



SYRIATEL

CUSTOMER CHURN

A telecommunications company



TABLE OF CONTENTS!



01

INTRODUCTION

02

OBJECTIVES

03

**DATA
UNDERSTANDING**



04

VISUALISATION

05

MODELLING

06

CONCLUSION



01

INTRODUCTION

INTRODUCTION!

- SyriaTel is a telecommunication company that prides itself in provision of a wide range of services including mobile, landline and internet connectivity.
- It however faces the menace of customer churn which is a problem witnessed across several other telecommunication companies.
- Since losing them would result in a large loss of revenue for the business, it is imperative to identify those clients who are likely to quit the organisation in the near future in advance.



OVERVIEW!

- Churn is the process by which customers leave a business or stop doing business with an entity.
- This is mainly brought about by dissatisfaction with the current services or better products by the competitors.
- In order to avoid this, we try to come up with a prediction model to identify the stongholds and weaknesses so as to retain our customers.



PROBLEM STATEMENT!

- In a fiercely competitive market, SyriaTel telecommunications company aims to enhance customer retention strategies to stabilize its market position.
- With the considerable expense of acquiring new customers and the financial impact of customer churn, SyriaTel seeks to implement effective measures to retain existing customers and mitigate losses.
- This project will explore strategies to improve customer retention, thereby ensuring sustainable growth and competitiveness in the telecommunications industry.





02 | OBJECTIVES

—Main Objective.

To predict customer churn using a machine learning classification algorithm model.



—Specific Objectives.

- To do exploratory data analysis on the dataset
- To fit different classification algorithm models to determine which one works best for churn prediction
- To select the best model
- To make predictions using the selected model
- To check the accuracy of the predicted variables





03

**DATA
UNDERSTANDING**

DATA SOURCE!

This data is from SyriaTel telecommunication company and it was obtained from kaggle Churn in Telecom's dataset, <https://www.kaggle.com/datasets/bekksddf/churn-in-telecoms-dataset> .

—Our Dataset...

The dataset has 21 columns and 3,333 rows.

The target column, churn, is a bool column where True means the customer did churn and False means the customer did not churn, making this a binary classification problem.

The other columns were; 6 categorical and 15 numerical columns.

COLUMNS WE ARE WORKING WITH!

- **State:** State the customer lives in.
- **Account Length:** How long the customer has had an account.
- **Area Code:** Area code of the customer's phone number.
- **Phone Number:** The customer's phone number.
- **International Plan:** Whether the customer has an international plan.
- **Voice Mail Plan:** Whether the customer subscribes to a voice mail plan.
- **Number of Voice Mail Messages:** Total number of voice mail messages left by the customer.
- **Total Day Minutes:** Total of daytime minutes used.
- **Total Day Calls:** Total of calls made during the day.

COLUMNS WE ARE WORKING WITH!

- **Total Day Charge:** Total charges obtained for daytime calls.
- **Total Eve Minutes:** Total minutes spent on calls in the evening.
- **Total Eve Calls:** Total of calls made during the evening.
- **Total Eve Charge:** Total charges for evening calls.
- **Total Night Minutes:** Total of night time minutes used.
- **Total Night Calls:** Total of calls made at night.
- **Total Night Charge:** Total charges obtained for night time calls.
- **Total Intl Minutes:** Total international minutes (day, evening, and night).
- **Total Intl Calls:** Total international calls (day, evening, and night).

COLUMNS WE ARE WORKING WITH!

