PHASE 3 PROJECT

CUSTOMER CHURN PREDICTION

BY: JULIET WANJA

OVERVIEW

What is customer churn?

- customer churn is the percentage of customers who stop using a company's products or services during a certain time period

- It's a common concern in industries such as telecommunications, subscription services, retail, and more

CONTENT

Business understanding

Data preparation

Modeling

Evaluation

conclusion

Recommendation

BUSINESS UNDERSTANDING

BUSINESS PROBLEM

➤ SyriaTel is facing a rising trend in customer churn, which is negatively impacting its revenue and market share.

➤ The business needs to understand the underlying reasons for churn and develop effective strategies
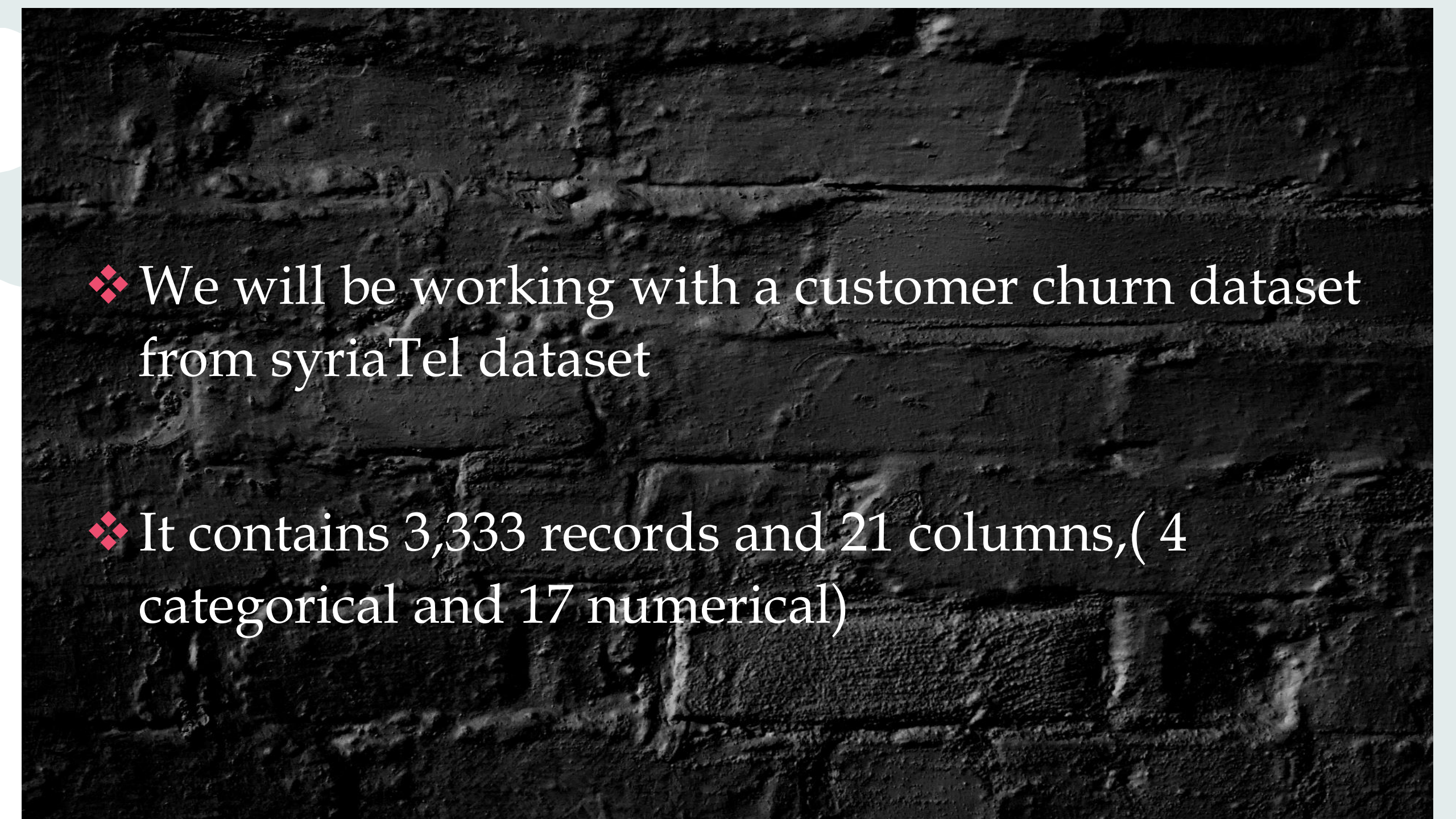
OBJECTIVES

- ❖ **Distinguish the factors that are leading and are more likely to lead to customer churn**

