

NEXUS BANK CASE STUDY

CUSTOMER SEGMENTATION AND
DEPOSIT DETECTION



INTRODUCTION

Nexus Bank is a financial institution dedicated to delivering unparalleled banking services to our clients.

Our mission is to establish enduring relationships with our customers by providing tailored financial solutions that align with their individual needs and goals.

At Nexus Bank, we believe that every individual deserves access to world-class financial products and services, regardless of their age, profession, or income level. That's why we offer a wide spectrum of banking solutions to accommodate your lifestyle, including term deposits, personal loans, and mortgage financing. Our team of seasoned banking professionals is committed to providing you with the utmost level of service, transparency, and honesty.



THE TASK

What is expected?

Objectives

Nexus bank has conducted campaigns with the goal of acquiring deposits. The directors were unsatisfied with their current situation and need to optimize the operations at Nexus bank.

- Leverage on data to gain insights into the bank and improve efficiency and identify patterns , trends in customer behaviour to decipher if customer demographics such as age, educational level etc. influences customer's attitude toward defaulting.
- Predict future customer behaviour and know the likelihood of deposits from customers. Understand how effective the bank campaigns are and thus develop marketing campaigns to reach specific customer segments.
- Analyse customer behaviours, loan trends, and marketing campaign effectiveness, Nexus wants to optimize its operations, mitigate risks/ loan defaults, and improve customer deposits.

DATA PRE - PROCESSING



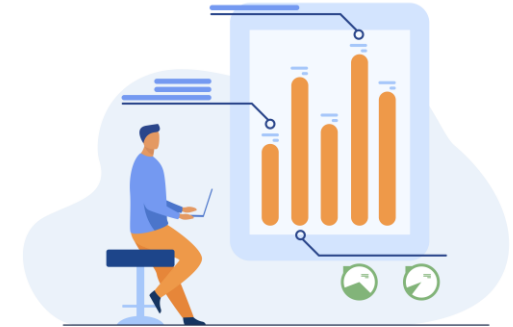
FUTURE ENGINEERING

- Drop unwanted features
- Handle missing values
- Handle categorical features
- Remove outliers



