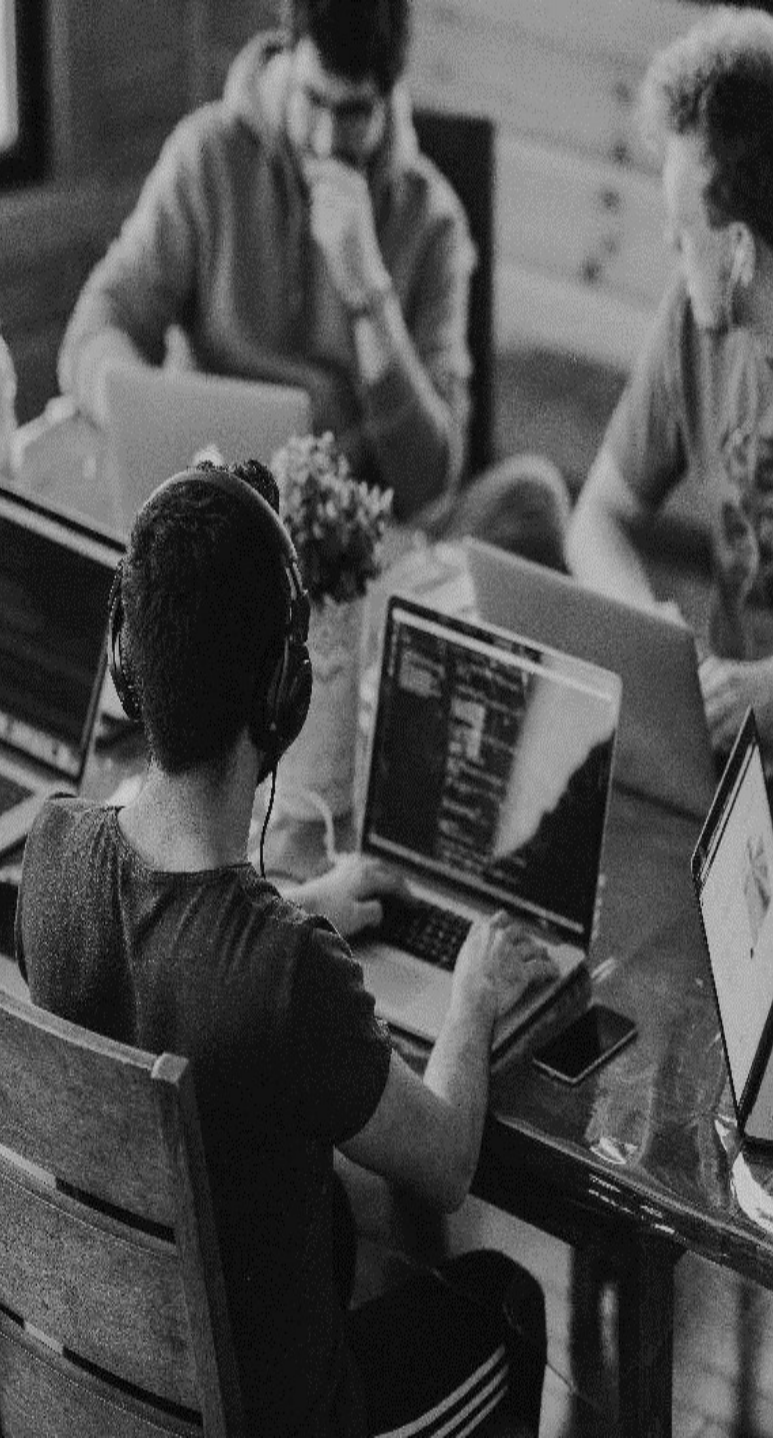
E-COMMERCE CUSTOMER BEHAVIOR ANALYSIS

NICO RAHN

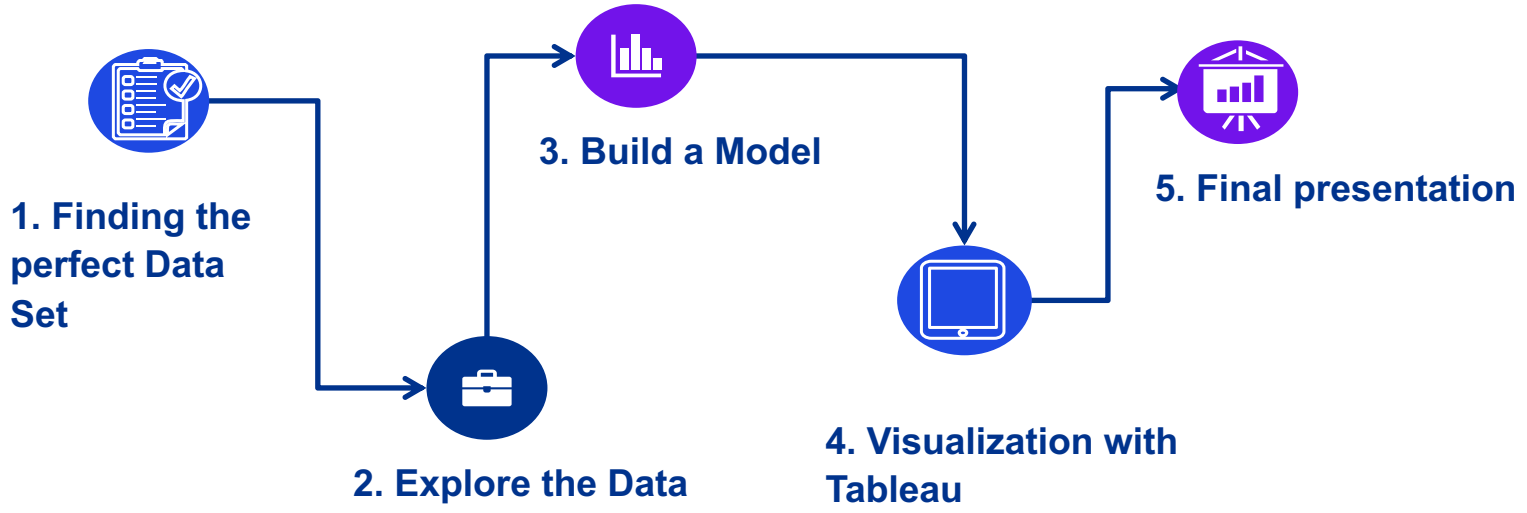
October 28, 2023

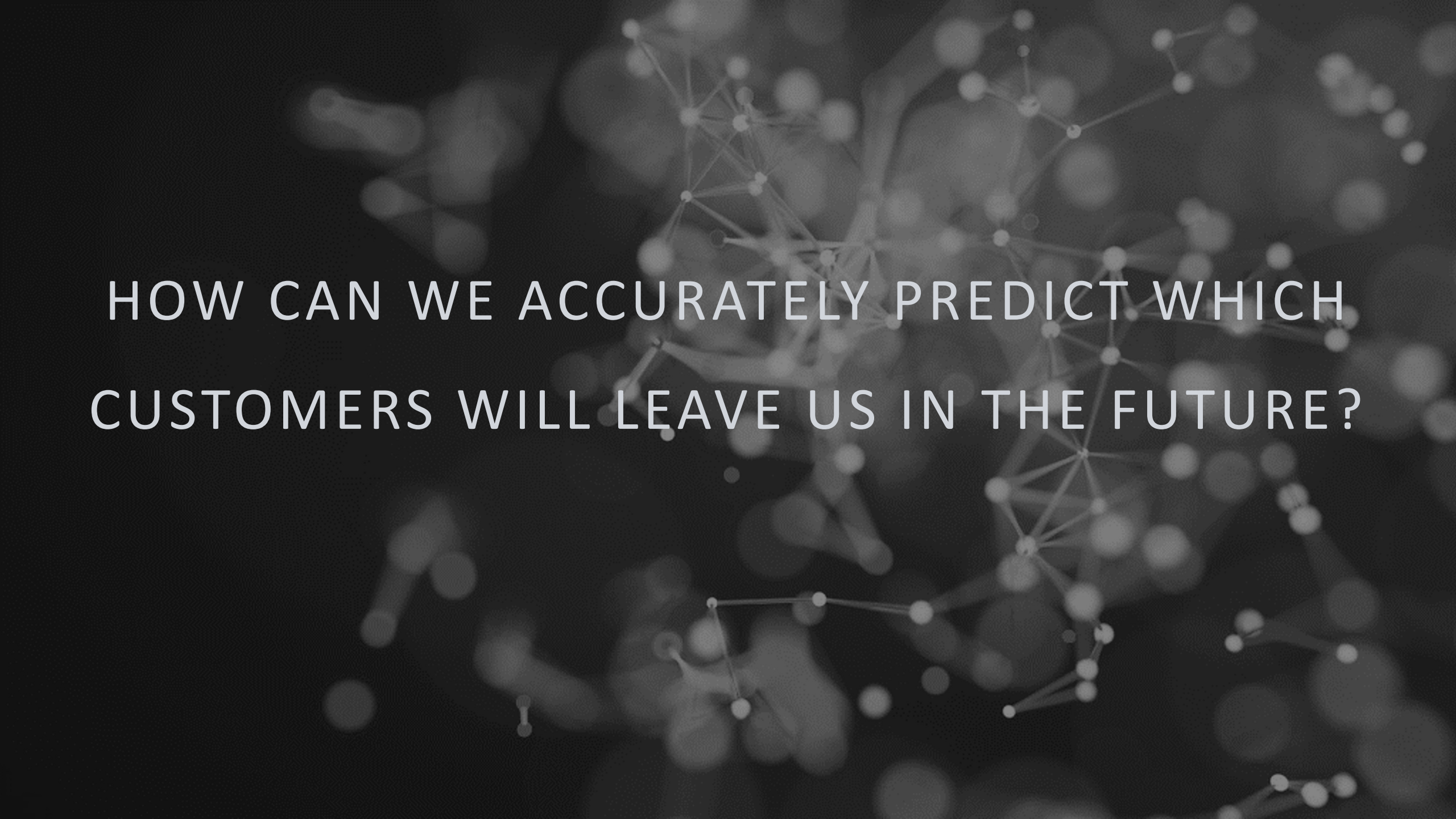
CONTENT



JOURNEY MAP

Timeline




HOW CAN WE ACCURATELY PREDICT WHICH CUSTOMERS WILL LEAVE US IN THE FUTURE?

EXPLORE THE DATA

EDA

EXPLORE THE DATA

1. **Explore the nature of the data**
2. **Data cleaning**
3. **Data pre – processing**
4. **Customer Segmentation Technique**

