Brunel University London

CS5803

Data Visualisation

Student Name: Ashkan Sheikhansari

Student ID: 2356272

1. Introduction

The provided dataset offers a comprehensive view of customer behavior in the telecommunications industry, particularly focusing on attributes related to service utilization and customer churn. Sourced from IBM Sample Data Sets, the dataset presents a granular look at customer interactions and subscriptions, which serves as a foundation for developing targeted customer retention strategies.

Each record within the dataset corresponds to an individual customer, encompassing a variety of attributes, from demographic information to detailed account usage. Key among the dataset's offerings are insights into customers who have discontinued their services within the last month, a category identified as 'Churn.' This pivotal metric allows for an analysis of attrition rates and customer loyalty, which are critical for business strategies.

1.1. Data Dictionary

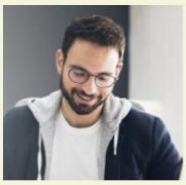
Field	Description	Format	Values
CustomerID	Unique identifier for the customer	Text	Example: 0002-ORFBO
Gender	Customer's gender	Text	Examples: Male, Female
SeniorCitizen	Indicates if the customer is a senior citizen	Integer	0 (No), 1 (Yes)
Partner	Indicates if the customer has a partner	Text	Examples: Yes, No
Dependents	Indicates if the customer has dependents	Text	Examples: Yes, No
Tenure	Number of months the customer has stayed with the company	Integer	0 to 72 (Assuming 6 years as maximum)
PhoneService	Indicates if the customer has a phone service	Text	Examples: Yes, No
MultipleLines	Indicates if the customer has multiple lines	Text	Examples: Yes, No, No phone service
InternetService	Type of internet service the customer has	Text	Examples: DSL, Fiber optic, No
OnlineSecurity	Indicates if the customer has online security service	Text	Examples: Yes, No, No internet service
OnlineBackup	Indicates if the customer has online backup service	Text	Examples: Yes, No, No internet service
DeviceProtection	Indicates if the customer has device protection service	Text	Examples: Yes, No, No internet service
TechSupport	Indicates if the customer has tech support service	Text	Examples: Yes, No, No internet service
StreamingTV	Indicates if the customer has streaming TV service	Text	Examples: Yes, No, No internet service
StreamingMovies	Indicates if the customer has streaming movies service	Text	Examples: Yes, No, No internet service
Contract	The term of the customer's contract	Text	Examples: Month-to-month, One year, Two year
PaperlessBilling	Indicates if the customer has paperless billing	Text	Examples: Yes, No
PaymentMethod	The customer's payment method	Text	Examples: Electronic check, Mailed check, Bank transfer, Credit card

Field	Description	Format	Values
MonthlyCharges	The amount charged to the customer every month	Numeric	18.25 to 118.75 (Assuming based on dataset)
TotalCharges	The total amount charged to the customer	Numeric	18.80 to 8684.80 (Assuming based on dataset)
Churn	Indicates if the customer left within the last month	Text	Examples: Yes, No

Table 1: Data description

Source of the dataset: https://www.kaggle.com/datasets/blastchar/telco-customer-churn

1.2. User Persona



Name: Jordan Ellis
Job Title: Customer Retention Manager
Age: 39 Years
Education: MBA in Marketing & Strategy
Industry: Telecommunication

Goals & Objectives:
Decrease the customer churn rate by identifying key factors that lead to customer turnover.
Develop targeted retention strategies tailored to different customer segments.
Enhance the customer experience through improved service offerings and personalized communication.
Increase customer loyalty and long-term revenue from existing customer base.

Tools They Need to Do Their Job:
Advanced analytics dashboard for churn prediction and segmentation.
Customer relationship management (CRM) software.
Marketing automation tools for personalized campaign execution.

Fig 1: User Persona Details

1.3. Persona Questions and Requirements

Simple Questions:

Q1. What percentage of churned customers are senior citizens, and how does this compare to the general customer population?

R1: Analyze the 'Senior Citizen's distribution' bar chart to compare the proportion of senior citizens in the churned customer segment against those who have not churned.

R2: Examine the 'Churn Distribution' pie chart to determine the overall percentage of churned customers.

Q2. Does having dependents impact the likelihood of churn?

R3: Inspect the 'Dependents' distribution' stacked bar chart to evaluate the churn rate among customers with and without dependents.

Complex Question:

Q3. Is there a correlation between customer tenure, monthly charges, and the likelihood of churn, particularly when examining different customer lifestyle segments?

R4: Observe the 'Tenure Distribution Histogram' to discern patterns in tenure among churned versus retained customers.

R5: Review the 'Monthly Charges vs. Total Charges' scatter plot to assess if higher monthly charges contribute to churn, focusing on the data points representing churned customers.

R6: Use the 'Combined's distribution by Churn' bar chart to investigate churn across different customer lifestyle segments (such as 'Child', 'Independent', 'Non-earner', and 'Earner').

R7: Evaluate the 'Tenure vs. Total Charges' scatter plot to understand if long-term customers with higher total charges are less likely to churn.

R8: For further granularity, look at the 'Tenure vs. Monthly Charges' scatter plot to see if there's a particular tenure period where customers with certain monthly charges are more inclined to churn.

Non-functional requirements:

NF1: The dashboard must allow for quick identification of key statistics and trends without overwhelming the user with information.

NF2: The visualization contrast and color-coding need to be clear enough to distinguish between different customer segments and churn status at a glance.

NF3: The interface should offer tooltip explanations for less technical users to understand the data presented without extensive background knowledge.

NF4: The dashboard should be optimized for quick loading and interaction, even on less powerful devices, to ensure no lag in analysis.

NF5: The tools for interaction, such as filters and hover details, must be intuitive and easily accessible for non-technical stakeholders.

NF6: Ensure that all visualizations are labeled accurately and legibly, facilitating easy interpretation and decision-making by the Customer Retention Manager.

2. Design

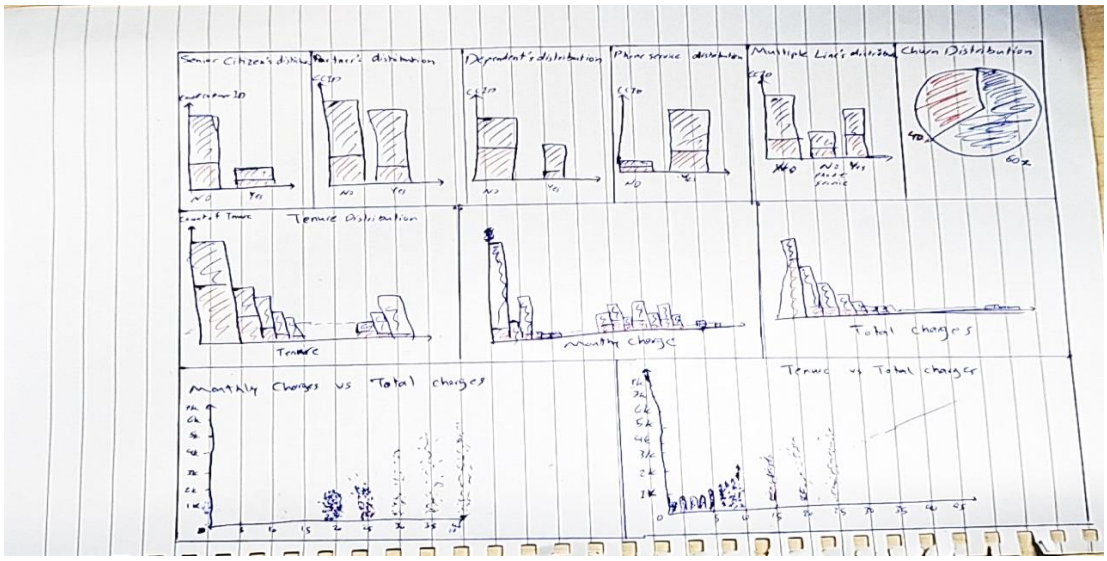


Fig 2: Prototype Design

During the initial design phase, I presented a hand-drawn sketch of the proposed dashboard to illustrate the planned visualizations. The sketches served as a blueprint for how the data would be presented in Tableau. Feedback from this stage pointed out discrepancies between the visual representations and the depth of analysis required by the research questions. This prompted a dual refinement process where both the research questions and the visualizations were fine-tuned to better align with each other.

Now the final view of dashboard:

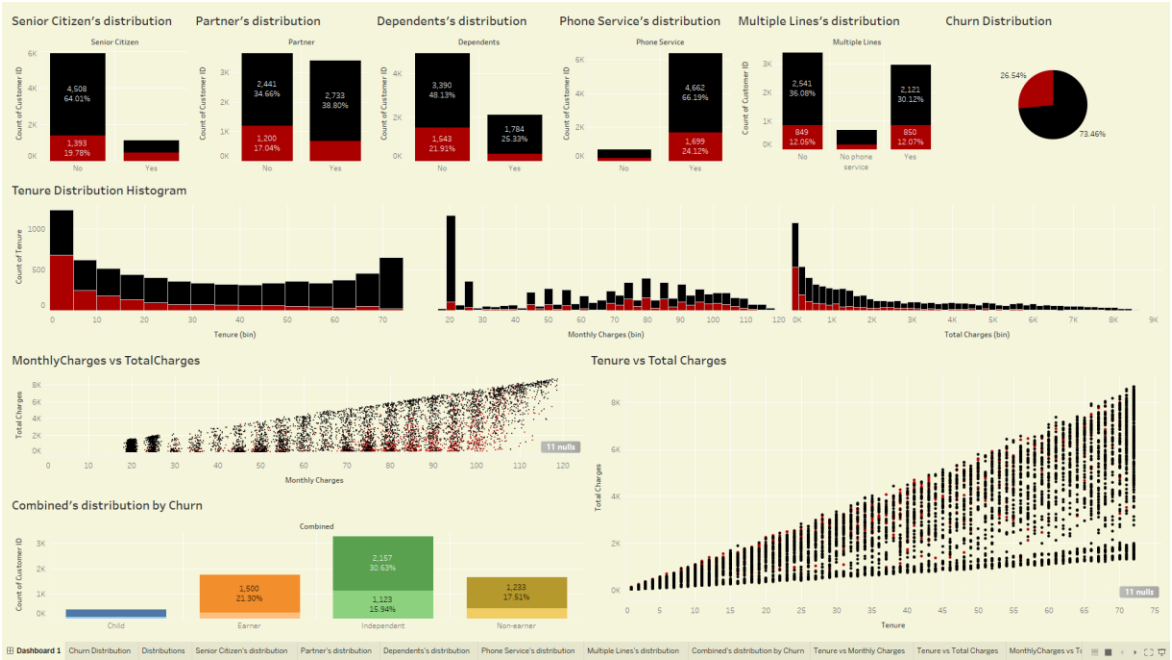


Fig 3: Final Dashboard

In the preliminary sketches for the Telecommunications Customer Churn Dashboard, the focus was on individual customer attributes like senior citizen status, partner, and dependents, which were plotted in isolation. The intent was to provide a straightforward depiction of these attributes' distributions in relation to customer churn.

However, moving from the prototype to the final implementation, the need for a more interconnected analysis became apparent. This led to the introduction of a composite visualization, 'Combined's distribution by Churn,' which merges several customer attributes to create a more segmented view of the customer base and their churn rates.

