

Srikandi

Final Project

E-COMMERCE CHURN PREDICTION

To Minimize Company Cost



Stages

00
01
02
03
04

- Background
- EDA & Insight
- Pre-Processing
- Modeling
- Recommendation



Srikandi

BACKGROUND

BACKGROUND

CONTEXT

sebuah e-commerce saat ini mengalami customer churn sebesar 16.8%. Hal tersebut menyebabkan e-commerce kehilangan banyak customer. Selain itu, biaya untuk mengakuisisi customer baru jauh lebih besar daripada biaya untuk mempertahankan customer lama. Oleh karena itu, e-commerce perlu melakukan identifikasi penyebab customer churn dan solusi untuk mengatasi permasalahan tersebut. Dengan memprediksi total cost yang dikeluarkan perusahaan sebelum menggunakan machine learning dan sesudah menggunakan machine learning dengan total cost (Acquisition cost + retain cost).

BACKGROUND

1 PROBLEM

Banyaknya customer yang churn sehingga Ecommerce perlu memprediksi pelanggan yang akan churn, dan pengaruhnya terhadap total cost.

2 GOALS

E-commerce mampu mengefisiensikan pengeluaran cost bagi customer churn menjadi lebih hemat sekitar 40% .

3 OBJECTIVE

Dapat mengefisiensikan cost yang dikeluarkan oleh e-commerce untuk customer churn.

4 BUSINESS METRICS

Bisnis metrics yang kami gunakan yakni total cost yang diperoleh dari perhitungan (Acqusition cost + retain cost)

Dr.
Srikandi

EDA & INSIGHT

OVERVIEW DATA

FEATURE & LABEL

19 FEATURE
1 LABEL

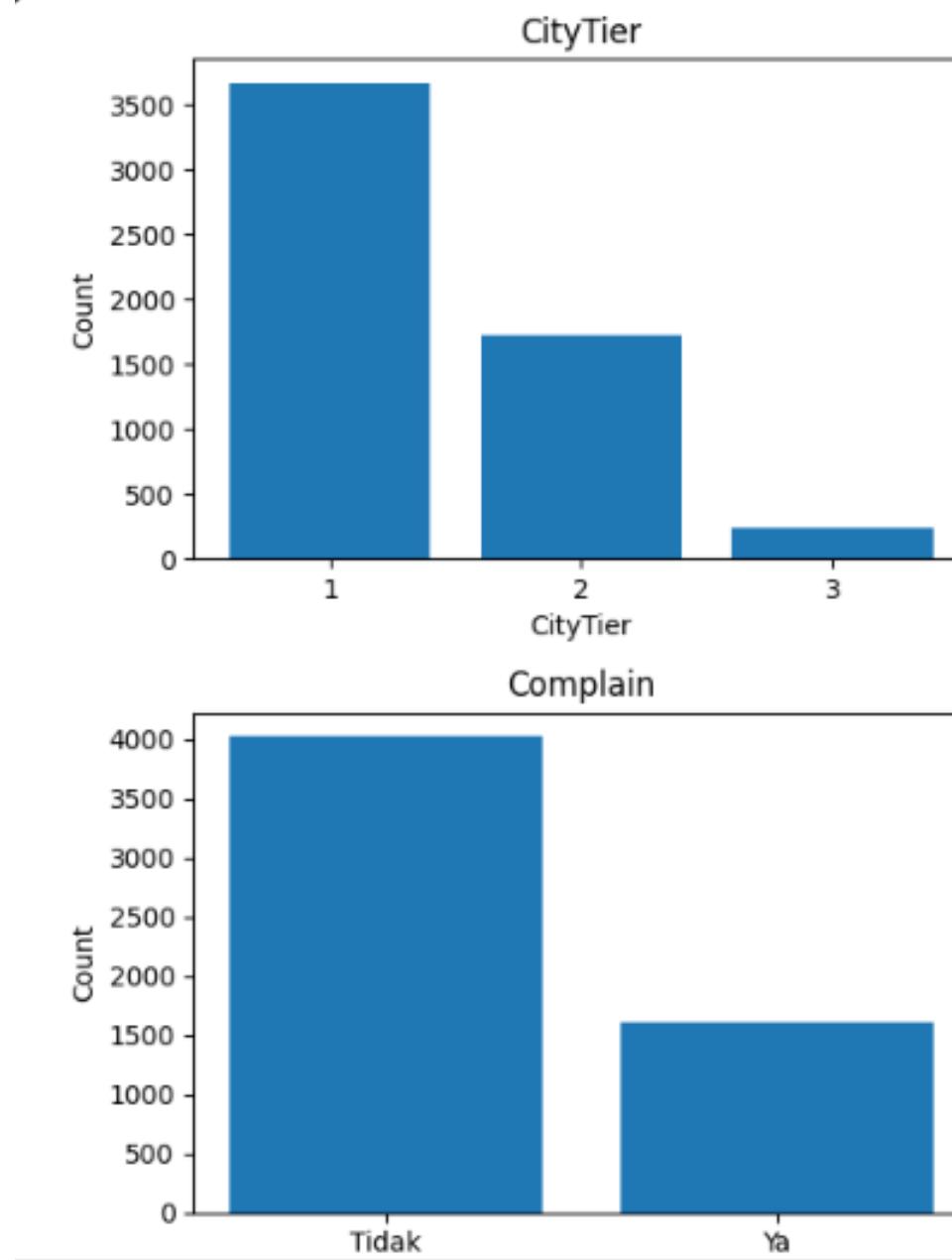
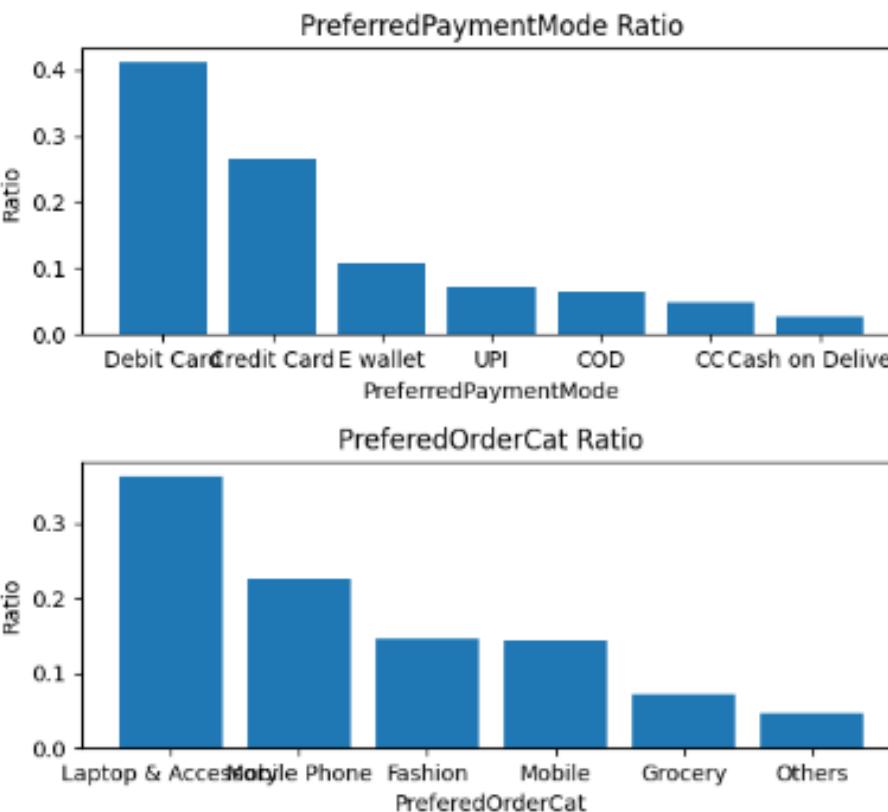
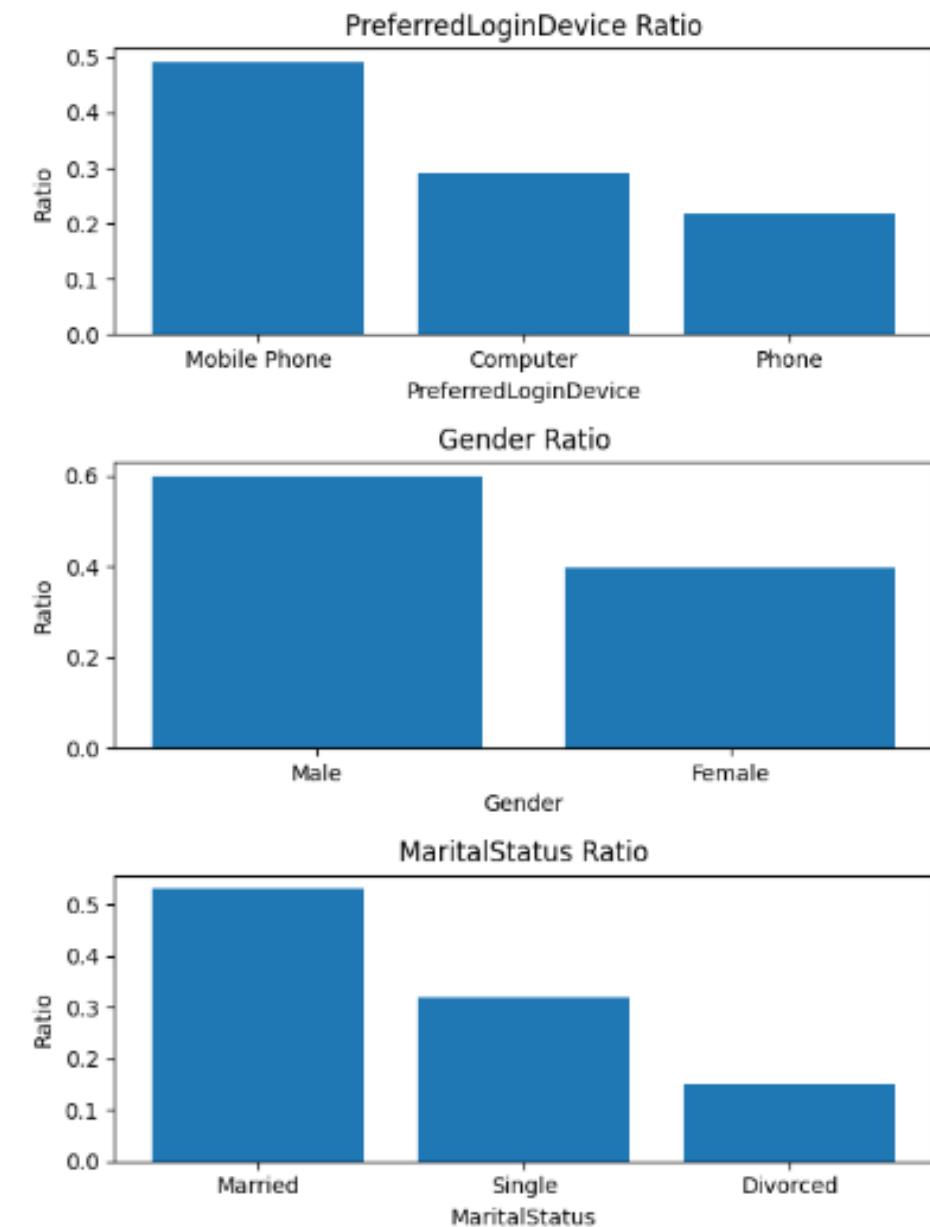
5360 BARIS
20 KOLOM