- ❖ **Develop an Accurate Predictive Model.**

- ❖ **Reduce Customer Churn for Syriatel**

- ❖ **Distinguish the factors that are leading and are more likely to lead to customer churn**
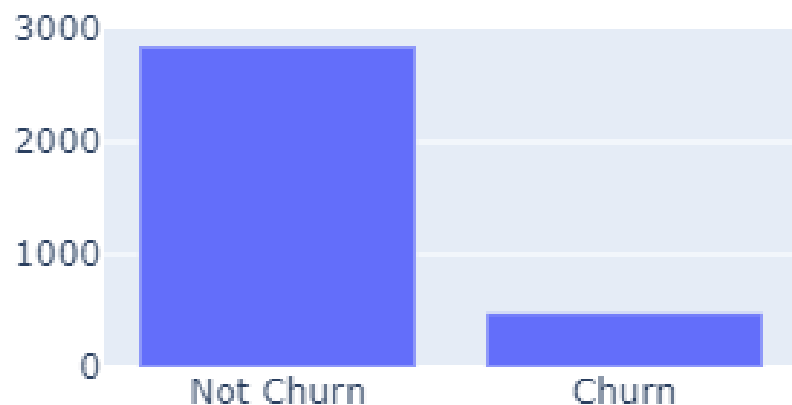
DATA UNDERSTANDING

❖ We will be working with a customer churn dataset from syriaTel dataset

❖ It contains 3,333 records and 21 columns,(4 categorical and 17 numerical)

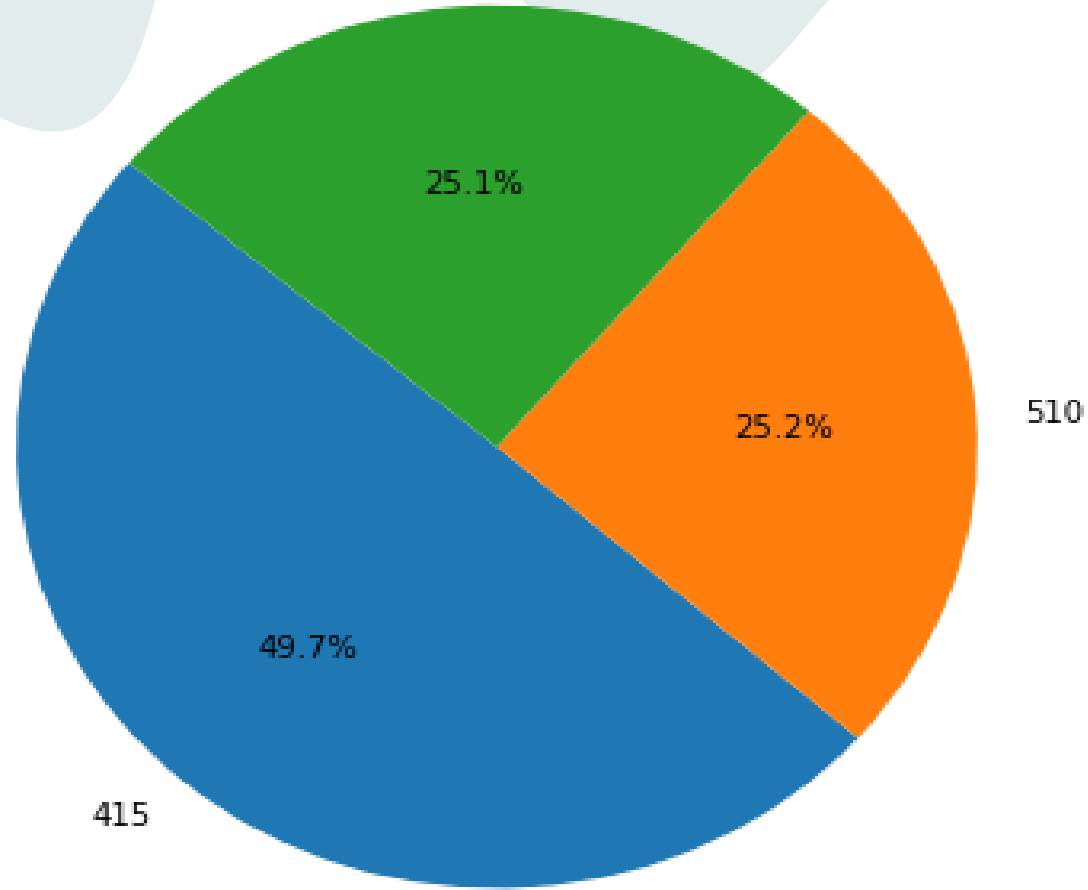# ANALYSIS RESULTS …

HOW MANY CUSTOMERS HAVE CHURNED?