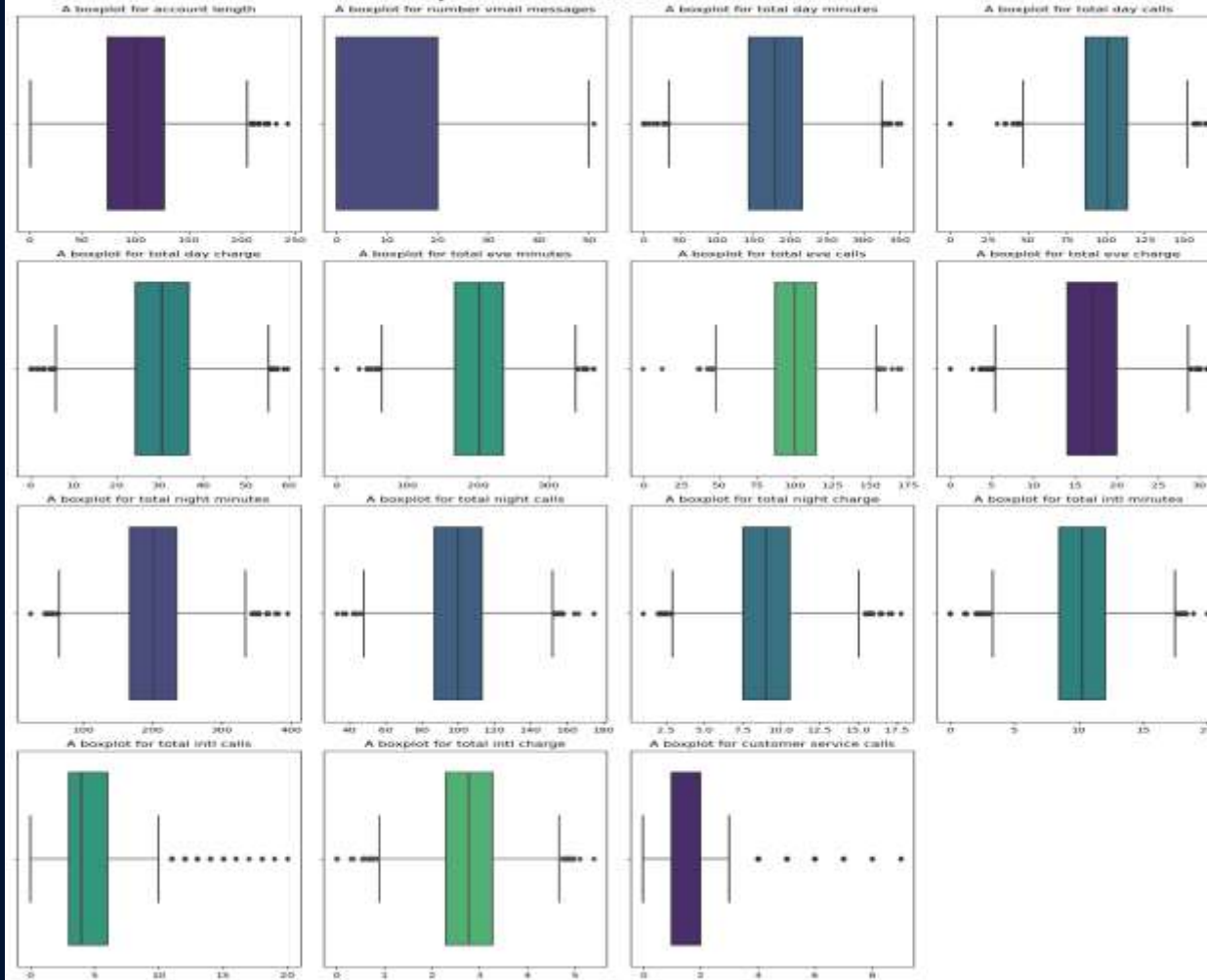
- ***Total Intl Charge:***
Total charges obtained for international calls (day, evening, and night).
- ***Customer Service Calls:*** Number of calls made to customer service by the customer.
- ***Churn:*** If the customer has churned (true or false)