TIPE DATA

15 NUMERIC
5 CATEGORICAL

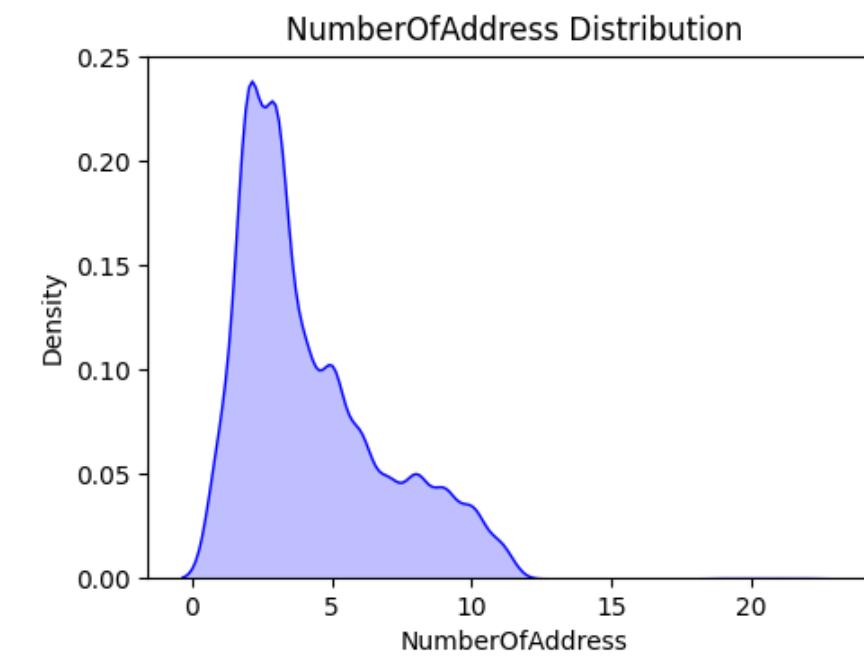
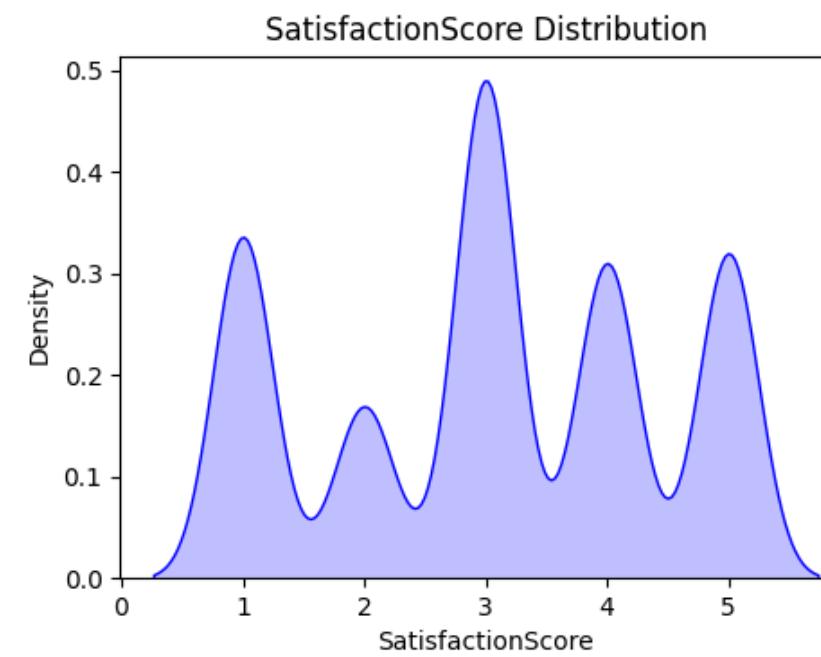
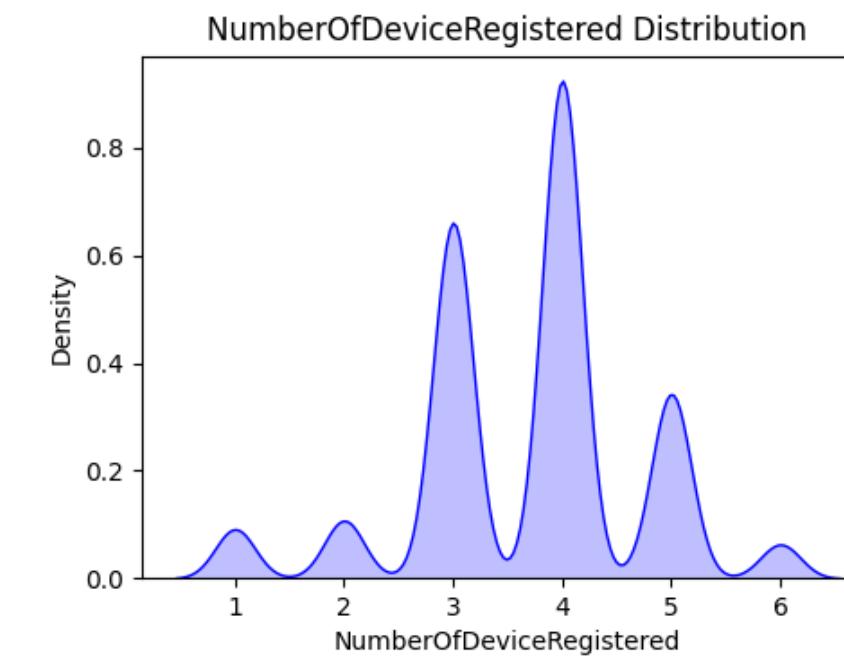
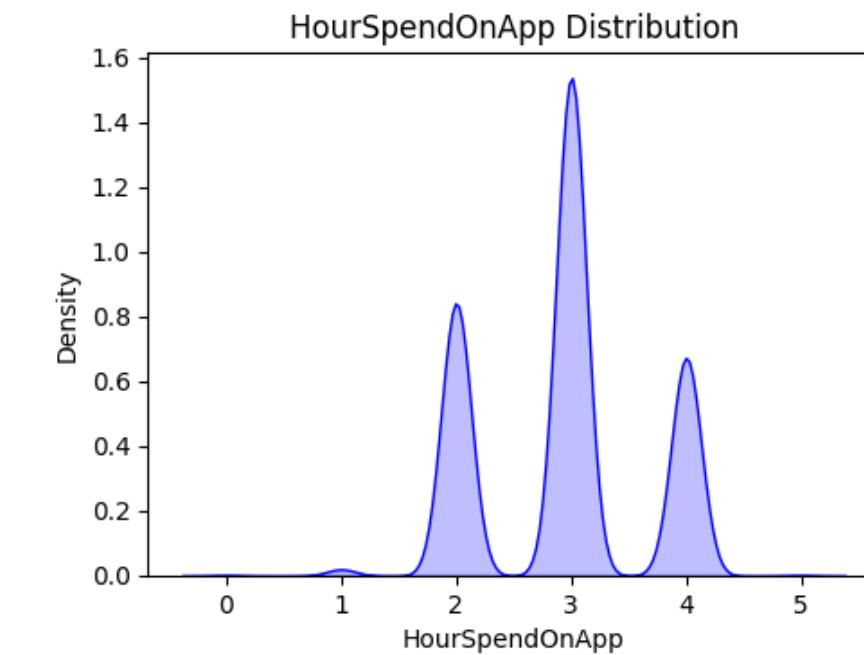
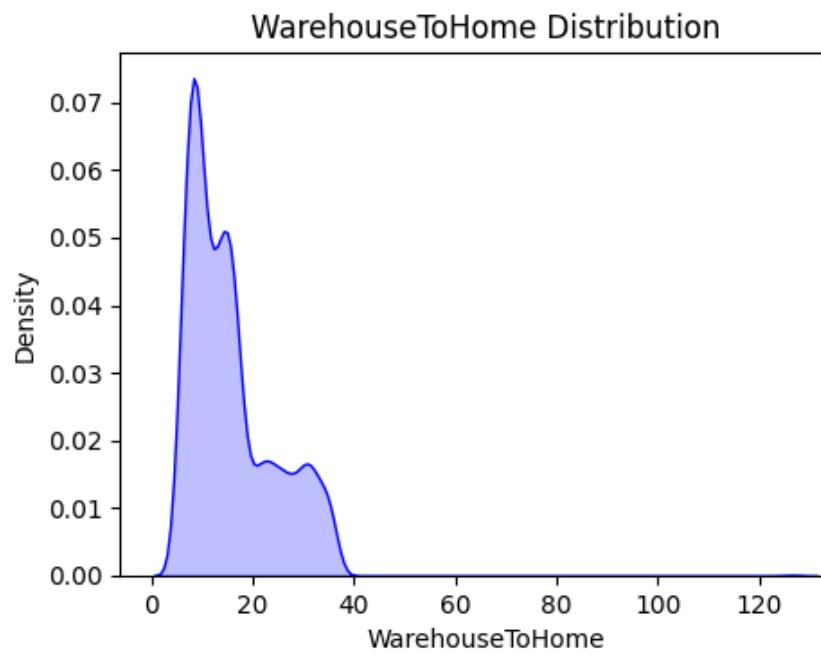
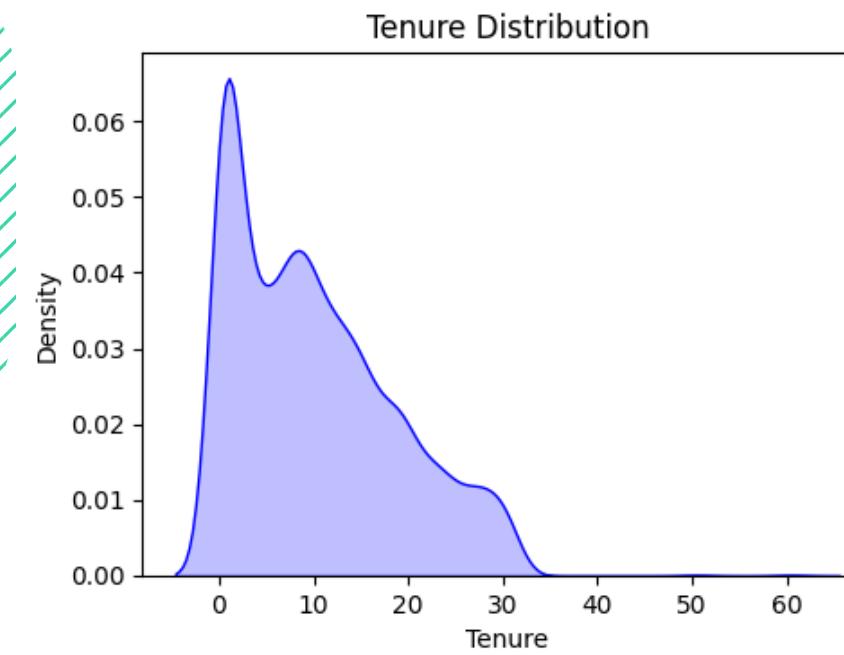
UNIVARIATE ANALYSIS

