

# Insights from Value-based Segmentation of Telecom Customer Data to Enhance Retention

## Executive Summary

### Introduction

The nature of today's telecommunications market is highly competitive, and customers are more informed than ever. This has given customers the ability to shop for telecom services more freely, moving from company to company to keep their costs down, while maintaining the services that are important to them. This is good for the customer, as it gives them the ability to shop around the market for highly customizable plans that fit their unique needs. However, this presents a challenge to telecom companies, as they struggle with customer retention due to the highly varied and unique needs of their customer base. This makes it imperative that telecom companies understand their customer base and can implement effective retention strategies to sustain growth and profitability. In this report, a value-based approach is used to segment and profile the telecom company's customer base to gain insights, that can support customer retention efforts.

Two approaches were initially considered for this segmentation analysis. A value-based approach and a decision tree model approach. Ultimately, the value-based approach was selected to complete analysis, as this method was more intuitive in terms of interpretability and allowed for more user control and customization. However, some analysis of the tree-based approach is also included in this report.

The following sections provide a brief overview of the methodology, and key findings of both approaches, as well as recommendations regarding customer retention.

### Decision Tree Approach

The decision tree approach focused on customer tenure with the company as the target variable and used selected features (Gender, Age, Years of Education, Household Income, Total Debt, Last Month Card Spending, Last Month Voice Spending, Last Month Data Spending, and Last Month Equipment Spending) in the dataset to segment customers into high, medium, and low retention levels. In addition to this segmentation, a list of each of the variable's importance values was also calculated.

### *Key Findings*

- All groups, high, medium, and low retention showed even distribution with 1849 customers in the high group, 1433 in the medium group, and 1599 in the low group.
- Amount spent monthly on voice services was determined to be the most important feature in determining retention. Customers that spent a lot of money on voice services tended to have a high retention level.
- The second most important feature was monthly spending with a card. However, this did not necessarily correlate directly with retention. Higher card spending did not correlate with higher retention. This may indicate that this feature had significant interactions with other features in the set.
- Age also appeared to play a factor in customer retention, as older customers tended to stick with the company instead of moving around. The average age of the high retention group of customers was 59 whereas the low retention group had an average age of 35.

## Value-based Approach

The value-based approach focused on three variables (Phone company tenure, total monthly spending with company, number of premium services) believed to be critical in understanding retention of the customer base. Variables underwent appropriate engineering steps to be used in segmentation analysis. The three variables were then segmented into two categories, either high or low, based on median value. A total of eight profiles were created (three variables, each with two segments). The profiles were then given names based on the characteristics of the profile. Distribution of customers amongst the profiles was studied to ensure meaningful distribution. Finally, analysis was performed on the profiles using various variables within the dataset to glean meaningful insights regarding customer retention.

### *Key Findings*

- Distribution amongst the customer profiles is fairly even, with the exception of the Savvy Savers and the Essentials Hunters, who are slightly under-represented.
- The Bargain Hunters are the largest represented group in the customer base.
- Older aged generations tended to have higher average household income and less loan default status percentage due to being more established in their careers.
- Younger generations, Gen-Z and Millennials, tended to have higher education levels and higher loan default status percentage, possibly due to student loans.
- Essentials Hunters, Platinum Patrons, and Possibly Platinums all value wireless data services with the Platinum Patrons and Possibly Platinums also valuing premium services.
- The Nomadic Navigators and Savvy Savors also value premium services, but this is likely out of necessity, as they also value cost efficiency.

## Recommendations

With the telecom company only having the budget to invest in retention marketing for two groups, the following are the recommendations of this analysis.

1. The first focus should be to market to our best customers, which are the Platinum Patrons. These customers are loyal, well established financial, unlikely to default on bills, and are willing to spend on basic and premium services. The customers in this group also tend to be older and have had time to establish a strong relationship with the company. It would be good to maintain strong relationships with these customers and remind them why they're using our services.
2. The second focus should be on our Savvy Savers. Once again, these are older customers who have had time to establish a strong relationship with the company. They are also well established financially and unlikely to default on bills. They do value premium services, as they likely need them for their job or other reasons. One area we could improve upon with this customer group, is with wireless data. These customers do like to be savvy about their spending, and while they don't like wasting money, they can't pass up a good deal. This could be a win-win type situation where we can offer these customers a great deal on wireless data and increase our profit and retention in this group by offering deals.
3. While the Platinum Patrons and Savvy Savers are the conservative recommendation, there is a risky option. The Possibly Platinums are a highly educated group that values quality basic services, premium add-ons, and wireless data. There is potential for this group to spend a lot with the company, however, this group has particularly low retention. With this being a younger, higher educated group, they value the latest technologies and upgrades. While them being younger means they haven't quite established financial security yet, we could see good ROI if we provide them with a good experience and the upgrades/technologies they're looking for. This may convince them to begin developing a strong relationship with the company and eventually become Platinum Patrons.