04

VISUALISATION

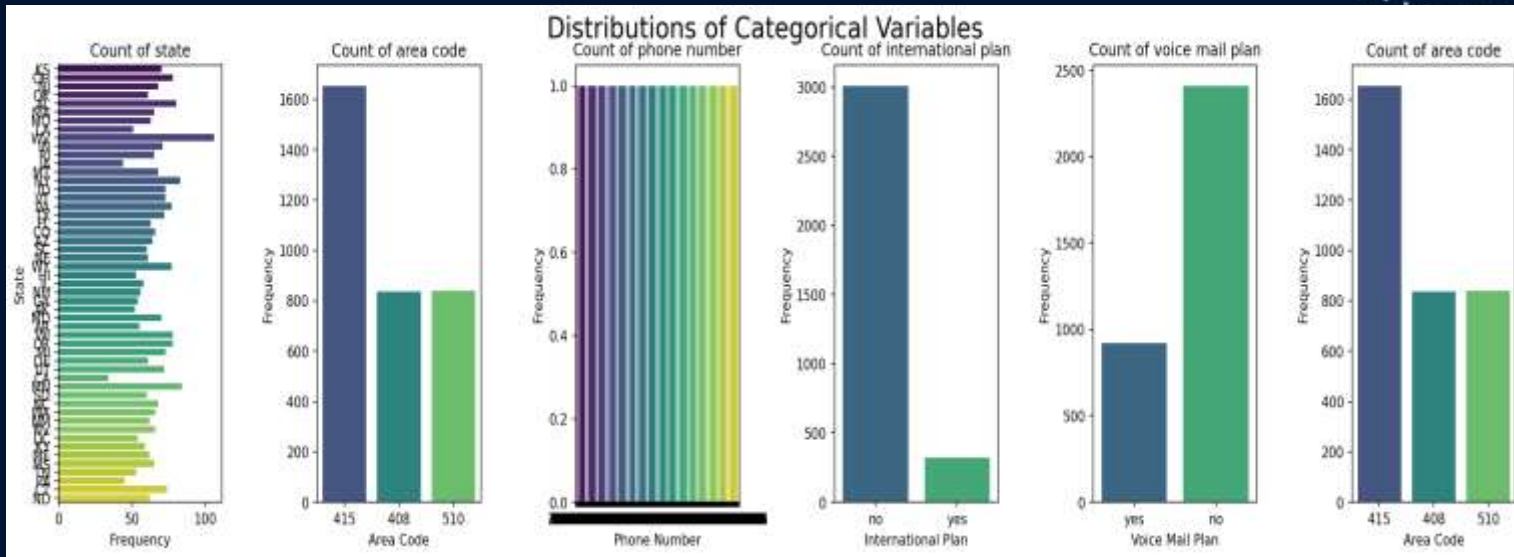
Boxplots for Numerical Variables



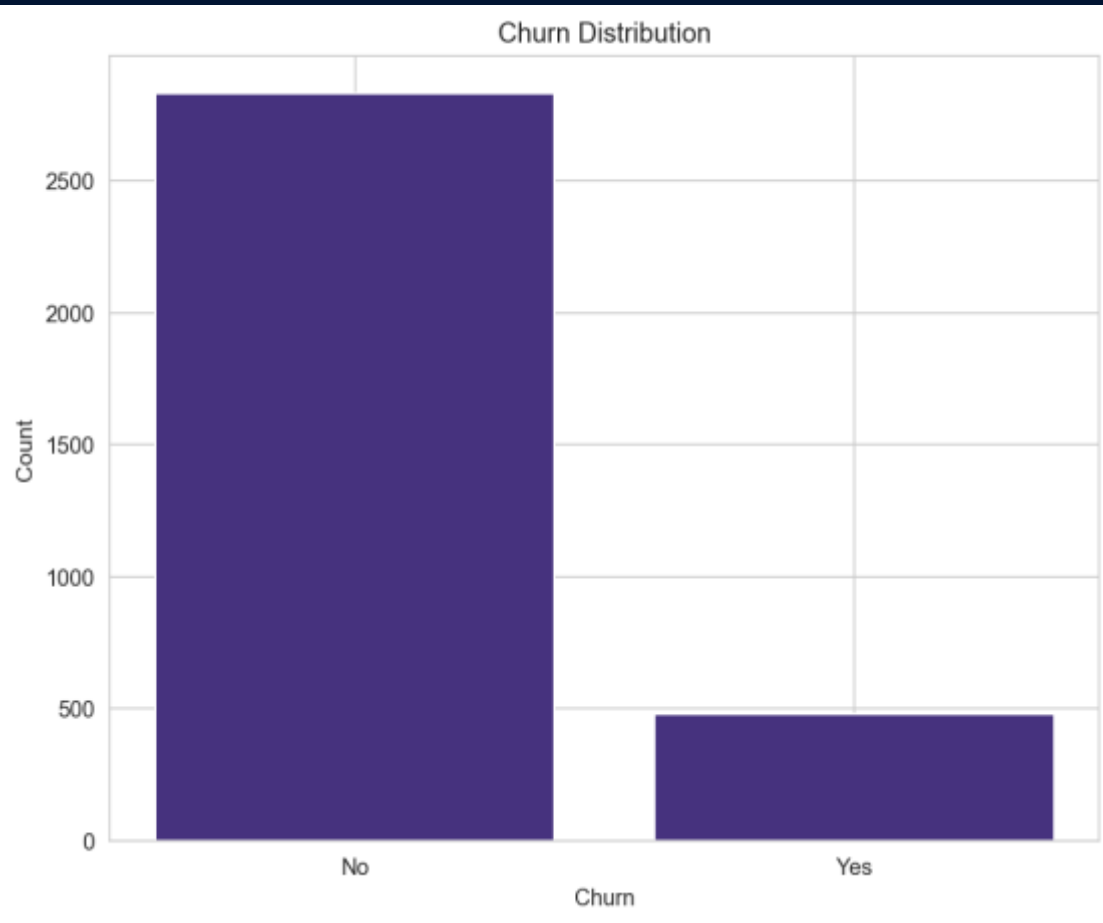
● Numerical Variables

These trends reveal different ways customers behave, like how long they've had their accounts, if they use voicemail, how often they call, how long their international calls last, and how often they ask for service.