Categorical Feature Plot



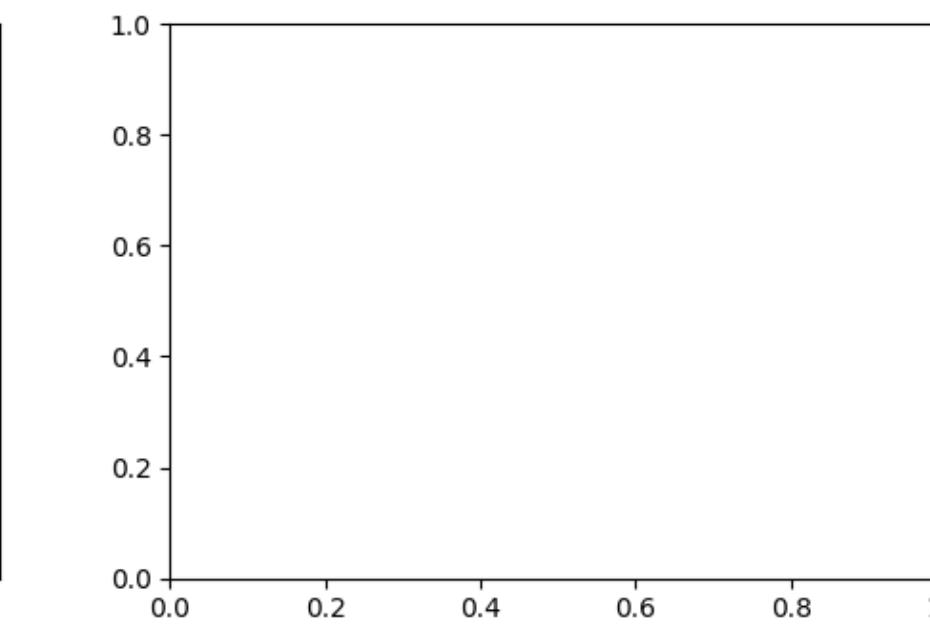
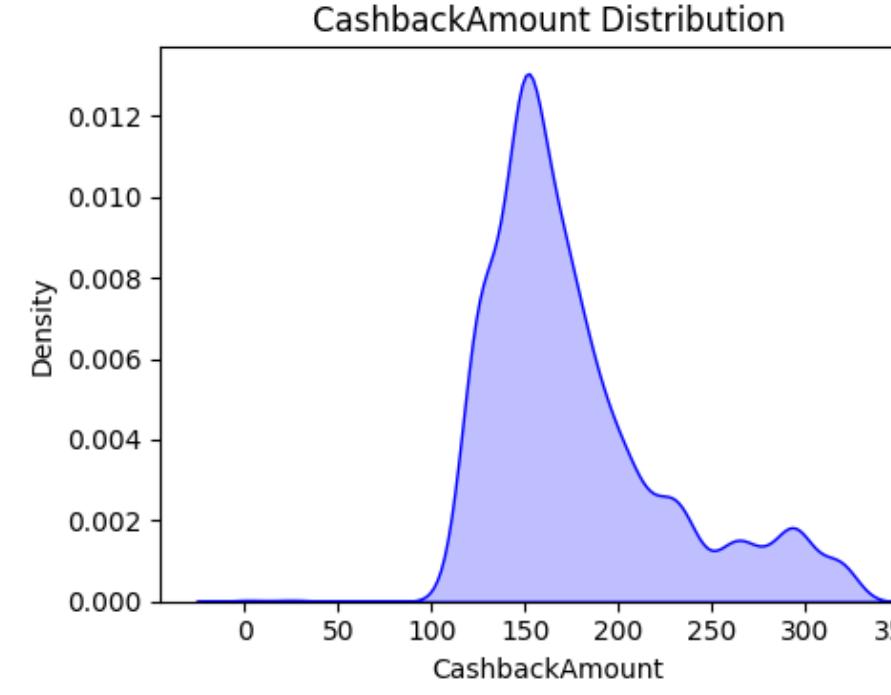
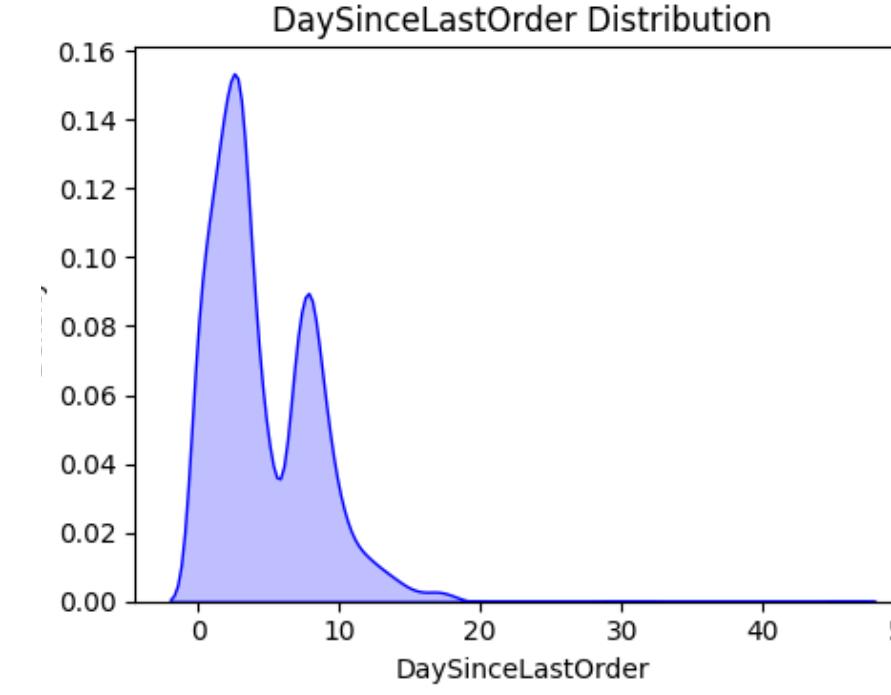