The histograms and bar charts initially conceived provided a foundational understanding of tenure and service distributions, but to achieve a deeper insight, these were expanded upon. For example, the 'Tenure Distribution Histogram' was augmented in the final dashboard to distinctly highlight segments of tenure where churn is most prevalent. This alteration gives an immediate visual cue into critical periods in customer tenure that may require targeted retention strategies.

Moreover, the scatter plots in the prototype, representing relationships between monthly charges, total charges, and tenure, were refined in the final dashboard to include color-coding that differentiates churned customers from those retained. This enhancement is particularly impactful in the 'Monthly Charges vs. Total Charges' scatter plot, allowing for rapid identification of at-risk customer segments based on their spending patterns.

Furthermore, the final implementation sees the introduction of a pie chart showing the overall churn distribution, which was not present in the initial sketches. This addition provides a quick, holistic view of the churn rate, crucial for benchmarking and goal setting in customer retention efforts.

3. Implementation

3.1 Tableau

The implementation journey for the Telecommunications Customer Churn Dashboard in Tableau was initiated with the intent to closely replicate the vision outlined in the original hand-drawn sketches. The objective was to bring to life a series of visualizations that could answer the research questions developed during the planning phase, using Tableau’s extensive suite of chart types, visual encodings, and interactive features.

As with any real-world data, the initial stage of implementation involved a meticulous data cleaning and preparation phase. The raw dataset presented issues such as incomplete records, potential outliers, and irregularities that could impact the accuracy and clarity of the visualizations. To address this, a methodical approach was taken to clean the dataset—removing any duplicates, filling in or omitting missing values where appropriate, and ensuring consistent formatting across all data fields. This groundwork was critical to ensuring the integrity and reliability of the subsequent analysis.

3.1.1. Customer Churn Implementation

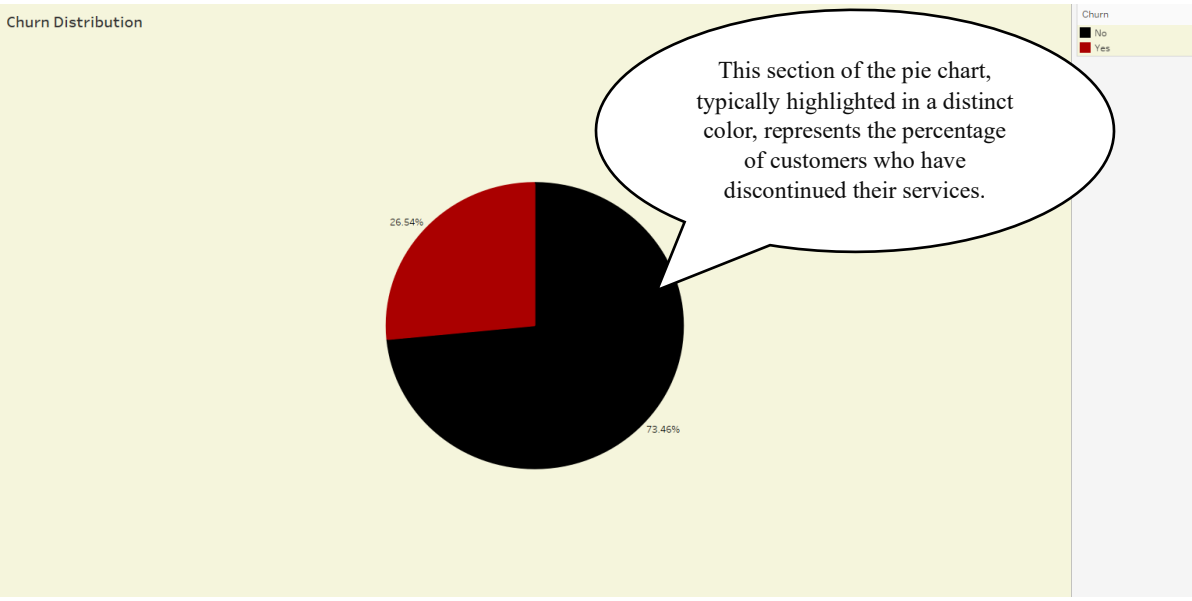


Fig 4: Churn Distribution Indicator (Top right part of Dashboard)

The figure illustrates the upper right segment of the dashboard, dedicated to providing a quick visual summary of the customer churn rate. This pie chart enables users to immediately grasp the proportion of customers who have churned within the last month. Tableau's robust visualization capabilities have been harnessed to allow users to interact with the chart, offering a drill-down into more detailed data such as churn rates by demographics or service types upon user interaction.

3.1.2. Distribution Analysis

Tenure Distribution Histogram Implementation

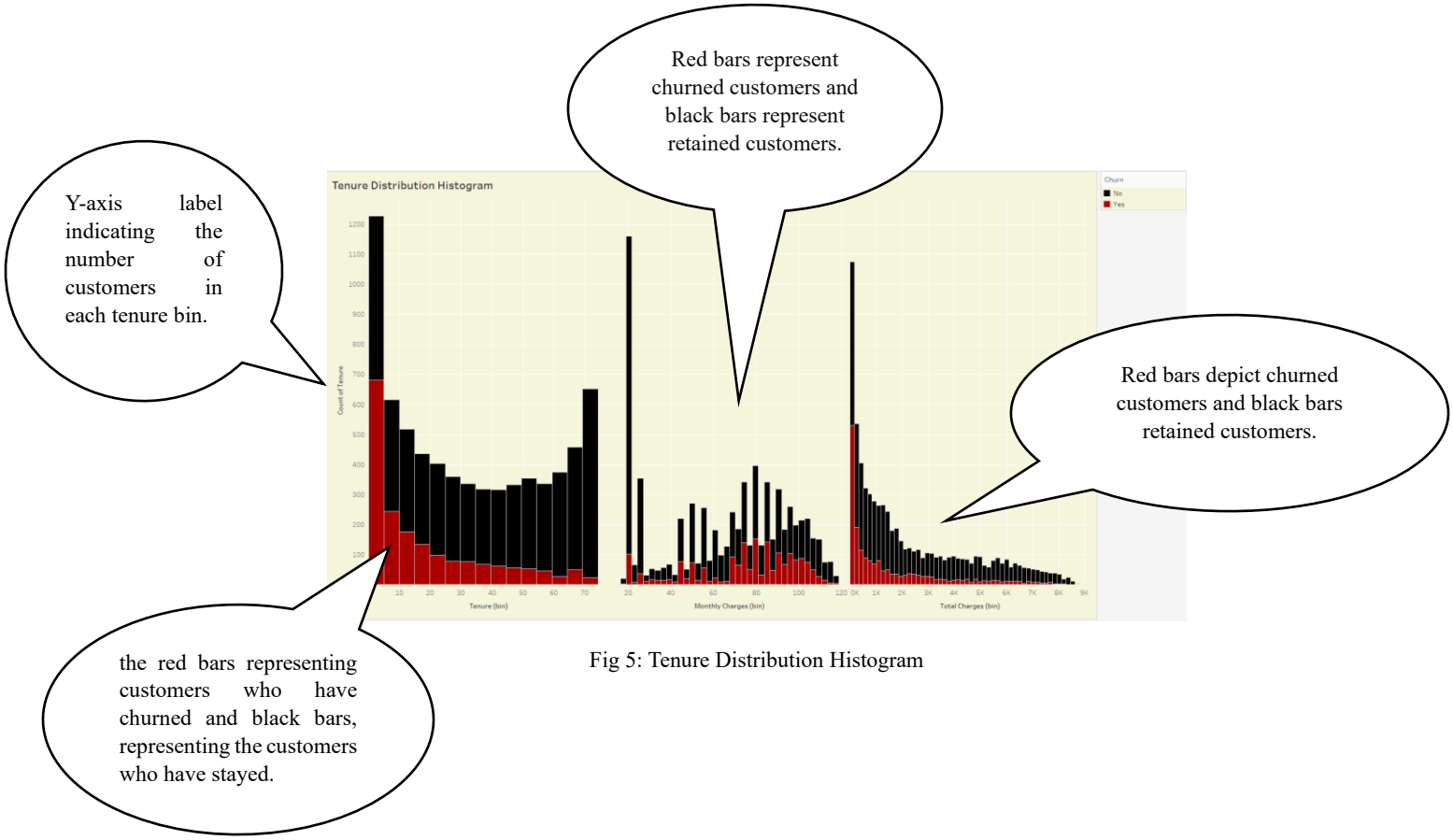


Fig 5: Tenure Distribution Histogram

The 'Tenure Distribution Histogram' dissects customer longevity with the service provider, differentiating between those who have churned and those who have remained. This histogram is an essential analytical tool for understanding at which points in the customer lifecycle churn is most prevalent.

To create this histogram in Tableau:

1. Tenure Binning: Divide the tenure length into discrete intervals (bins) to analyze the distribution effectively.

2. Churn Status Overlay: Use color to differentiate between customers who have churned and those who are still active, enabling a visual comparison within the same tenure bins.

Monthly Charges Distribution Histogram Implementation

The histogram displaying the distribution of monthly charges offers insights into the spending behavior of customers and its potential impact on churn. Each bar within the histogram corresponds to a range or 'bin' of monthly charges and the count of customers within that range.

Steps to create this histogram in Tableau include:

1.Defining bins for monthly charges to segment the data into meaningful intervals.

2. Using a color scheme to differentiate between churned and retained customers within each monthly charge bin.

3. Implementing a hover action to provide additional details on demand.

Total Charges Distribution Histogram Implementation

The histogram analyzing total charges unveils the cumulative financial commitment of customers and how this relates to their decision to stay or leave. It's particularly useful in identifying if higher overall charges lead to increased churn.