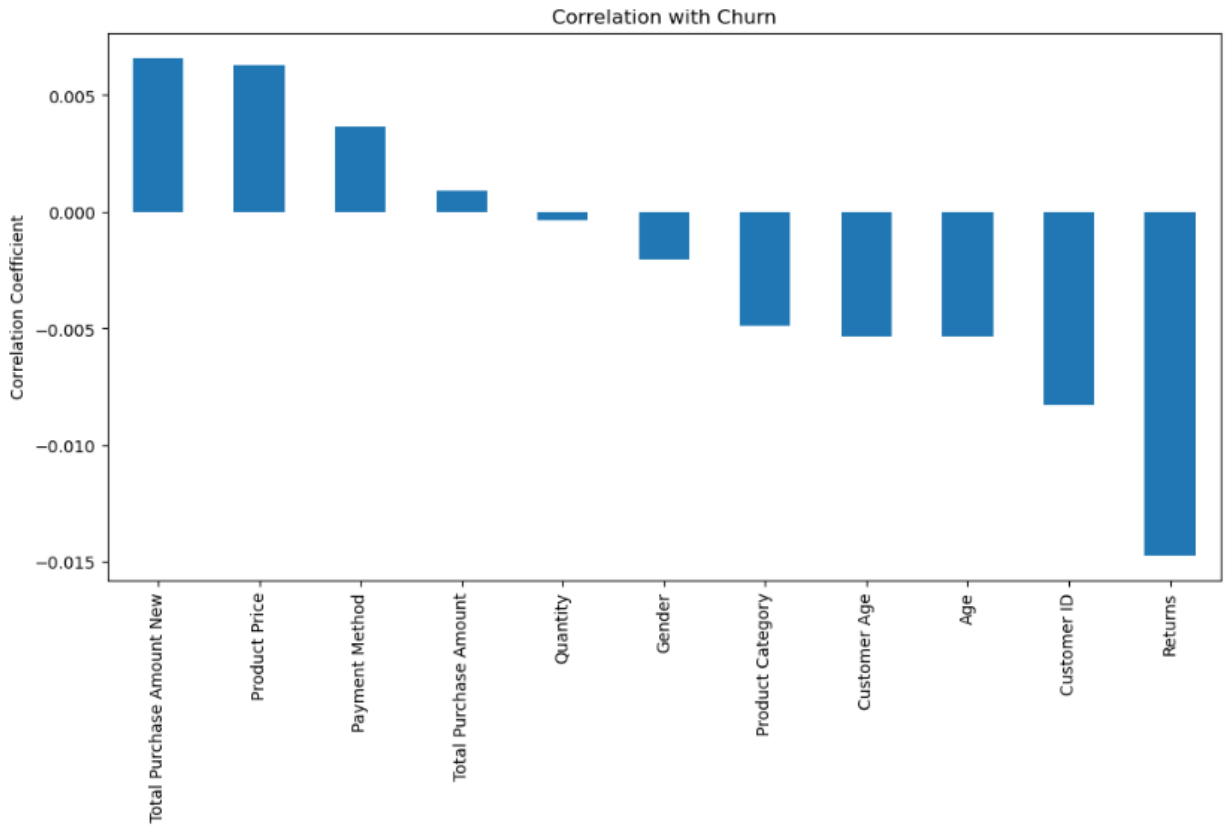
30,513 Customers

47,142 Rows
13 Columns

80%

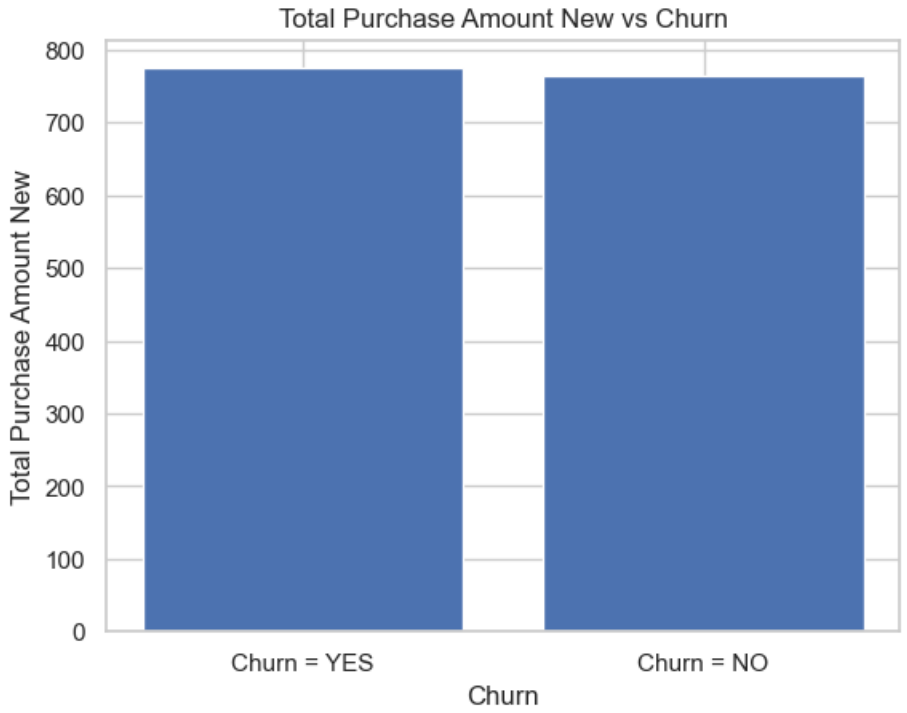