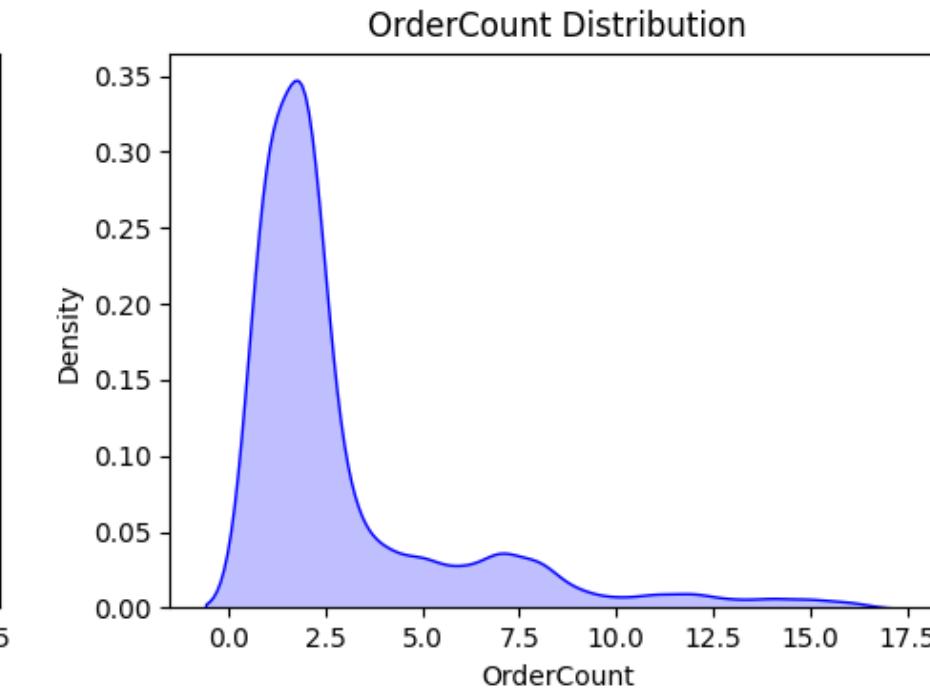
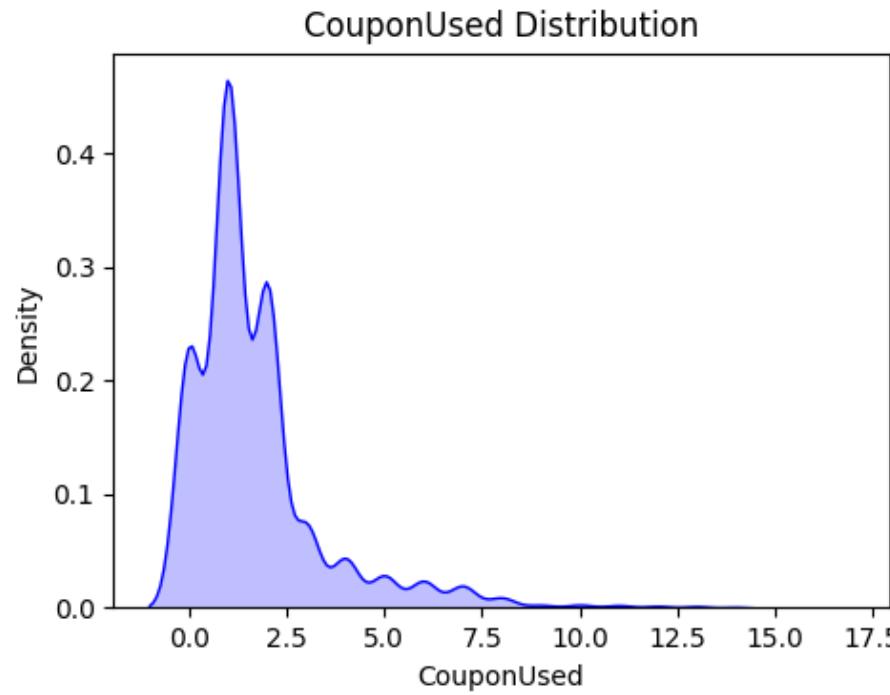
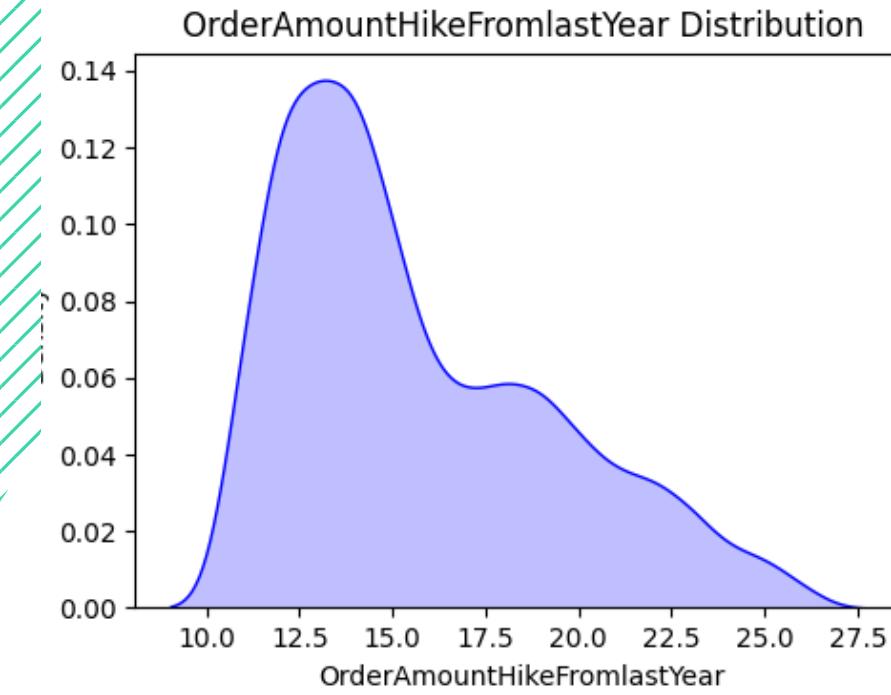
UNIVARIATE ANALYSIS