To produce this histogram in Tableau:

1. Create bins for total charges to categorize customer data based on their total expenditure.

2. Apply a dual-axis to visually compare the churn data directly against the total customer data.

3. Integrate interactive elements to allow users to explore specific bins more closely.

3.1.3. Categorical Analysis

Senior Citizen's Churn Bar Chart Implementation

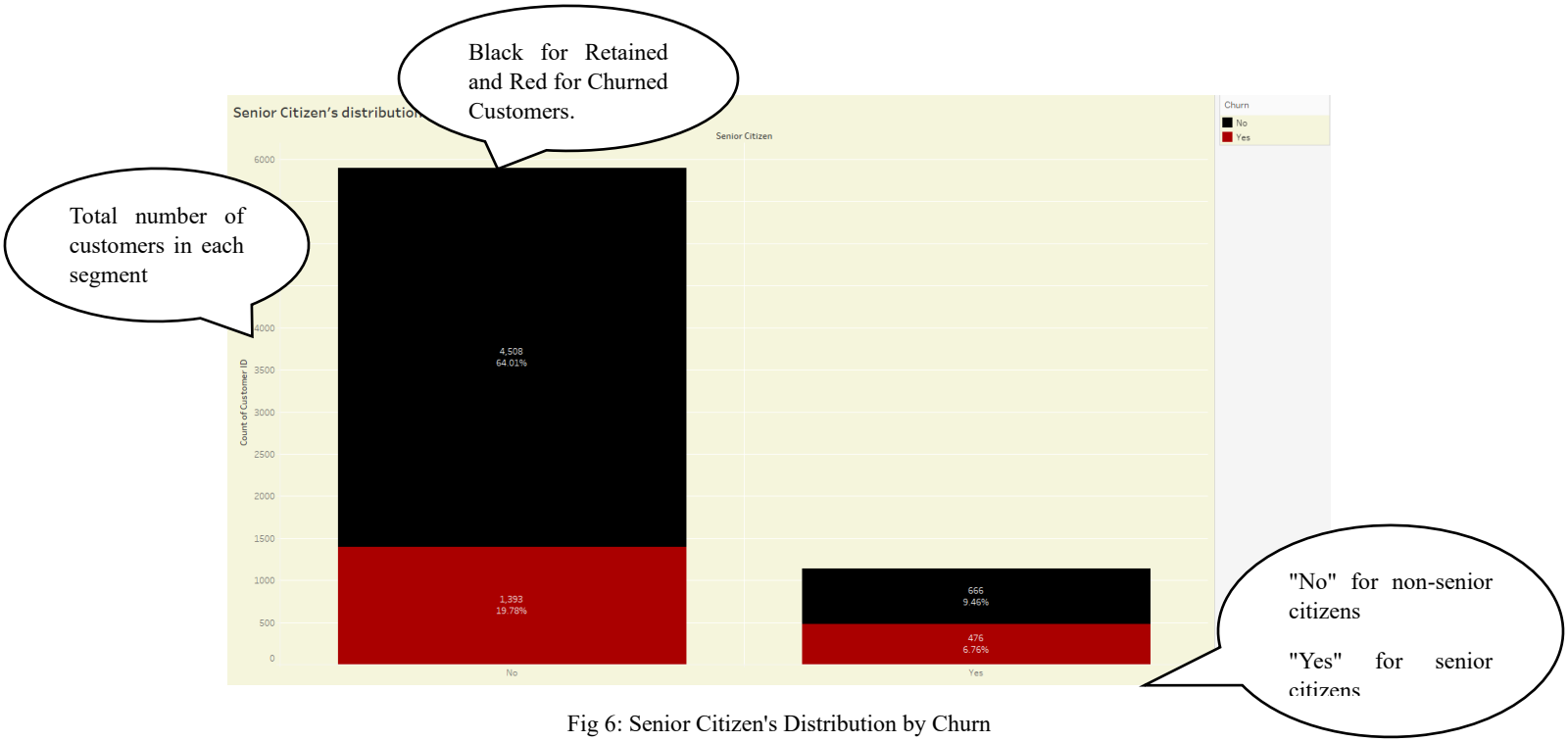


Fig 6: Senior Citizen's Distribution by Churn

The bar chart presents a comparative analysis of churn within the senior citizen customer segment against the overall customer base. It's strategically positioned to immediately inform stakeholders about the retention health of this demographic, which can be a critical factor in targeted marketing and service adjustments.

The implementation in Tableau was a straightforward process:

1. A boolean field identified senior citizens within the dataset.

2. Color-coding was applied, with red for churned customers and black for retained, allowing for an instant visual differentiation.

Partner's Churn Bar Chart Implementation

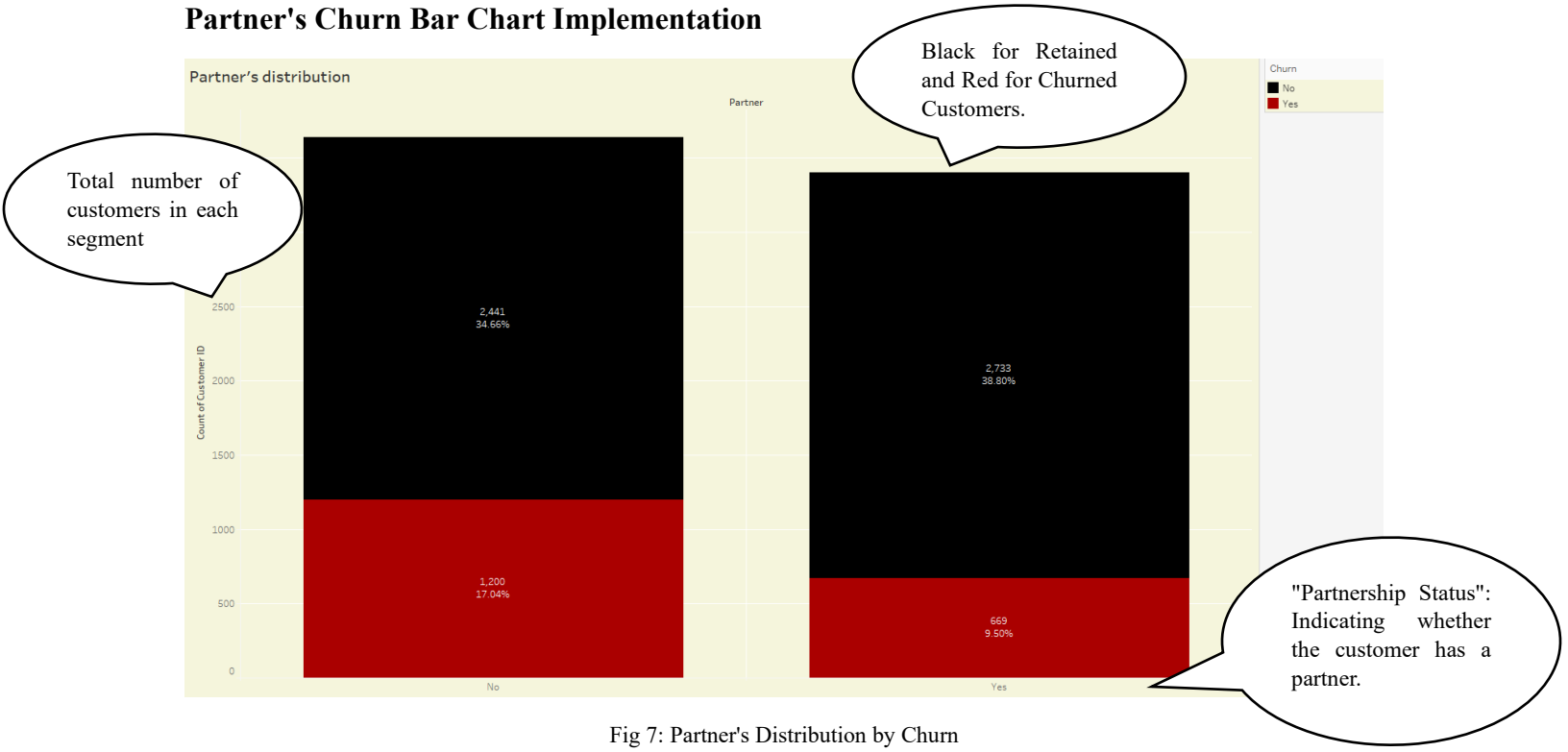


Fig 7: Partner's Distribution by Churn

This bar chart serves to clarify the influence of a customer's partnership status on their likelihood to churn. The visualization distinguishes between those with and without partners, employing a red and black color scheme to signify churned versus retained customers, respectively.

In Tableau, the chart was constructed as follows:

1. The dataset was first segmented based on the 'Partner' attribute to create two distinct groups.

2. Two separate bar charts were created, one for each group ('Yes' for customers with partners and 'No' for those without), and then the data regarding churn was layered on top using a distinct color.

Dependents' Churn Bar Chart Implementation

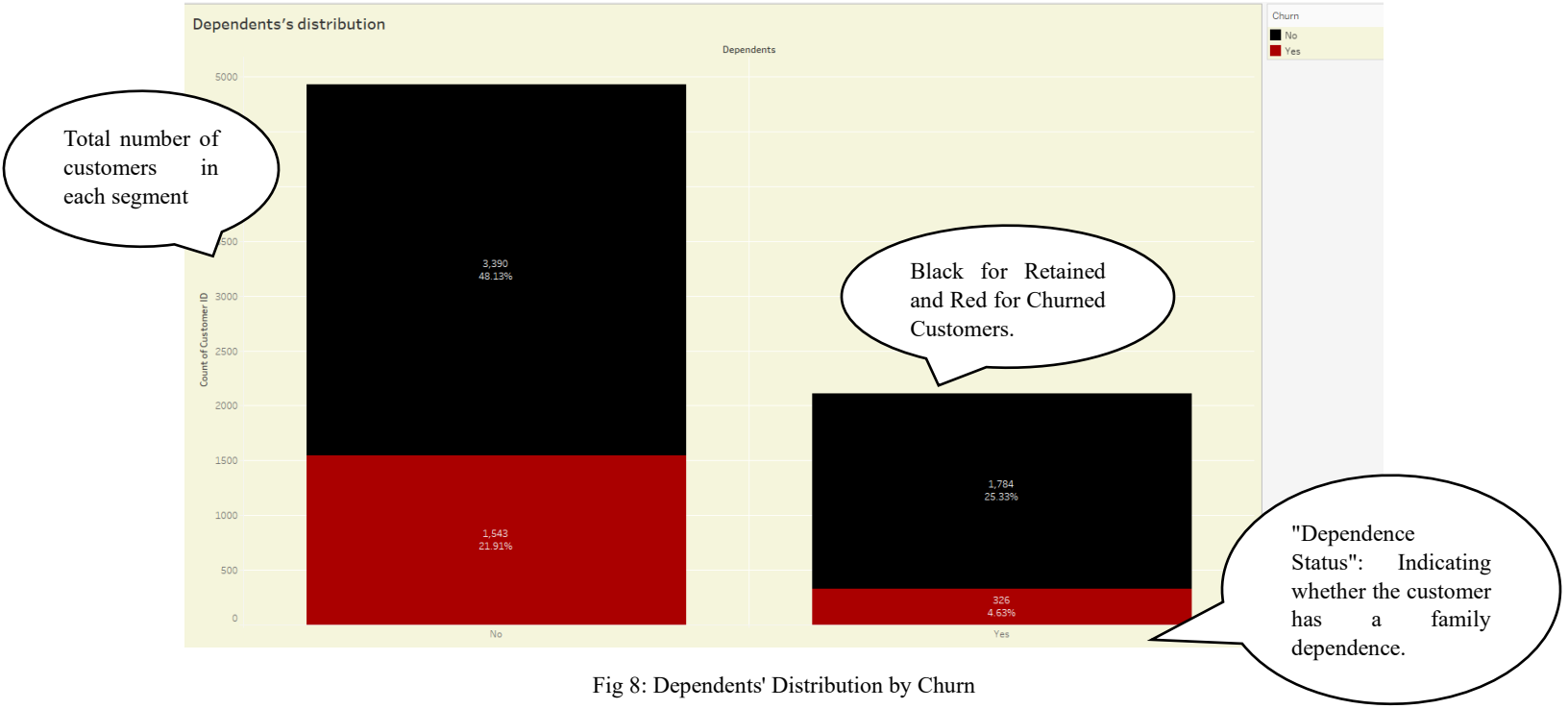


Fig 8: Dependents' Distribution by Churn

The Dependents' Distribution bar chart illustrates the comparison between customers with and without dependents in relation to their churn rates. This bar chart is pivotal in understanding how family obligations might influence customer decisions to stay with or leave the service provider.

To implement this bar chart in Tableau, the following steps were taken:

1. Classification of customers into those with dependents and those without was done based on the 'Dependents' attribute in the dataset.

