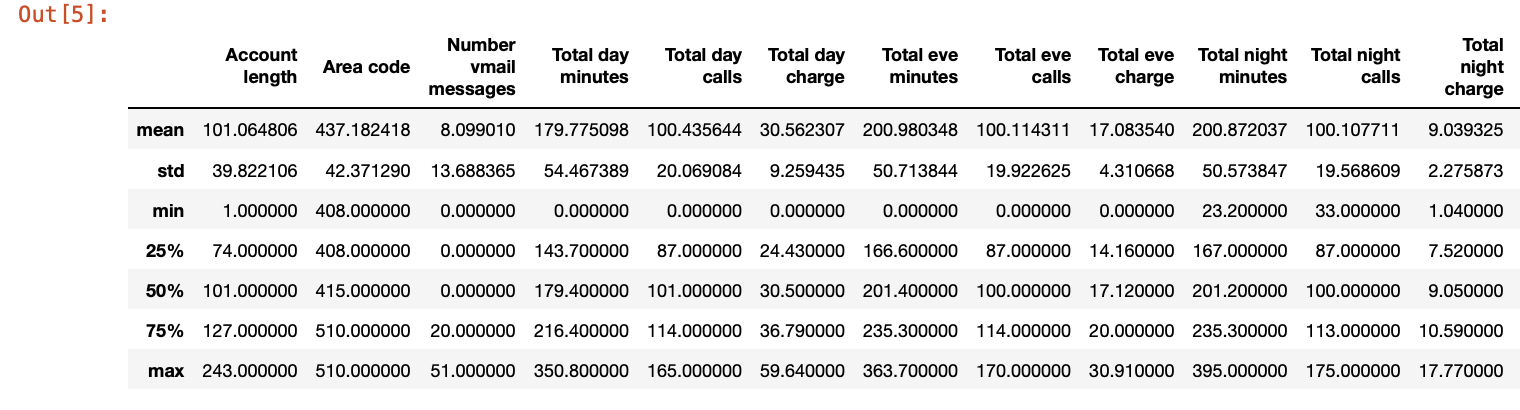
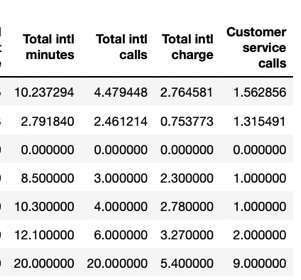
Telecom Market Analysis

The following information was achieved that would prove to be valuable to the boss and the clients.

### Descriptive Statistics

Descriptive statistics and visualizations are powerful tools for communicating complex data and enabling informed decision-making. They provide concise summaries and intuitive representations of data, highlighting key insights and trends. By comparing variables and identifying patterns, these techniques support evidence-based decisions and enhance data accessibility.