Numerical Feature Plot

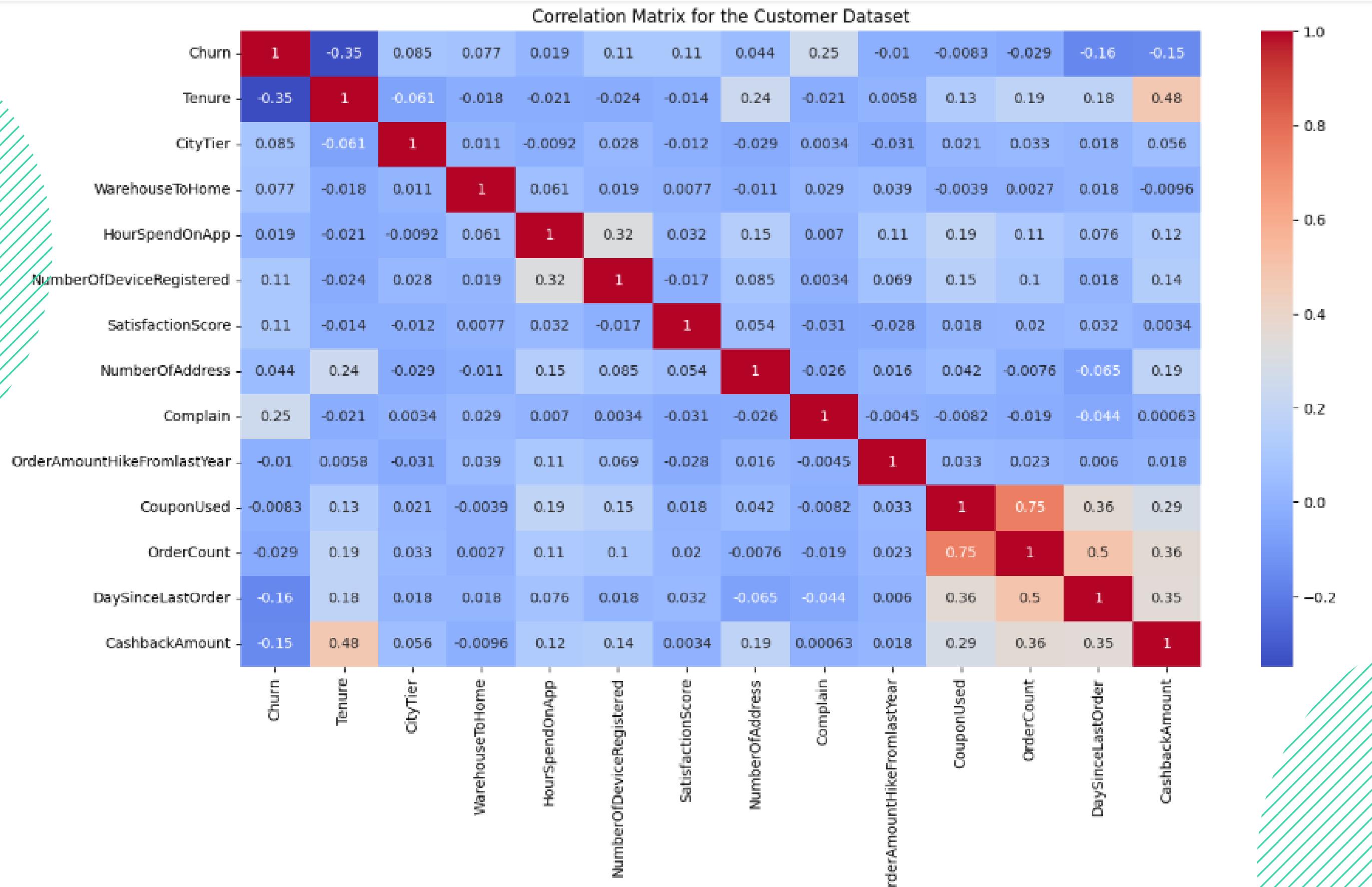


UNIVARIATE ANALYSIS

Numerical Feature Plot

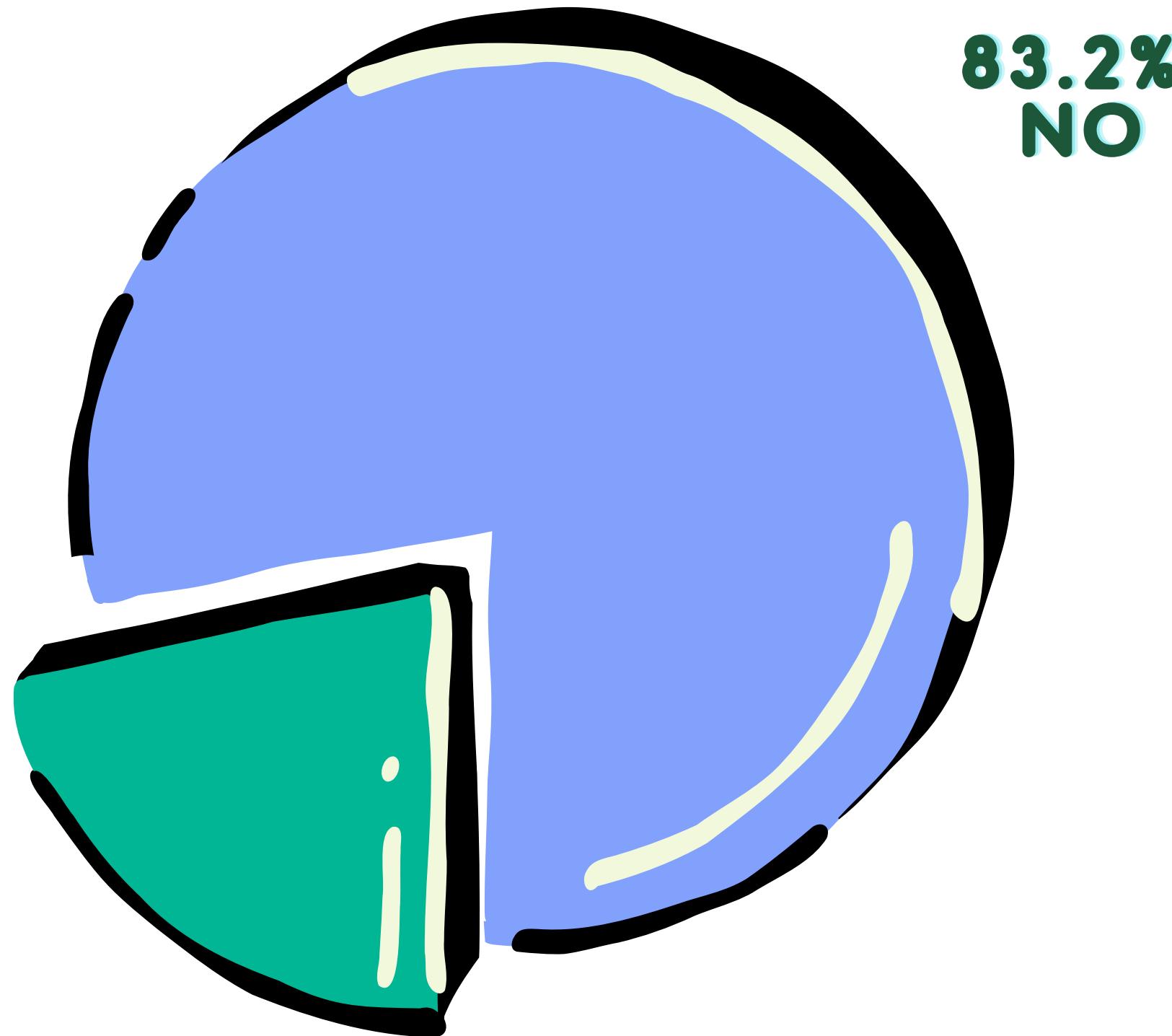


MULTIVARIATE ANALYSIS



Srikandi

PERSENTASE CHURN



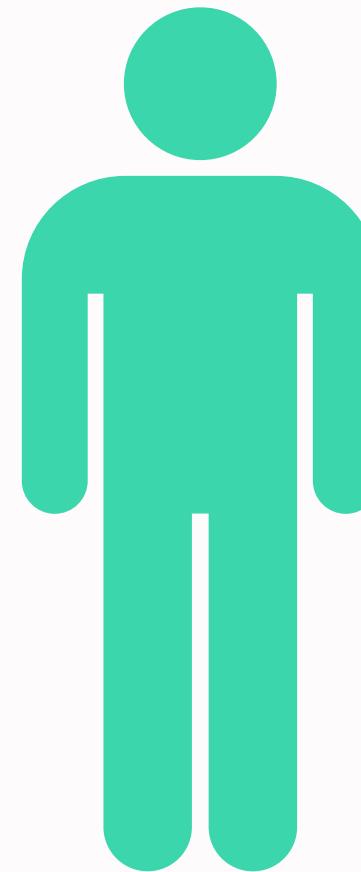
CHURN BERDASARKAN GENDER

FEMALE



46.63%

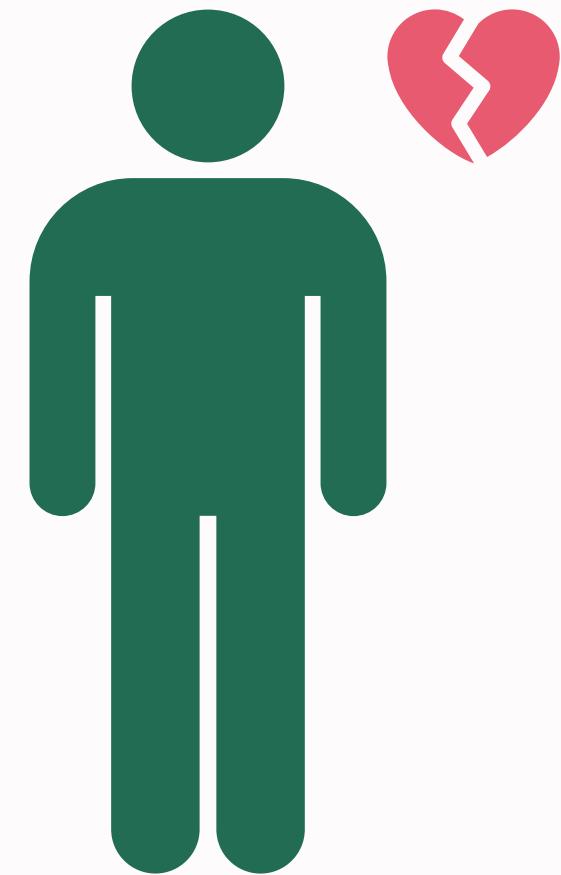
MALE



53.37%

CHURN BERDASARKAN STATUS

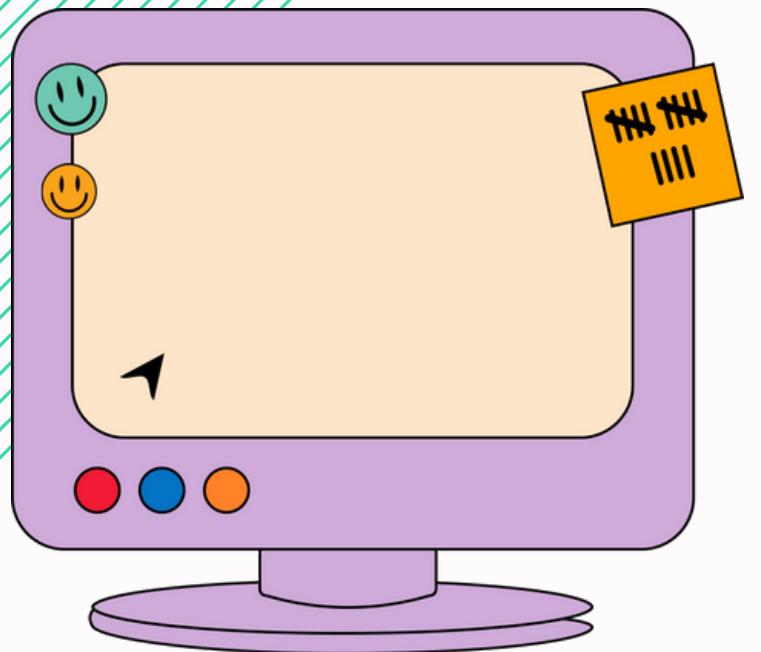
SINGLE



50.57%

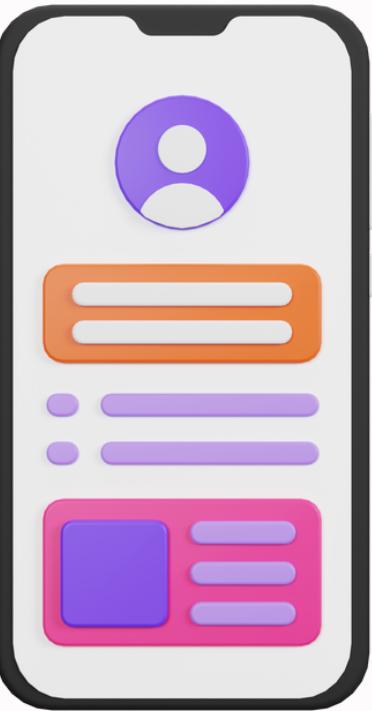