DATA ENCODING

- Use label encoder to transform categorical data to numerical variables



DATA SCALING

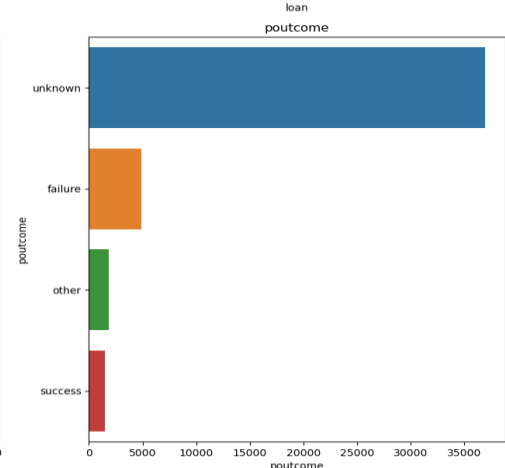
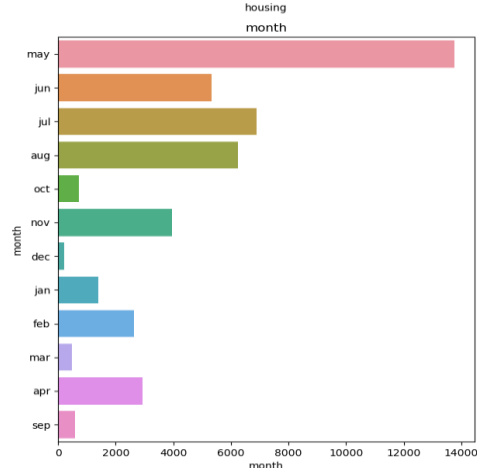
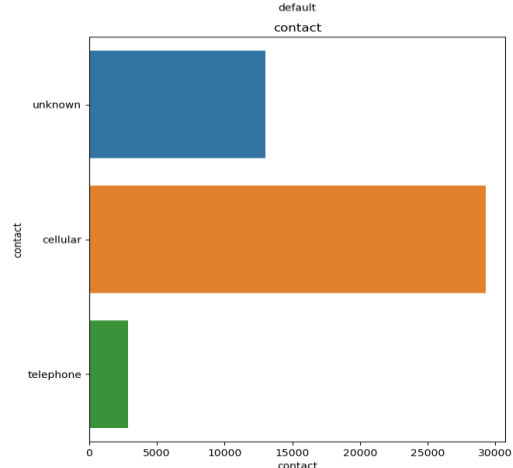
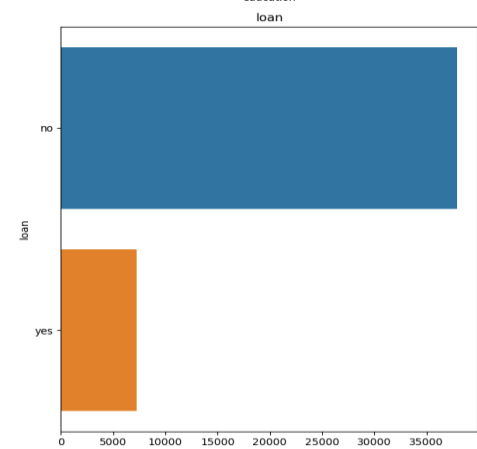
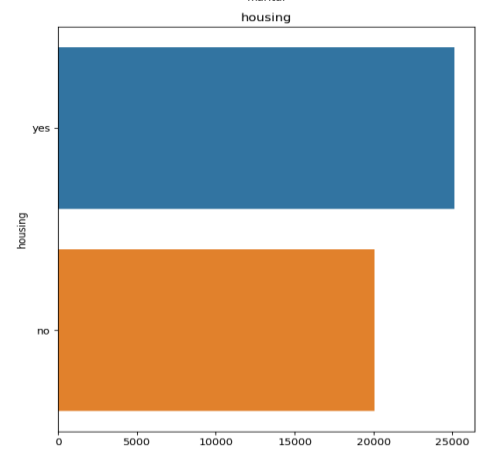
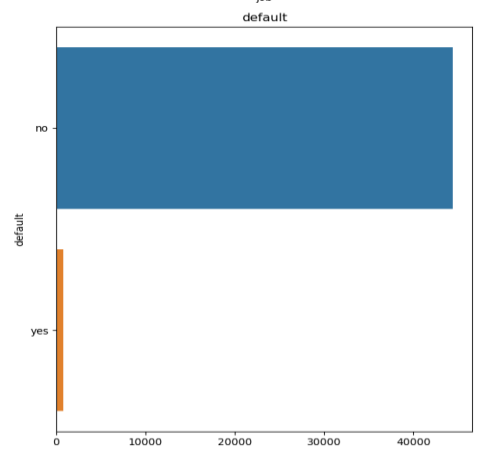
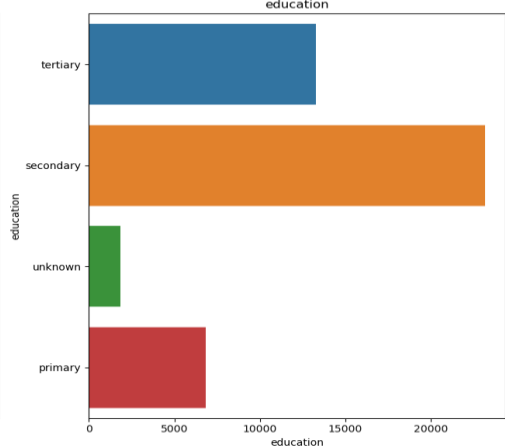
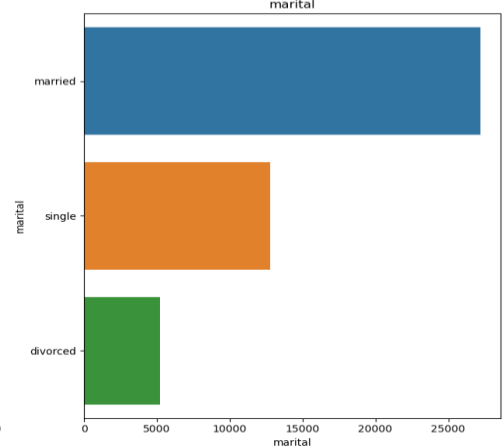
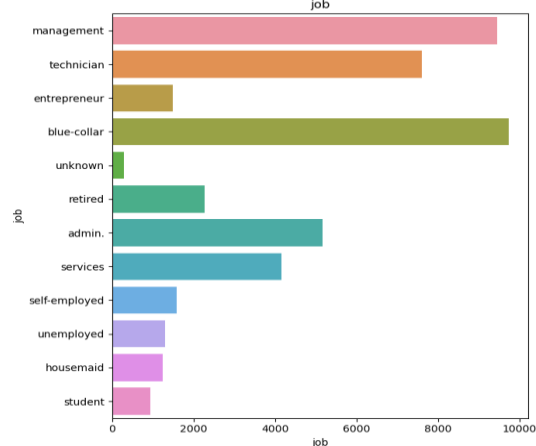
- Use the Min - Max scaler to scale the dataset – compress to balance and duration in the dataset

EXPLORATORY DATA ANALYSIS:

Univariate Analysis

Observation:

- Customers with job type as management and blue-collar records are high in the given dataset and house are less
- Married customers are high and divorced are less
- The dataset shows customers with secondary educational background are higher
- Default from the dataset does not play any significant role as it shows number of 'no' in high ratio compared to the number of 'yes'
- Monthly distribution in the data shows that May contributed high number while Dec shows lesser contribution