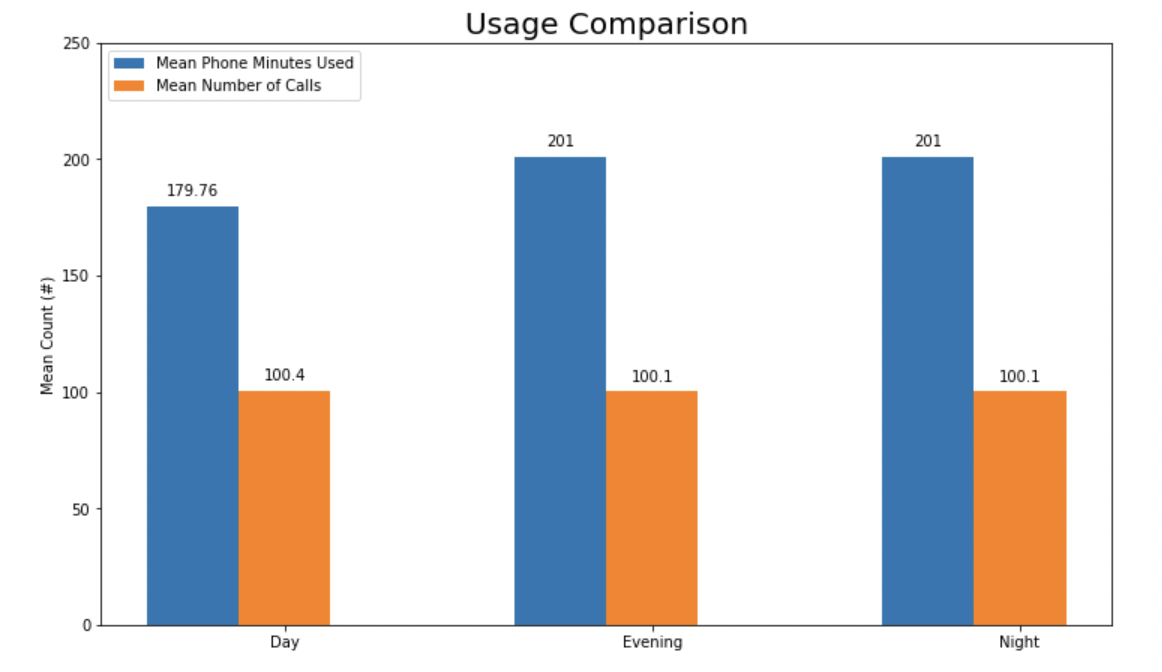


Key insights from the observed patterns include:

* Phone usage is highest during the evening and night.
* The average number of calls is consistent throughout the day, night, and evening.
* Charges vary based on the time of day, with higher rates during the day.
* International calls have lower usage and minutes compared to their charges.
* Customer service calls are relatively low, indicating successful operations.
* Call volume is not strongly correlated with charges.
* Daytime phone usage is lower despite higher charges.
* There is potential to promote and increase international call usage.

### Usage Patterns and Trends

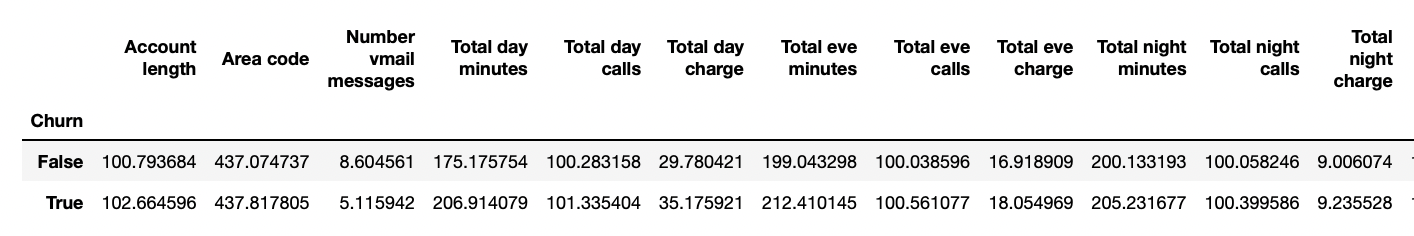
Analyzing usage patterns in the telecom market provides insights for pricing strategies, service plans, and infrastructure investments. Usage data helps determine appropriate pricing structures, offer competitive rates, and introduce bundled plans. Aligning service plans with customer usage patterns enhances satisfaction and attracts a larger customer base. Furthermore, analyzing usage patterns guides infrastructure investments, optimizing network coverage and capacity. Data-driven insights enable telecom companies to optimize revenue generation and enhance the customer experience.

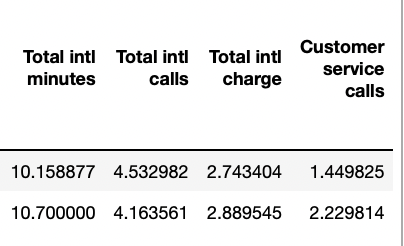


Analyzing usage patterns in the telecom market guides pricing decisions, service plan customization, and infrastructure investments. It helps set competitive rates, offer bundled plans, and introduce pricing incentives based on specific usage patterns. Tailoring service plans to meet customer needs enhances customer attraction and satisfaction. Furthermore, analyzing usage patterns identifies high-demand time periods, informing infrastructure investments to improve network coverage and capacity.

