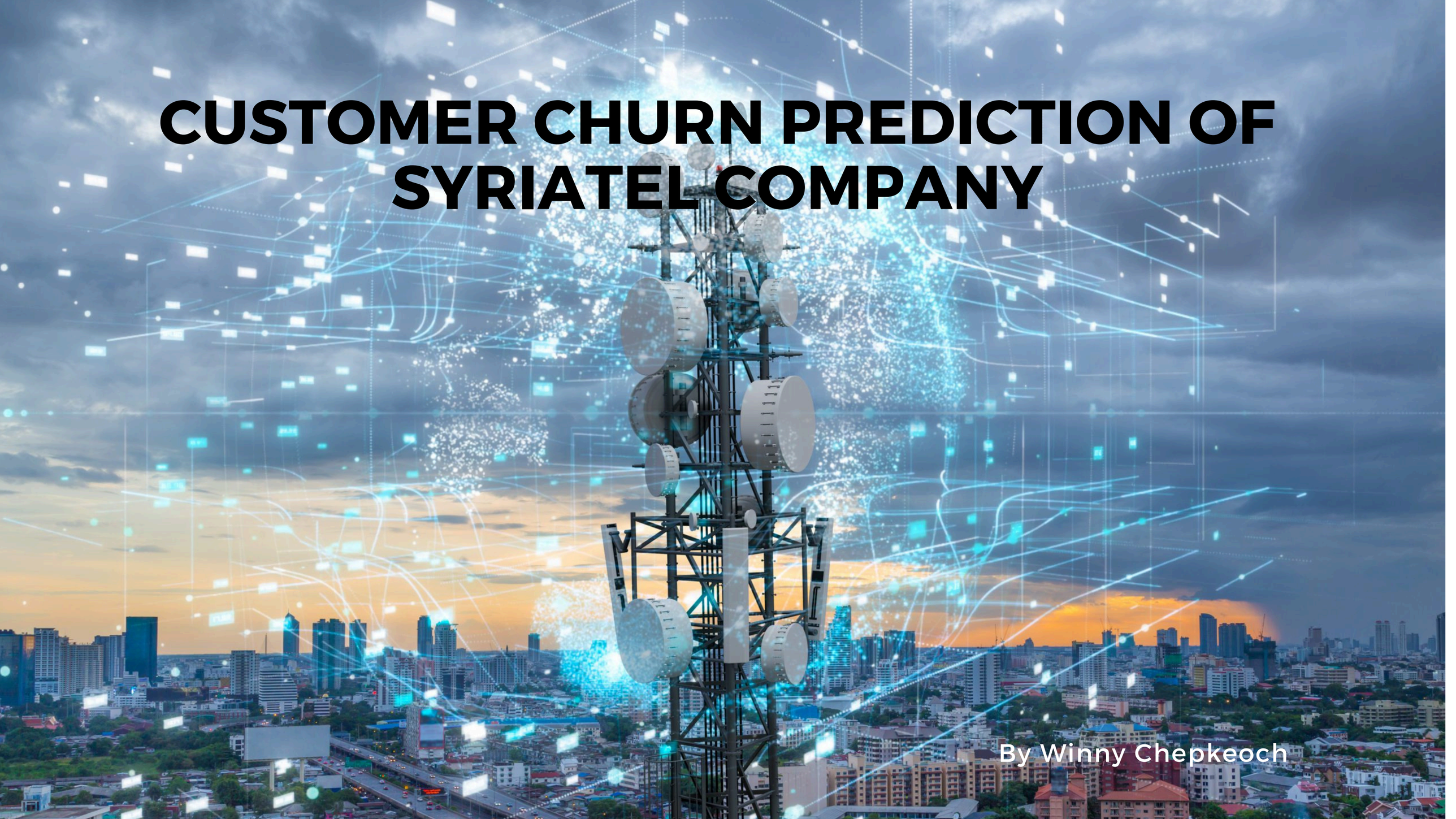


CUSTOMER CHURN PREDICTION OF SYRIATEL COMPANY

By Winny Chepkeoch

LIST OF CONTENTS

OVERVIEW

- Predictive Model Development: SyriaTel is creating machine learning models to predict customer churn by analyzing data like demographics and usage patterns, aiming to identify customers at risk of leaving.
- Feature Analysis & Recommendations: By analyzing the key factors driving churn, SyriaTel will develop targeted strategies to improve customer retention and reduce churn rates.
- Continuous Adaptation & Profitability: The models will be regularly updated to reflect changing customer behaviors, ensuring SyriaTel stays competitive, reduces churn, and enhances profitability.

BUSINESS UNDERSTANDING

1. Predicting Churn: Crucial for SyriaTel to maintain a competitive edge; retaining customers is more cost-effective than acquiring new ones.
2. Predictive Analytics: Analyzes customer data (demographics, usage, call details) to identify churn indicators and develop accurate prediction models.
3. Continuous Refinement: Regularly updates models to adapt to changing customer behaviors, enabling effective retention strategies and sustaining market position.