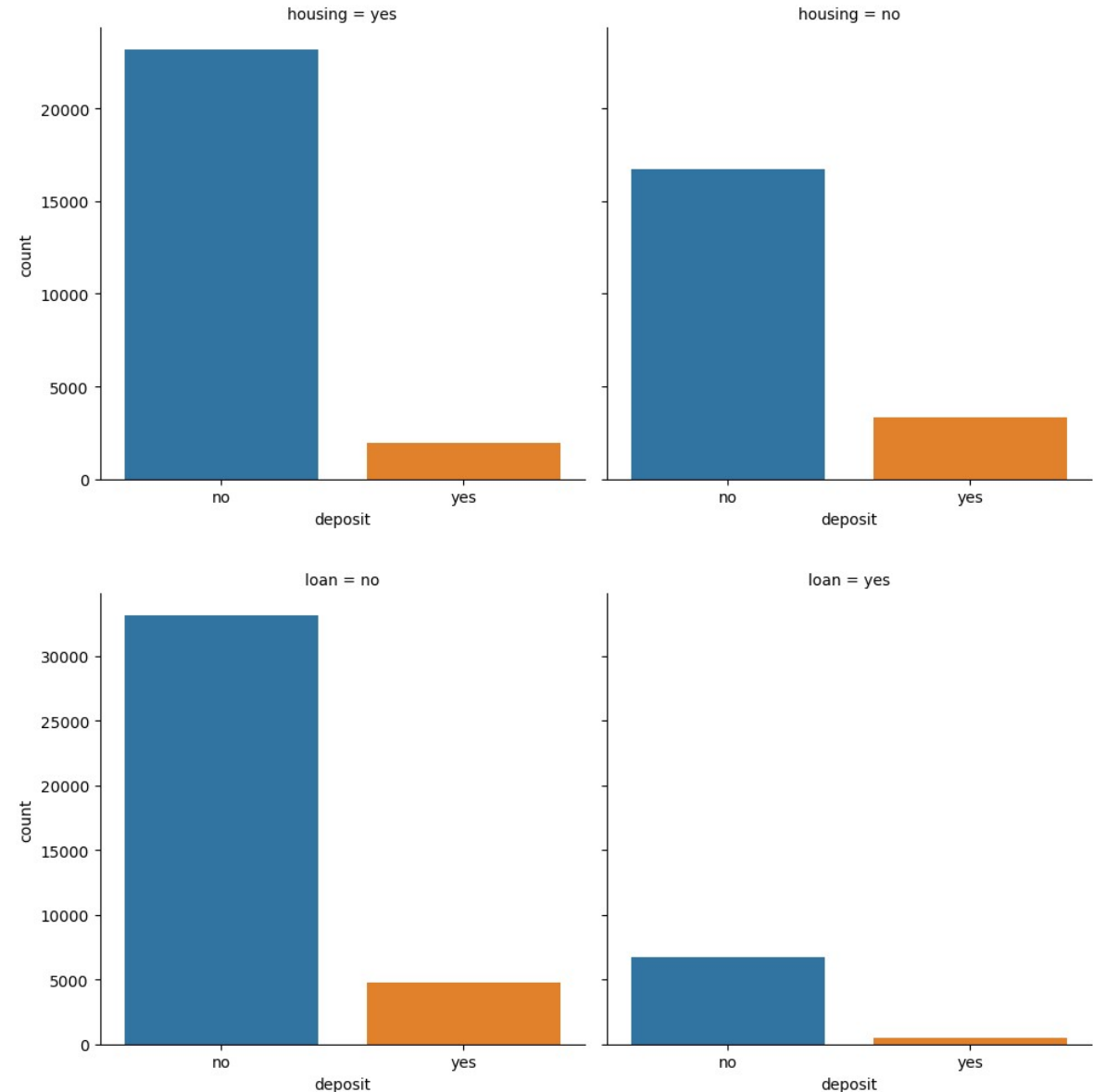


EXPLORATORY DATA ANALYSIS

Univariate Analysis

Observation:

- From the dataset retired customers are interest in bank deposit
- The month of May records high number with low ratio of interest in deposit
- Customers with a housing loan to be uninterested in making a bank deposit
- Campaign outcome with poutcome = success have high tendency of making deposit
- March , September, October, and December customers showed much interest on deposit

