ISSN Online 3108-026X, ISSN Print 3139-4833
International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 3, Issue 7, July 2026 | DOI 10.17148/IMRJR.2026.030703

Customer Churn Analysis of a Telecom Company

Dr. Emmanuel Udoh[1], Evelyn Ehinofe Iyere[2]

Subhani Department of Computer Information Systems, Texas A&M University Central, Killeen, USA[1]

UdoCenter Integrated Services, 1 Eko Road, Ekpechor Ebu, Delta, Nigeria[2]

Abstract: Customer churn refers to the number of customers who end their relationship with a company within a given period, and it is an important metric for businesses, especially those that rely heavily on subscriptions or recurring revenue streams. Using statistical methods such as binary logistic regression, this paper analysed a publicly available dataset containing information about a telecom provider offering home phone and internet services to 7,043 customers in California in the third quarter. The churn factors were visualized, enabling telecommunications companies to identify customers at significant risk of departing and improve customer loyalty. Month-to-month contracts have the highest churn rate at 42.7%, followed by one-year contracts at 11.3% and two-year agreements at 2.8%. The results demonstrate that customers who sign longer-term contracts exhibit greater loyalty, presenting an opportunity for retention programs to move month-to-month customers into longer-term agreements.

Keywords: Churn, Prediction, Profitability, R-Programming Visualization Tools, Telecommunication

I. INTRODUCTION

Businesses need steady revenue to survive. The subscription model is a premier strategy that helps companies move away from sporadic one-time sales to predictable, recurring income. By fostering long-term relationships and compounding growth, this model has expanded far beyond its foundational roots in software, entertainment, and telecom to reshape industries like healthcare, manufacturing, and consumer goods. This paper focuses on the customer churn phenomenon in the telecommunications industry. Customer churn results in significant revenue losses and costly replacement marketing for telecommunications firms [1]. Since retaining clients is far cheaper than acquiring them, churn management is a top strategic priority. Companies must actively engage both current and potential customers to survive in today's competitive business landscape. However, as market competition intensifies, companies devote more time and effort to maintaining relationships with existing clients than to attracting new customers [9]. Using the Kaggle dataset [10], this project investigates customer attrition by analyzing and predicting key variables like tenure, contract type, monthly charges, and churn scores. Through advanced visualization and statistical modeling, the project uncovers the root causes of customer turnover and provides strategic, practical solutions to boost long-term retention.

While telco customer churn happens both voluntarily and involuntarily, it severely threatens a firm's profitability. High churn rates drive up acquisition costs, drain revenue, erode brand reputation, and slash customer lifetime value. Managing it is critical to a telecom company's long-term market viability. Losing clients directly results in declining revenue streams, which heavily impacts telecom profits because these companies depend on steady monthly service payments. Replacing these lost clients can be costly, as telecom companies often spend more on marketing and sales to acquire new customers than on retaining existing ones. Furthermore, a high churn rate can severely damage a brand's reputation by signaling service dissatisfaction, poor customer experience, or network deficiencies, making it increasingly difficult to attract new customers. Ultimately, this customer attrition reduces customer lifetime value, a key metric for assessing long-term profitability, because premature departures cut short the total value a provider can derive from a customer over time.

An investigation into mobile customer churn using subscriber demographics, billing records, and network usage data from a leading telecom provider was conducted [8]. By tracking active users over an eight-month period, the study demonstrated that customer retention is significantly influenced by switching costs and loyalty programs. Beyond these membership benefits, customer experience, pricing, and network performance are primary drivers of customer churn, as bad support, unexpected price increases, and service interruptions easily drive consumers to competitors. Additionally, aggressive rival promotions—such as bundled packages or lower rates—and personal life events, such as relocating or financial shifts, often prompt users to switch providers.