● Categorical Variables

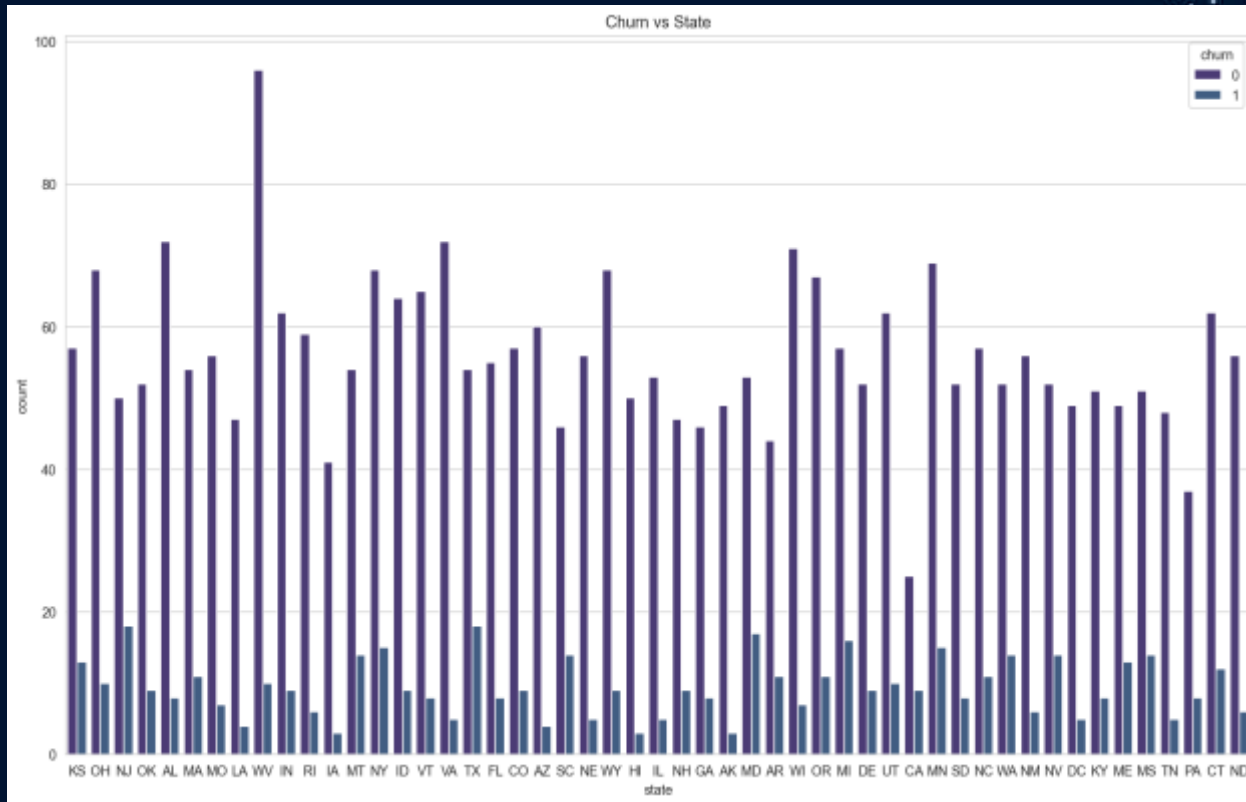


These trends reveal different ways customers behave, like what states they reside in, their telephone area codes, whether or not they have an international plan and their voice mail plan.