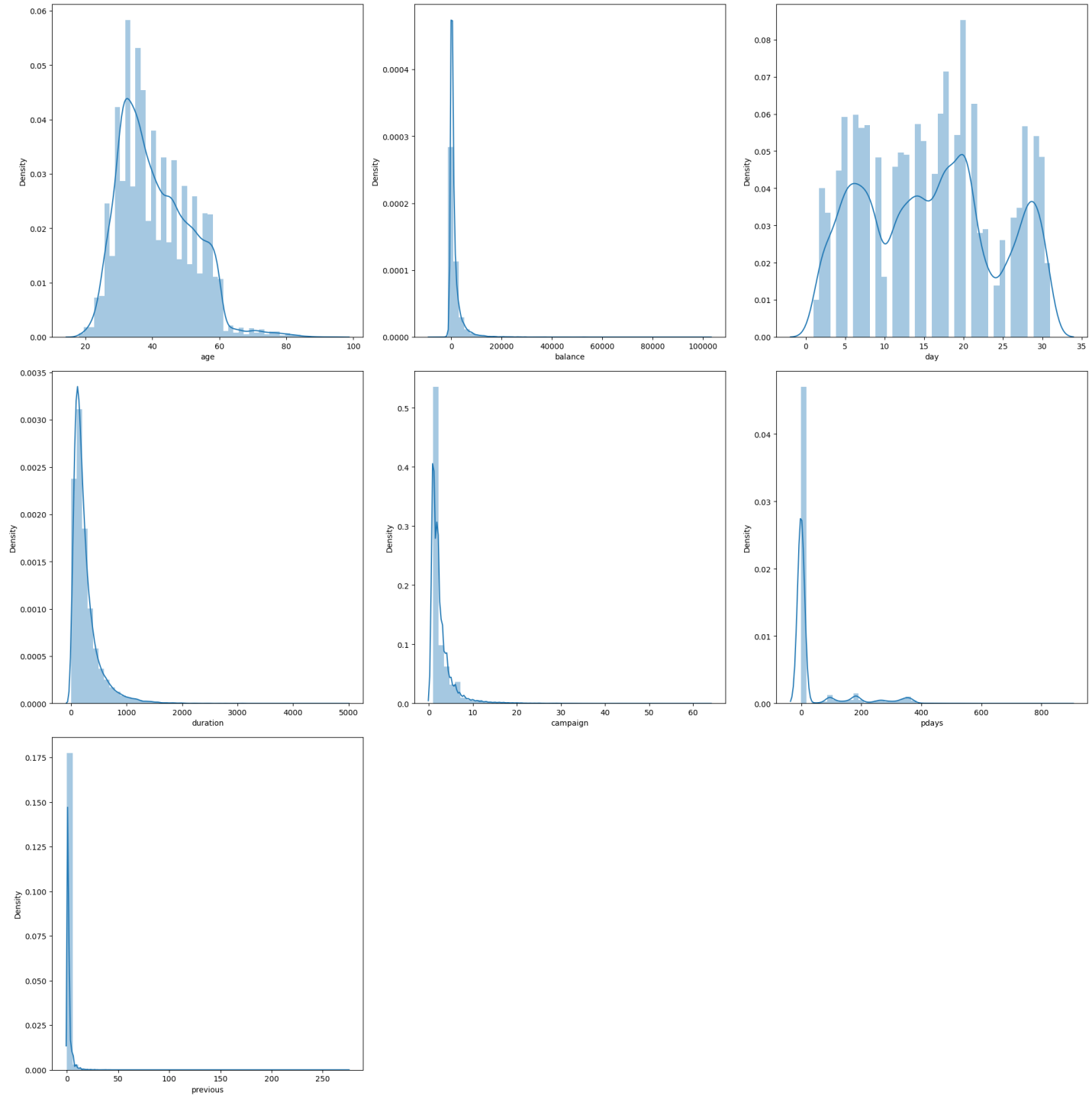


EXPLORATORY DATA ANALYSIS

Univariate distribution of continuous features

Observation:

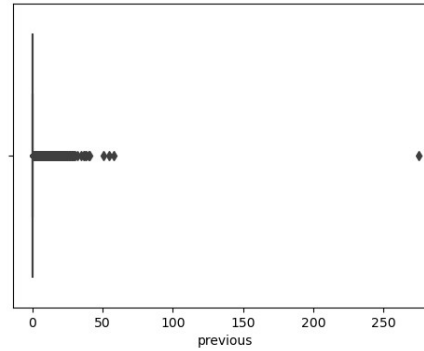
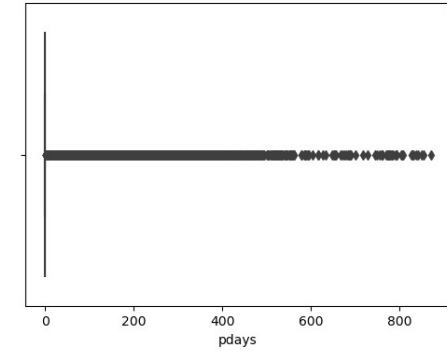
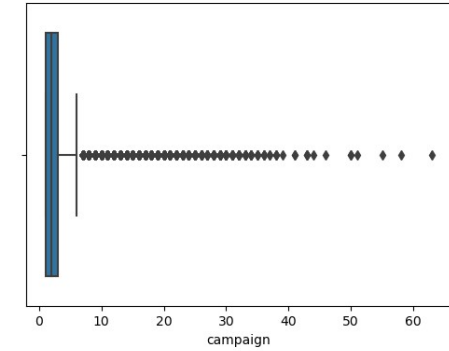
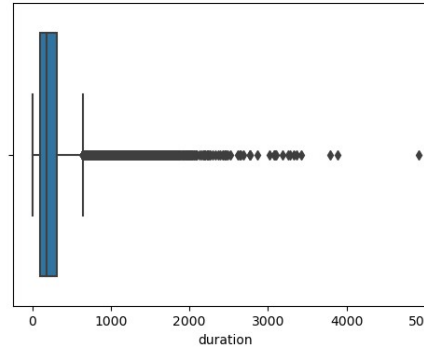
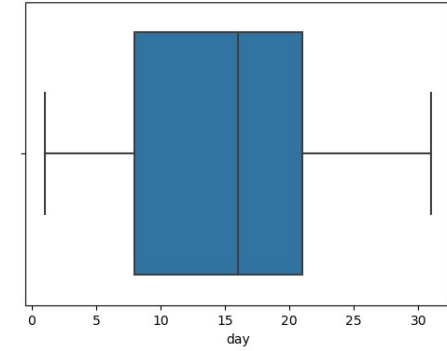
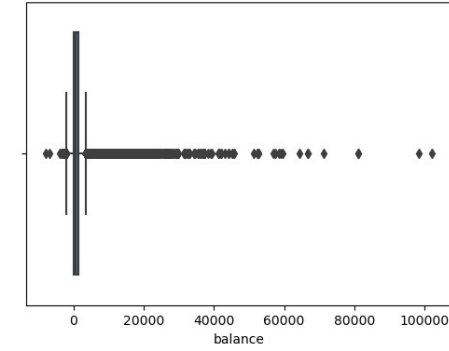
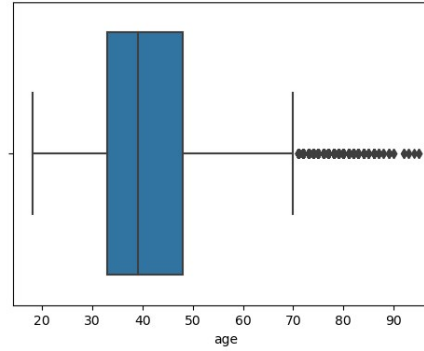
- The dataset indicate that balance, campaign, pdays, and previous are highly skewed towards the left hand and show number of outliers
- Age and days are normally distributed



EXPLORATORY DATA ANALYSIS

Observation:

- Previous, campaign, age, duration, pdays, and balance shows an outlier in the dataset



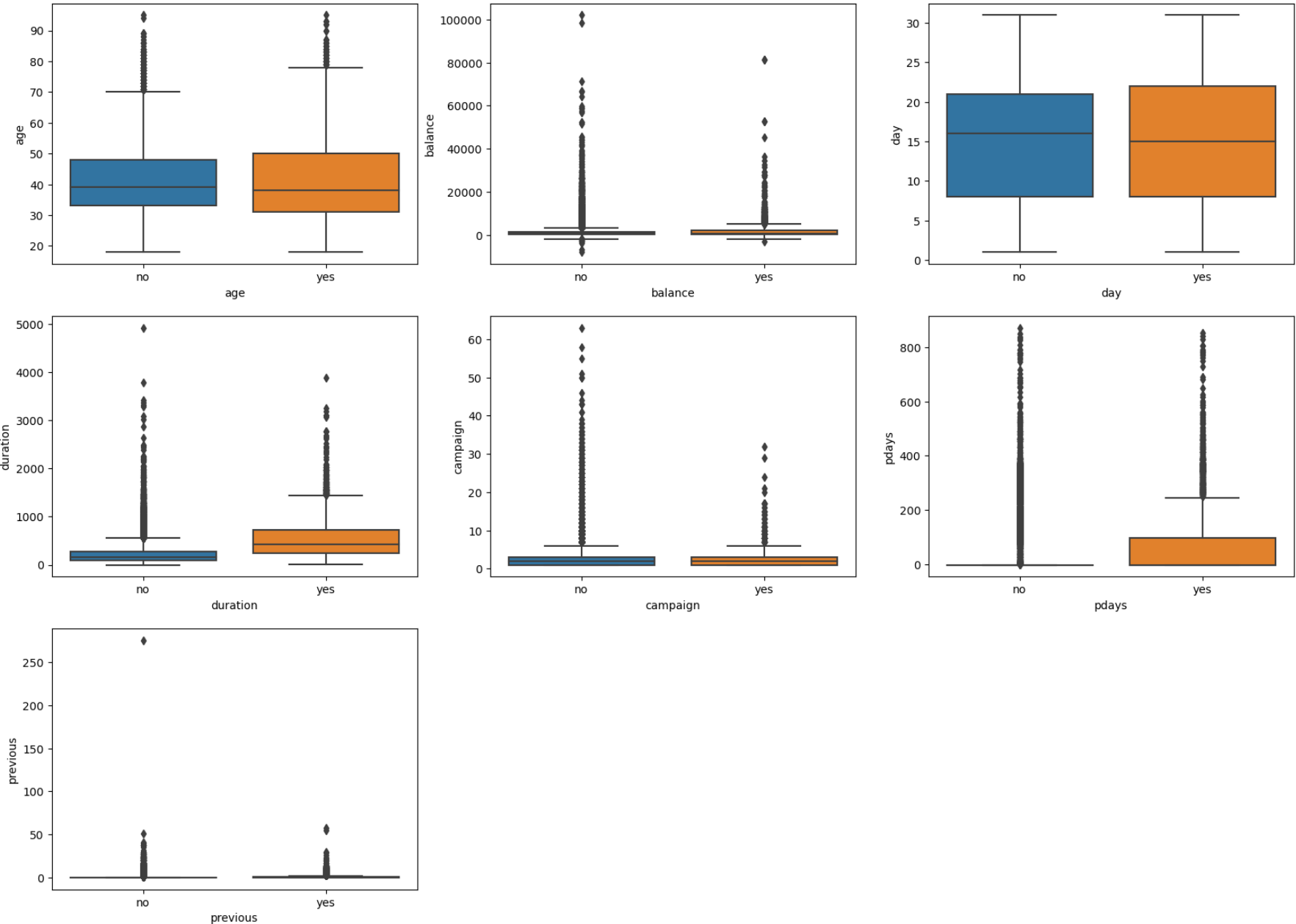
EXPLORATORY DATA ANALYSIS

Multivariate Analysis:

Relationship between continuous numerical features and label

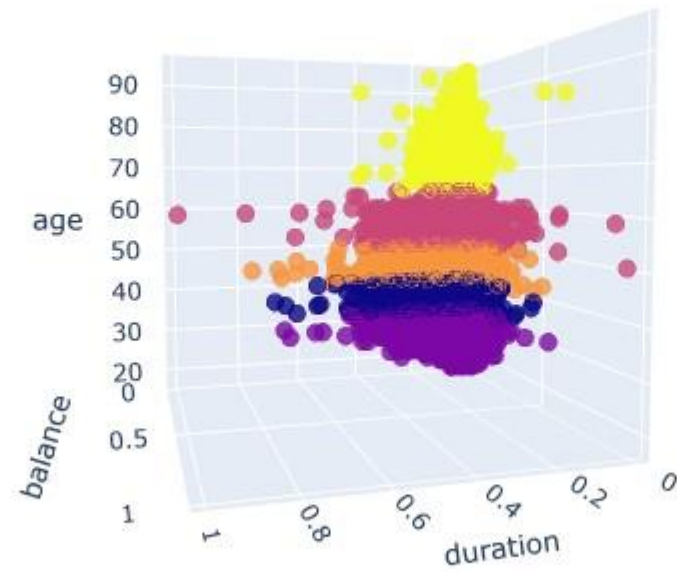
Observation:

- Long days engagement and communication shows that customers indicate interest on deposit.