ISSN Online 3108-026X, ISSN Print 3139-4833
International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 3, Issue 7, July 2026 | DOI 10.17148/IMRJR.2026.030703

II. METHOD

The dataset (Table 1) for this project is sourced from Kaggle.com (a publicly available online data hub) and primarily consists of five core components detailing customer profiles and their likelihood of leaving the service. First, it captures customer demographics by tracking gender, senior citizen status, and whether the individual has a partner or dependents. Second, it compiles account information, including contract types, tenure, payment methods, and financial metrics such as monthly and total charges. Third, the data records specific services used, including internet connectivity, online security, tech support, and streaming options for both television and movies. Finally, it uses predictor variables—specifically churn scores, tenure, contract terms, and monthly charges—to predict the target variable, whether the customer churned. R visualizations were used to uncover trends and relationships within the dataset.

TABLE 1 A SNIPPET OF THE CUSTOMER CHURN DATASET

1	City	Zip Code	Lat Long	Latitude	Longitude	Gender	Senior Citize	Partner	Dependen	Tenure Month
3	Los Angeles	90005	34.059281, -118.30742	34.059281	-118.307420	Female	No	No	Yes	2
4	Los Angeles	90006	34.048013, -118.293953	34.048013	-118.293953	Female	No	No	Yes	8
5	Los Angeles	90010	34.062125, -118.315709	34.062125	-118.315709	Female	No	Yes	Yes	28
6	Los Angeles	90015	34.039224, -118.266293	34.039224	-118.266293	Male	No	No	Yes	49
10	Los Angeles	90028	34.099869, -118.326843	34.099869	-118.326843	Male	No	Yes	Yes	47
13	Los Angeles	90039	34.110845, -118.259595	34.110845	-118.259595	Male	No	No	Yes	5
14	Los Angeles	90041	34.137412, -118.207607	34.137412	-118.207607	Male	No	No	Yes	34
15	Los Angeles	90042	34.11572, -118.192754	34.115720	-118.192754	Female	No	No	Yes	11
16	Los Angeles	90056	33.987945, -118.370442	33.987945	-118.370442	Female	No	No	Yes	2
17	Los Angeles	90061	33.92128, -118.274186	33.921280	-118.274186	Female	No	No	Yes	15
18	Los Angeles	90063	34.044271, -118.185237	34.044271	-118.185237	Female	Yes	Yes	Yes	8
19	Los Angeles	90065	34.108833, -118.229715	34.108833	-118.229715	Male	Yes	No	Yes	18
20	Beverly Hills	90211	34.063947, -118.383001	34.063947	-118.383001	Female	No	No	Yes	9
21	Huntington Pa	90255	33.97803, -118.217141	33.978030	-118.217141	Female	No	No	Yes	1
22	Lynwood	90262	33.923573, -118.200669	33.923573	-118.200669	Female	No	Yes	Yes	7
23	Marina Del Re	90292	33.977468, -118.445475	33.977468	-118.445475	Female	No	No	Yes	12
25	Inglewood	90303	33.936291, -118.332639	33.936291	-118.332639	Female	Yes	Yes	Yes	25
27	Torrance	90504	33.867257, -118.330794	33.867257	-118.330794	Male	No	Yes	Yes	55
28	Whittier	90602	33.972119, -118.020188	33.972119	-118.020188	Female	Yes	Yes	Yes	37

To analyze the data and ensure the dataset was suitable for analysis, the following preprocessing steps were undertaken using R programming tools. We assessed the extent of missing data by calculating the total number of missing values in each column using colSums(is.na(telcom_churn_data)). Based on this assessment, we removed rows with missing values in the 'Total Charges' column using telcom_churn_data <- telcom_churn_data %>% filter(!is.na(total_charges)). This approach used listwise deletion, a method that excludes entire records with missing data [2, 3]. This method was appropriate for our analysis because the 'Total Charges' predictor variable only contained 11 missing values. Additionally, we dropped the 'Churn Reason' column entirely since it lacked data for 73% of the records. While listwise deletion effectively simplifies a dataset, it can also significantly reduce the overall sample size. Consequently, if other critical predictor variables exhibit substantial missing data in future analyses, alternative approaches such as mean imputation or multiple imputation should be considered [10], 11, 12]. After preprocessing, the data were imported into R for analysis using the readxl package, which facilitates reading Excel files, and various statistical methods were applied to identify features and patterns within the dataset. Churn analysis was implemented using R, based on several packages designed for data visualization and manipulation, such as ggplot2, plotly, and viridis. These packages ensure that data are efficiently processed for statistical tests and visualized as histograms, bar charts, boxplots, correlations, and binary logistic regression. The binary logistic regression places data on a clear 0-to-1 scale using an S-shaped curve. It is perfect for making yes-or-no decisions, such as identifying churned/non-churned customers.