CHURN BERDASARKAN DEVICE

COMPUTER



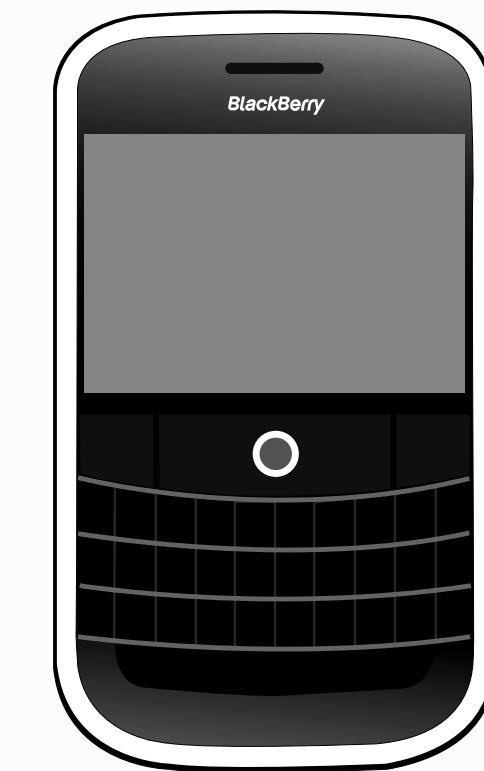
36.25%

MOBILE PHONE



22.97%

PHONE



40.78%

Srikandi

PRE-PROCESSING

DATA CLEANSING

**MISSING VALUES
HANDLING**

7 FEATURE

**DUPLICATED
DATA HANDLING**

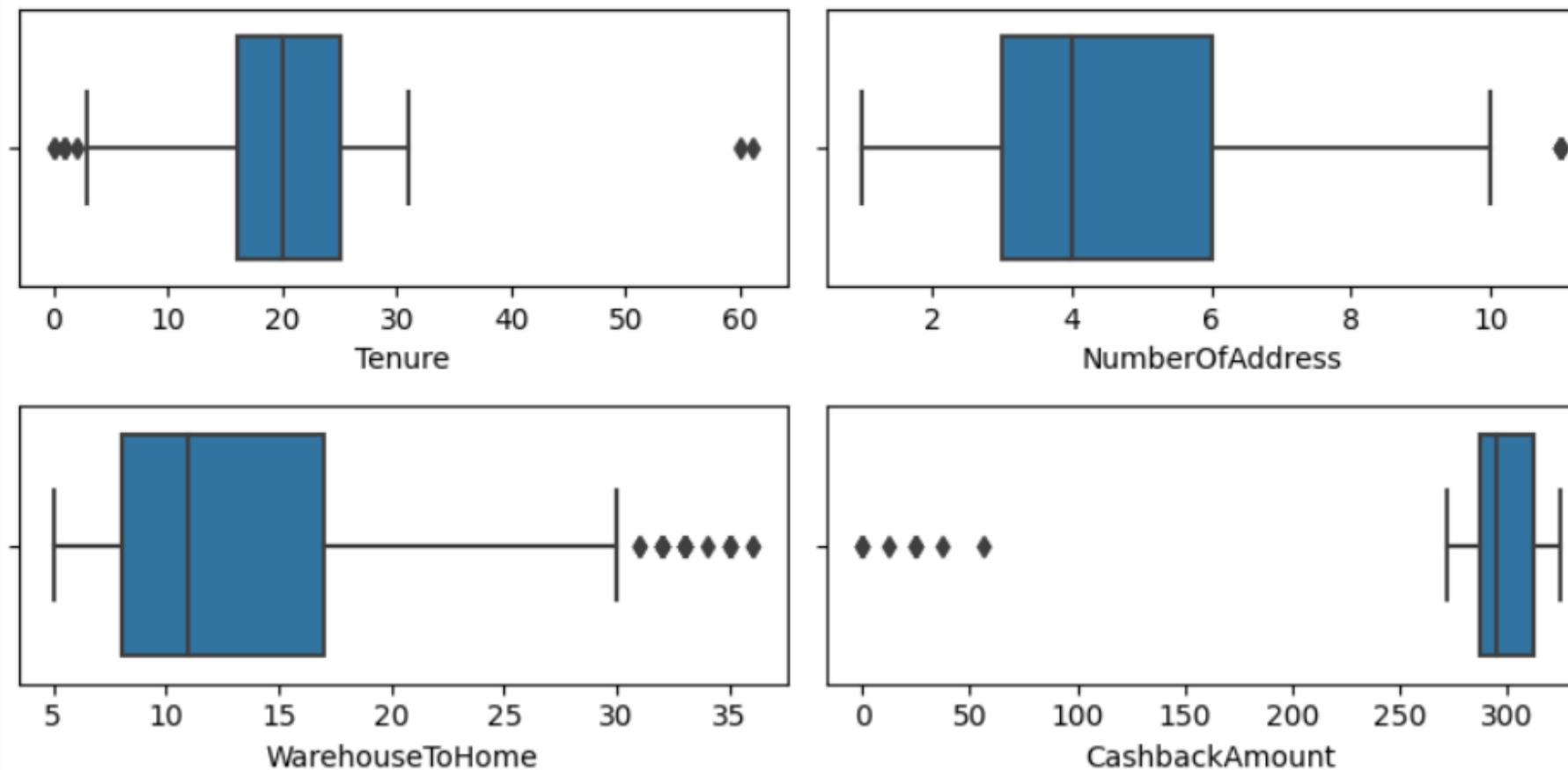
0

Srikandi

Before

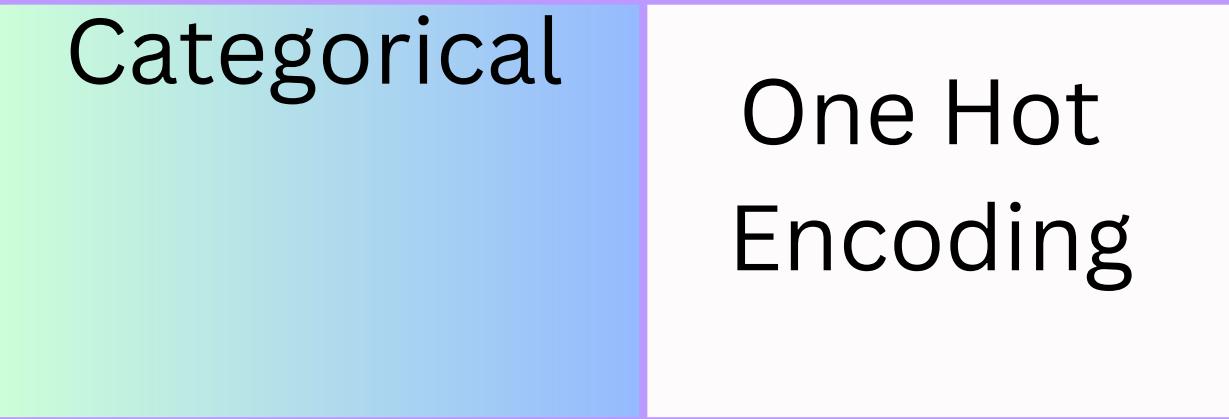


After

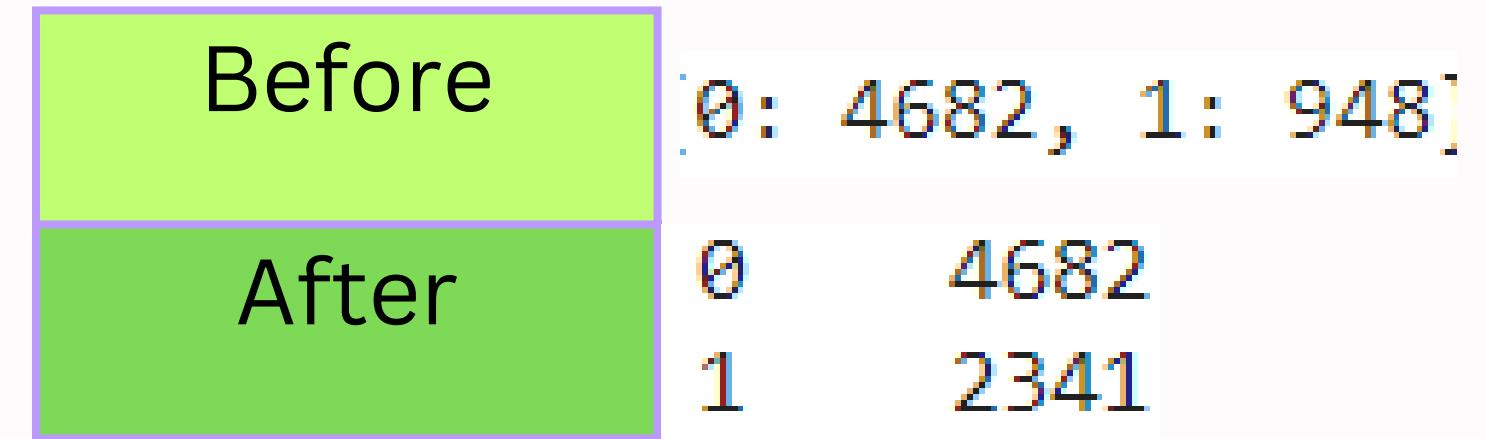


FEATURE ENGINEERING

01 Feature Encoding



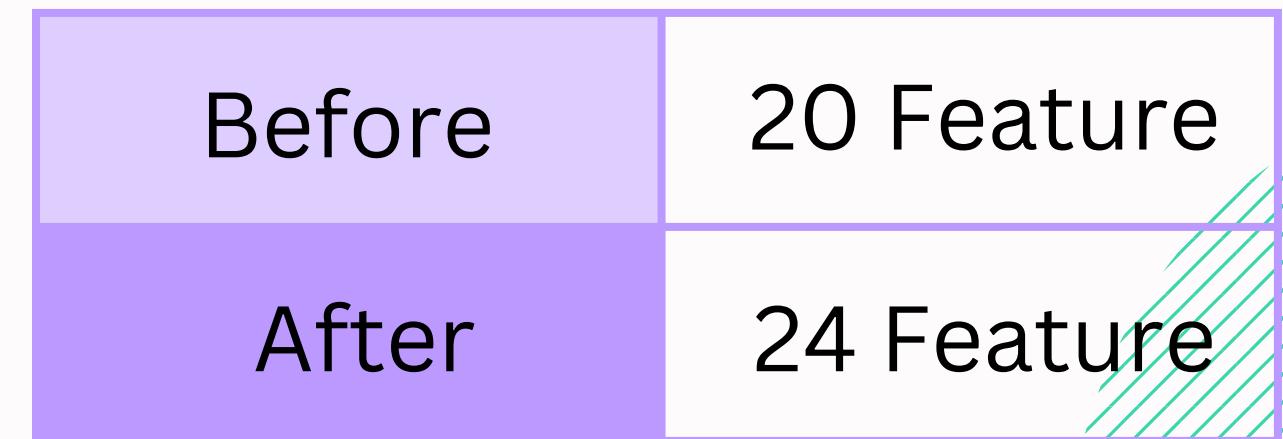