3D Customer Clustering

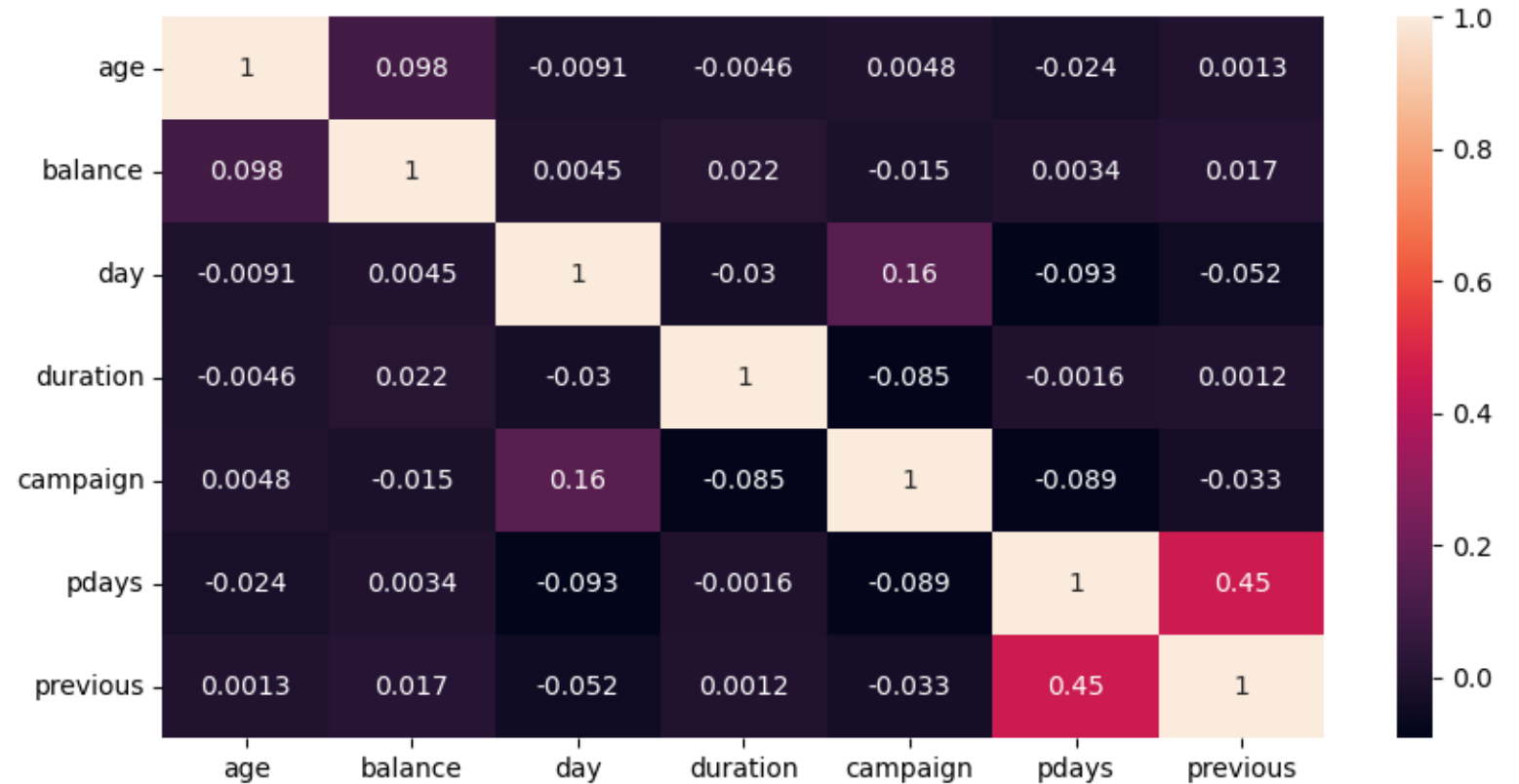
3 variable
visualization for
Age, balance and
duration to show
clusters.