III. RESULTS

This study utilized R programming and advanced data visualization tools to analyze telco churn datasets. Fig. 1 shows the proportion of churned and non-churned customers. This pie chart, a visual format for showing the percentage of churned customers, shows the balance between churned and non-churned customers, indicating a substantial churn rate for this company.

ISSN Online 3108-026X, ISSN Print 3139-4833
International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 3, Issue 7, July 2026 | DOI 10.17148/IMRJR.2026.030703

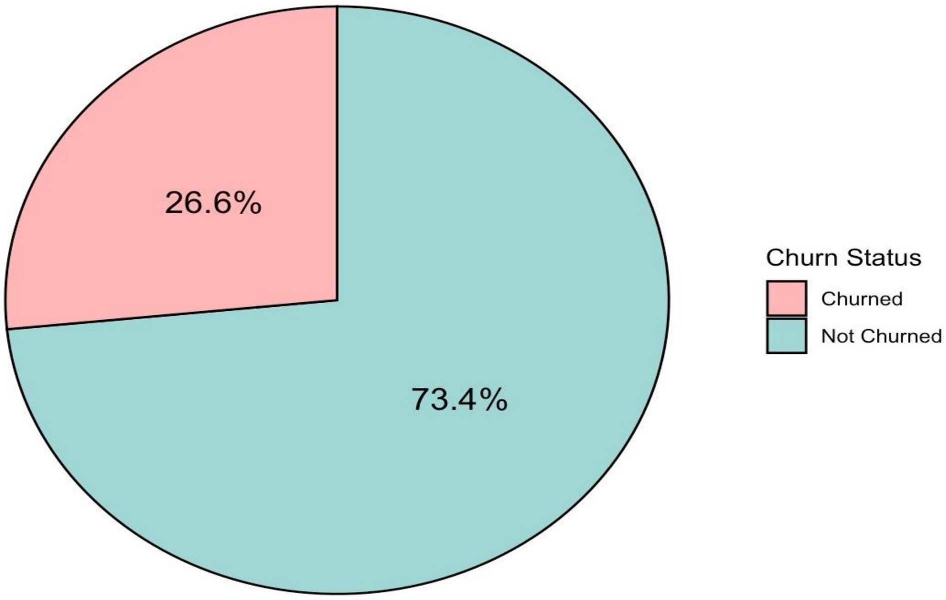

Fig 1. Customer churn Distribution

To analyze the percentage change in contract type, a stacked bar chart was created. Fig. 2 shows the contract types and churn for the month-to-month, one-year, and two-year durations. As the stacked bar chart reveals, the churn rate for month-to-month customers is 42.7%, exceeding the much lower churn rates for customers with one-year or two-year contracts. Longer-term contracts tend to have the lowest churn rates, helping retain customers; that is, contract length dramatically prevents customer loss [4, 7].