OBJECTIVES

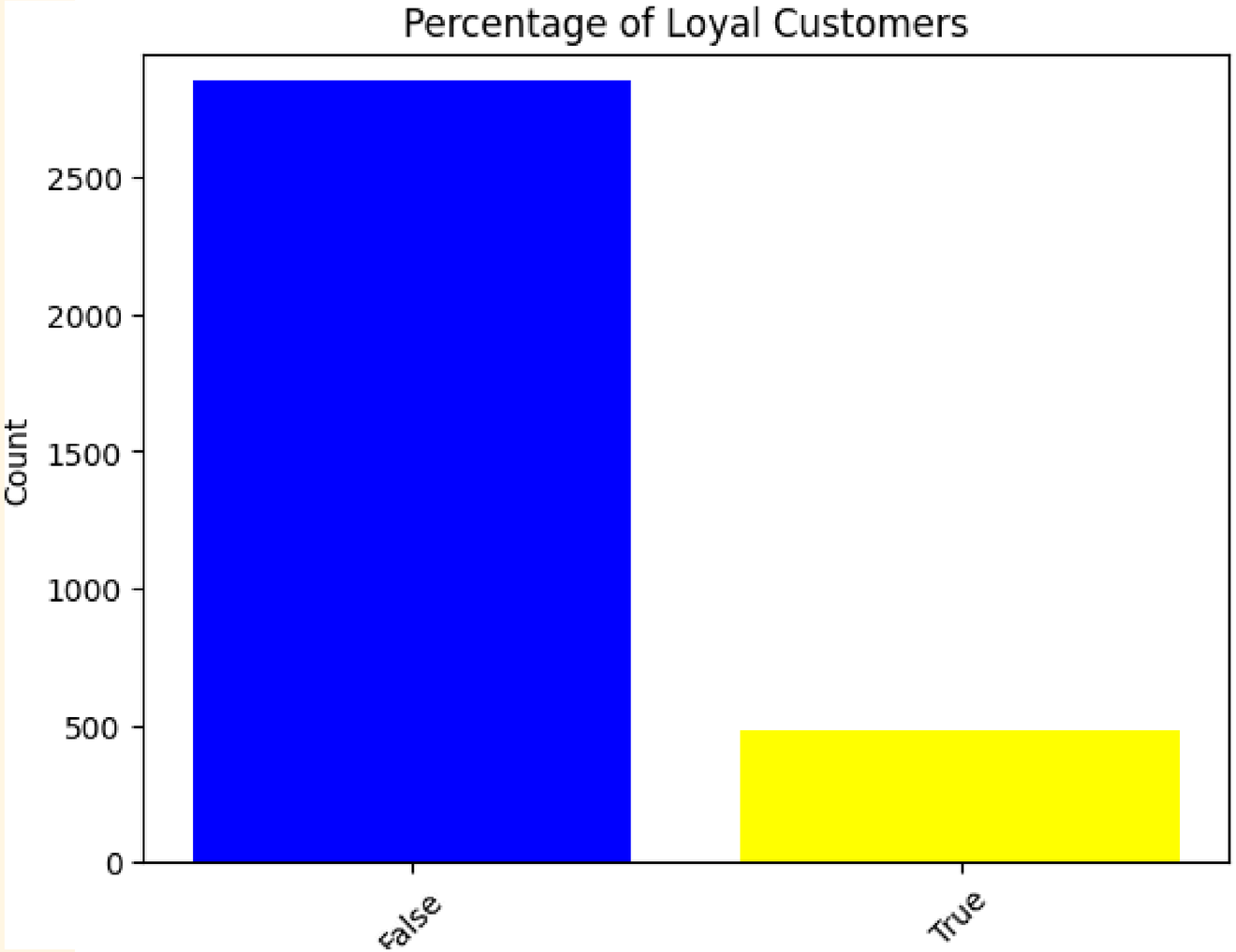
- Develop Models: Build machine learning models to predict customer churn using demographics and usage data.
- Evaluate Accuracy: Compare models to find the most accurate one for predicting churn.
- Analyze Features: Identify key features that impact churn to understand retention drivers.
- Recommend Actions: Provide strategies to reduce churn based on analysis.
- Optimize Strategies: Refine retention strategies using predictive insights and adapt to customer behavior changes.

DATA UNDERSTANDING

- Data Source: Utilized SyriaTel’s customer churn dataset, providing valuable insights into customer behavior and churn patterns.
- Data Information: The dataset includes customer features such as demographics, usage patterns, and service details, aiding in a comprehensive analysis of churn factors.
- Data Structure: The dataset comprises 5 rows and 21 columns, offering a solid foundation for data visualization and exploratory data analysis.