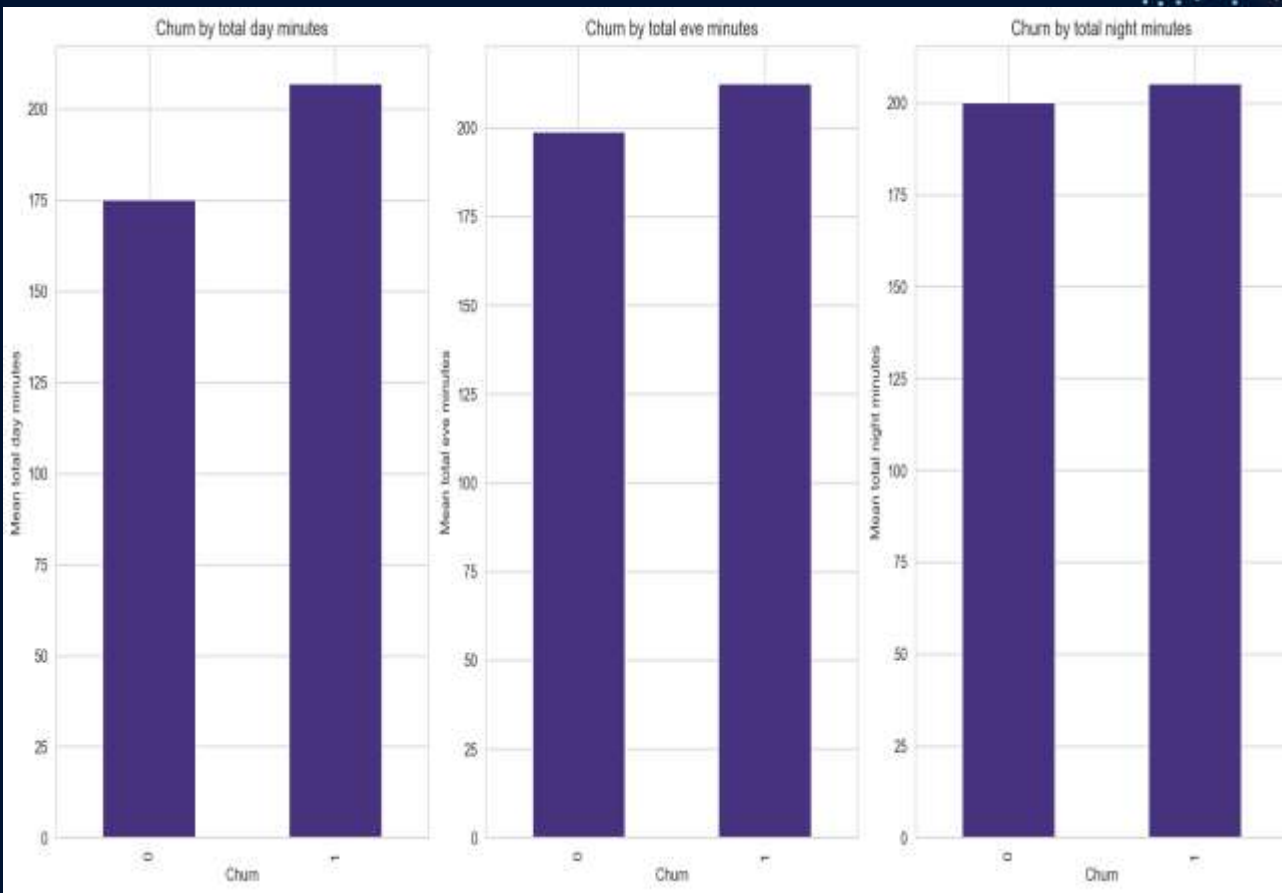
● Churn Distribution

This bar graph shows that most of the customers are sticking with Syria Tel company while a few of the customers are getting churned.



States Analysis

WV, (West Virginia) has the highest customer churn rate while CA, (California) has the lowest customer churn rate.



● Minutes Analysis

Churned customers exhibit slightly higher minute usage across all categories compared to non-churned customers.



05

MODELLING

RANDOM FOREST MODEL

The model achieved an accuracy of 92% on the test data, and overall model performance of 60% as the F1 score indicating a good balance between recall and precision.

Train Accuracy: 1.0
Test Accuracy:
0.9215686274509803

Train Precision: 1.0
Test Precision: 0.9512195121951219

Train Recall: 1.0
Test Recall: 0.43820224719101125

Train F1 Score: 1.0
Test F1 Score: 0.6

LOGISTIC REGRESSION

The logistic regression model showed decent accuracy and precision on the training dataset with the given hyperparameters. However, its performance dropped on the test dataset, suggesting potential overfitting.

Train Accuracy:
0.9173315602836879
Test Accuracy:
0.8793363499245852

Train Precision:
0.9631087063453025
Test Precision: 0.6153846153846154

Train Recall: 0.8679078014184397
Test Recall: 0.2696629213483146

Train F1 Score: 0.9130333411051527
Test F1 Score: 0.375

GRADIENT BOOST MODEL

In conclusion, Gradient Boosting outperforms the other two models on the test set, with an accuracy of approximately 94%, achieving high accuracy while maintaining a well-balanced precision-recall choice.

Train Accuracy: 1.0
Test Accuracy:
0.9441930618401206

Train Precision: 1.0
Test Precision:
0.8333333333333334

Train Recall: 1.0
Test Recall: 0.7303370786516854

Train F1 Score: 1.0
Test F1 Score: 0.7784431137724551