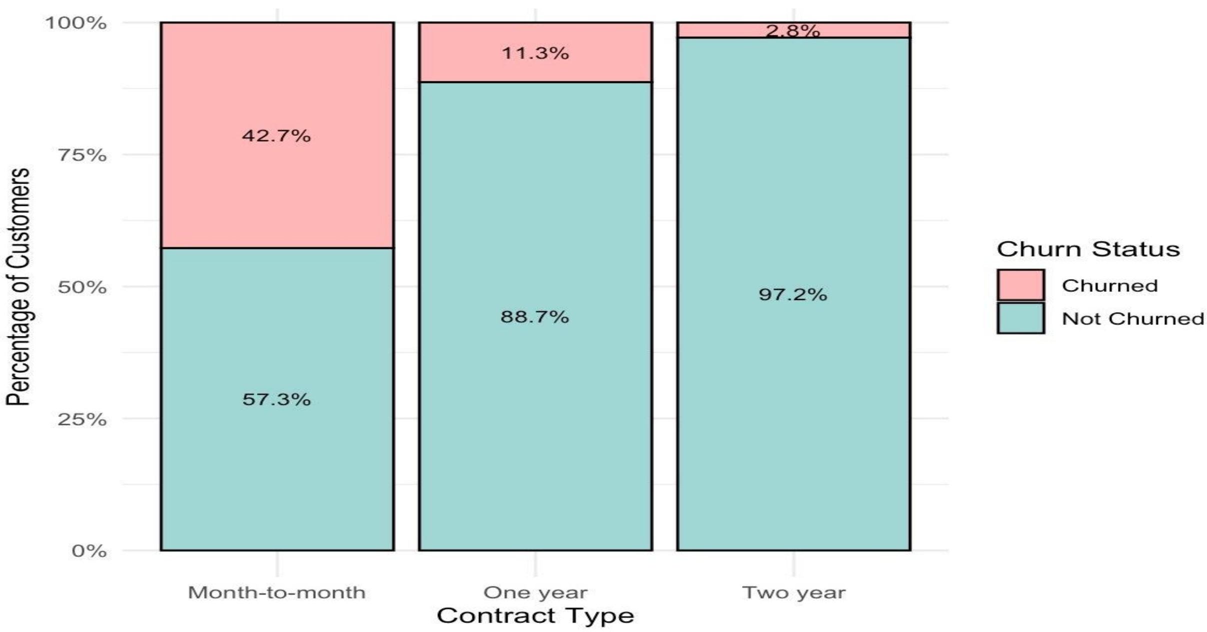

Fig 2. Churn Rate by Contract Type

To capture the tenure distribution on churn, a histogram represented in Fig. 3 was constructed. Data visualization uses color-coded bars to show churned customers in red and non-churned customers in blue. Data organization into 5-month intervals reveals important patterns. The initial tenure stages experience the greatest customer churn rates, as represented by the majority of red bars in the first bins of the graph. The data indicates that customers who have recently joined the service tend to stop using it earlier. Customers who stay with the service beyond 20 months exhibit higher loyalty as churn rates fall and blue-colored portions dominate their tenure representation. Customers who have stayed with the service for over 60 months exhibit low churn rates, indicating high retention. The initial months' successful onboarding and customer engagement patterns proved vital because they reduced customer departure rates and built long-lasting loyalty. Fig. 3 shows that the highest churn rates occur in the initial months of customer tenure, indicating that early customer experiences are crucial to retention.