03 Class Imbalance using ROS



02 Feature Transformation



04 Feature Selection



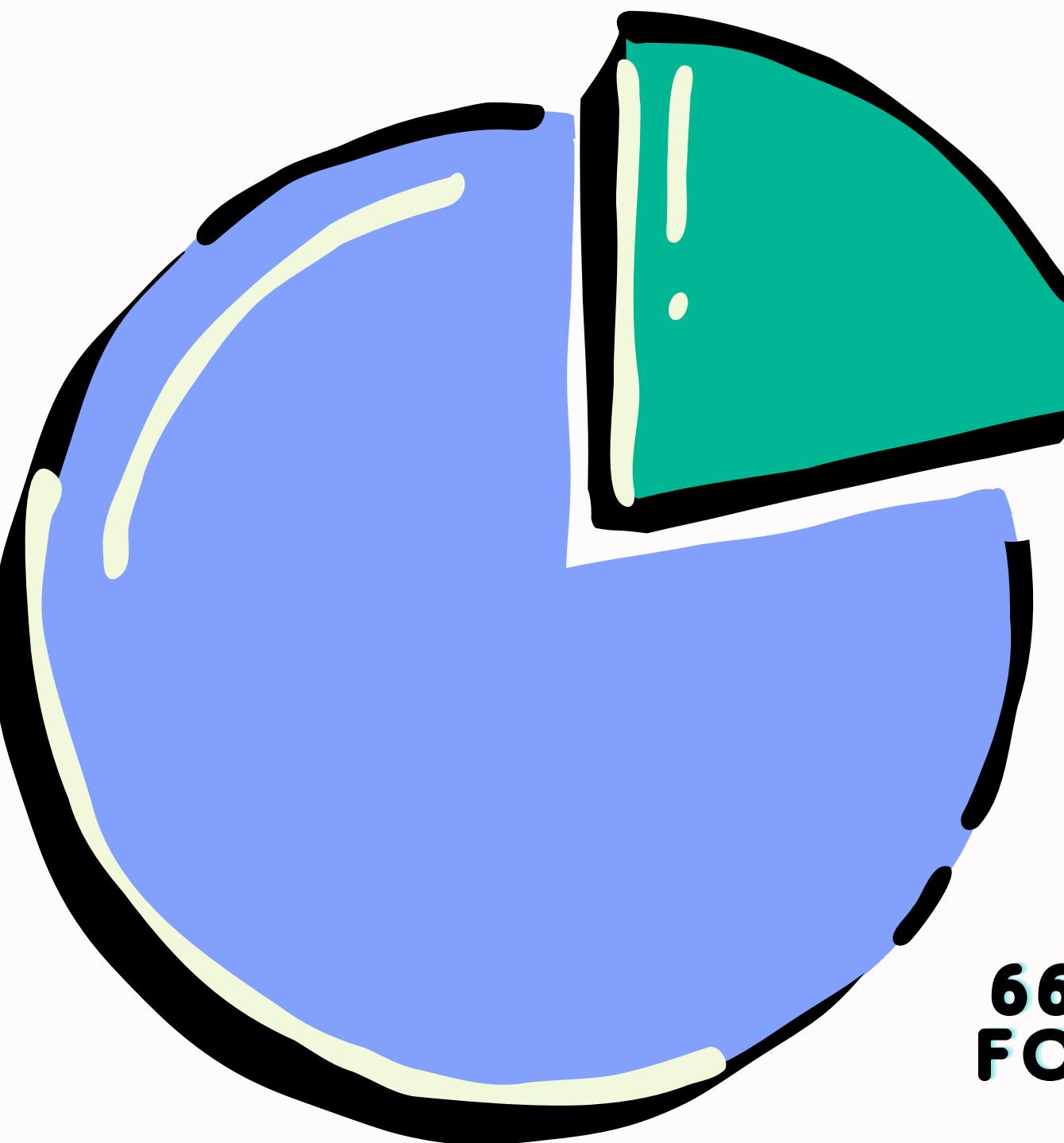
Srikandi

MODELING

A photograph of a man with a beard and dark hair, wearing a light-colored button-down shirt, sitting at a wooden desk in a library or study room. He is looking down at a silver laptop. On the desk next to him is a small orange notebook. In the background, there are bookshelves filled with books and some abstract paintings on the wall.