2. The bar chart was then constructed, with separate bars for customers with and without dependents, and differentiated by color to show the churn rates within these groups.

Phone Service's Churn Bar Chart Implementation

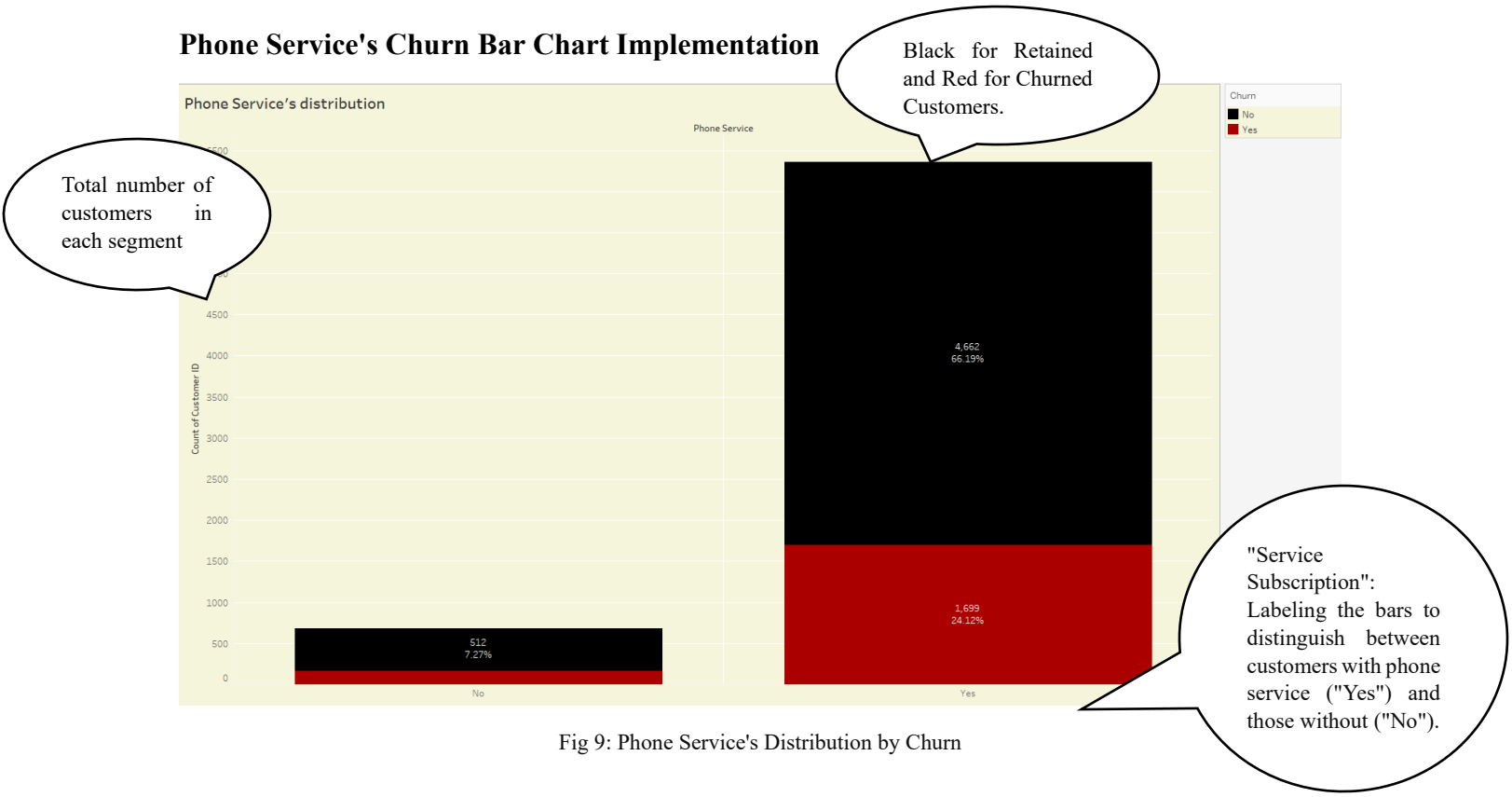


Fig 9: Phone Service's Distribution by Churn

The bar chart provides an overview of how the subscription to phone services correlates with customer churn. It juxtaposes customers with and without phone services, marked by color differentiation to reflect churn status.

The following method was employed to create this bar chart in Tableau:

1. Segmentation of the customer base was conducted based on whether they subscribed to phone services.

2. A bar chart was then plotted, with phone service subscription status on the x-axis, customer count on the y-axis, and a color distinction applied to indicate churn rates.

Multiple Lines' Churn Bar Chart Implementation

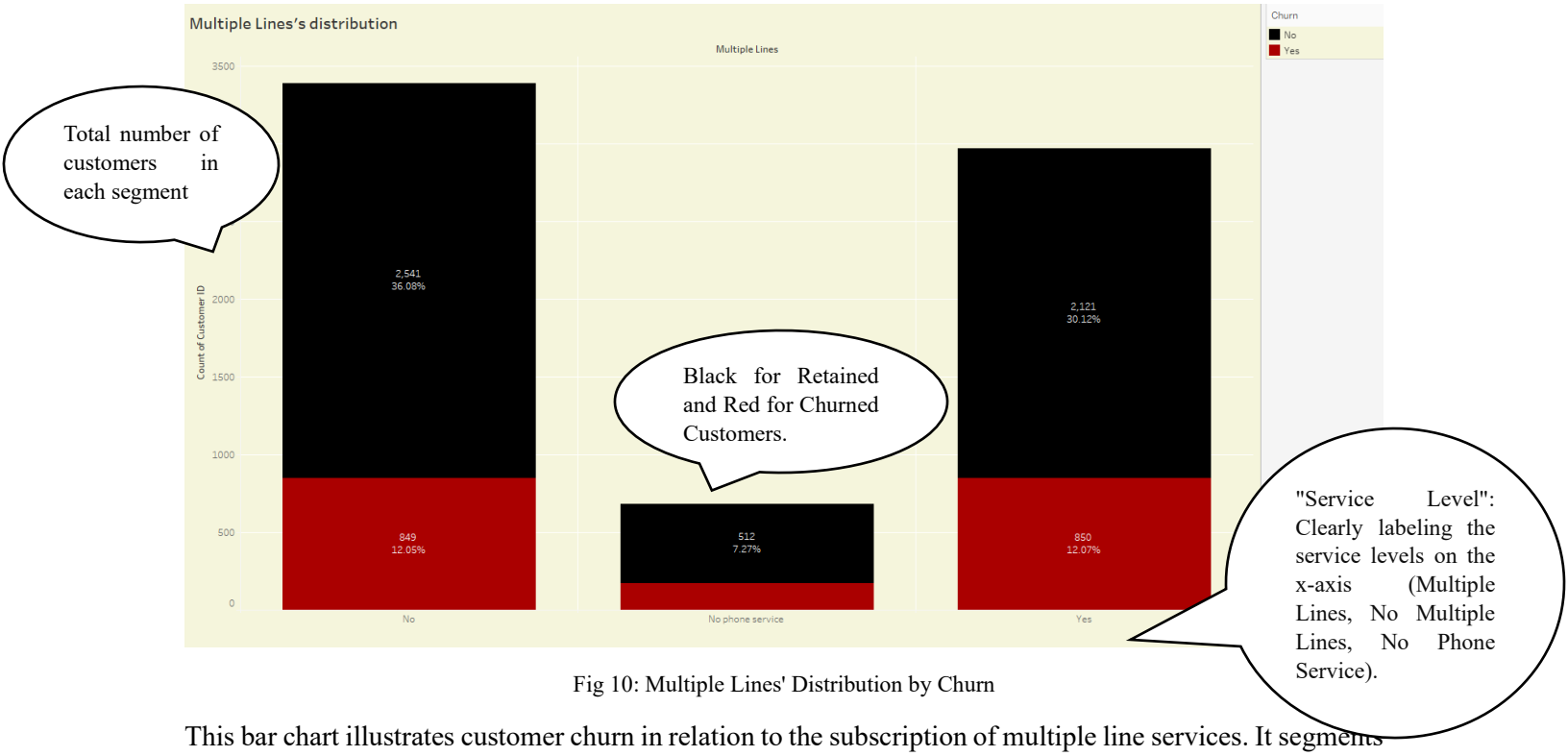


Fig 10: Multiple Lines' Distribution by Churn

This bar chart illustrates customer churn in relation to the subscription of multiple line services. It segments customers into three categories: those with multiple lines, those with no multiple lines but with phone service, and those without any phone service.

To construct this visualization in Tableau:

1. Customers were first categorized based on their multiple lines subscription status.

2. Three bar charts were created for each category with the x-axis representing the subscription status and the y-axis showing the count of customers. Churned customers were then highlighted using red color, while retained customers were shown in black.

3.1.4. Combined Lifestyle and Churn Distribution Analysis

Combined Lifestyle's Churn Bar Chart Implementation

"The orange stack corresponds to customers who have both partners and dependents, often considered earners. It is one of the larger customer segments, indicating a higher customer count and an associated churn rate of 3.54%."

"The green stack is the most substantial, indicating the largest segment of independent customers without partners or dependents. Despite its size, this group has a churn rate of 5.96%, suggesting effective retention strategies may be in place for this segment."

Total number of customers in each segment

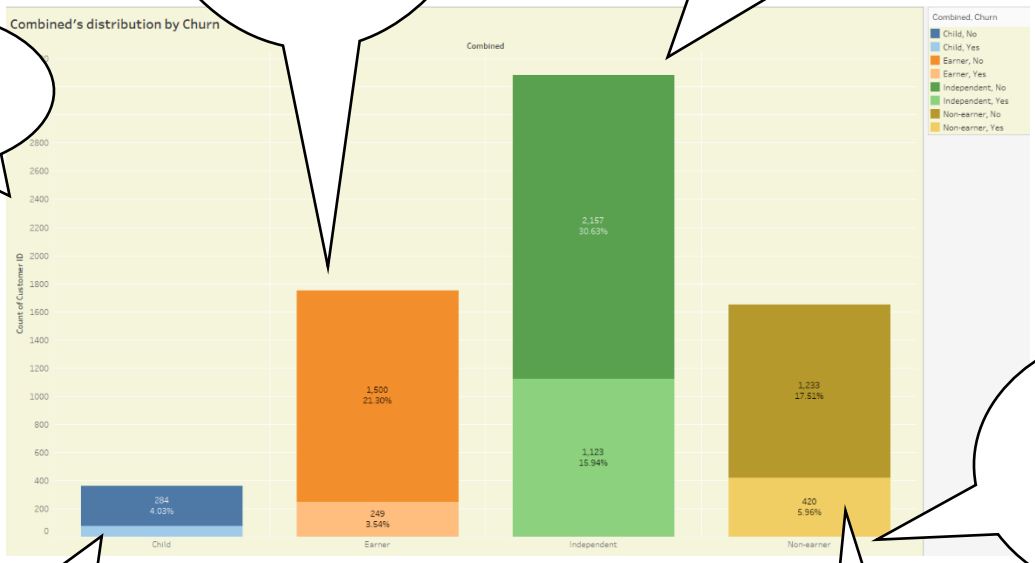


Fig 11: Combined Lifestyle's Distribution by Churn

"Customer Segments": Labeling each bar to identify the customer segment it represents.