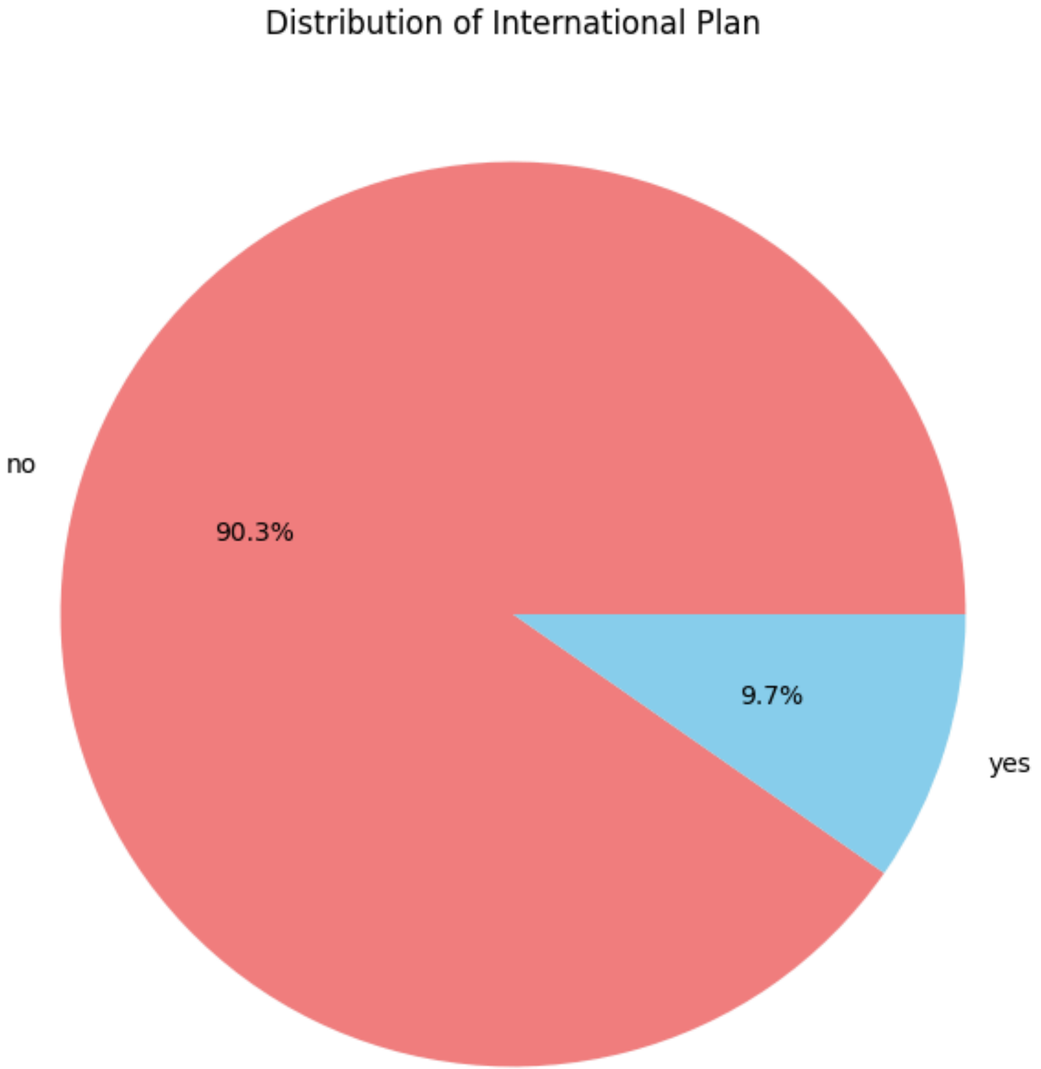
DATA EXPLORATION: PERCENTAGE OF LOYAL CUSTOMERS

- .From above visualization we can see that higher percentage of customers in the dataset did not churn, indicating good customer retention.
- False churn has the highest count while true churn has the least count