## **Detailed Methodology**

**Note:** The customer dataset has previously undergone cleaning, pre-processing, and engineering (refer to Due Diligence Report for details). Any additional engineering will be highlighted in this section when appropriate.

### **Decision Tree Method**

The first step in creating the decision tree model was to define the target variable and the independent features. The target variable chosen was tenure with the phone company, as the purpose of the model is to determine what factors are the most important for customer retention and segment the existing customer data based on these features. The supporting features of the model were chosen to be used in the model as they were believed to either contribute directly or indirectly to the model's ability to segment customer retention levels. Below is a list of the features used in the model:

- Gender (Encoded to Male = 0 and Female = 1)
- Age
- Years of Education
- Household Income
- Total Debt
- Card Spending Last Month
- Voice Spending Last Month
- Equipment Spending Last Month
- Data Spending Last Month

After establishing the target variable and the features the model was trained on the data. Then, the model was used to "predict" tenure, or retention, of each customer in the dataset and categorized the tenure into either a low, medium, or high segment (low is 0 to <=24 months, medium is 25 to <=48 months, and high is 49+ months). Finally, tree was pruned using cross-validation and list of feature importances was produced.

### **Value-Based Method**

Three critical variables were considered for segmentation and profiling of the customer base. The variables considered, along with a brief description is provided below:

- Phone Company Tenure: How many months the customer has been with the company.
- Total Monthly Spending: How much the customer spends with the company on voice, data, and equipment services monthly (based on the customer's spending last month).
- Number of Premium Services: How many premium services the customer uses (Multiline, Voicemail, Pager, Internet, Caller ID, Call Waiting, Call Forwarding, and Three-Way Calling).

Prior to segmentation, the "Number of Premium Services" variable had to be engineered. This was done by summing the Multiline, Voicemail, Pager, Internet, Caller ID, Call Waiting, Call Forwarding, and Three-Way Calling variables. These columns are represented as binary "Yes" and "No" values. To sum these variables for each row, the "Yes" values were counted across each row and put into the new column "Number of Premium Services".

Next, segmentation of each variable was performed. This was done about the median of each variable, as median represents true middle and isn't sensitive to outliers. The median for "Phone Company Tenure" is 38 and was segmented: >38 is high and <=38 is low. The median for "Total Monthly Spending" is 49.85 and was segmented: >=50 is high and <50 is low. There is a total of eight premium services where >=4 is high and <4 is low. It was chosen to include 4 as high to even out the distribution of high to low premium service customers, as the majority of customers have less than 4 premium services.

Following segmentation was the creation of profiles. The eight combinations of high and low variable segments were given names based on their characteristics. A guide to the combinations and their respective names is provided below:

(Phone Company Tenure, Total Monthly Spending, Number of Premium Services)

- High, High, High = Platinum Patron
- High, High, Low = Essentials Enthusiast

- High, Low, High = Savvy Saver
- Low, High, High = Possibly Platinum
- High, Low, Low = Frugal Follower
- Low, Low, High = Nomadic Navigator
- Low, High, Low = Essentials Hunter
- Low, Low, Low = Bargain Hunter