06

CONCLUSION

CONCLUSION!

- ❖ Contributors of High Customer Churn are:
 - ✓ Customers who have an international plan
 - ✓ High number of international minutes
 - ✓ High number of day minutes
 - ✓ High number of night minutes
 - ✓ High number of customer service calls



CONCLUSION!

- ❖ Certain states exhibited a higher churn rate than others, potentially attributed to coverage issues.
- ❖ In the dataset, it was evident that SyriaTel charged customers based on their minutes of usage. However, high customer churn seemed to be driven by factors leading to increased bills, discouraging customers from continuing their phone plans.
- ❖ The model achieves a fine balance between precision and recall, effectively identifying positive class instances while keeping false positives and false negatives low. With an accuracy of 94%, it's well-suited for predicting churn.



RECOMMENDATION!

- ✓ Focus on customer retention strategies in states with higher churn rates.
- ✓ Given its importance in predicting churn, it would be beneficial to review the structure and pricing of the international plan to ensure it meets customer needs.
- ✓ Customers with a higher total charge are more likely to churn. A review of pricing strategies and structures could help to ensure they are competitive and provide value to customers. Evaluate the pricing structure for day, evening, night, and international charges.
- ✓ Enhance the value proposition of the voicemail plan to increase adoption among customers.



An abstract background on the left side of the slide. It features a bright orange and yellow light streak that curves upwards and to the right, surrounded by a trail of small, glowing blue and white particles. The background is a solid dark blue.

THANK YOU!

By yours truly,
THIGA MARGARET WANJIRU.