"This stack represents customers who have dependents but no partner. The blue color indicates the segment's overall smaller presence in the dataset, and the size of the stack shows a relatively lower count of customers, with a churn rate of 4.03%."

"The yellow stack denotes customers with a partner but no dependents, often labeled as non-earners. While not the largest group, it shows a significant churn rate of 5.9%, highlighting a segment that could

The bar chart provides an analysis of customer churn within defined combined segments based on partnership and dependent status, offering a nuanced view into how different life stages or household types may affect churn behavior.

Implementation steps in Tableau for this visualization:

Creation of a calculated field named 'Combined' to classify customers into segments—Child, Independent, Earner, and Non-earner—based on partner and dependent status.

The calculated field logic in Tableau:

```
IF [Partner] = "No" AND [Dependents] = "No" THEN "Independent"
ELSEIF [Partner] = "No" AND [Dependents] = "Yes" THEN "Child"
ELSEIF [Partner] = "Yes" AND [Dependents] = "No" THEN "Non-earner"
ELSEIF [Partner] = "Yes" AND [Dependents] = "Yes" THEN "Earner"
END
```

Code 1: Calculation field for combining categories

A dual-axis bar chart was then created, with one axis for churn ('Yes') and another for retention ('No'), applying distinct colors to each segment within these categories.

Justification for this approach includes:

Reasoning for Segment Division: Dividing customers into life stage segments provides clearer insights for personalized marketing strategies and service adjustments.

Color-Coded Segments: Different colors were used for each segment to make the distribution across various customer types easily distinguishable, aiding in quick visual analysis.

3.1.5. Tenure vs Total Charges Scatter Plot Analysis

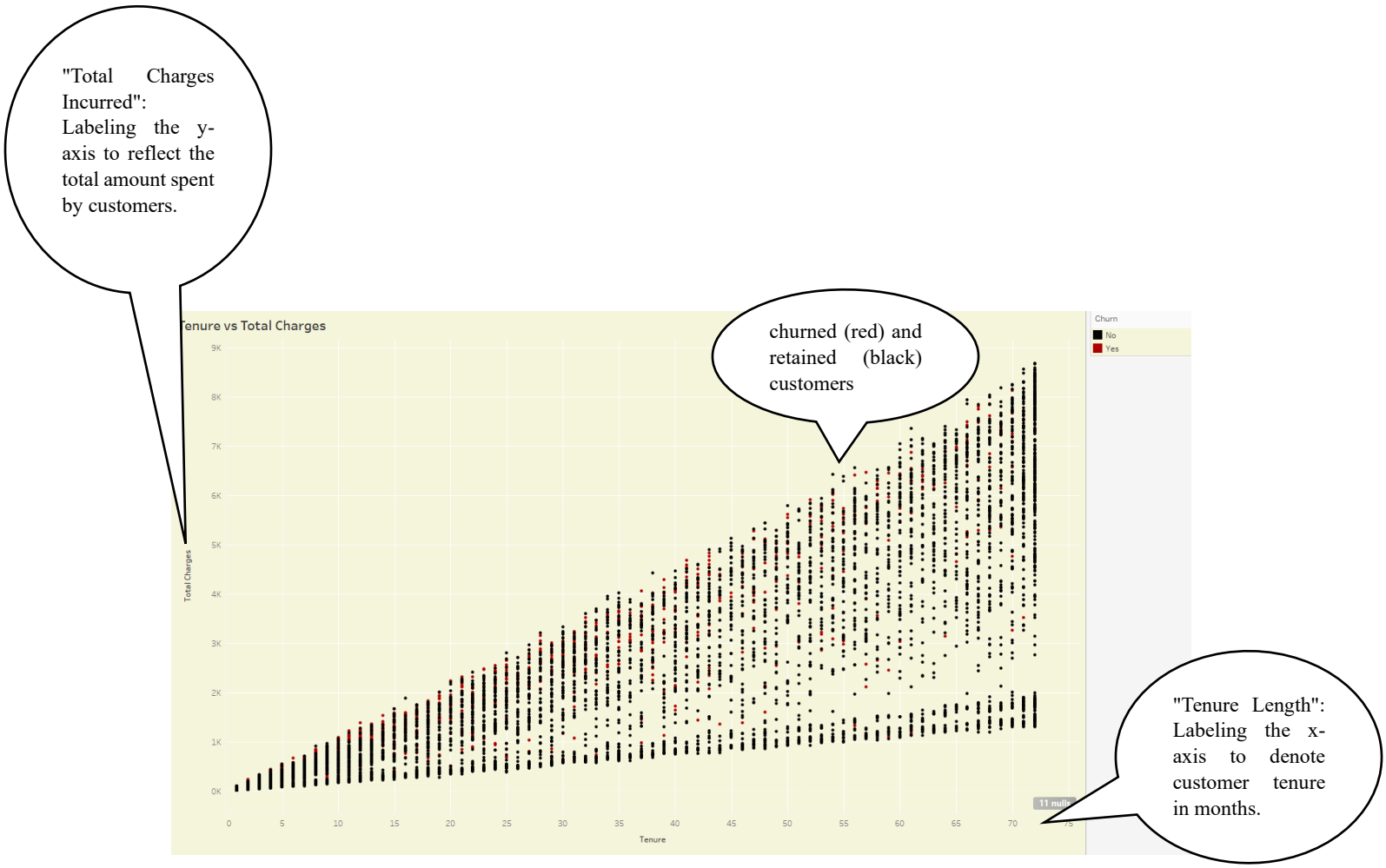


Fig 12: Tenure vs Total Charges

This scatter plot provides a visual correlation between the tenure of customers and their total charges, with the additional layer of churn indicated by color. It's an analytical tool to observe if longer tenure correlates with higher total charges and to assess at which points customers are more likely to churn based on financial commitment.

The process in Tableau included:

1. Plotting tenure on the x-axis and total charges on the y-axis to create a foundational scatter plot.

2. Calculating churn instances as a separate measure and using color to differentiate churned customers (red) from those who have not churned (black).

3. Applying a filter to exclude null values, ensuring a clean and accurate visualization.

Justification for Scatter Plot Usage:

Exploration of Two Continuous Variables: Scatter plots are ideal for showing the relationship between two continuous variables, such as tenure and total charges.

Churn Visualization: Color-coding adds a dimension of churn status, making it possible to identify any patterns between financial commitment and customer retention.

Strategic Insight: This visualization helps in understanding customer lifecycle value and identifying potential thresholds of service cost that might contribute to churn.

3.1.6. Monthly Charges vs Total Charges Scatter Plot Analysis

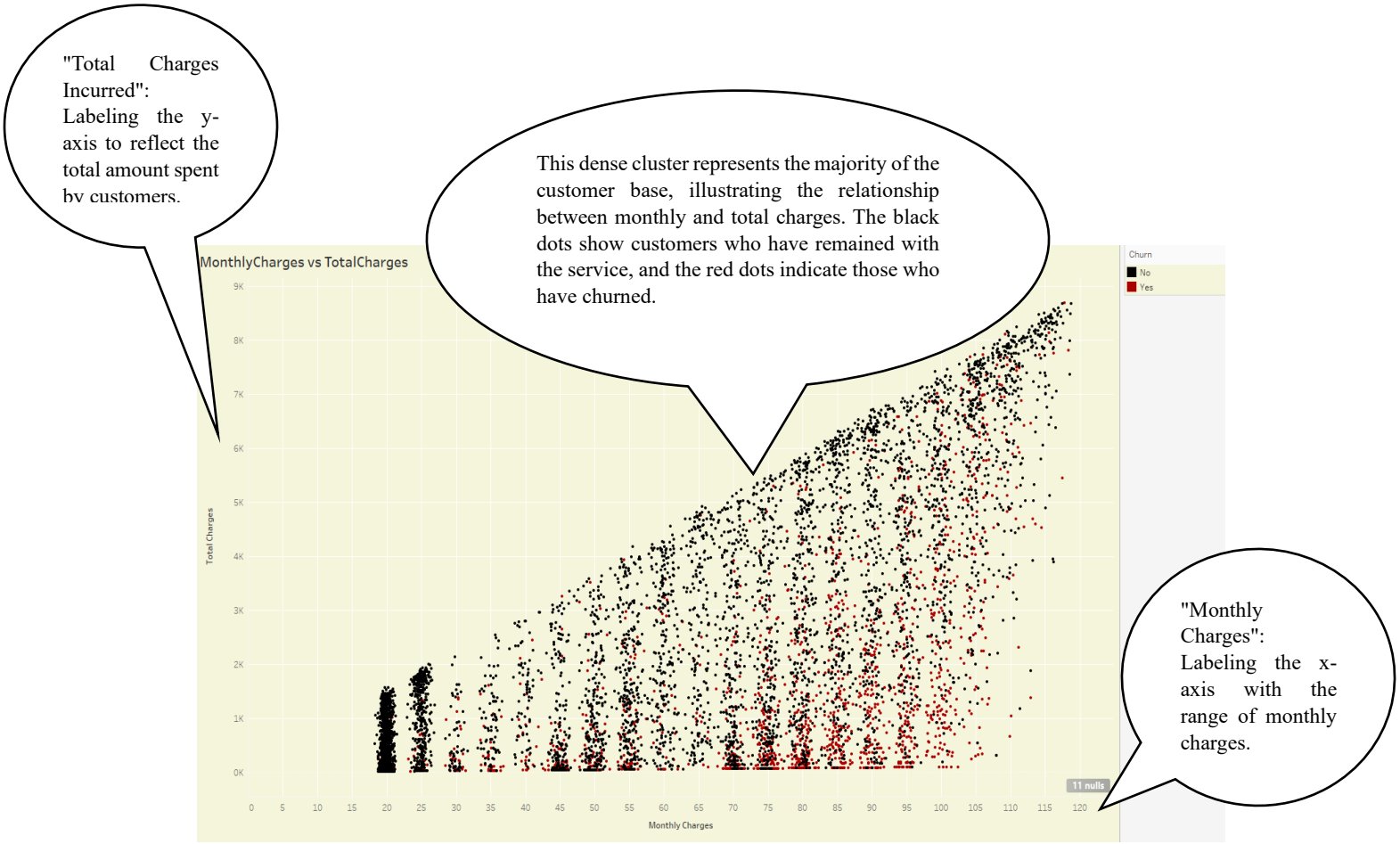


Fig 13: Monthly Charges vs Total Charges

This scatter plot maps the relationship between monthly charges and total charges accrued by customers, with an additional dimension of churn represented in color. This visualization is key for understanding customer payment patterns and identifying churn trends related to pricing.

For this plot in Tableau, the approach included:

1. Positioning monthly charges on the x-axis and total charges on the y-axis to construct a foundational scatter plot.

2. Differentiating churn status using color; with red dots representing customers who have churned and black dots for those who have not.