CORRELATION ANALYSIS

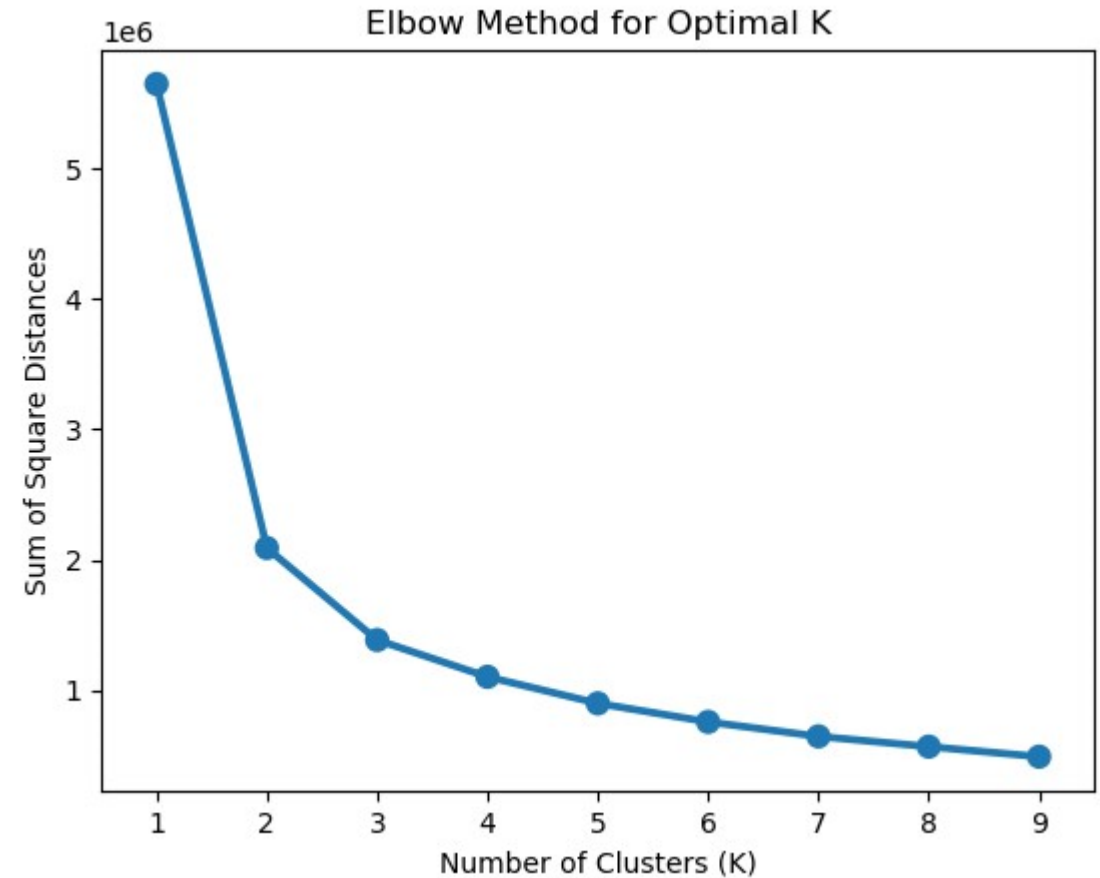
Observation:

- There is no significant correlation observed among the variables in the dataset.



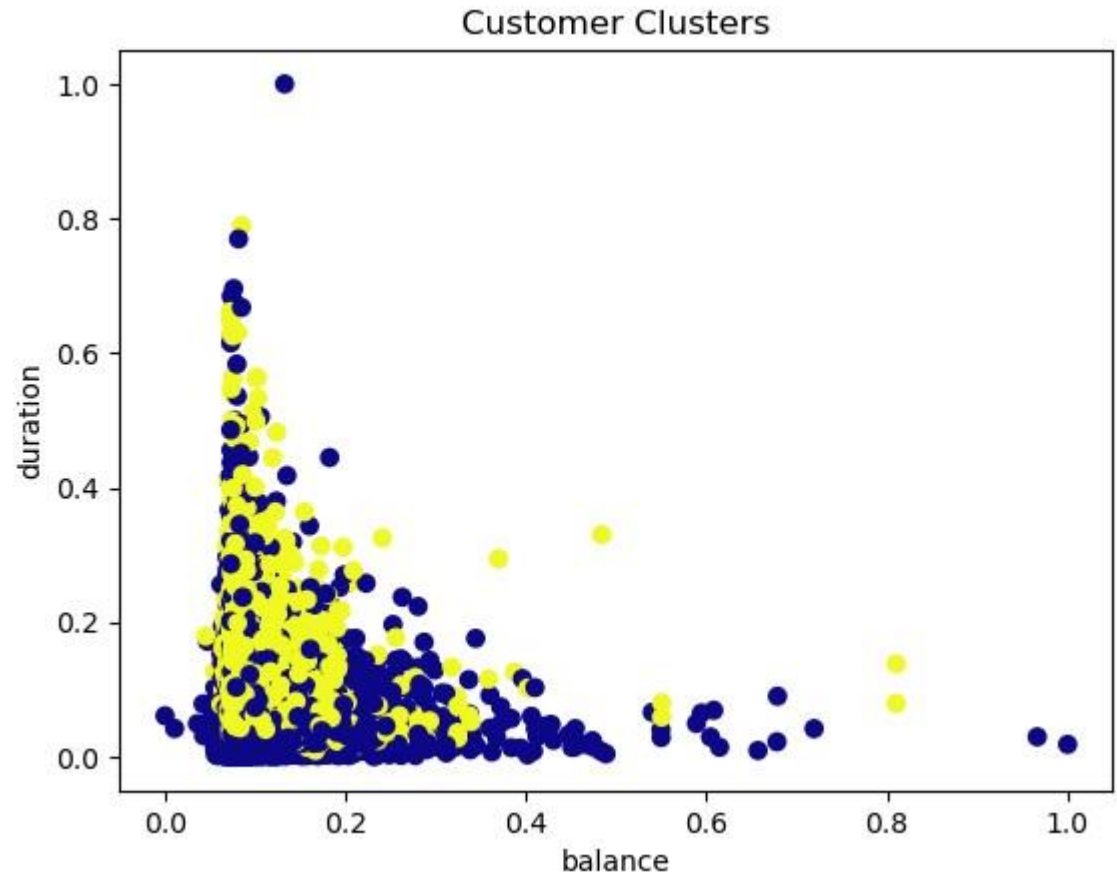
MODEL BUILDING - ELBOW METHOD

- Determine the optimal number of clusters for customer segmentation.
- The inertia starts decreasing from $k=4$