Churn Distribution

- 483 customers that have terminated their contract
- Churn percentage 14.49%
- Those who did not churn 85.51%.

AREA CODE DATA DISTRIBUTION

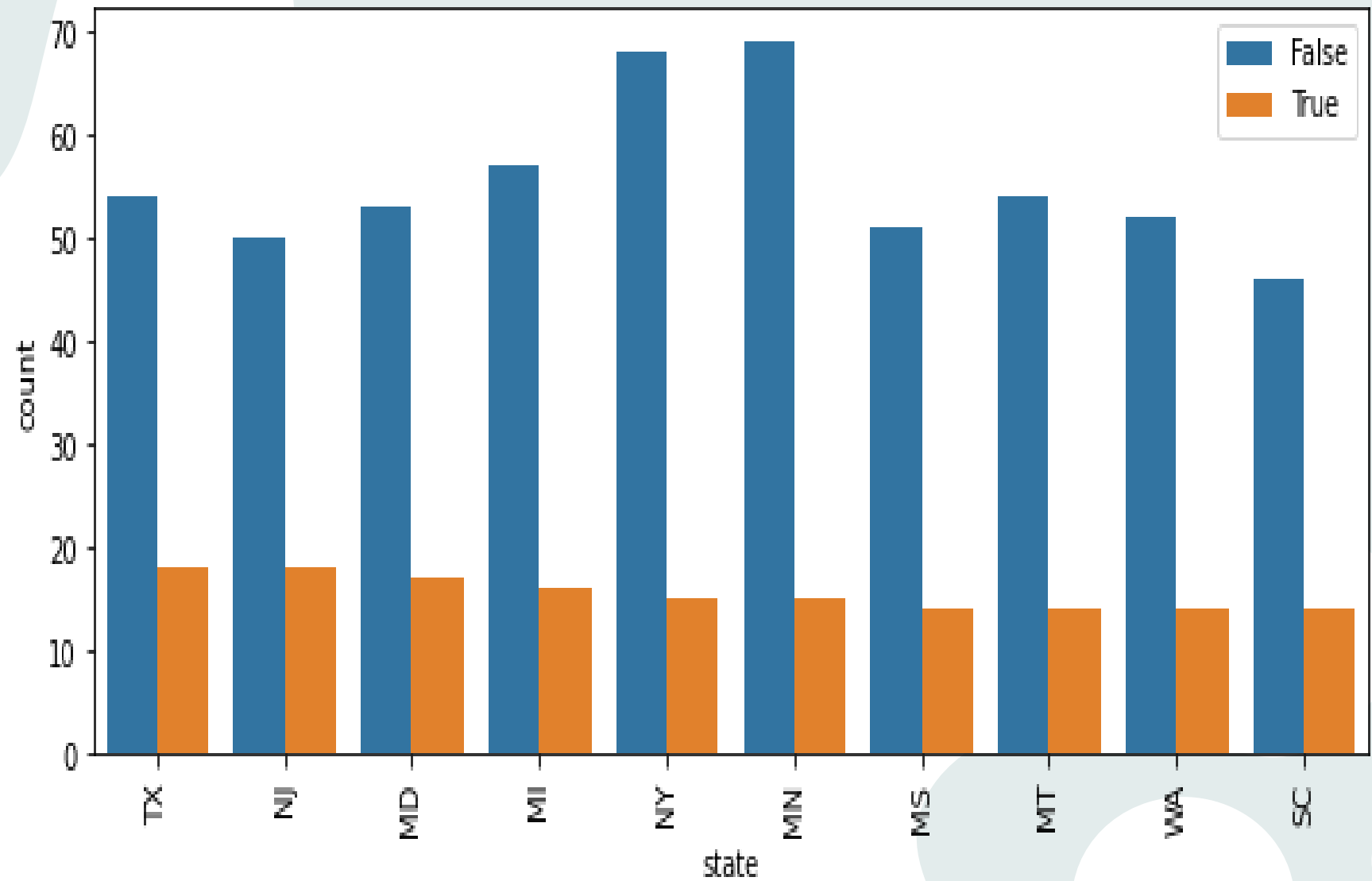Distribution of Area Code Feature

408
25.1%
25.2%
510
49.7%
415

o Approximately half (50%) of the customers are in area code 415.

o About one-fourth (25%) of the customers are in area code 510.