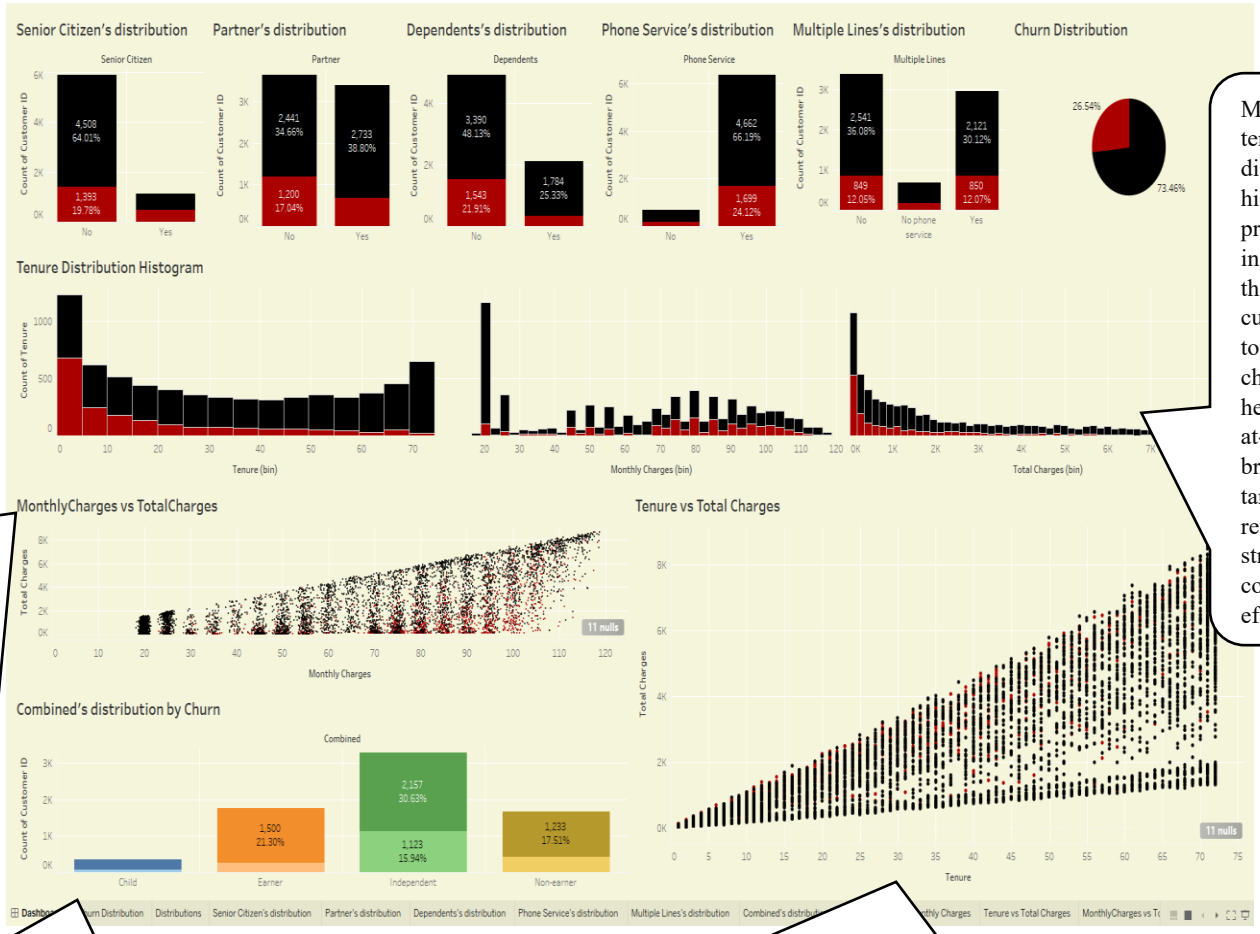
3. Applying interactive filters to handle null values and to allow users to engage with the data selectively.

4. Walkthrough

The dashboard presents a series of bar charts that break down churn by key customer demographics—Senior Citizens, Partners, Dependents, Phone Service, and Multiple Lines. Each chart offers a comparative look at the churn rates, aiding in the identification of demographics that are more susceptible to leaving.

The dashboard initiates with a clear visual of the overall churn rate, represented by a pie chart. This immediately orients users to the scale of churn, setting the stage for a deeper dive into the specifics.

The dual scatter plots of monthly vs. total charges draw a correlation between service costs and churn. They allow users to spot clusters where churn is most prevalent, highlighting critical price points that may influence customer decisions.



Moving on, the tenure distribution histogram provides insights into the duration customers tend to stay before churning. This helps identify at-risk tenure brackets where targeted retention strategies could be most effective.

A composite bar chart then analyzes churn within customer lifestyle segments, providing a multifaceted view of how combinations of demographics impact churn rates.

Throughout the dashboard, a consistent color scheme is employed to indicate churn, with annotations and a legend to guide interpretation. This consistency ensures that users can easily track churn across different visualizations.

Fig 14: Walkthrough of Dashboard

The walkthrough ensures that users can effectively navigate the dashboard and extract meaningful insights to inform data-driven decision-making to mitigate churn.

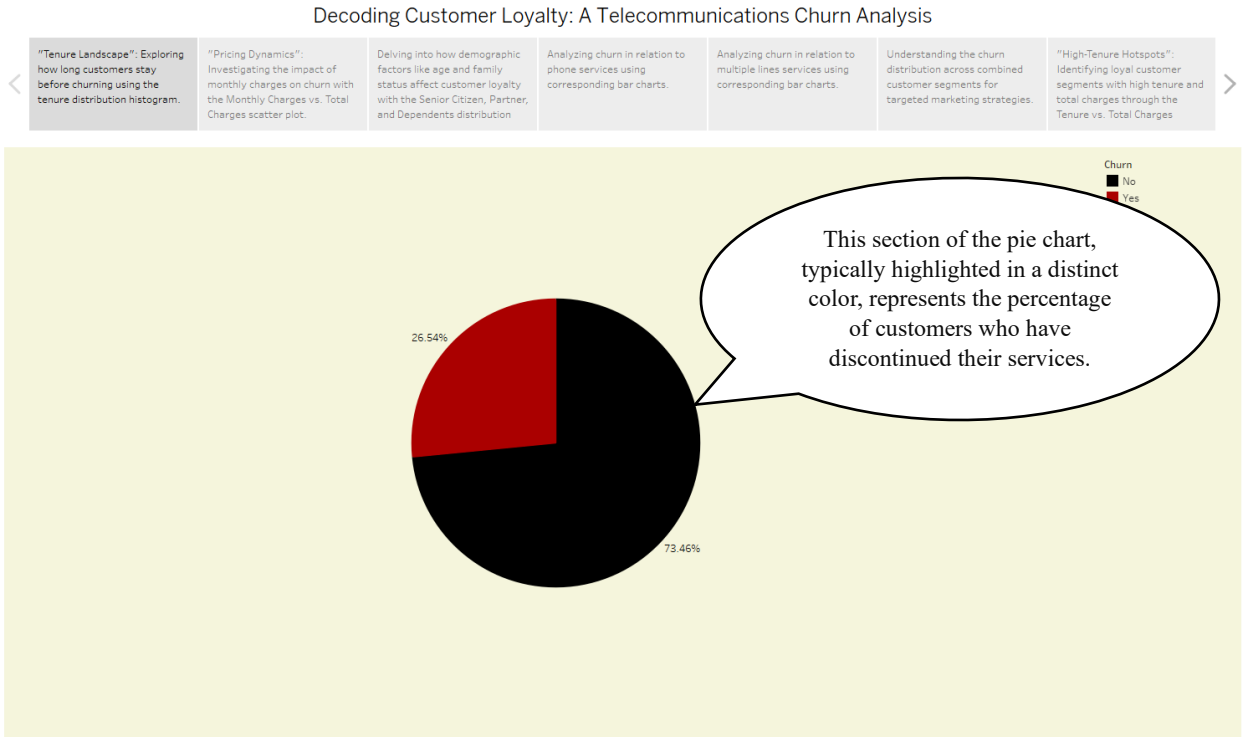


Fig 15: Story-1 view

This initial view presents a clear and immediate picture of customer churn through a pie chart. It provides a stark representation of the churn rate, setting the stage for a detailed analysis. This visual directly addresses our core question, allowing us to gauge the extent of churn at a glance.

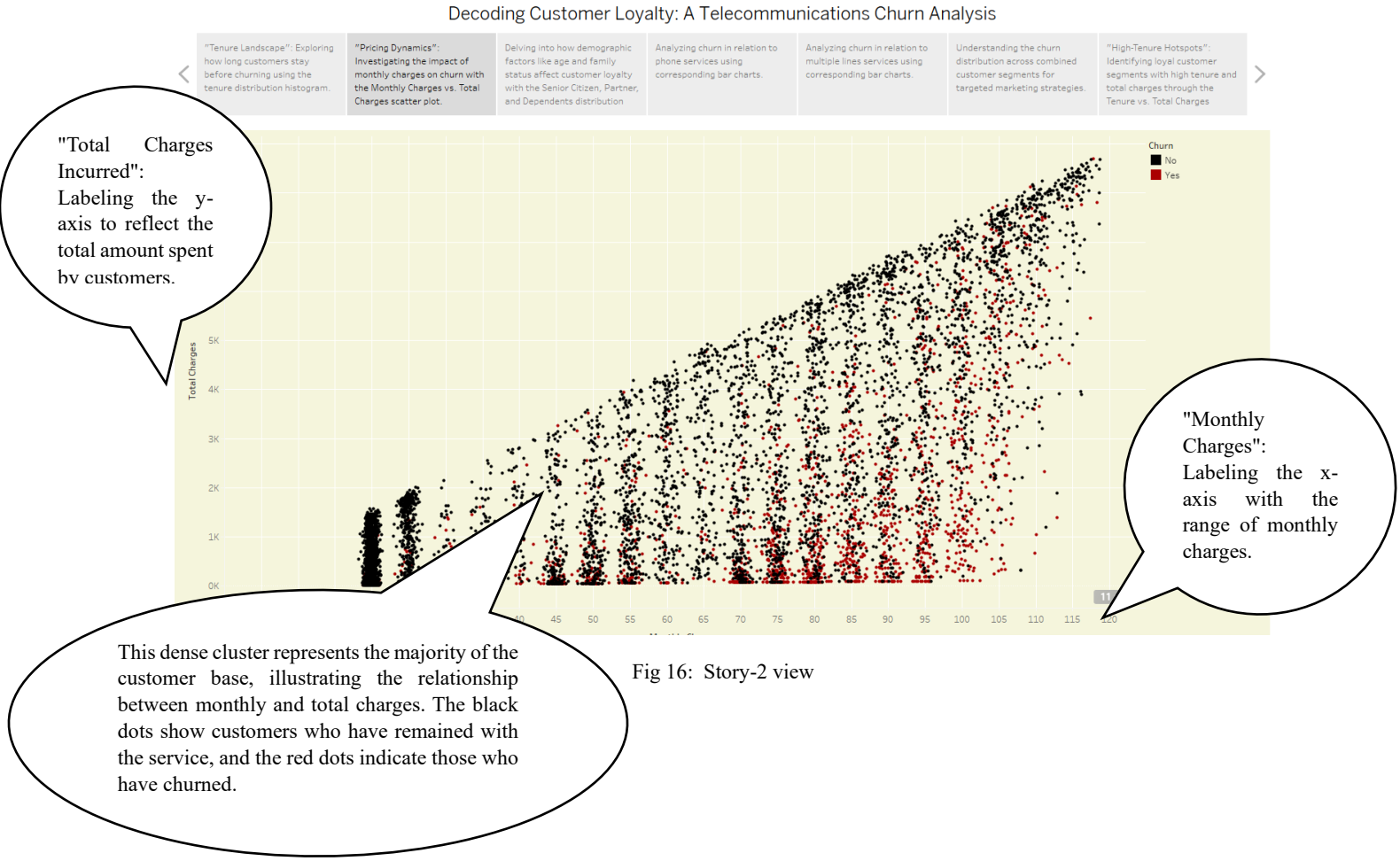


Fig 16: Story-2 view

The scatter plot examining Monthly Charges vs. Total Charges brings to light the influence of pricing on customer churn. It allows us to pinpoint specific charge ranges that are most associated with high churn rates.