Clean Data

Correlation

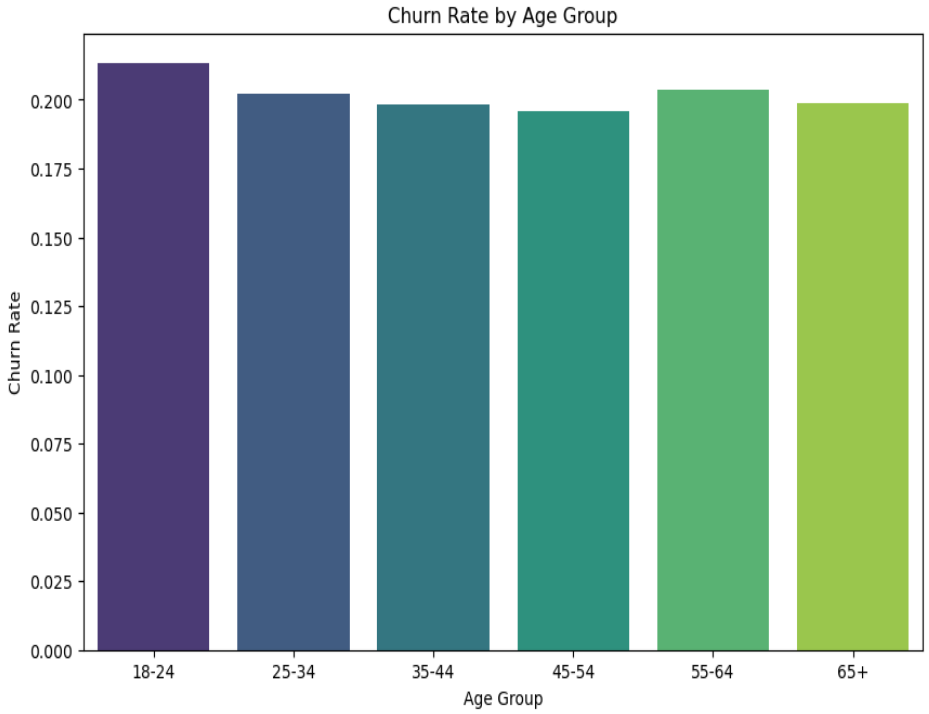


EXPLORE THE DATA

Total Purchase Amount vs. Churn