o Another one-fourth (25%) of the customers are in area code 408.
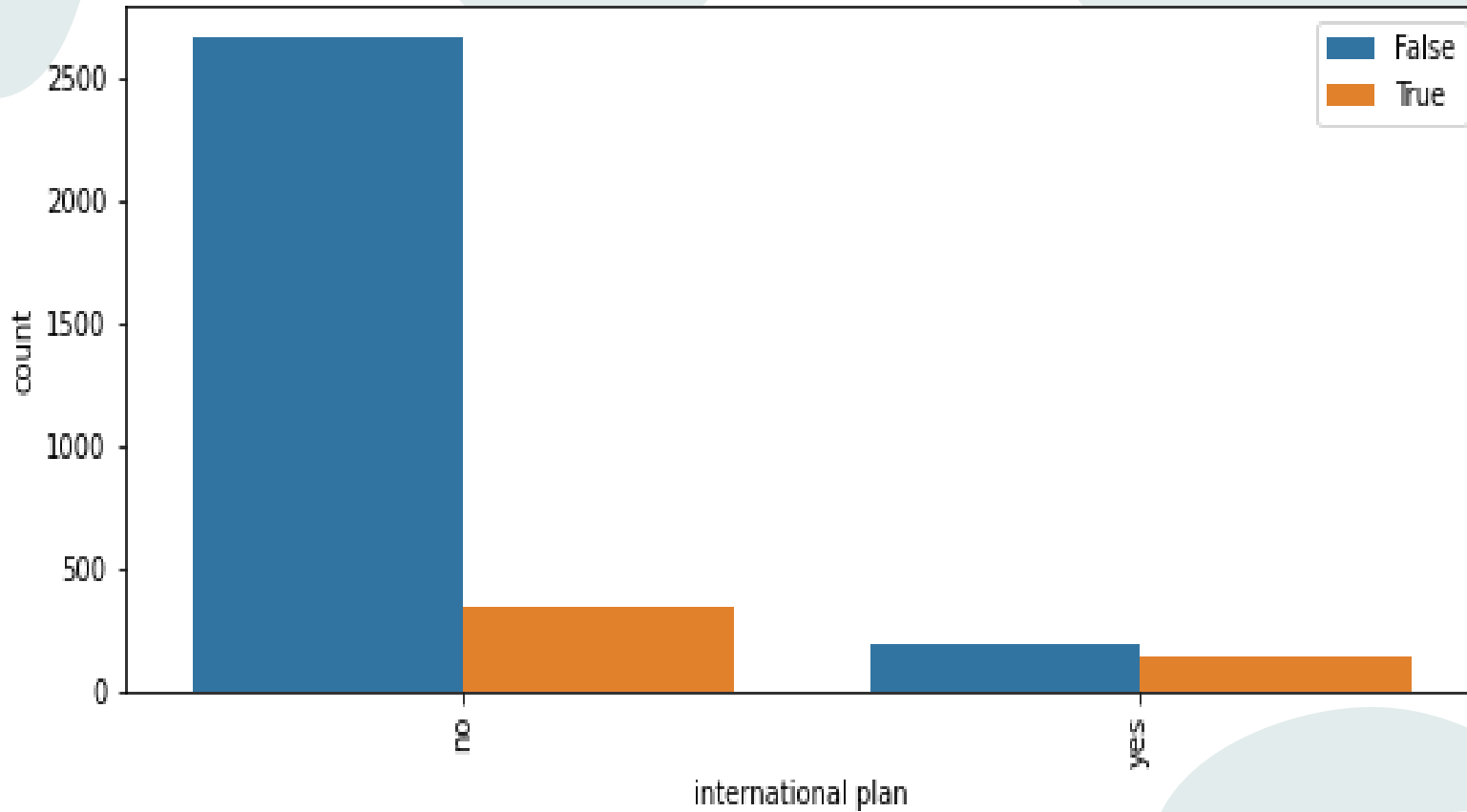
STATES WITH THE HIGHEST CHURN RATE.