DATA EXPLORATION: DISTRIBUTION ON INTERNATIN PLAN

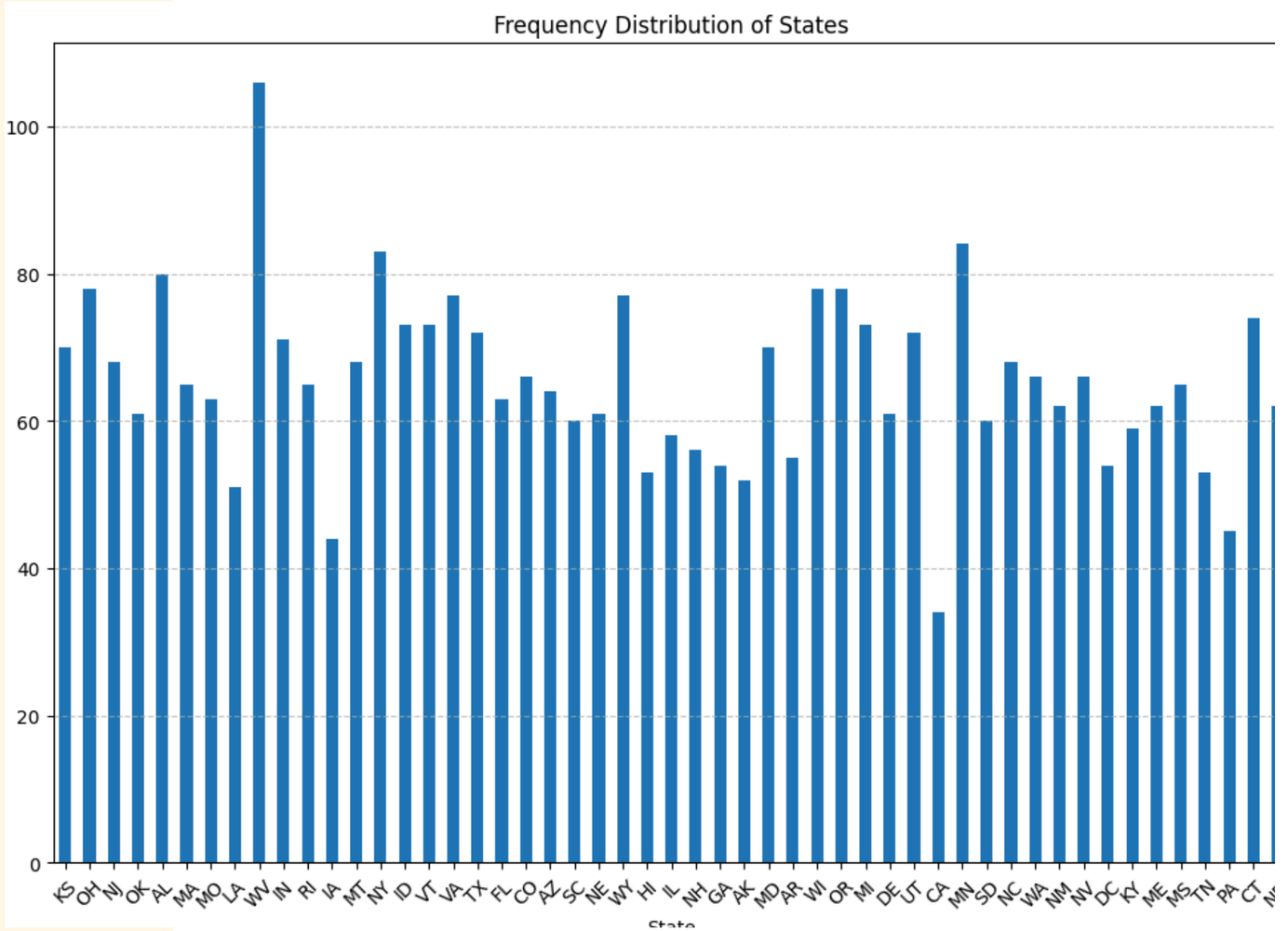
- .The pie chart shows that majority of customers do not have an international plan, with a significantly larger portion of the chart which is `90.3%` representing those without it. Therefore indicating that the international plan is less popular among the customer base in syriatel Company.



DATA EXPLORATION: FREQUENCY OF EACH STATE

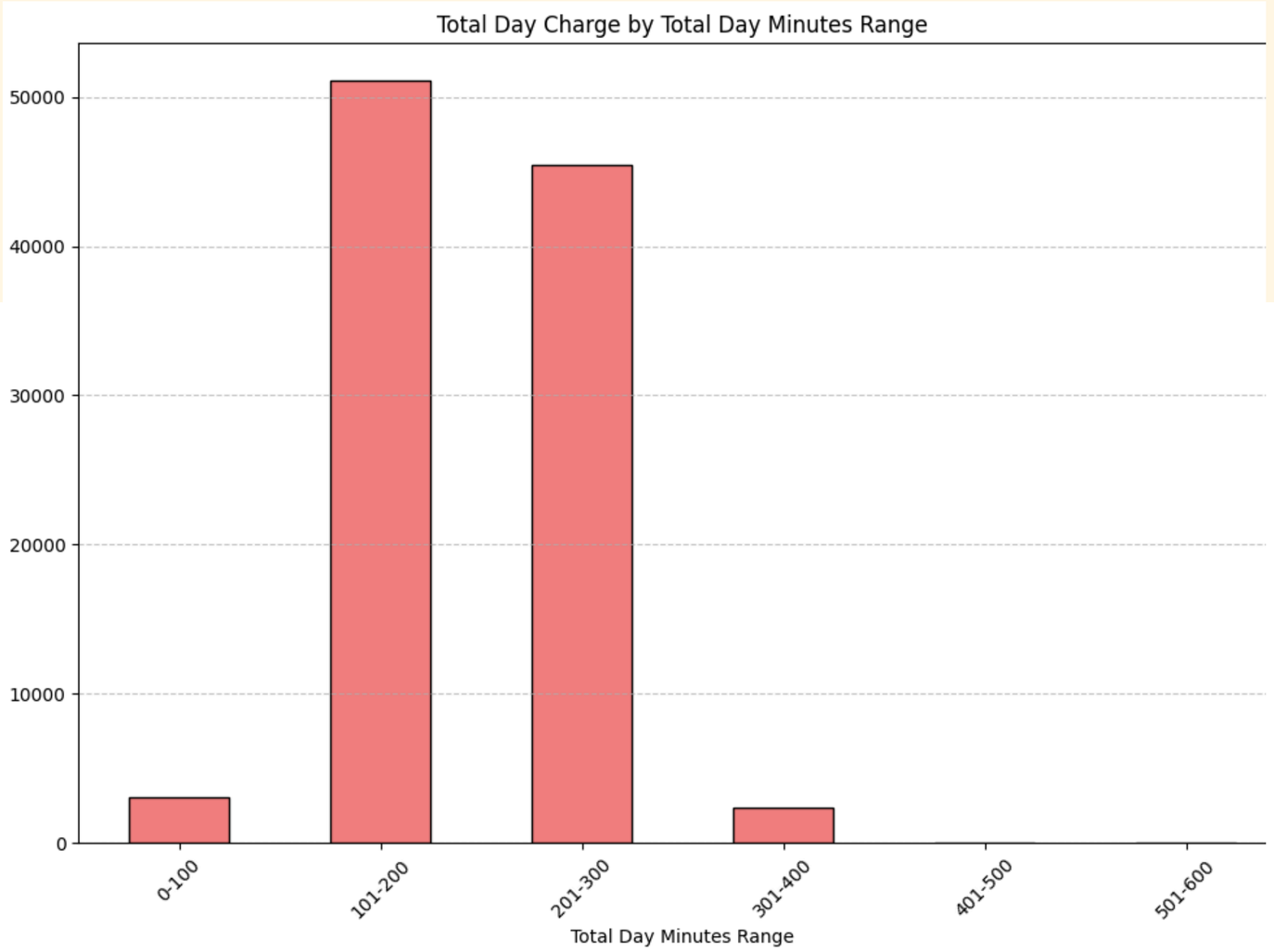
- `WV` is the highest state with the highest frequency while `CA` has the smallest frequency.
- `WV` which is the state of Western Virginia.

