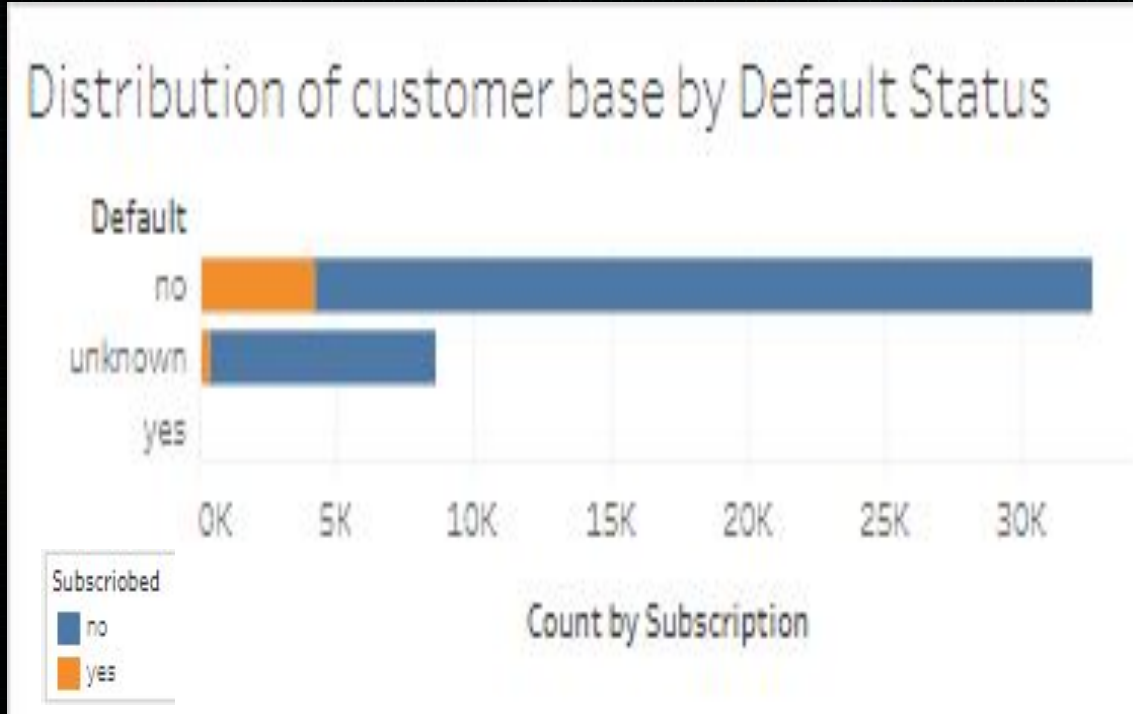


Marital	Y	
	no	yes
married	22,396	2,532
single	9,948	1,620
divorced	4,136	476

Around 50% of customer base comes under married status. But still more number of married customers are not opting to subscribe term deposit product.

Interestingly when we compare no of customers subscribing to product vs Not subscribing product, more single customers prefer to buy term deposit.

# Customer base analysis by Default Status



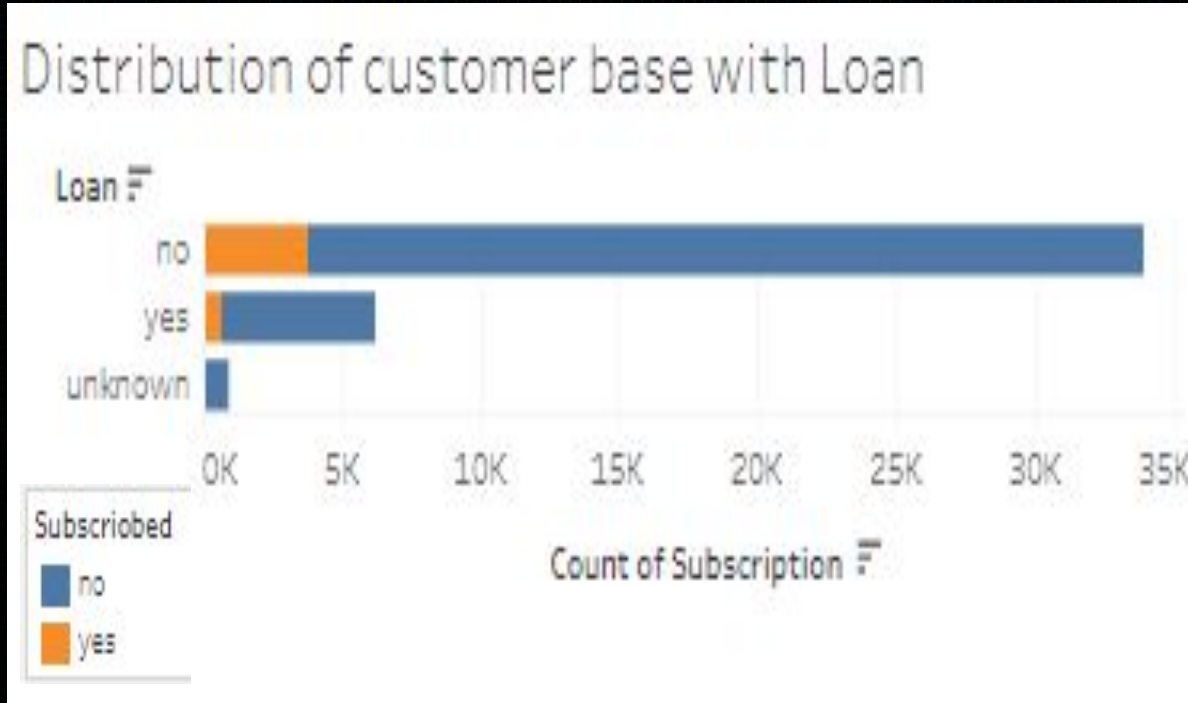
Default	Y	
	no	yes
no	28,391	4,197
unknown	8,154	443
yes	3	