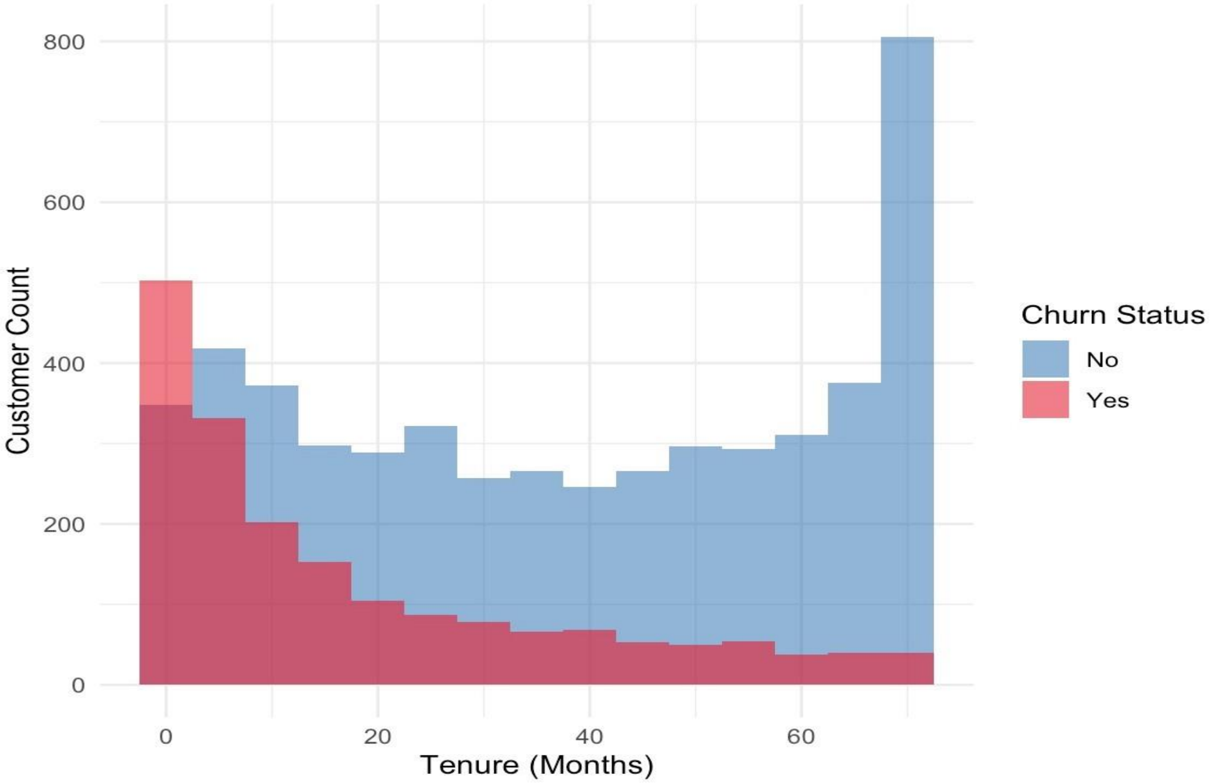

Fig 3. Tenure Distribution by Churn

The boxplot shows the churn score as a predictor (Fig. 4). It illustrates differences in Monthly Charges between customers who stayed and those who left, highlighting pricing patterns that affect customer churn. The x-axis divides customers into two categories: "Not Churned" (blue) and "Churned" (pink), while the y-axis shows their Monthly Charges. Data from the plot indicates that churned customers usually pay higher monthly charges, as the median line in the pink box exceeds that in the blue box. The broader interquartile range (IQR) for churned customers indicates greater price variability than for non-churned customers, who maintain more stable pricing. The R package produced the visualization using geom_boxplot() to create the box plots and scale_fill_manual() to assign teal for "Not Churned" and pink for "Churned." At the same time, labs() added the necessary title and axis labels, and theme_minimal() provided a simple layout. The findings demonstrate a link between elevated monthly fees and greater churn risk and recommend that retention rates can be improved through price stabilization and incentive programs for expensive customers [5, 6]. Fig 4 reveals that high churn risk peaks between scores of 75 and 100**.** This validates the metric as an accurate early warning system. Risk multiplies with higher spending**.** Furthermore, analysis through both box and scatter plots shows a strong relationship between increased monthly charges, higher churn scores, and elevated churn rates. The likelihood of leaving increases significantly for customers who pay higher prices under month-to-month contracts. These visuals highlight critical customer churn and tenure trends against monthly billing data, providing clear directions for any retention strategy.

The scatter plot (Fig. 5) shows how Monthly Charges relate to the Churn Score by placing Monthly Charges on the x-axis and plotting the Churn Score on the y-axis. Churn Status determines the color scheme for points: non-churned customers appear in red, and churned customers appear in teal. The plot includes a linear trend line (dashed) created using geom_smooth(method = "lm", se = TRUE) to show the general relationship between these variables. The scatter plot shows that churned customers tend to have higher churn scores, particularly as their monthly charges increase. The initial code maps Monthly Charges and Churn Scores to the axes and uses Churn Label for color differentiation between points. The labs() function sets the plot title and axis labels, while theme_minimal() ensures a sleek, professional appearance. Higher monthly charges correlate with higher churn risk, as evidenced by higher churn scores among customers who churned. The correlation between the Churn Score and Monthly Charge shows that the Churn Score rises proportionally with increased Monthly Charges. The data indicates that customers with higher monthly charges are more likely to leave the service. Based on the above correlation, businesses can refine pricing strategies or offer promotional deals to high-risk customers.