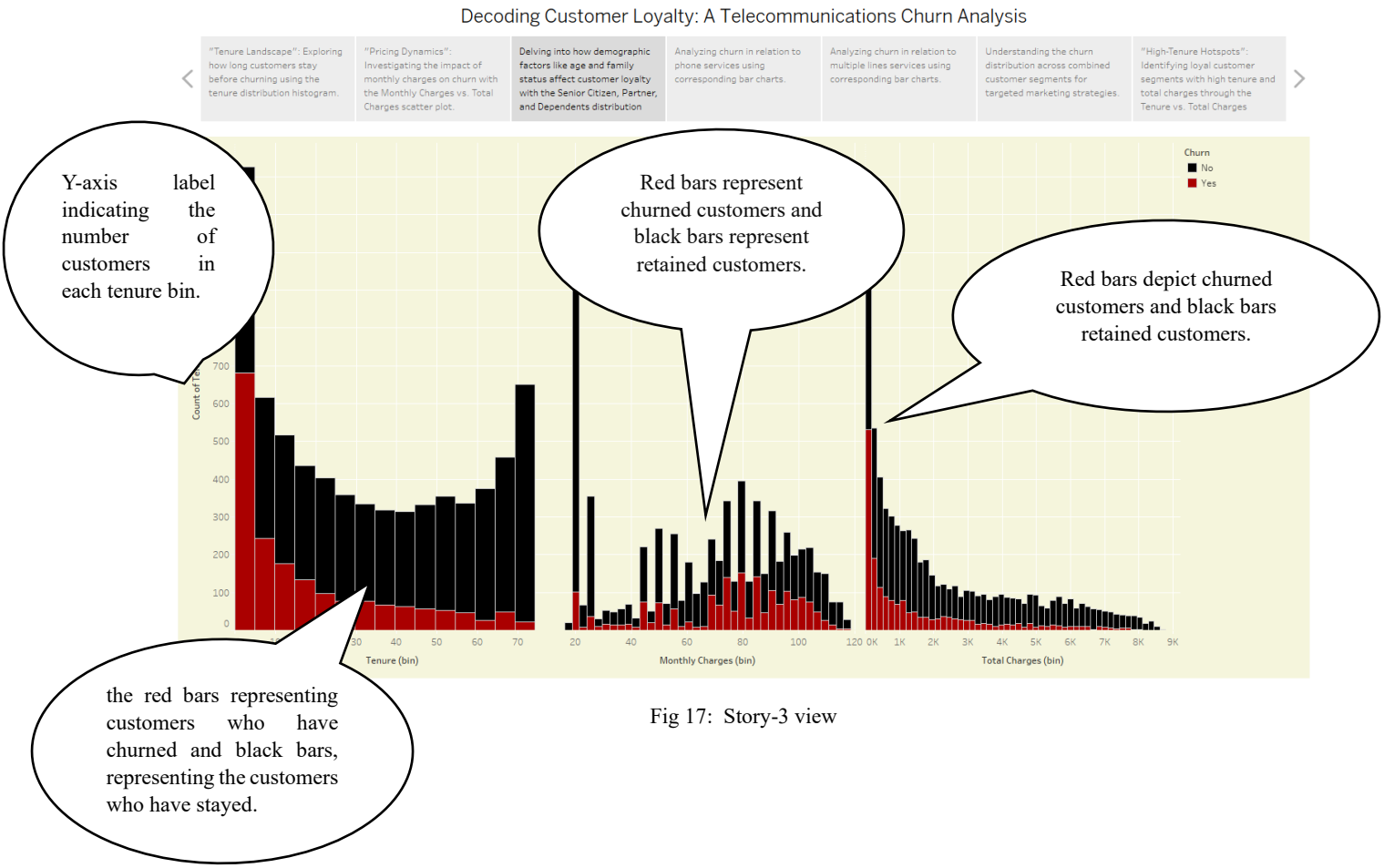


Fig 17: Story-3 view

In this view, we delve into customer longevity through a tenure distribution histogram. It reveals at which service duration customers are more likely to churn, informing strategies to improve customer retention during these critical periods.

By breaking down churn according to customer demographics like age and family status, this view sheds light on patterns and trends that could inform targeted customer engagement strategies.

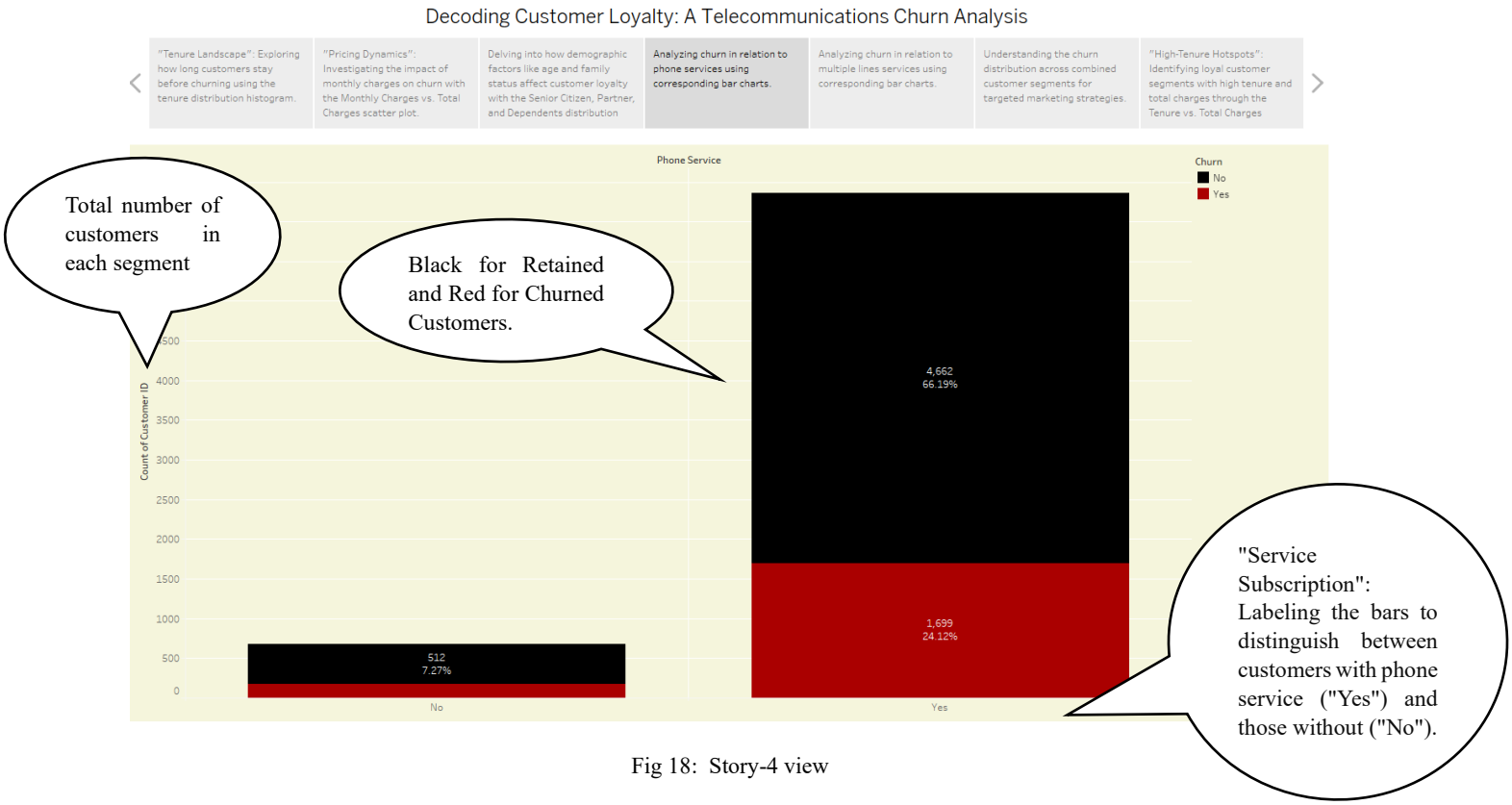


Fig 18: Story-4 view

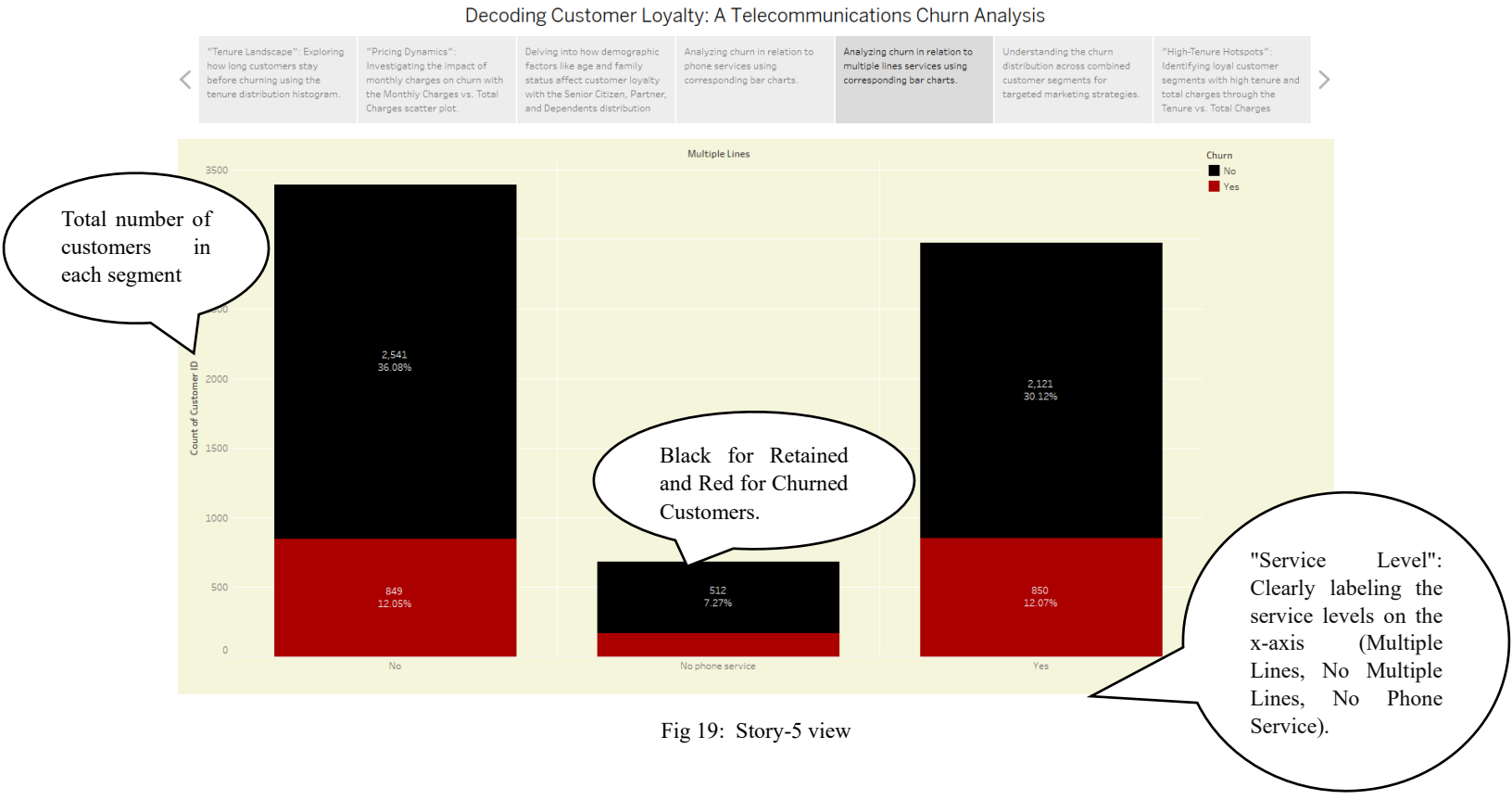


Fig 19: Story-5 view

Exploring churn in relation to service subscriptions, theses view utilize bar charts to show how different services correlate with customer retention or churn, providing actionable insights for service improvement or promotional offers.