Approximately 79% of customers have default credit but only 12.9% subscribed to Term Deposit

Almost 8500 customers have not disclosed their default status.



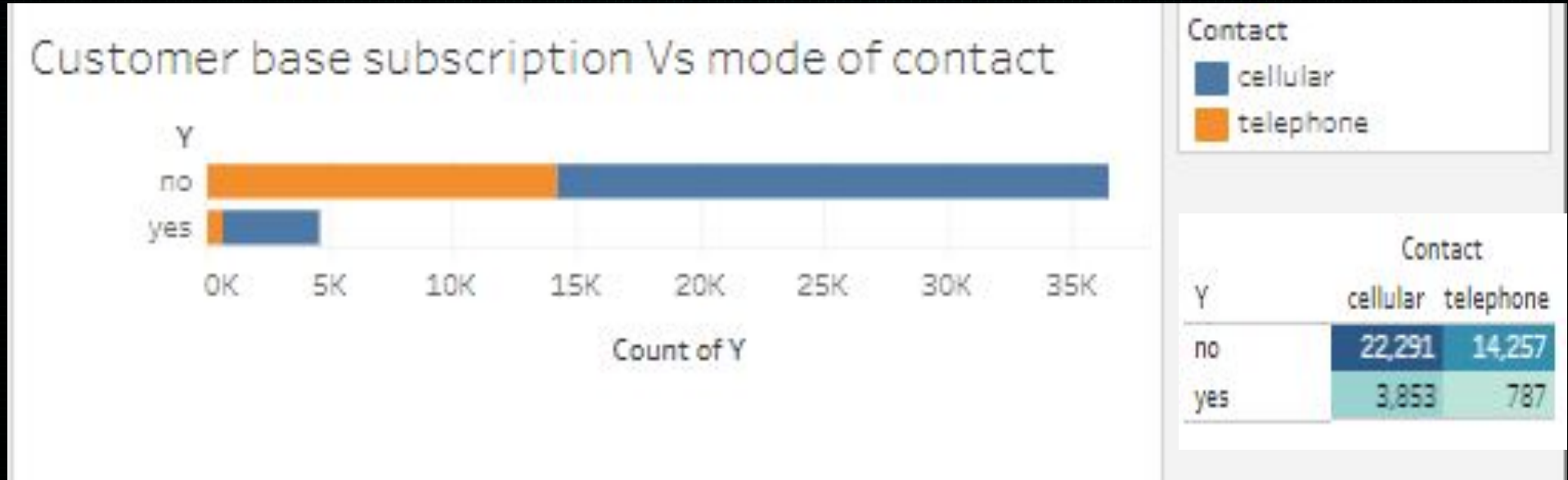
# Customer base with loan analysis



Loan	Y	
	no	yes
no	30,100	3,850
unknown	883	107
yes	5,565	683

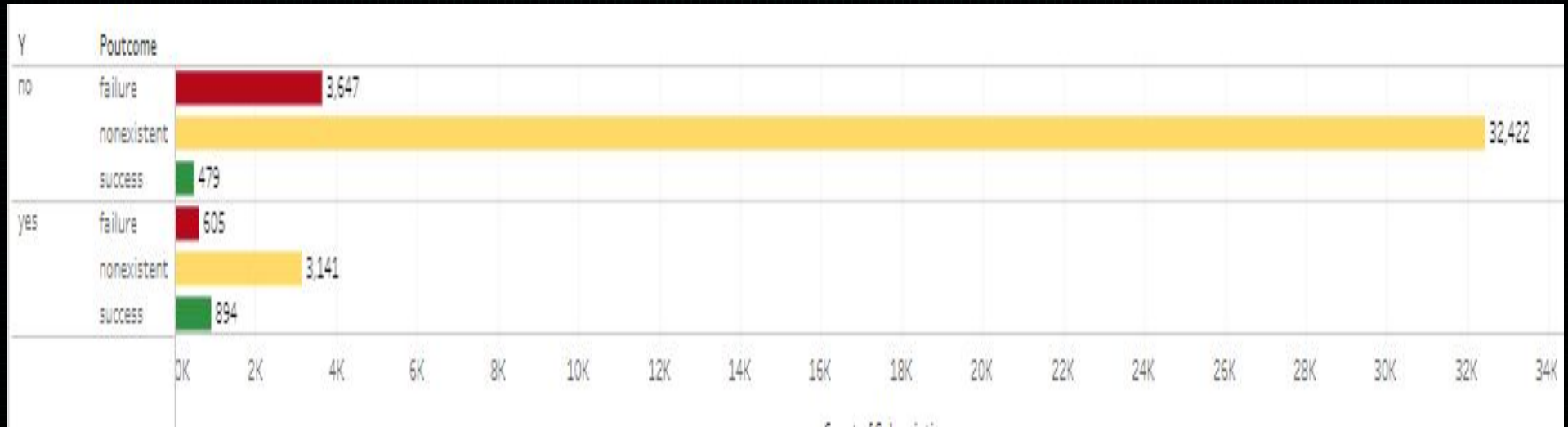
Approximately 83% of customers who opted for term deposit do not have person loan. But when we compare that to no of customers with no loan is just 3%

# Customer base analysis by Mode of Contact



Out of 36548 customers not opting for the term deposit plan, approximately 22000 are contacted through cellular communication.