Distribution of state based on Churn

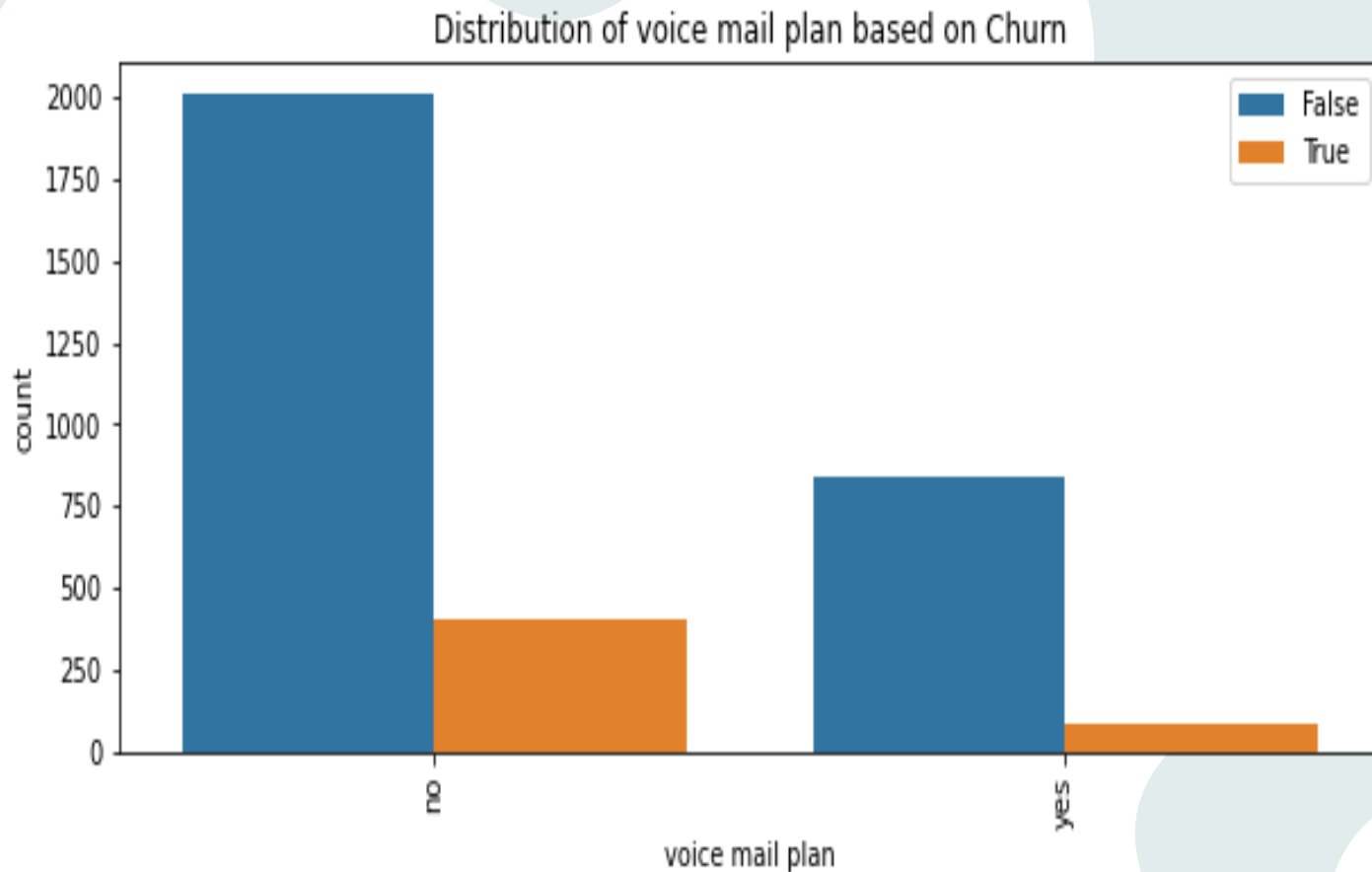❑ Texas and New Jersey are taking the lead in highest churn rate

International plan impact on churn

Distribution of international plan based on Churn

❑ From the customer who churned most did not have international plan

Voice-mail effect to churn.