- `CA` which is the state of carlifonia.



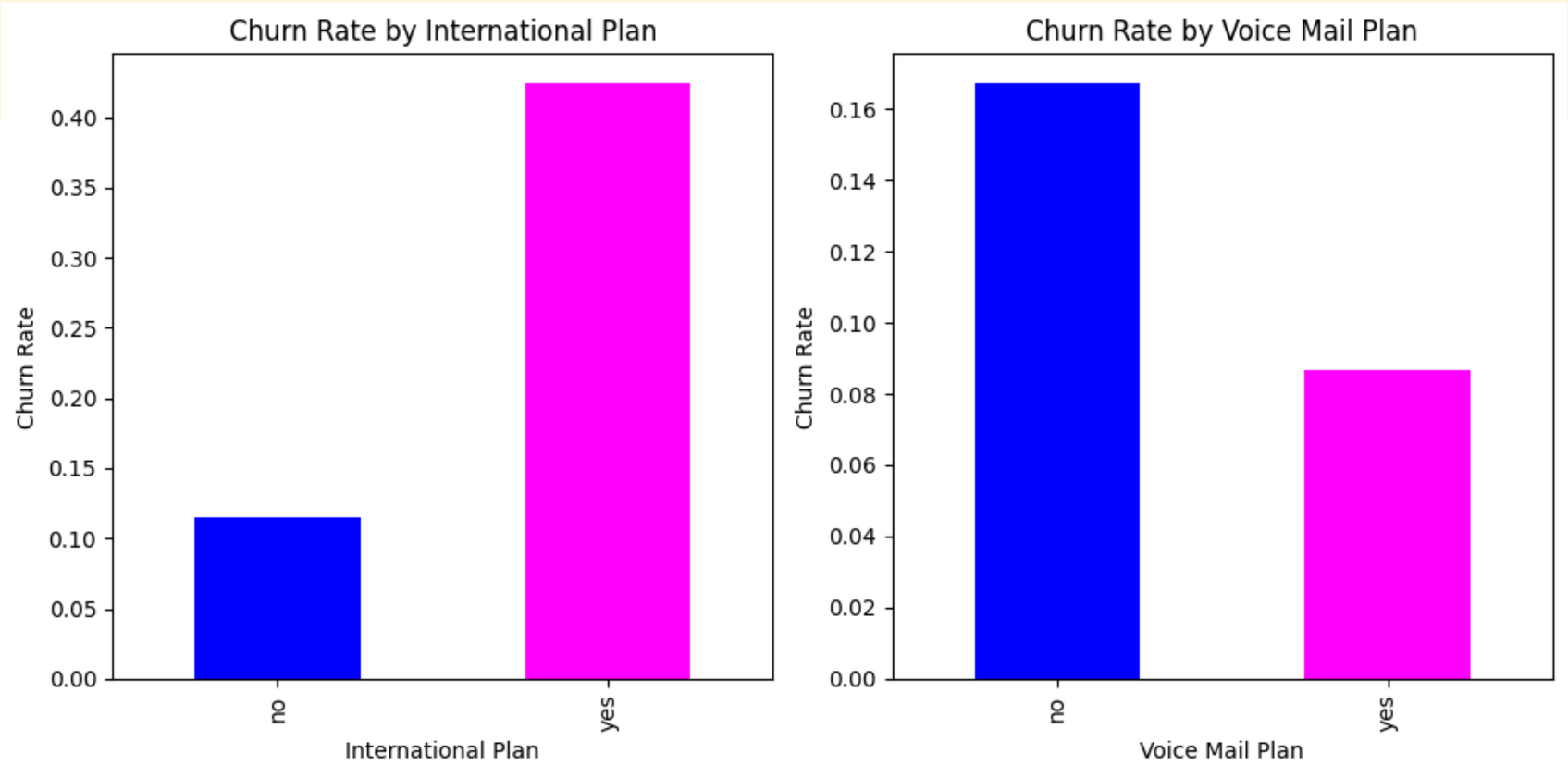
DATA EXPLORATION: TOTAL DAY CHARGE BY TOTAL DAY MINUTES RANGE

- `Highest Charges:` The `101-200` minute range has the highest total day charges, indicating it's the most common usage category among customers.
- `Lower Charges:` The `0-100` minute range shows the lowest total day charges, and charges for usage above `401` minutes also decrease, possibly due to pricing strategies or caps.