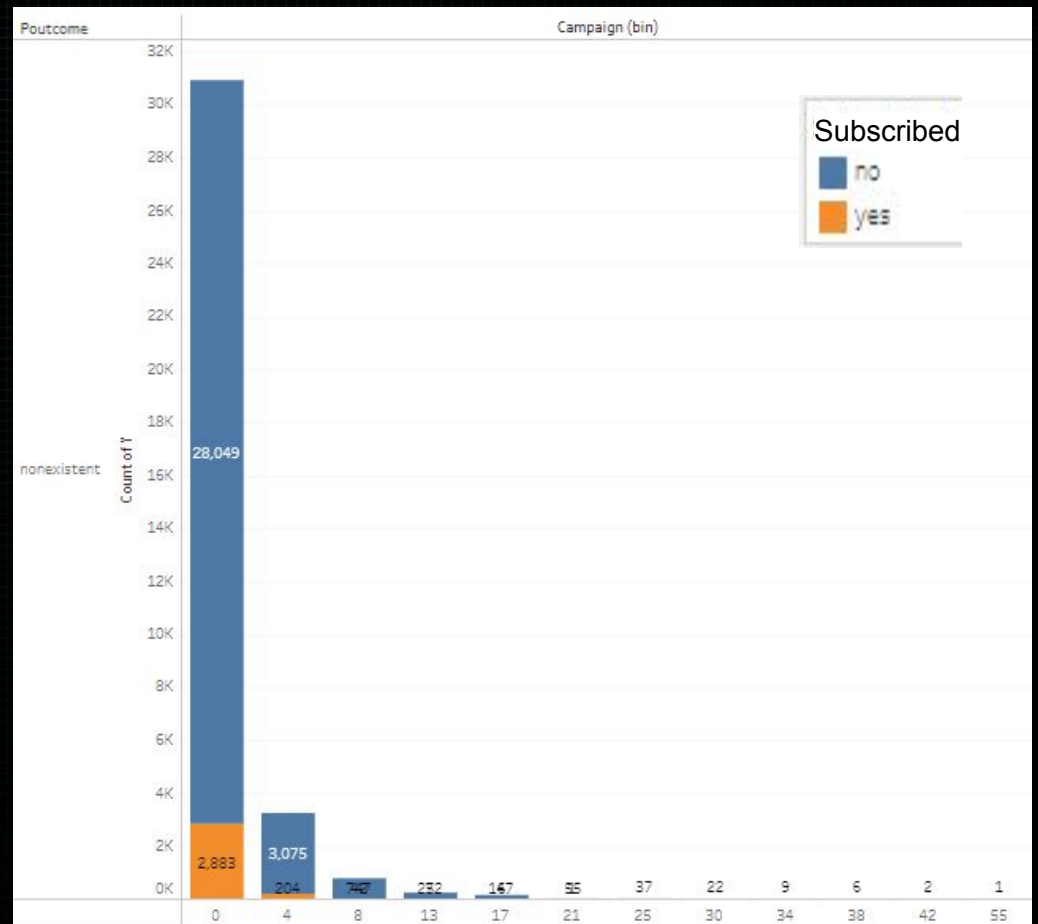
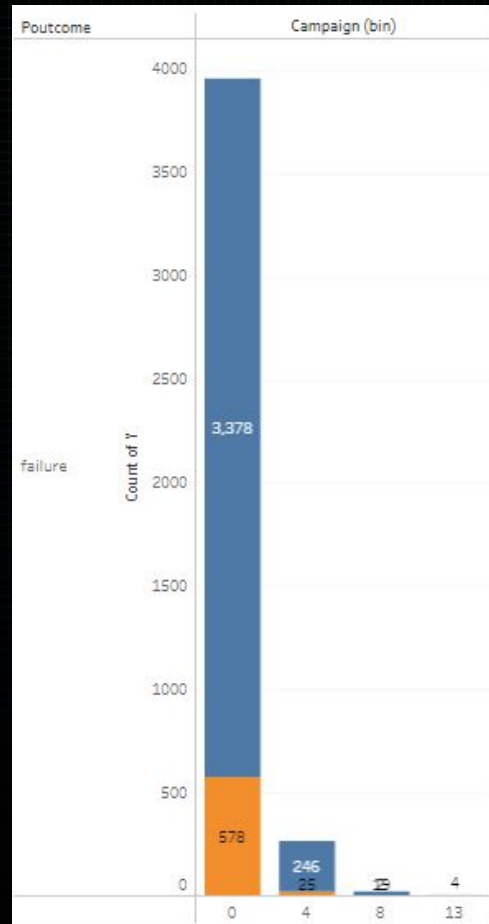
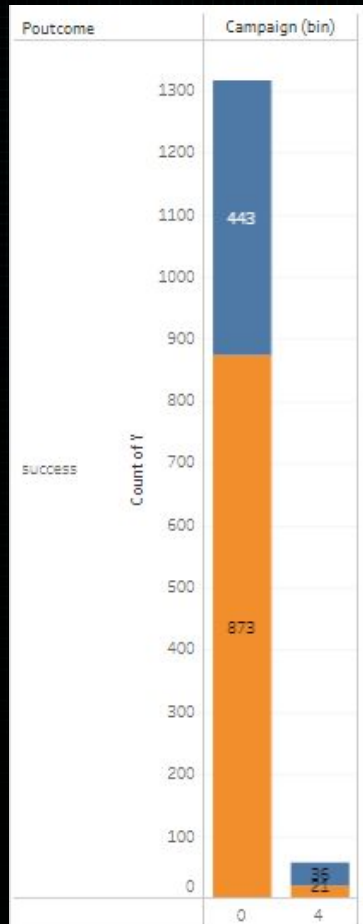
# Customer base analysis by Previous Campaign Outcome



Out of current customer base, 86.4% of customers previous campaign outcome is nonexistent.

894 customers who subscribed during previous campaign also opted for term deposit this season, and approximately 3100 nonexistent status opted during this campaign

# Customer contact analysis by Previous Campaign Outcome





# Model Recommendation

## Random Forest Model:

- Ensemble learning method that combines multiple decision trees to make predictions
- Good for handling non-linear and complex relationships in the data
- Provides feature importance measures to help identify important variables
- Robust to missing data and outliers

## Extreme Gradient Boosting (XGBoost) Model:

- Gradient boosting algorithm that combines weak models to make predictions
- Optimized for speed and scalability
- Handles missing values and imbalanced datasets well
- Provides feature importance measures and regularization to prevent overfitting

# Model Recommendation

## Naive Bayes:

- Probabilistic algorithm that uses Bayes' theorem to make predictions
- Fast and efficient, making it well-suited for large datasets
- Good for text classification and other high-dimensional data
- Handles irrelevant features and missing values well

# Thank You

Group Name: LimitLess Team

Name: Chitra S Chaudhari

Email: [chitraksonawane@gmail.com](mailto:chitraksonawane@gmail.com)

Country: USA

Specialization: Data Science

Name: Nuri Öztürk

Email: [ozturknuri8@gmail.com](mailto:ozturknuri8@gmail.com)

Country: Turkey

Specialization: Data Science

Batch Code: LISUM 17

Github Repo link: <https://github.com/ChitraChaudhari/Bank-Marketing-campaign>