Srikandi

PEMBAGIAN DATASET



66.67% DATA
FOR TRAINING

33.33% DATA
FOR TESTING

PEMILIHAN MODEL

LOGISTIC REGRESSION



Recall Logistic Regression: 0.52

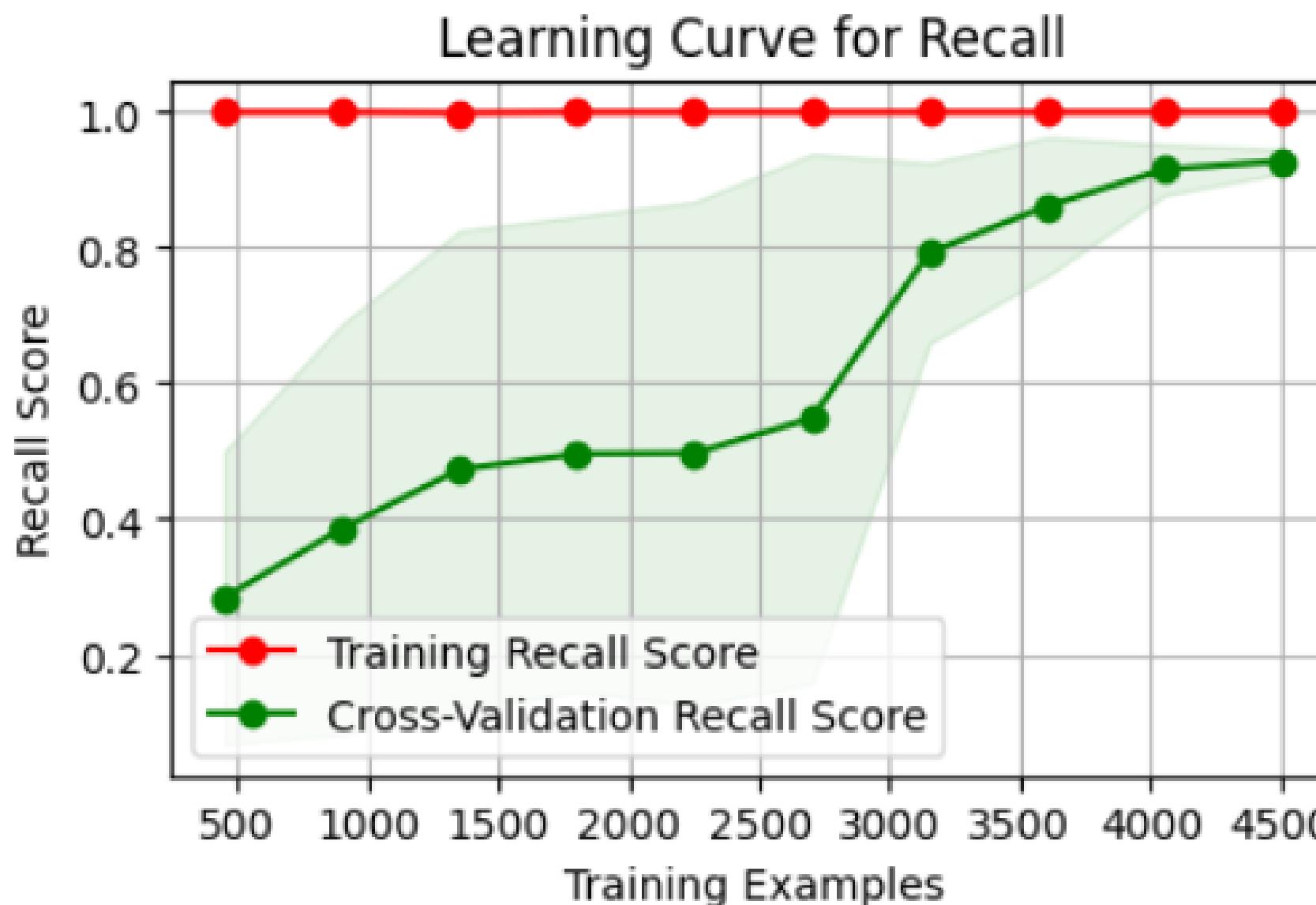
RANDOM FOREST



**BEST
MODEL!**

Recall Random Forest: 0.79

LEARNING CURVE

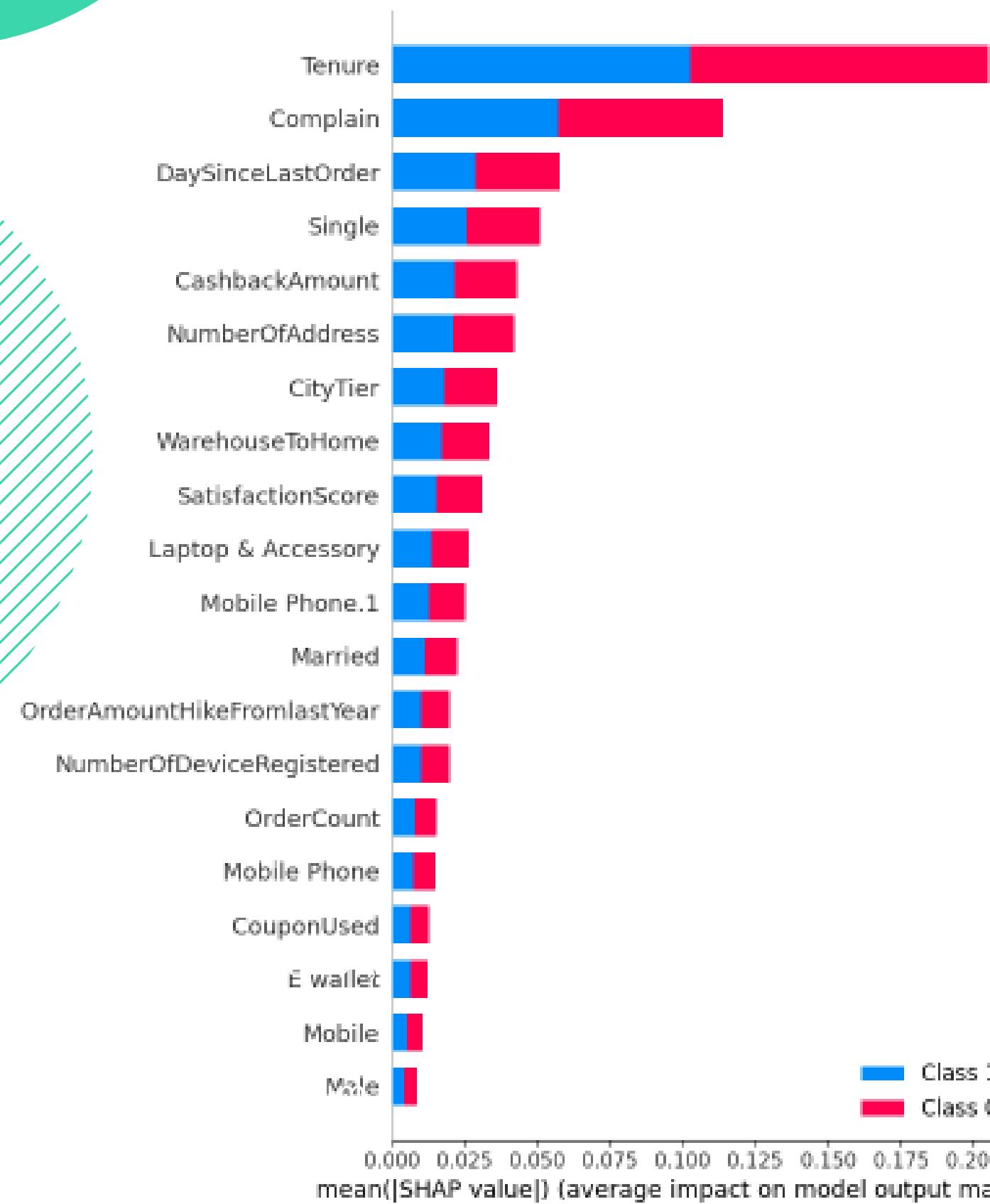


BEST HYPERPARAMETERS

bootstrap	False
max_depth	20
min_samples_leaf	2
min_samples_split	2
n_estimators	100

Srikandi

RECOMMENDATION



Feature Tenure memiliki pengaruh yang paling signifikan menunjukkan bahwa Feature tenure adalah fitur yang paling penting dalam membuat prediksi. dan dapat melakukan Penekanan pada Retensi Pelanggan.



Feature Complain: dapat melakukan analisis lebih lanjut untuk memahami jenis keluhan apa yang paling umum dan apa yang paling memengaruhi keputusan pelanggan.

Rekomendasi Bisnis

Tenure

- Tawarkan program loyalitas atau keuntungan khusus kepada pelanggan yang telah setia selama periode yang lama.

Cashback Amount

- Berikan informasi yang jelas kepada pelanggan tentang program cashback dan manfaatnya untuk memotivasi mereka untuk tetap menggunakan layanan .

Day Since Last Order

- Identifikasi pelanggan dengan lama waktu sejak pesanan terakhir yang tinggi dan kirimkan penawaran khusus atau insentif untuk mendorong pemesanan ulang.

Complain

- analisis lebih lanjut untuk memahami jenis keluhan apa yang paling umum dan apa yang paling memengaruhi keputusan pelanggan.



INTERPRETASI

Asumsi

Dalam kasus ini, kami akan menganggap biaya retensi sebesar \$10, sementara biaya akuisisi adalah \$70 (7 kali lipat lebih besar).

Biaya Retensi

mengacu pada pengeluaran yang dikeluarkan oleh sebuah perusahaan untuk mempertahankan pelanggan yang sudah ada dan mencegah mereka meninggalkan atau beralih ke penyedia layanan lain.

Biaya Akuisisi

pengeluaran yang dikeluarkan oleh perusahaan untuk mendapatkan pelanggan baru. mencakup biaya untuk meyakinkan calon pelanggan agar melakukan pembelian pertama atau mulai menggunakan produk atau layanan dari perusahaan .

79%
RECALL

WITH MACHINE LEARNING



WITHOUT MACHINE LEARNING



50%
AKURASI

WITH MACHINE LEARNING

Ket	Cost	Jumlah Cust	Total
Acquisition Cost	70	$948 \times 21\% = 199$	13.930
Retain Cost	10	$948 \times 79\% = 749$	7.490
			21.420

WITHOUT MACHINE LEARNING

Ket	Cost	Jumlah Cust	Total
Acquisition Cost	70	$948 \times 50\% = 474$	33.180
Retain Cost	10	$948 \times 50\% = 474$	4.740
			37.920

Finally...


$$\begin{aligned}\text{Selisih : } & 37.920 - 21.420 \\& = 16.500\end{aligned}$$

$$\begin{aligned}\text{Saving : } & 16.500 / 37.920 \times 100 \\& = 44\%\end{aligned}$$

Srikandi



E Commerce

THANK YOU



Rakamin
Academy