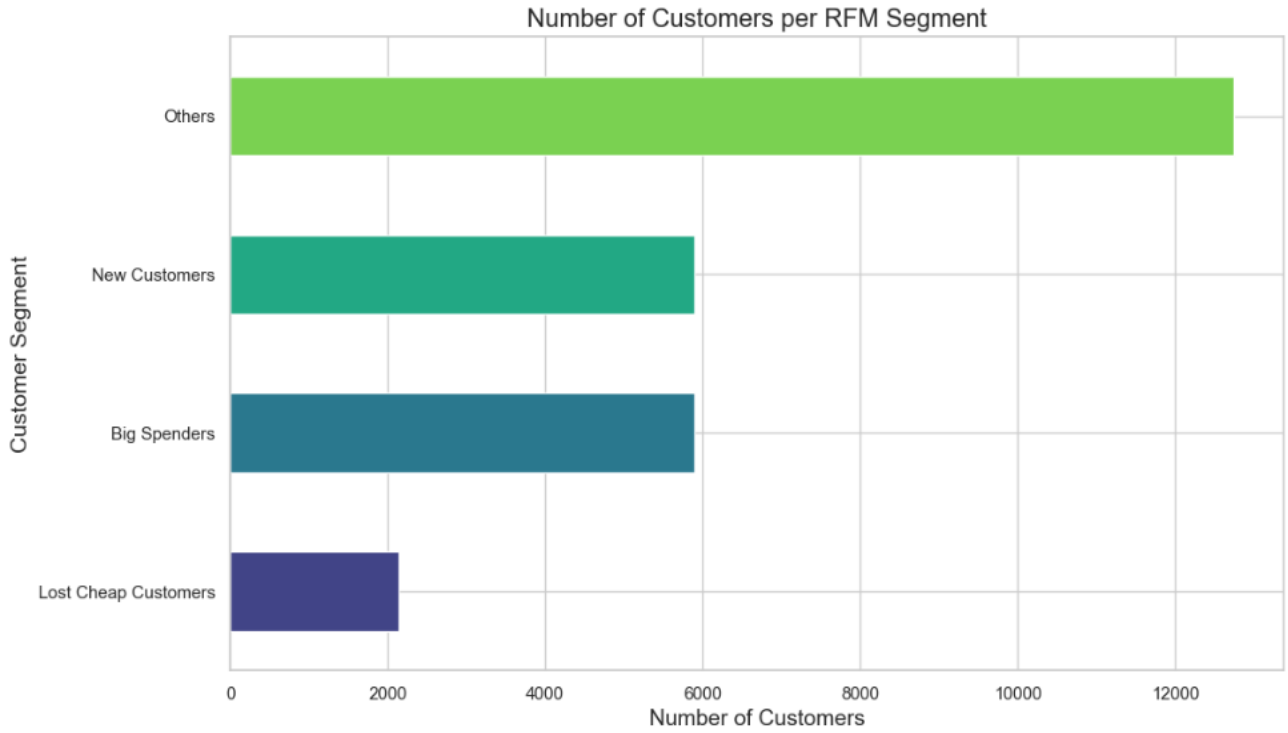
Building Age Groups



CUSTOMER SEGMENTATION

RFM Analysis:

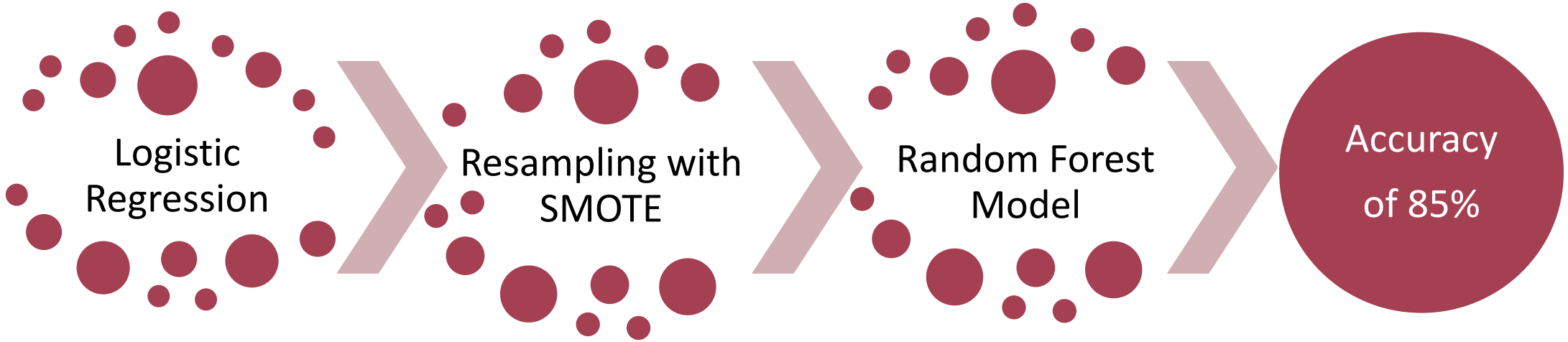
- Recency: How recently a customer made a purchase (timestamp of each purchase)
- Frequency: How often a customer makes a purchase (quantity or number of unique purchase dates)
- Monetary: How much money a customer spends (Total Purchase Amount New)