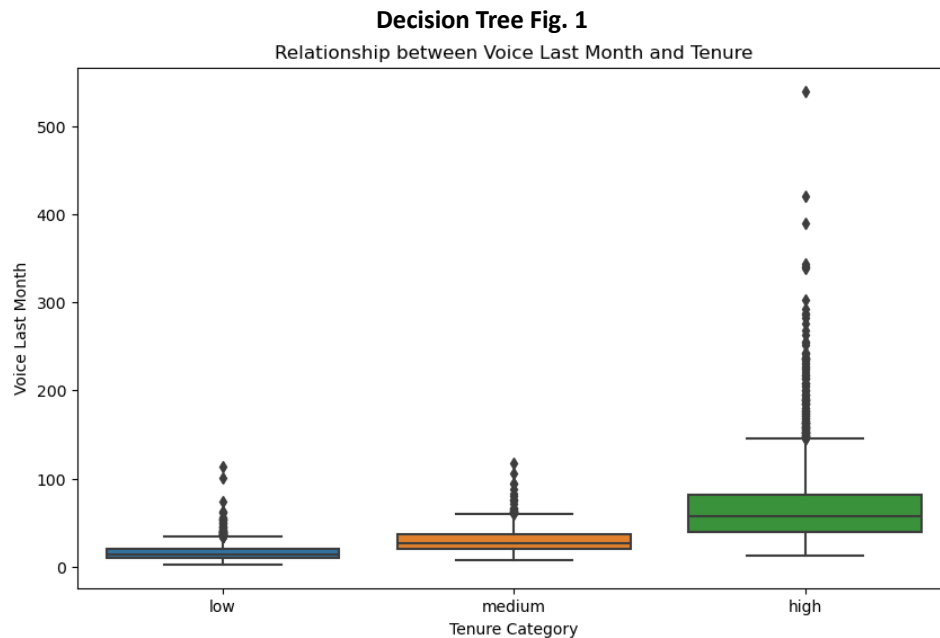
## **Detailed Findings**

### **Decision Tree Approach**

The distribution of customers amongst the low, medium, and high segments were fairly even at 1433, 1599, and 1849 respectively. List of feature importance values provided below:

- Gender: 0.042
- Age: 0.136
- Years of Education: 0.109
- Household Income: 0.132
- Total Debt: 0.140
- Card Spending Last Month: 0.152
- Voice Spending Last Month: 0.187
- Equipment Spending Last Month: 0.057
- Data Spending Last Month: 0.046

Customer spending on voice services monthly was the most important feature in determining customer retention.



Boxplot of spending on voice last month by level of tenure shows that the high tenure group spends the most on voice services.

Card spending last month also appeared to be an important factor for the model when segmenting tenure groups. However, there does not appear to be a direct relationship between card spending and tenure, which

indicates that it's the complex relationships card spending has with other features in the model that deemed it important in segmentation.

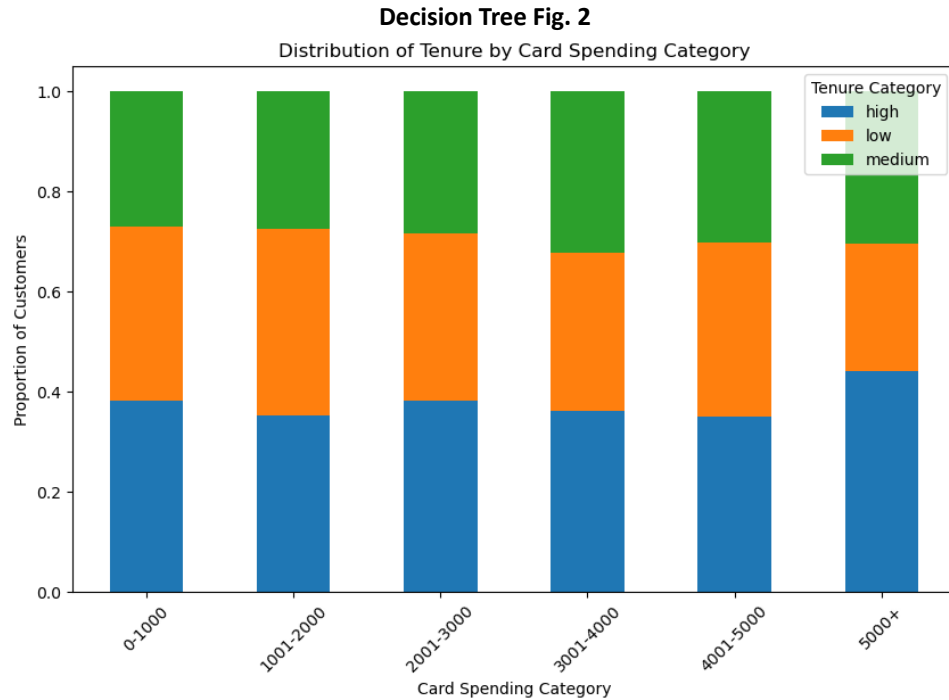


Figure 2 showing the distribution of card spending by tenure category. There does not appear to be a significant direct relationship between last month card spending and tenure category.