K-Means Clustering

- Fit the data model
- Determine the centroids
- Visualize the clusters

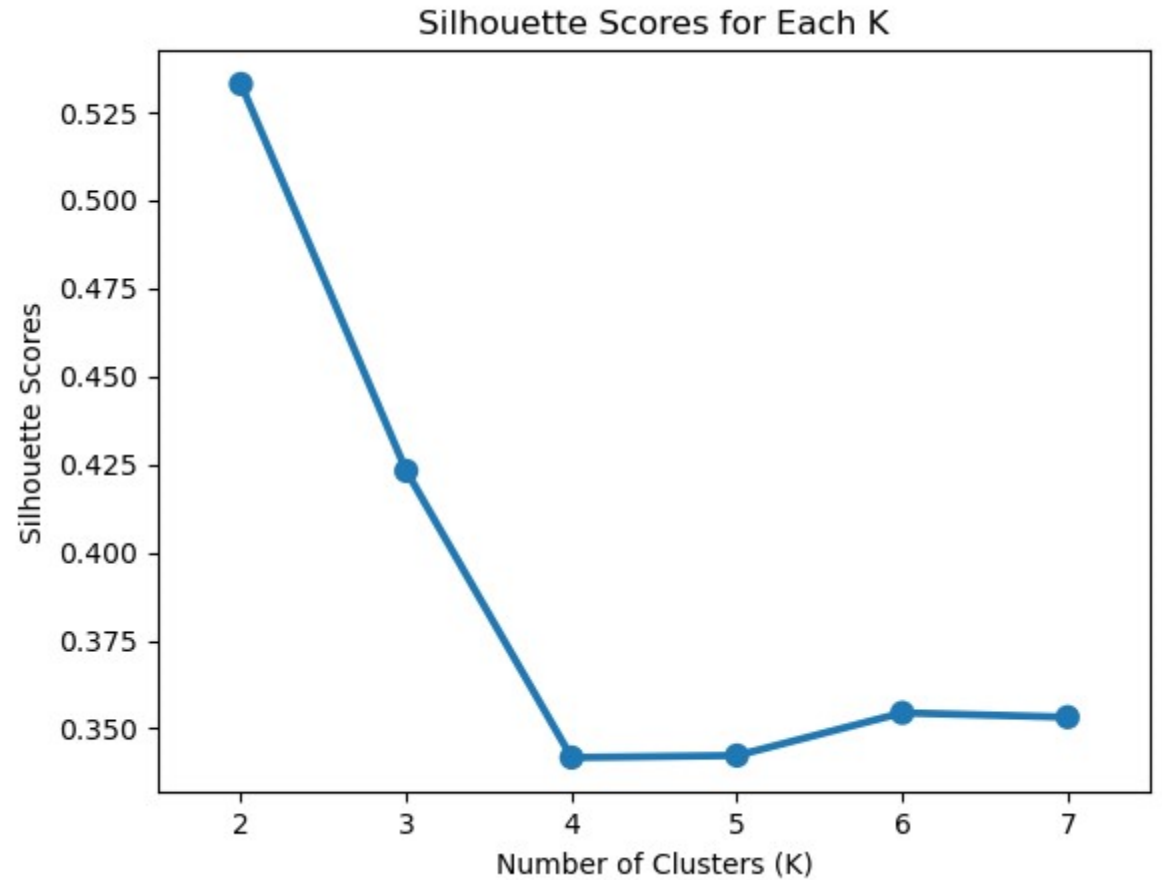


SILHOUETTE EVALUATION

Silhouette scores:

0.34110636768885627

- The silhouette score analysis, shows that the best and optimal cluster for the customer segmentation is $k=4$



CONCLUSION

Key Insights for Customer Segmentation and Deposit Marketing Strategy

a. Retired Customers

- Retired customers have shown a strong interest in bank deposits.

b. Month of May: Low Conversion Ratio

- May records a high number of interactions but a low ratio of interest in deposits.
- Investigate reasons behind this trend and consider campaign adjustments for May.

c. Impact of Housing Loans

- Customers with housing loans are generally uninterested in making bank deposits.
- Develop targeted strategies to address concerns and promote deposit offerings to this customer segment.

d. Successful Campaign Outcomes

- Campaign outcomes marked as "success" have a higher tendency to result in deposits.
- Prioritize and replicate successful campaign strategies to increase overall deposit conversion rates.

e. Seasonal Patterns of Interest

- Customers exhibit greater interest in deposits during March, September, October, and December.
- Leverage these months for targeted marketing campaigns and promotional offers.

CONT.....

f. Understanding Customer Data

- The dataset includes 7 continuous numerical features.
- Certain features, such as balance, campaign, pdays, and previous, are skewed and contain outliers.

g. Normal Distribution and Outliers

- Age and days follow a normal distribution, while other features have outliers.
- Consider outliers in age, duration, pdays, campaign, previous, and balance for further analysis or data preprocessing.

h. Dataset Balance

- The dataset exhibits balance between positive and negative responses for deposits.
- Ensure equal representation of both deposit and non-deposit customers in future analysis and campaigns.

i. Optimal Customer Segmentation

- Silhouette score analysis show the optimal number of clusters for customer segmentation is $k=4$ with silhouette score of **0.3444014756318102**.
- Utilize these segments to tailor marketing messages and strategies to specific customer groups.

CONT.....

j. No Correlation Between Variables

- There is no significant correlation observed among the variables in the dataset.
- Rely on other factors and insights to drive deposit marketing strategies rather than relying on inter-variable relationships.

These key insights provide actionable information to improve deposit conversion rates and develop targeted marketing strategies. By understanding customer behaviour, seasonality, and campaign outcomes, the bank can enhance its marketing efforts and drive customer engagement and deposit growth.



Kingsley Jolly Jackson

Data Science Consultant 10Alytics

