**Project Documentation: Customer Churn Analysis in the Telecom Industry**

*Document Version: 1.0*

**Table of Contents:**

1. Introduction
2. Problem Statement
3. Data Description
4. Data Exploration and Preprocessing
5. Descriptive Analysis
6. Key Findings and Insights
7. Visualizations
8. Recommendations
9. Conclusion
10. References

**1. Introduction:**

The purpose of this project is to analyze customer churn rates in the telecom industry and identify potential factors influencing customer attrition. By understanding the underlying patterns and trends, telecom analysts can develop effective strategies to reduce churn and improve customer retention.

**2. Problem Statement:**

The problem statement for this project is as follows:

"As a telecom analyst, I want to analyze customer churn rates to identify potential factors influencing customer attrition. The goal is to gain insights into the telecom dataset and provide recommendations to reduce churn and improve customer retention."

**3. Data Description:**

The dataset used for this project is the Telecom Customer Churn dataset. It contains various telecom industry factors such as revenue, minutes of use, recurring charges, voice calls, data calls, and other customer-related information. The dataset consists of 100,000 records and 100 columns, including the churn status of customers.

**4. Data Exploration and Preprocessing:**

In this phase, we performed the following tasks:

* Examined the structure of the dataset using functions like **head()**, **info()**, and **describe()** to get an initial understanding of the data.
* Checked for missing values in the dataset using functions like **isnull()**, **isna()**, or **info()** and handled them appropriately using methods like **fillna()** or **dropna()**.
* Identified and handled outliers in the dataset using techniques such as box plots, z-scores, or the interquartile range (IQR) method.
* Formatted the data if needed, such as converting categorical variables into the appropriate data types or transforming variables into a more suitable format for analysis.

**5. Descriptive Analysis:**

In this phase, we conducted various descriptive statistics to gain insights into the telecom dataset. We calculated measures such as mean, median, mode, standard deviation, etc., to understand the distribution of variables. Additionally, we explored the relationships between variables using correlation analysis and visualizations such as scatter plots.

**6. Key Findings and Insights:**

Based on our analysis, we identified the following key findings and insights:

* The churn rate in the telecom industry is approximately 49.56%.
* Variables such as monthly revenue, minutes of use, recurring charges, and customer care calls showed significant correlations with customer churn.
* Customers with higher revenue, higher minutes of use, and more customer care calls were more likely to churn.
* There were noticeable differences in customer behavior between churned and non-churned customers, indicating potential factors influencing customer attrition.

**7. Visualizations:**

To present our findings and trends related to customer churn, we implemented various visualizations such as bar plots, line charts, histograms, and heatmaps. These visualizations effectively communicate the relationships between variables and provide insights into customer churn patterns.

**8. Recommendations:**

Based on our analysis and findings, we recommend the following strategies to reduce customer churn and improve customer retention in the telecom industry:

* Implement proactive customer care initiatives to address customer issues and concerns promptly.
* Offer personalized incentives and discounts to customers at risk of churning.
* Analyze the impact of pricing strategies and consider adjusting prices to improve customer satisfaction and loyalty.
* Enhance the quality and reliability of network services to reduce customer dissatisfaction.
* Conduct targeted marketing campaigns to engage and retain high-value customers.

**9. Conclusion:**

In conclusion, this project aimed to analyze customer churn rates in the telecom industry and identify potential factors influencing customer attrition. Through data exploration, descriptive analysis, and visualizations, we gained valuable insights into customer behavior and patterns. The findings provide a foundation for developing effective strategies to reduce churn and improve customer retention in the telecom industry.

**10. References:**

[List any references or data sources used in the project]

This concludes the documentation for the Customer Churn Analysis project in the telecom industry. The documentation provides a comprehensive overview of the project, including the problem statement, data analysis, findings, recommendations, and visualizations.

rev\_Mean mou\_Mean ... eqpdays Customer\_ID

count 99643.000000 99643.000000 ... 99999.000000 1.000000e+05

mean 58.719985 513.559937 ... 391.932309 1.050000e+06

std 46.291677 525.168140 ... 256.482193 2.886766e+04

min -6.167500 0.000000 ... -5.000000 1.000001e+06

25% 33.260000 150.750000 ... 212.000000 1.025001e+06

50% 48.195000 355.500000 ... 342.000000 1.050000e+06

75% 70.750000 703.000000 ... 530.000000 1.075000e+06

max 3843.262500 12206.750000 ... 1823.000000 1.100000e+06

[8 rows x 79 columns]

rev\_Mean 357

mou\_Mean 357

totmrc\_Mean 357

da\_Mean 357

ovrmou\_Mean 357

kid11\_15 1732

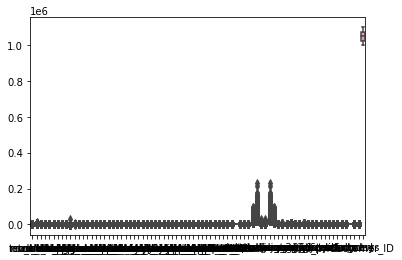
kid16\_17 1732

creditcd 1732

eqpdays 1

Customer\_ID 0

Length: 100, dtype: int64



rev\_Mean mou\_Mean ... eqpdays Customer\_ID

count 9.999500e+04 100000.000000 ... 100000.000000 1.000000e+05

mean -inf 513.559937 ... 391.932309 1.050000e+06

std NaN 524.229868 ... 256.480910 2.886766e+04

min -inf 0.000000 ... -5.000000 1.000001e+06

25% 3.506008e+00 151.500000 ... 212.000000 1.025001e+06

50% 3.879035e+00 357.500000 ... 342.000000 1.050000e+06

75% 4.257455e+00 701.250000 ... 530.000000 1.075000e+06

max 8.254077e+00 12206.750000 ... 1823.000000 1.100000e+06

correlation\_matrix = dataset\_TSAP.corr()

Churn Rate: 49.56%

A black square with white text

Description automatically generated with low confidence

A picture containing text, screenshot, diagram, font

Description automatically generated

A picture containing line, text, diagram, plot

Description automatically generated

A picture containing text, screenshot, diagram, rectangle

Description automatically generated