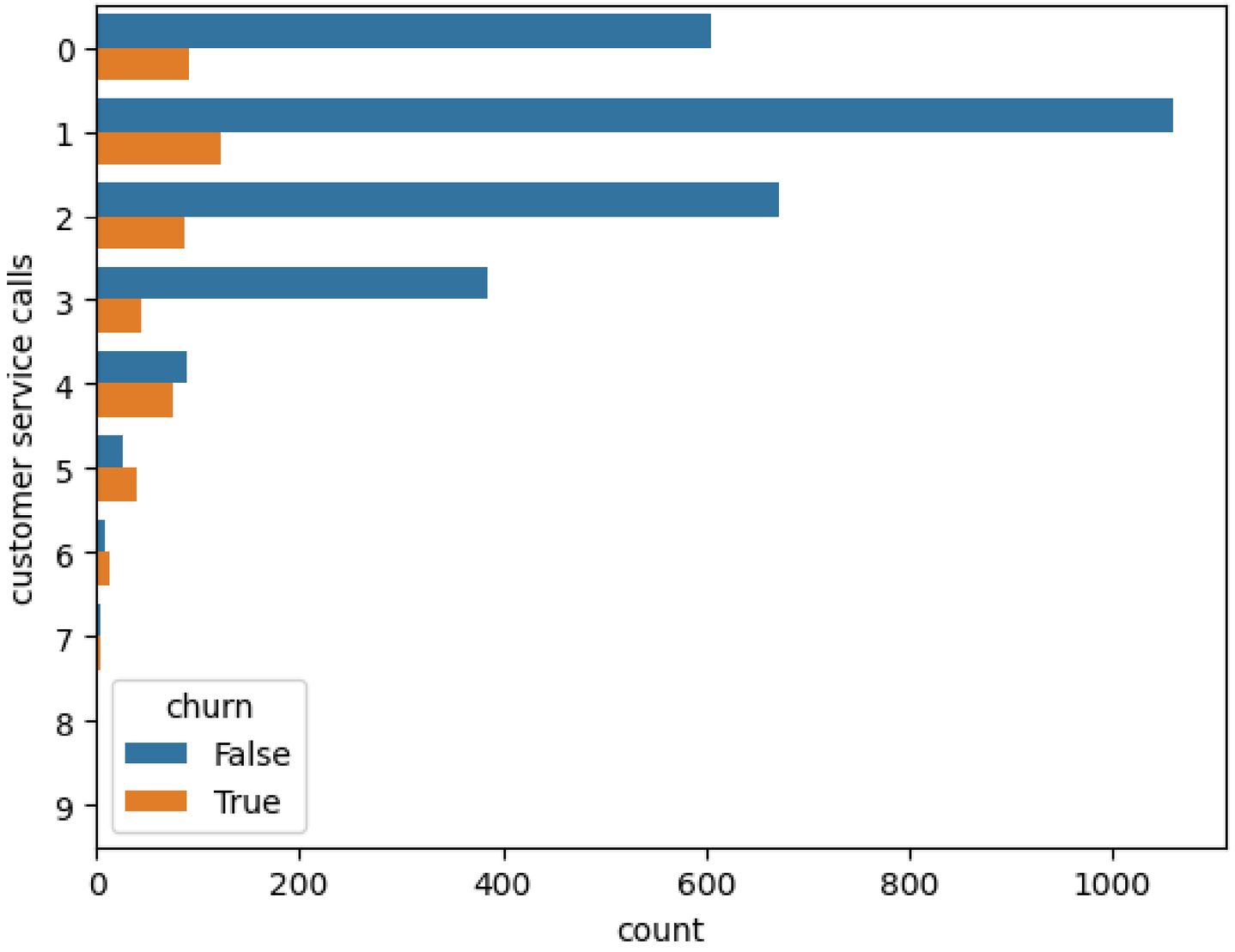
DATA EXPLORATION: INTERNATIONAL CALLS MADE BY PEOPLE

- `Voice Mail Plan Subscribers` has lower churn rate indicates higher customer loyalty.
- `International Plan Subscribers` has Hhigher churn rate suggests dissatisfaction and a higher likelihood of switching.
- `Customer Preference `: Voice Mail plan is preferred over the International Plan based on lower churn and higher satisfaction.