### Churn Rate Analysis

Analyzing churn rate data provides insights into customer retention strategies' effectiveness. A churn rate of 14.5% suggests successful retention efforts, prompting further analysis of successful strategies. Customer segmentation based on attributes enables targeted retention efforts tailored to specific needs and preferences. Thus, on grouping data among the customers who churned and who did not, following figure was achieved:

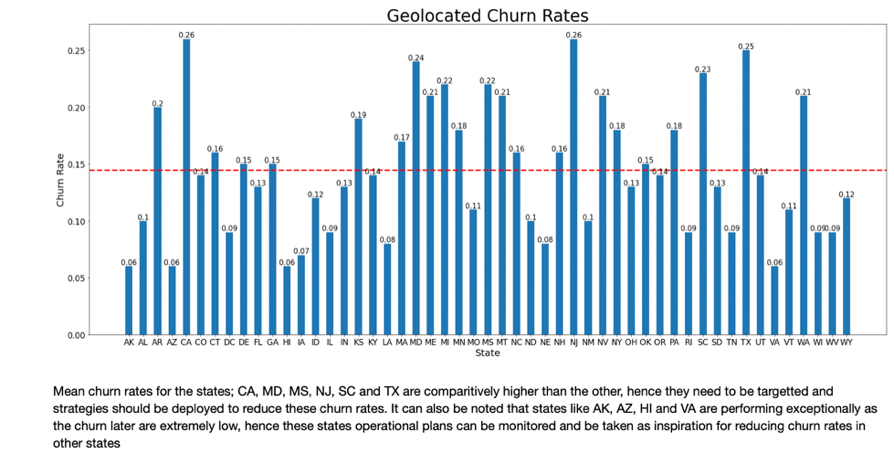


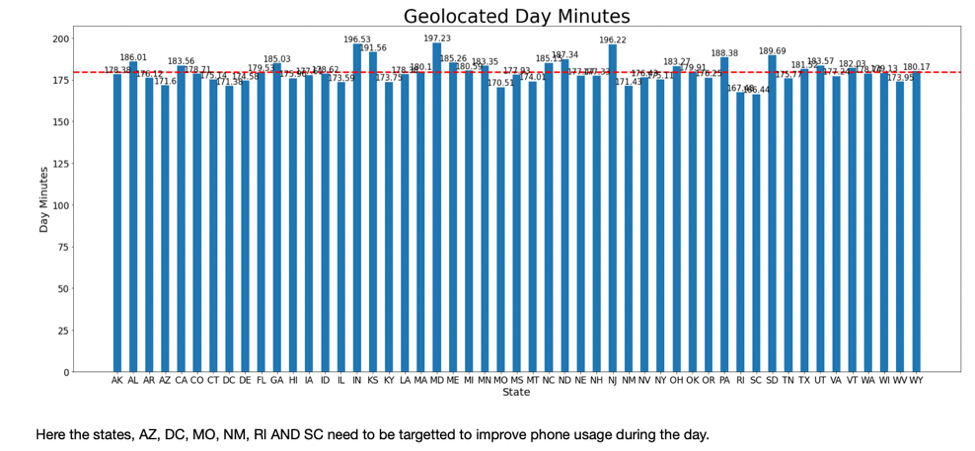


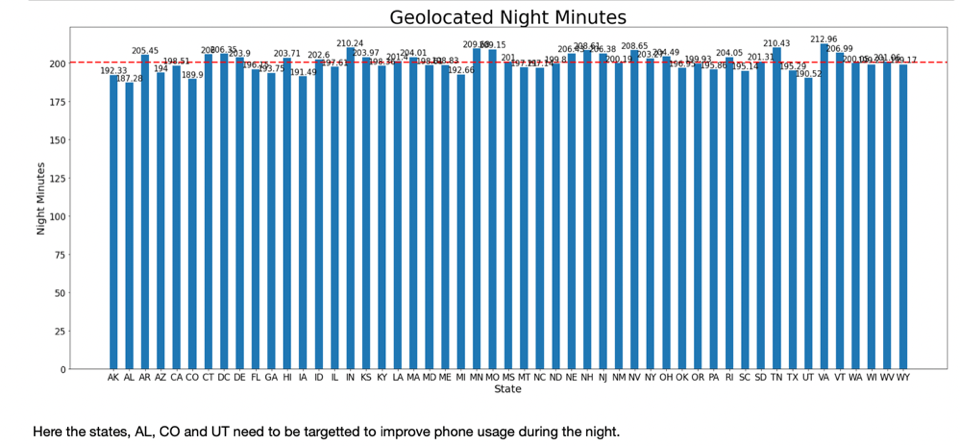
This information reveals that the customers who churn do not use voice mail messages as much and also have a higher number of mean customer service calls which probably suggests that these calls are made due to complains. Given these insights, voice mail messages need to be promoted and customer service needs to be improved to provide efficient and quick solutions and help to customers to further lessen the churn rate.