Distribution of voice mail plan based on Churn

❑Most customers from the 483 customers most did not have a voice mail plan

MODELING

The choice of modeling
techniques and
algorithms used

* Logistic Regression

* Decision Tree

* Random Forest

* K-Nearest Neighbors
(K-NN)
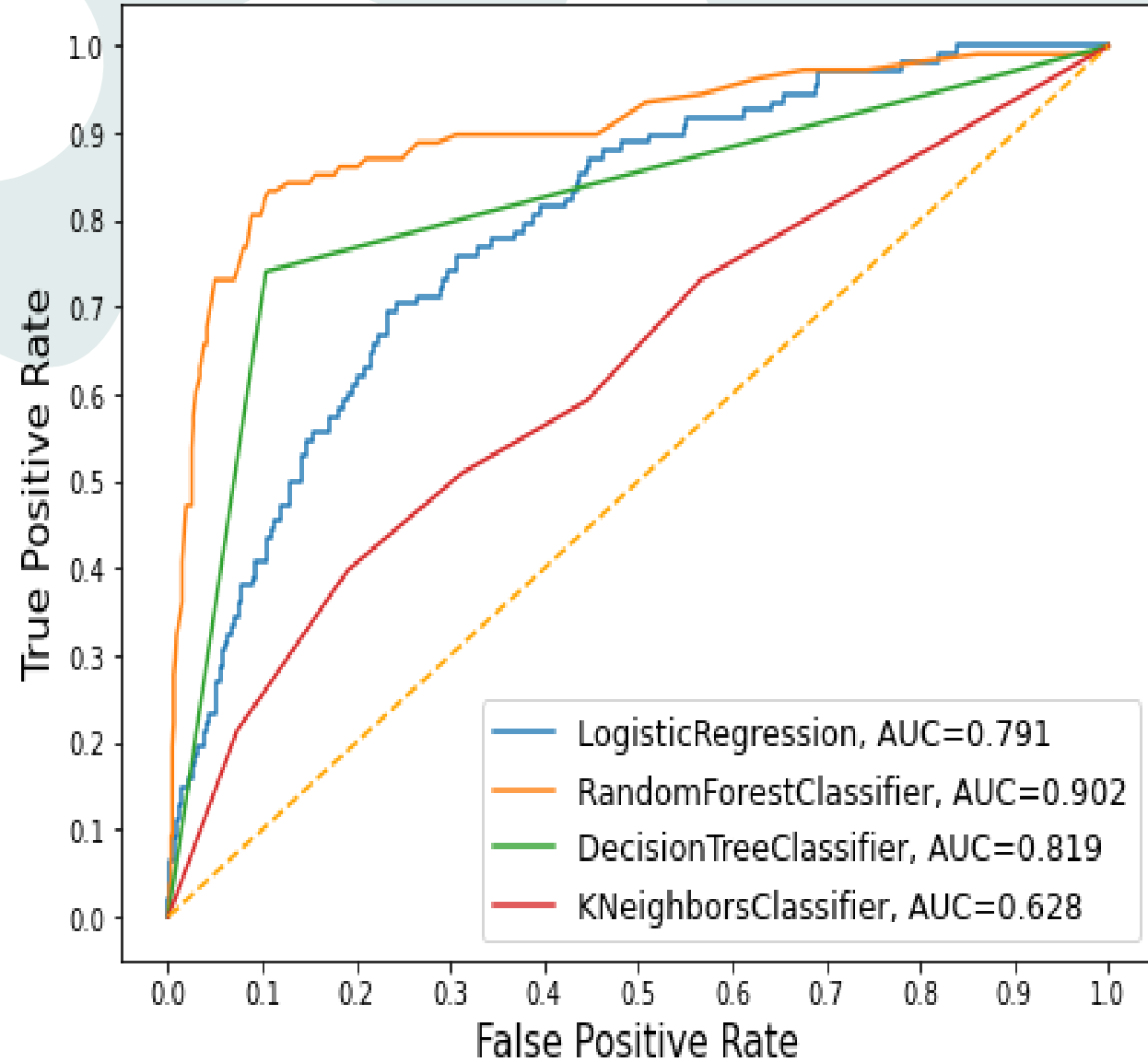
ROC Curve Analysis

BASED ON RECALL

- Logistic Regression   0.703704
- Random Forest Classifier   0.731481
- Decision Tree Classifier   0.740741
- K Neighbors Classifier   0.509259

BASED ON ROC CURVE ANALYSIS

- Logistic Regression   0.792
- Random Forest Classifier   0.902
- Decision Tree Classifier   0.819
- K Neighbors Classifier   0.628

MODEL TUNING RESULTS

TUNED RANDOM FOREST MODEL

Random forest model was doing better than the others and after tuning below was the analysis