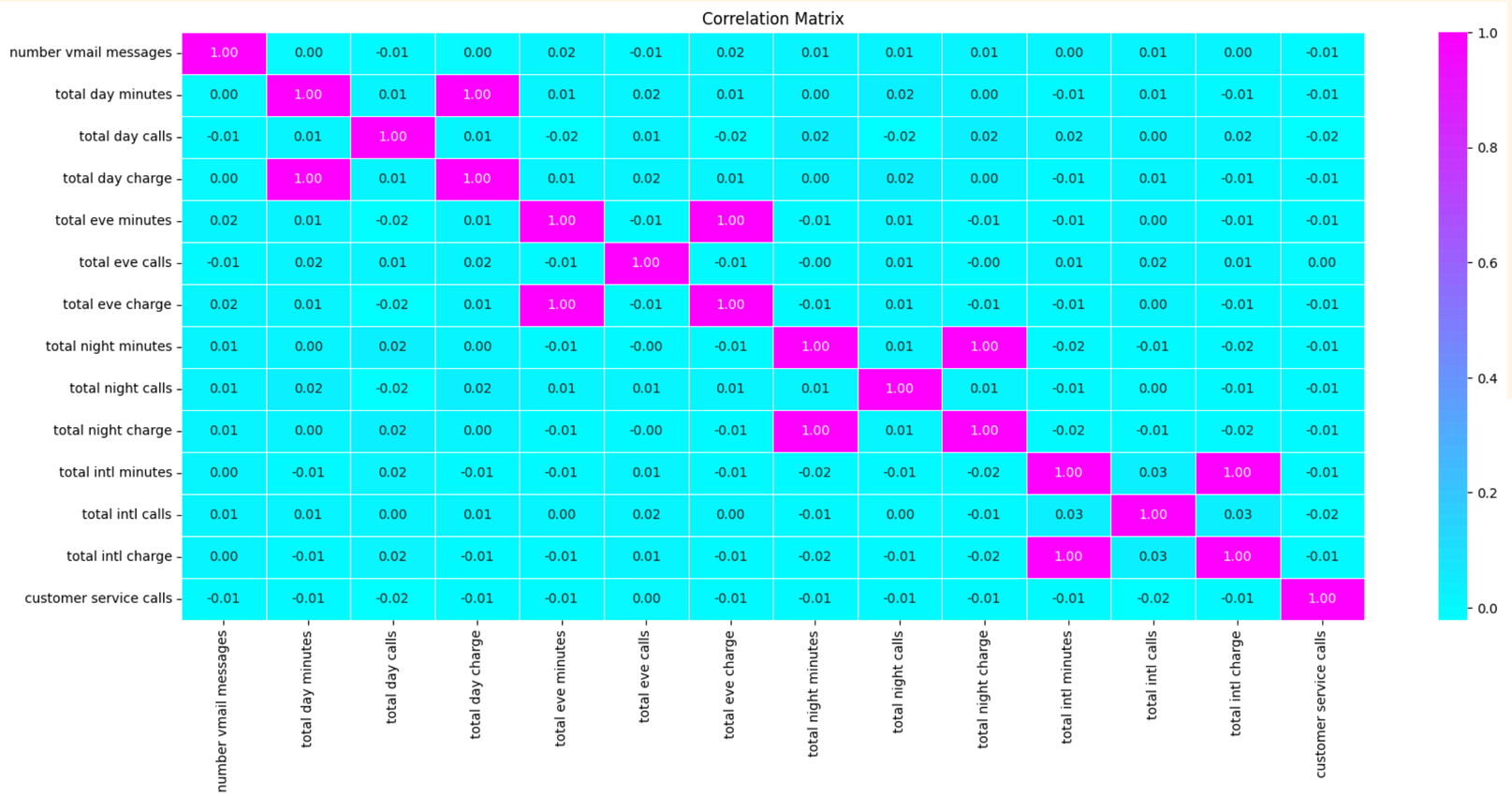


DATA EXPLORATION: RELATINSHIP BETWEEN NUMBER OF CALLS AND LOYALTY

- The chart above depicts the relationship between the number of calls made to the call center and loyalty. This shows that the loyalty has a relatioship with calls as syriatel company recieves most calls from loyal customers which ensures that their is low probability of these customers to use other products from different compnaies.