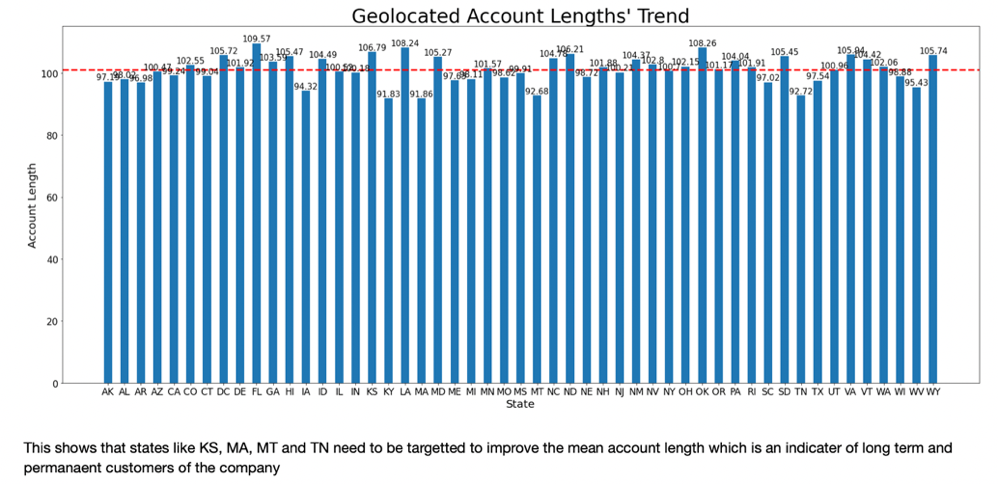
### Comparative Analysis Based on Geolocation

Comparative analysis between telecom market regions provides insights into customer behavior, allowing for segmentation, target market identification, and expansion planning. It helps understand unique customer characteristics, guide marketing efforts, and prioritize infrastructure investments. The analysis also reveals competitive dynamics, aiding in retention strategies. Overall, it facilitates data-driven decision-making for effective market targeting, resource allocation, and service customization. By grouping all features of dataset based on state, the mean for each feature respective to the state was achieved. On further analysis, standard deviation suggested the features that were to be further analyzed to improve churn rates, minutes used and account length. The following results and conclusions were achieved:









## CHURN RATE PREDICTION- SUMMARY

The following models were implemented to predict if the given customer will churn or not. The accuracy for each model is also mentioned.

* Logistic Regression



* Random Forest Classifier



* Support Vector Machine



* Decision Tree



As seen from above accuracy scores, Random Forest Classifier performs the best among all the others. When comparing Random Forest with other models (logistic regression, decision trees, and SVM) for predicting churn in this dataset, Random Forest emerges as the better and more useful choice due to the following reasons:

* Accuracy: Random Forest achieves high accuracy due to its ensemble approach, reducing overfitting and improving predictions.
* Nonlinear Relationships: Random Forest captures complex nonlinear relationships, crucial for predicting churn with interactions and non-linear patterns.
* Robustness: Random Forest is robust to noise and outliers, providing reliable and stable churn predictions.
* Feature Importance: Random Forest offers feature importance measures, identifying key factors for churn and aiding decision-making.
* Handling High-Dimensional Data: Random Forest handles high-dimensional data effectively without overfitting, unlike logistic regression and SVM.
* Generalization and Scalability: Random Forest performs well with different data distributions and sizes, with good generalization and scalability for real-world applications.