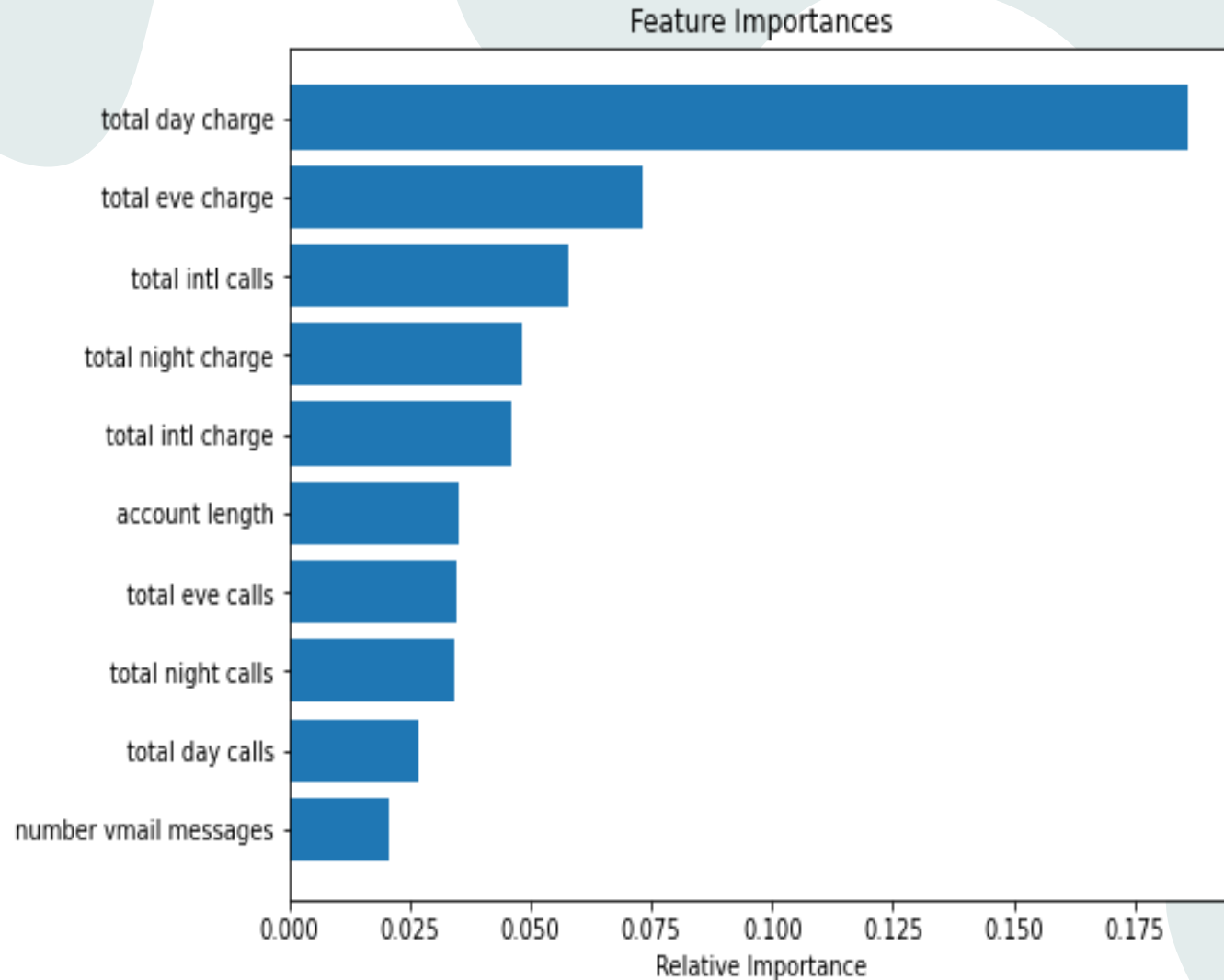|	precision	recall	f1-score	support
0.0	0.96	0.94	0.95	685
1.0	0.68	0.77	0.72	108
accuracy			0.92	793
macro avg	0.82	0.86	0.84	793
weighted avg	0.92	0.92	0.92	793

RANDOM FOREST MODEL
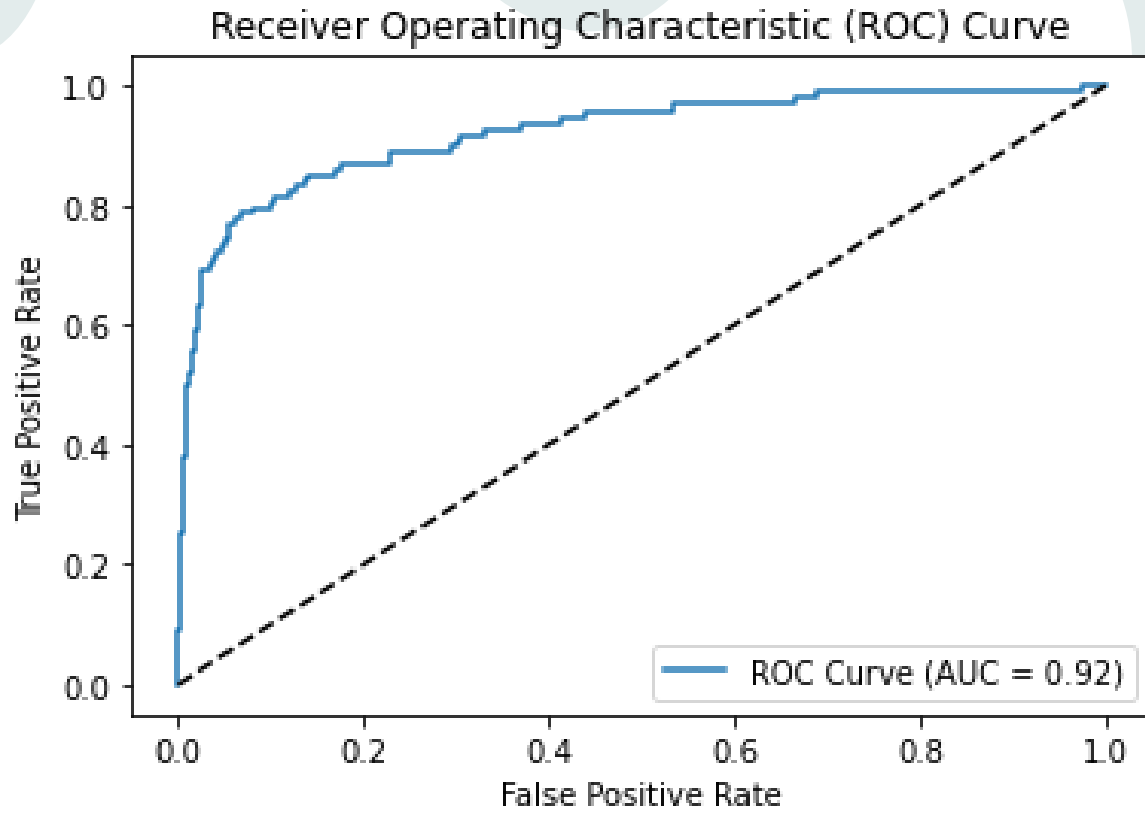
- Recall score = 0.77(77%)