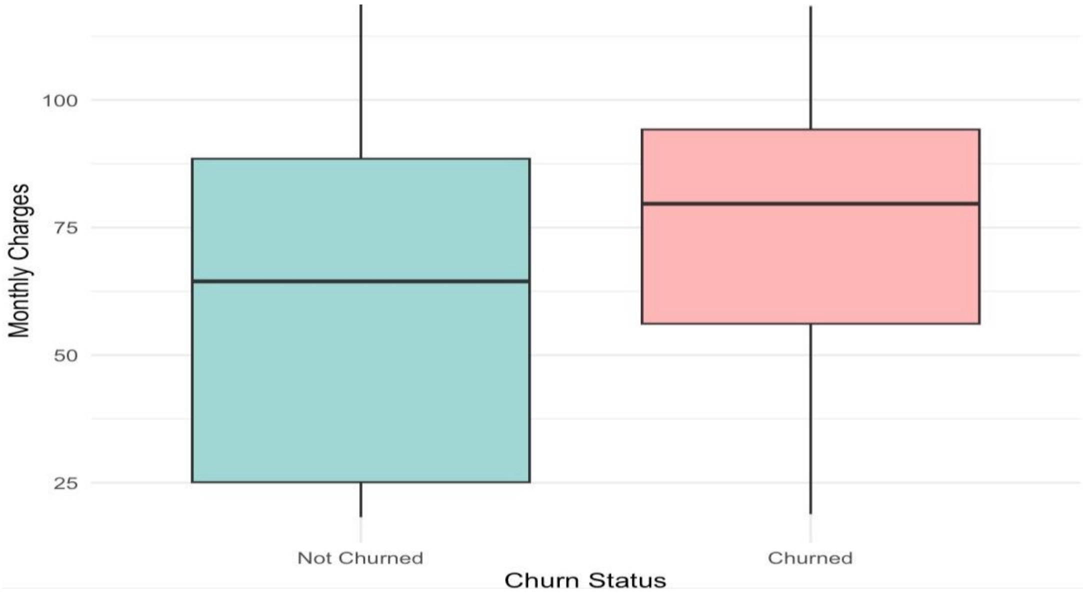

Fig 4. Monthly Charges vs Churn

Fig 5. Relationship Between Monthly Charges and Churn Score.

The histogram (Fig 6) shows the distribution of churn scores across churned and non-churned customers, providing insights into the relationship between churn scores and customer churn status. The horizontal axis shows the Churn Score (Predicted Probability), and the vertical axis shows the number of customers. The distribution of non-churned customers, shown in teal, spans mainly lower churn scores below 75, while churned customers, shown in pink, dominate the higher-score range from 75 to 100. The position "dodge" was set to geom_histogram() to separate the bars by churn status, and setting bins to 30 provided the necessary granularity, scale_fill_manual() with custom colors and alpha = 0.7 for transparency are set, while labs() provided clear labels and titles. The visualization shows a strong relationship between high churn scores and customer churn, enabling these scores to serve reliably as indicators of customer risk. The combined histogram and scatterplot analysis reveals multiple perspectives on how Monthly Charges and Churn Scores relate to Churn Status. The histogram shows that customers with churn scores between 75 and 100 experience the highest churn, indicating that churn scores are reliable predictors of impending churn. This aligns well with the scatter plot, which helps refine understanding by showing that customers with higher monthly charges tend to have higher churn scores,

ISSN Online 3108-026X, ISSN Print 3139-4833
International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 3, Issue 7, July 2026 | DOI 10.17148/IMRJR.2026.030703

indicating a greater likelihood of churn. These visualizations highlight a strong relationship: increased monthly charges are associated with higher churn scores, demonstrating a clear link to customer churn behavior. To retain the customers, the pricing strategies need to be modified for customers with high churn scores