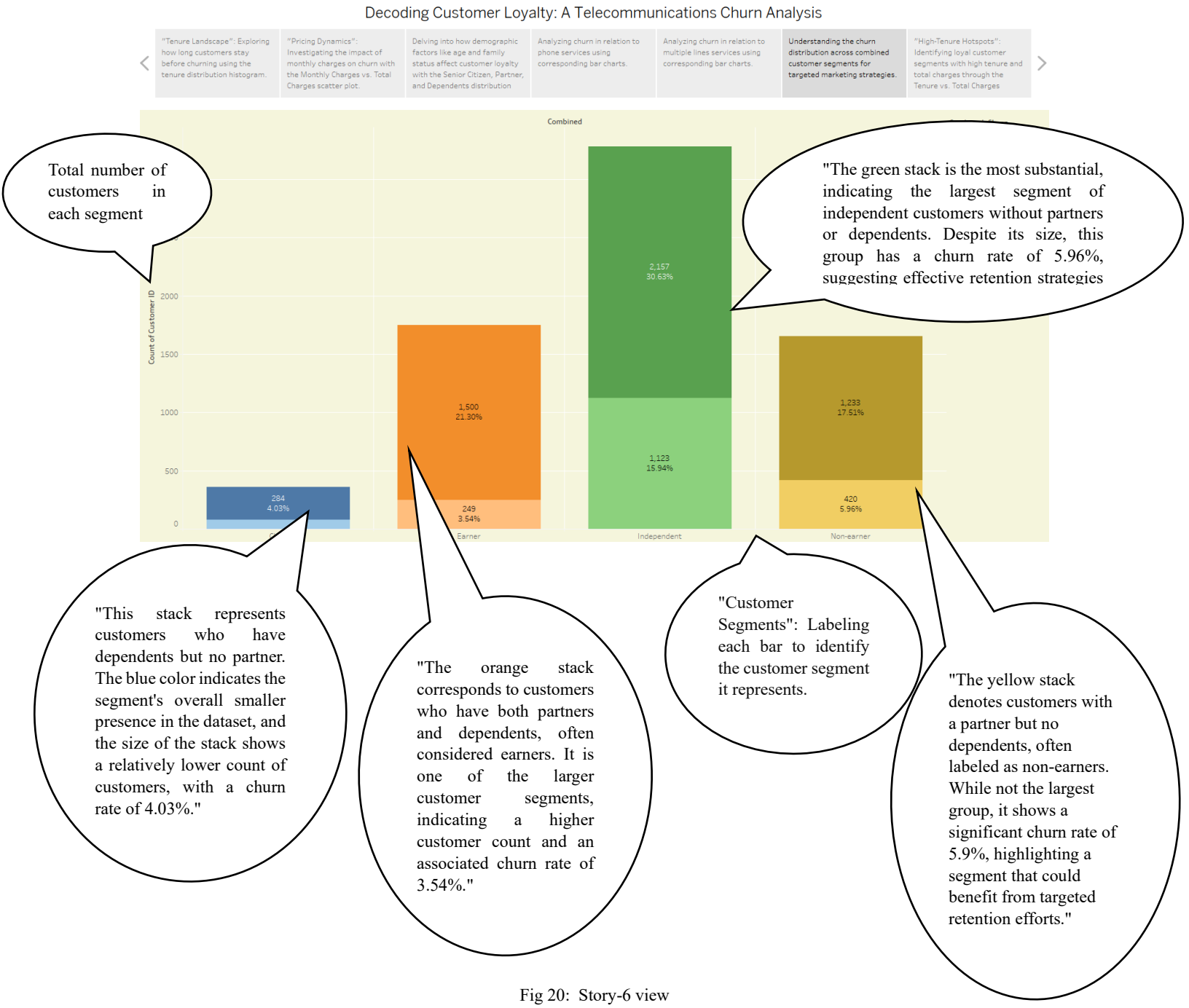


Fig 20: Story-6 view

This composite view focuses on the churn distribution across combined customer segments, using a bar chart to highlight which segments are more prone to churn, thereby enabling more focused marketing and retention efforts.

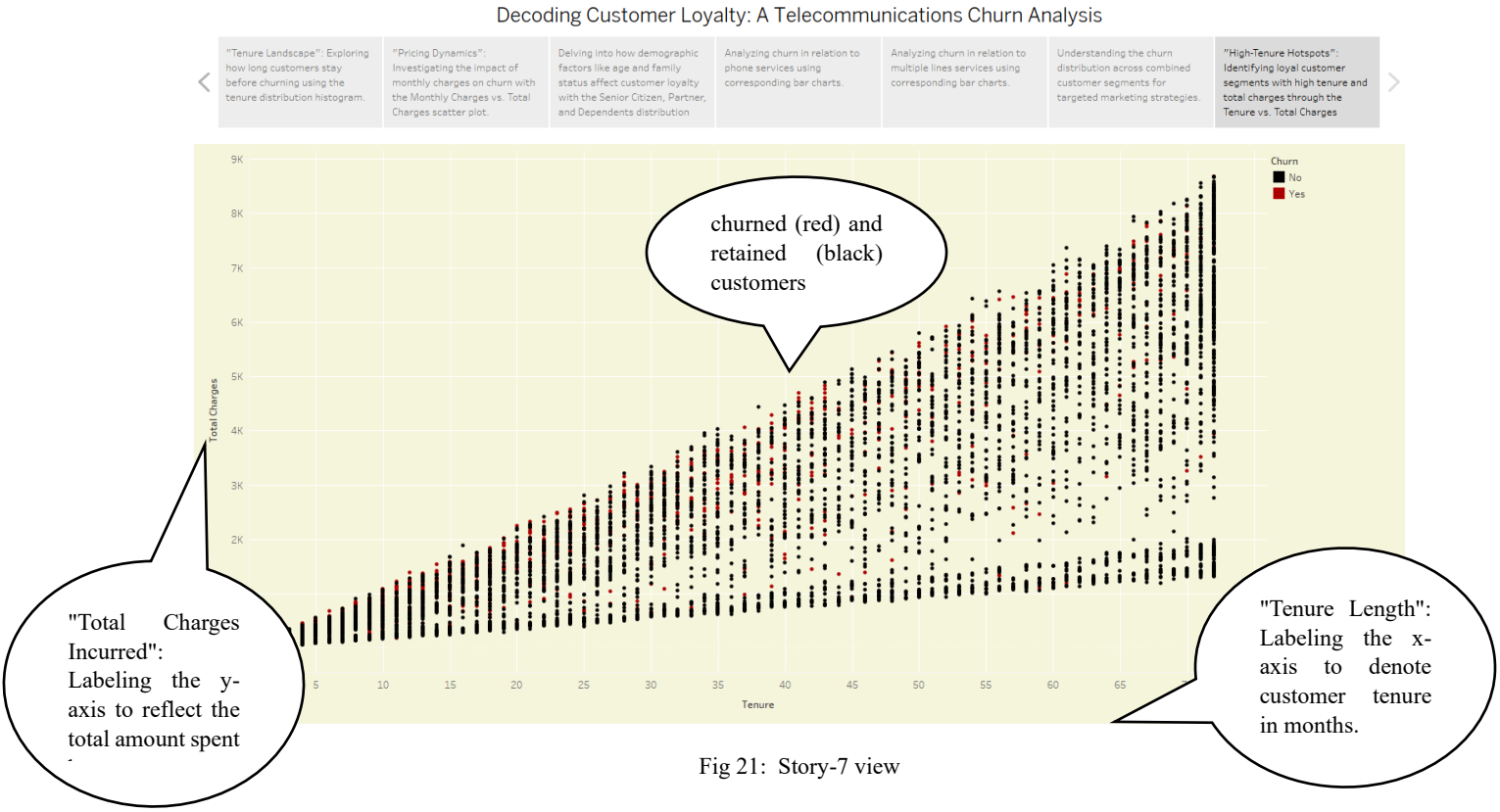


Fig 21: Story-7 view

Identifying segments with high tenure and total charges, this view points out our loyal customer base, offering an opportunity to understand and replicate success factors across other segments.

Each of the above views is designed to systematically answer the overarching questions related to customer churn and to empower stakeholders with the knowledge to make informed decisions aimed at reducing churn and fostering customer loyalty.

5. Reflective Discussion

Throughout the process of developing the Telecommunications Customer Churn Dashboard, I experienced firsthand the challenges and rewards of a user-centric design approach. The project underscored the critical nature of having clearly articulated research questions as a foundation. These questions became the compass that guided all subsequent design decisions, from the selection of visualizations to the interactive features integrated into the dashboard.

The journey from concept to execution illuminated the often significant divide between initial ideas and their practical application. Bridging this gap was not a straight path but rather an iterative cycle of creation, feedback, and refinement. Engaging with users and stakeholders provided invaluable perspectives that steered the dashboard's evolution, leading to a tool that not only informs but empowers decision-making.

I gained a deeper appreciation for the selection of visual encodings and chart types. This experience cemented my understanding that the power of a visualization lies not just in the data it represents but in how intuitively it communicates that data. The subtle art of choosing the right interaction methods and color schemes can make complex data accessible and insights more actionable.

Reflecting on the overall experience, I find the medium of Tableau to be moderately user-friendly, offering a balance between flexibility and structure that, while at times constraining, also provided a scaffolded environment within which to design and implement the dashboard.

6. Conclusion

In conclusion, the Telecommunications Customer Churn Dashboard project has been an enlightening journey into the practicalities of data visualization. Navigating through the stages of data preparation, visualization design, and execution within Tableau has culminated in a comprehensive tool that directly addresses critical business questions regarding customer churn.

Despite the challenges, particularly with data cleanliness and navigating Tableau's limitations, the project has notably advanced my proficiency in data visualization. It has underscored the importance of clear, actionable insights in data-driven decision-making.