ROC_AUC score = 0.92(92%)

Which features impacted churn according to our model?

Feature Importances

- According to the model, total day charge, total eve charge, total int call are the top three most important features

- **Total Day Charge**:
 represents the total charges incurred by a customer for daytime usage of the telecom services.
- **Total Evening Charge**:
 represents the total charges incurred by a customer for evening (evening hours) usage of telecom services.
- **Total International Calls**:
 represents the total number of international calls made by a customer.

CONCLUSION

- Our Random Forest classifier achieved a recall score of 77%, indicating a good predictive model for customer churn.

- In summary, we successfully achieved our objective of predicting customer churn with an acceptable level of recall
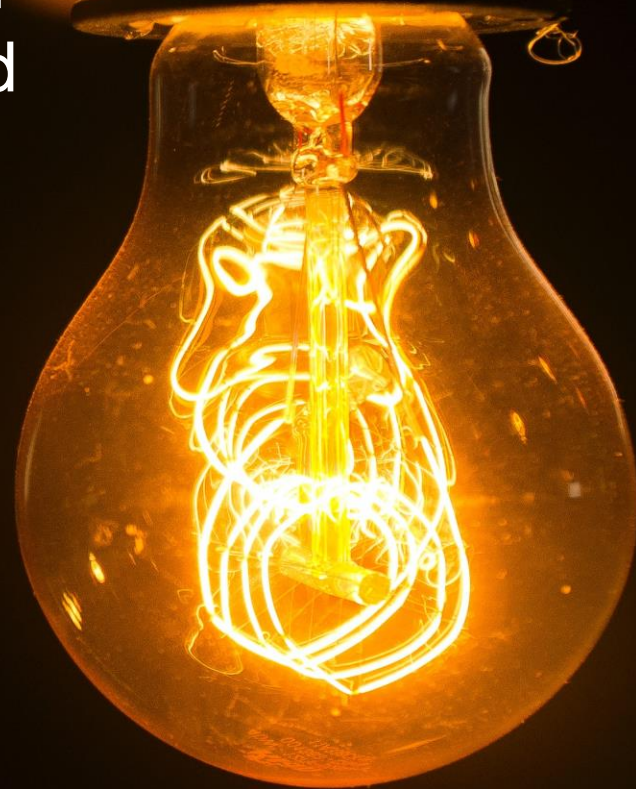
RECOMMENDATION

Recommendations:

1. Pricing Structure Evaluation:
Evaluate the pricing structure for day, evening, night, and international charges. Consider adjusting pricing plans or introducing discounted packages to address the pricing concerns associated with customers who are more likely to churn.

2. Improve Customer Service:
Enhance the quality of customer service and reduce the number of customer service calls. Implement comprehensive training programs for customer service representatives to ensure efficient issue resolution, leading to increased customer satisfaction and reduced churn.

3. Targeted Incentives:
Offer discounts or promotional offers to customers in area codes 415 and 510, which experience a higher churn rate. These incentives can motivate customers to stay with the company.

4. Voicemail Plan Enhancement:
Enhance the value proposition of the voicemail plan to increase adoption among customers. Highlight the benefits and convenience of voicemail services and consider offering additional features or discounts to encourage customer sign-ups.tates.

5. State-Specific Strategies:
Focus on customer retention strategies in states with higher churn rates, such as Texas, New Jersey, Maryland, Miami, and New York. Tailor marketing campaigns, offers, and support to meet the specific needs and preferences of customers in these states

Thank You!

For more information reach out
: wanjajuliet2007@gmamil.com