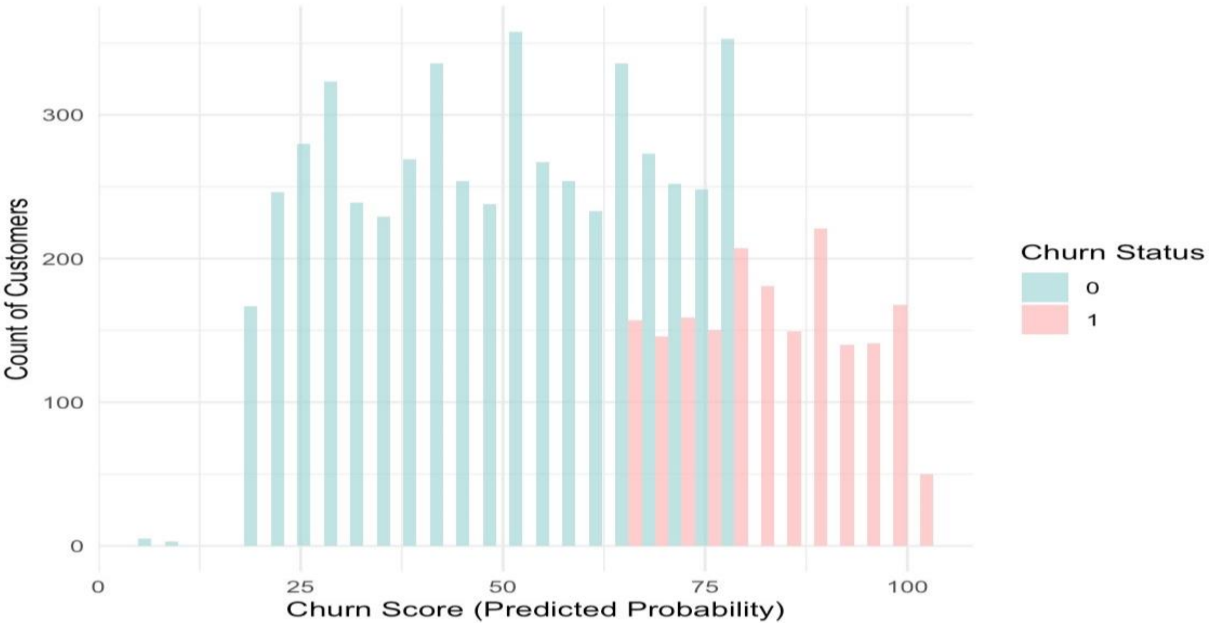

Fig. 6 Distribution of Churn Scores by Churn Status

DISCUSSION

The visualizations display essential patterns and insights related to customer churn and tenure alongside monthly charges and churn scores, which indicate potential business strategy actions. As regards churn rivers, the tenure distribution histogram shows that the highest churn rates occur in the initial months of customer tenure, indicating that early customer experiences are crucial to retention [5, 6, 7]. Analysis through both box and scatter plots shows a strong relationship between increased monthly charges, higher churn scores, and elevated churn rates. The likelihood of leaving increases for customers who pay higher prices under month-to-month contracts. Using contract types and churn, the stacked bar chart shows that the churn rate for month-to-month customers reaches 42.7%, exceeding the much lower churn rates for customers with one-year or two-year contracts. Longer-term contracts tend to have the lowest churn rates, helping retain customers. As a churn score predictor, the churn score histogram shows that customers with churn scores between 75 and 100 have a high likelihood of churn, indicating that churn score is an effective predictor of at-risk customers. The scatter plot shows that customers who churn tend to cluster in the upper ranges of the churn score, particularly those who pay higher monthly fees.

With these analyses, these recommendations will help with retention strategies to keep customers and provide them with high-quality service, thereby keeping churn low. These are some of the strategies Telco can use to keep its customers. To improve onboarding and early engagement, it is important to enhance the initial customer experience during the first half-year to reduce the risk of customer churn. To encourage ongoing engagement, businesses could implement proactive customer support systems, personalized welcome offers, and loyalty rewards. The pricing strategies could be optimized. Flexible pricing plans, discounts, or bundle packages should be introduced for customers with higher monthly charges to keep them loyal. In this vein, it could be vital to promote long-term contracts by suggesting year-long contracts to monthly clients, offering reduced rates and additional benefits. The strategy enhances revenue consistency while improving customer retention. To leverage churn scores for retention campaigns, it may be reasonable to use customized retention strategies to reach customers with high churn risk. Customers who achieve churn scores above 75 should be offered loyalty rewards or upgraded service packages. Furthermore, customer insights could be enhanced by conducting continuous evaluation to understand the interactions among monthly fees, customer churn scores, and customer churn tendencies.