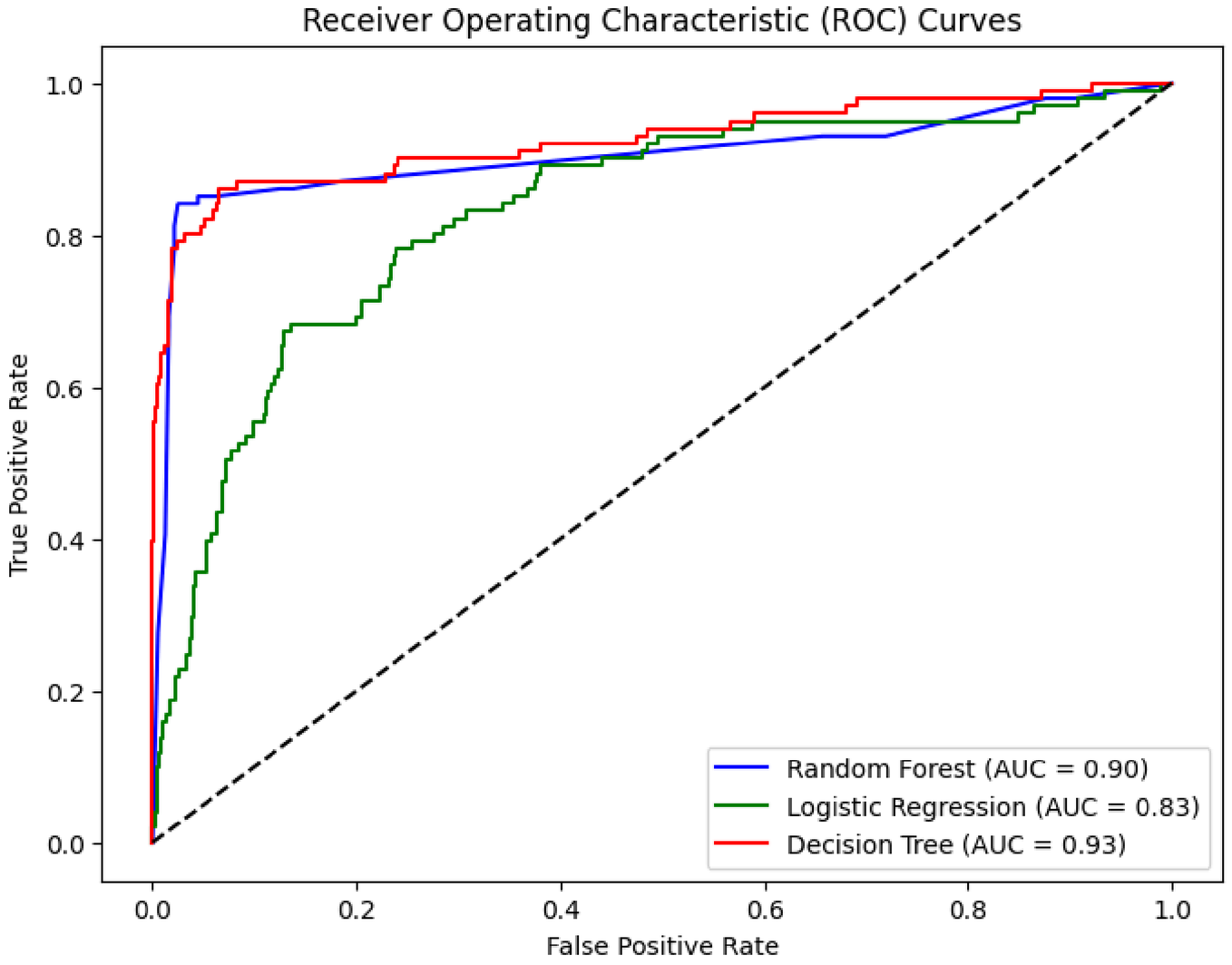
DATA EXPLORATION: CHECKING FOR CORRELATION USING HEATMAP



- As seen from our heatmap there are some variables that have `1.00` which is a perfect correlation with other variables.
- Multicollinearity occurs if this predictor variables are highly correlated, therefore leading to instability of estimates of the model parameters.
- Hence we will remove one of the correlacetd factors which is total day minutes, total eve minutes, /total night minutes and total intl minutes
- since we have `1.00` which indicates a Perfect positive correlation meaning the variables increase together in perfect proportion.
- `0.00`to -`0.29`: Weak negative correlation meaning the variables have a weak tendency to decrease together.

USING ROC CURVE TO CHECK THE BEST MODEL

- Decision Tree (AUC 0.93): Best model, excellent at distinguishing classes with high TPR and low FPR.
- Random Forest (AUC 0.88): Slightly less accurate than Decision Tree with higher false positives and negatives.
- Logistic Regression (AUC 0.77): Lowest accuracy, poorer discrimination with more false positives and negatives.



CONCLUSION

1. Improved Customer Retention: By accurately predicting which customers are most likely to churn, Syriatel can proactively implement targeted retention strategies, reducing churn rates and enhancing customer loyalty.

2. Cost Efficiency: The model will enable Syriatel to allocate resources more effectively, focusing on customers who need the most attention, thereby optimizing marketing and customer service efforts while minimizing unnecessary costs.

3. Data-Driven Decision-Making: The insights provided by the model will empower Syriatel to make more informed, strategic decisions that align with customer needs and business goals, leading to better overall performance.

4. Enhanced Customer Experience: By identifying factors that contribute to churn, Syriatel can improve its services and customer interactions, leading to higher satisfaction and a stronger, more loyal customer base.

RECOMMENDATION

1. Implement Targeted Retention Campaigns: Use the model to identify high-risk customers and develop personalized retention strategies, such as special offers or loyalty programs, to reduce churn and increase customer satisfaction.

2. Optimize Resource Allocation: Focus resources on customers who are most likely to churn, ensuring that retention efforts are cost-effective. This will help in maximizing the return on investment for customer retention initiatives.

3. Enhance Customer Experience: Analyze the factors contributing to churn, such as service issues or pricing concerns, and make improvements to address these areas. By enhancing the overall customer experience, Syriatel can boost loyalty and reduce churn.

4. Leverage Data for Strategic Decision-Making: Use the insights gained from the model to inform broader business strategies, such as product development or market expansion. Data-driven decisions will help Syriatel stay competitive and better meet customer needs in the long term.

NEXT STEP

- **Launch Targeted Campaigns: Develop and execute personalized retention strategies, such as special offers and loyalty programs, for customers identified as high-risk by the Decision Tree model. This will help reduce churn and increase satisfaction.**
- **Optimize Resource Use: Direct resources and efforts towards customers who are most likely to churn, ensuring that retention initiatives are cost-effective and provide the best return on investment.**
- **Enhance Services: Improve service quality and address factors contributing to churn, such as service issues or pricing concerns, to increase overall customer satisfaction and build stronger loyalty.**

LettrLabs

Thank you