MACHINE LEARNING

PYTHON

MACHINE LEARNING



0.7985084390945963

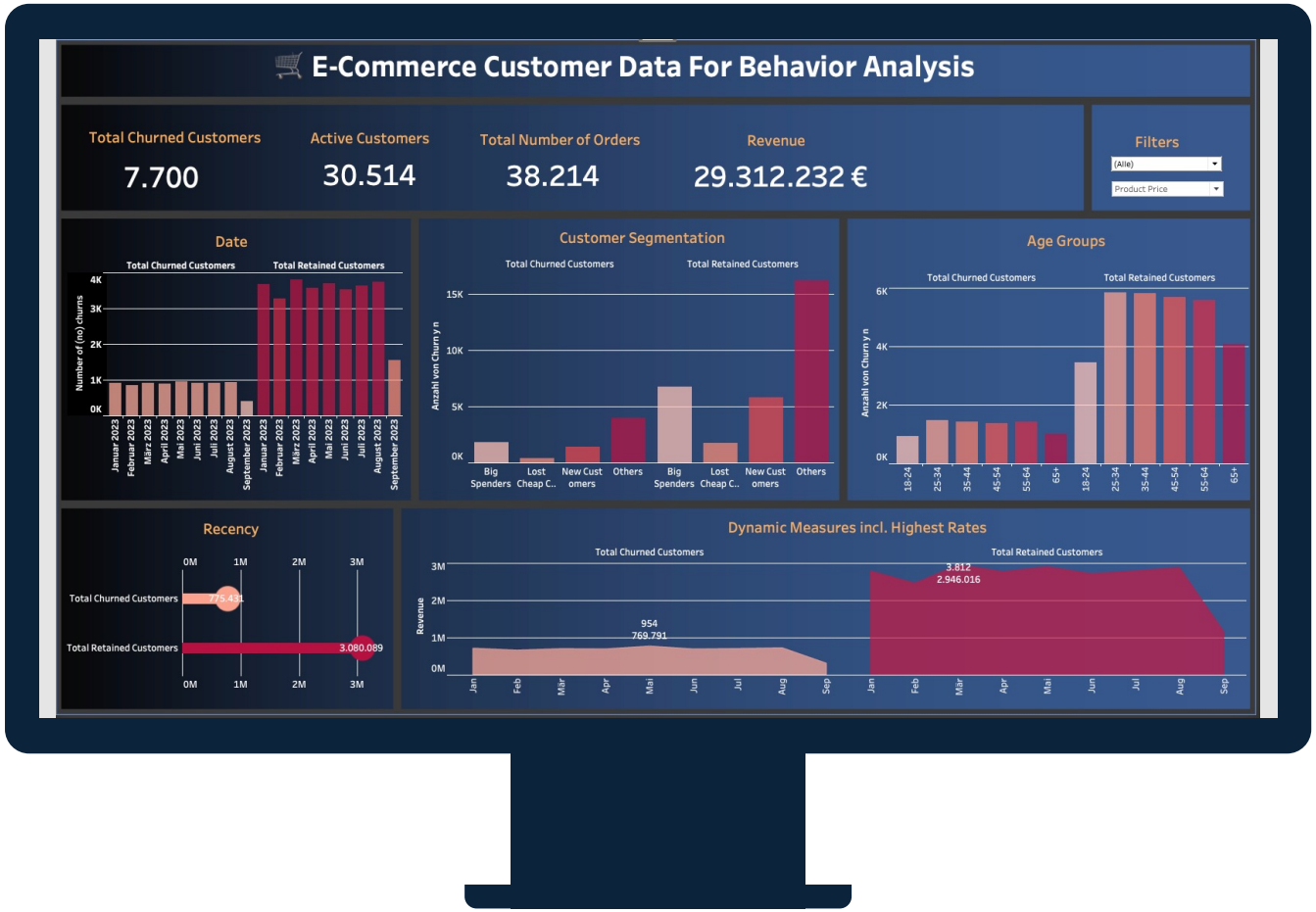
	precision	recall	f1-score	support
0	0.80	1.00	0.89	6103
1	0.00	0.00	0.00	1540
accuracy			0.80	7643
macro avg	0.40	0.50	0.44	7643
weighted avg	0.64	0.80	0.71	7643

	precision	recall	f1-score	support
0	0.79	0.96	0.86	6053
1	0.95	0.75	0.84	6153
accuracy			0.85	12206
macro avg	0.87	0.85	0.85	12206
weighted avg	0.87	0.85	0.85	12206

VISUALIZATION WITH TABLEAU

DASHBOARD

VISUALIZATION WITH TABLEAU



1. Customer Segmentation Visualization:

- Categorizes customers into segments such as 'Lost Cheap Customers', 'Best Customers', 'Loyal Customers', etc.
- Enables tailoring of marketing and retention strategies.

2. Recency Plot:

- Histogram showing the recency of customer purchases.
- Analyzes customer engagement and loyalty over time.

3. Dynamic Metrics with High/Medium/Low Segmentation:

- Metrics like 'Total Active Customers' and 'Total Churned Customers' segmented into 'High', 'Medium', and 'Low'.
- Helps identify active, at-risk, or churned customer groups for targeted interventions.

SUMMARY

NEXT STEPS

SUMMARY

Exploratory Data Analysis (EDA):

- Purpose: Discover patterns, anomalies, or relationships in data
- Methods: Descriptive statistics, visualization, and correlation analysis

Machine learning:

- Logistic Regression: bad accuracy.
- SMOTE: Addressed class imbalance.
- Random Forest: Improved model with 85% accuracy after resampling

Tableau Visualization

- Display of key metrics
- Segmentation
- Recency Plot
- Dynamic Metrics

Next Steps:

- Marketing Campaign for each customer group
- Deeper understanding on Machine learning

THANK YOU!

Nico Rahn

+1 (589) 555-0199

nico@ironhack.com