CONCLUSION

This analysis underscores that customer churn is not an inevitable loss, but a manageable business metric that directly impacts long-term profitability. By applying binary logistic regression to the California telecom dataset, the study clearly demonstrates that contract duration is a primary driver of retention, with month-to-month subscribers exhibiting a critical 42.7% churn rate compared to the stable 2.8% seen in two-year agreements. These findings provide a clear roadmap for telecommunications companies to transition vulnerable, short-term accounts into secure, long-term relationships through proactive, data-driven incentives. By shifting from reactive troubleshooting to targeted, customer-first attrition management, telecom providers can successfully mitigate risk, cultivate enduring brand loyalty, and secure a competitive advantage in a high-stakes market.

ACKNOWLEDGMENT

The authors express their gratitude to colleagues at Texas A&M University–Central Texas and Udocenter Integrated Services for their continuous support and encouragement throughout this research. Additionally, the departmental resources provided were instrumental in successfully completing the project.

REFERENCES

[1]. V. Chang et al., "Prediction of Customer Churn Behavior in the Telecommunication Industry Using Machine Learning Models", *Algorithms,* vol. 17, no. 6, p. 231, 2024, https://doi.org/10.3390/a17060231.
[2]. A. K. Ahmad, A. Jafar and K. Aljoumaa, Customer churn prediction in telecom using machine learning in big data platform", *J Big Data,* vol. 6, no. 28, 2019, https://doi.org/10.1186/s40537-019-0191-6.
[3]. A. Amin et al., "Customer churn prediction in the telecommunication industry using data certainty", *Journal of Business Research*, vol 94, pp. 290-301, 2019, https://doi.org/10.1016/j.jbusres.2018.03.003.
[4]. M. Maduna et al., "Smart customer churn management system using machine learning" *Procedia Computer Science*, vol. 237, pp. 552-558, 2024.
[5]. Y. Feng et al, "Enhancing e-commerce customer churn management with a profit-and AUC-focused prescriptive analytics approach", *Journal of Business Research*, vol. 184, pp. 114872, 2024.
[6]. M. Shahabikargar et al., M., "A comprehensive survey on customer churn analysis studies", *Journal of Information and Telecommunication,* vol. 10, no. 1, pp. 24-70, 2026.
[7]. P. Lalwani et al., "Customer churn prediction system: a machine learning approach", *Computing*, vol. 104, no. 2, pp. 271-294, 2022.
[8]. J. H. Ahn, S. P. Han and Y. S. Lee, Customer churn analysis: Churn and mediation effects of partial defection in the Korean mobile telecommunications service industry", *Telecommunications Policy,* vol. 30, no. 10, pp. 552–568, 2006, https://doi.org/10.1016/j.telpol.2006.09.007.
[9]. K. Dahiya and S. Bhatia, "Customer churn analysis in telecom industry, 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)", *Noida, India, pp. 1–6, 2015.* https://doi.org/10.1109/ICRITO.2015.7359318.
[10]. Telco Customer Churn. IBM Dataset. https://www.kaggle.com/datasets/yeanzc/telco-customer-churn-ibm-dataset/data.
[11]. S. Douville, et al., "Data visualization expert lessons learned: Implications for program evaluations", *American J. of Evaluation,* vol. 46, no. 3, https://doi.org/10.1177/10982140241290744, 2025.
[12]. H. Wang and C. Zhang, "Visualization analysis of literature in the design issues", Journal from the Perspective of Sustainable Development. In *International Conference on Human-Computer Interaction*, pp. 103-114, 2025

BIOGRAPHY

Dr. Emmanuel Udoh is a professor of Information Technology at Texas A&M University Central, Texas, USA. He holds two doctoral degrees, one in Information Technology from Capella University, USA, and one in Geology from the University of Erlangen-Nuremberg, Germany.

Evely Ehinofe Iyere is the CEO of the UdoCenter Integrated Services, Nigeria. She is a graduate of computer science from the Ignatius Ajuru University of Education, Rumuolumeni, Port Harcourt, Nigeria.