Finally, age appeared to factor heavily on the model's decision-making during segmentation. Average age amongst the low, medium, high segments was 35, 46, and 59 years old, respectively. Older customers tended to have higher retention level where younger individuals appeared to come and go more frequently.

According to the model, older customers who spend on the higher end for voice services monthly tended to have the best retention. This could be due to long-standing relationships older customers have built with the company, and communication needs of the older customers may not exceed much further than voice services. Younger customers who perhaps care more about data services may be more likely to continuously shop around for the best deals on newer technologies and have also not had the time to develop long-standing relationships with the company.

### Decision Tree Fig. 3

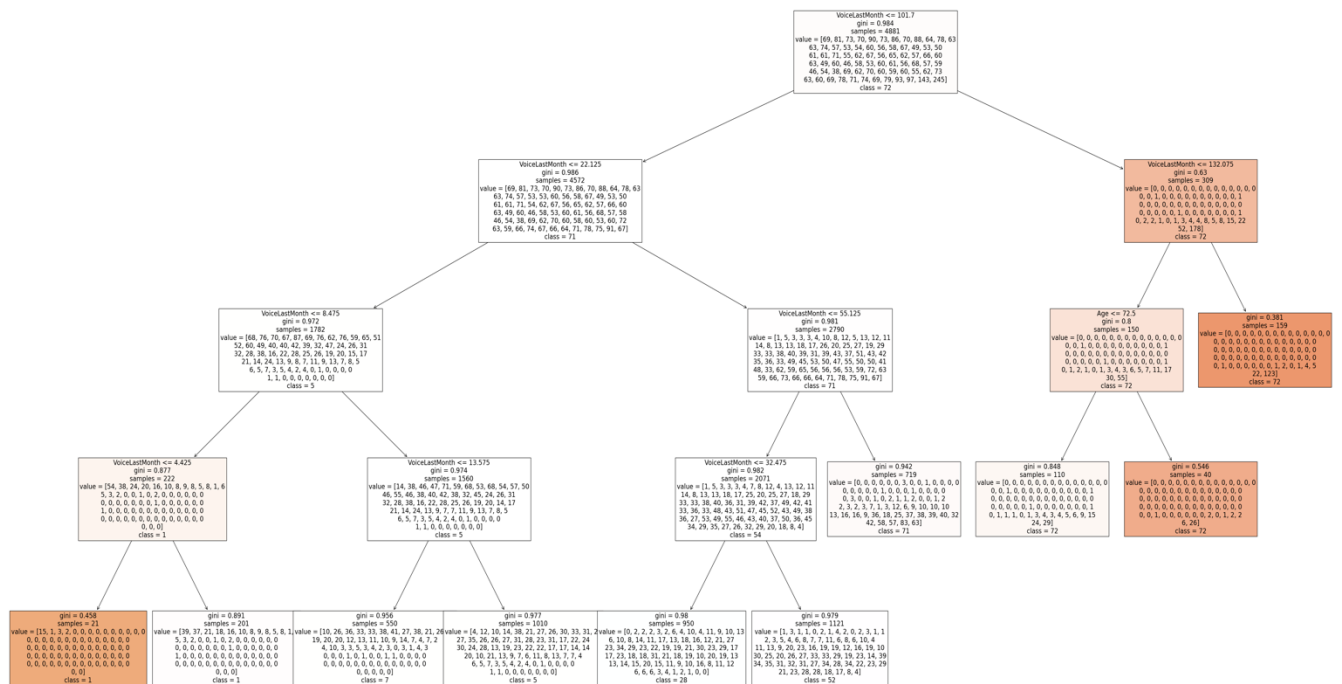


Figure 3 showing the model's emphasis on the importance of voice spending and age.

## Value-Based Approach

**Value-Based Fig. 1**









Customer Segment	Description	Avg HHIncome		
<b>Bargain Hunter</b>	A younger group of customers with a low level of income and low level of education. Looking for the best deals.	\$38,135		1,197 Customers 24.52% of Total
<b>Essentials Enthusiast</b>	An older, loyal customer but does not have as much income to spend on premium services. However, has no issue spending a lot on the basic services.	\$55,944		785 Customers 16.08% of Total
<b>Essentials Hunter</b>	A younger, highly educated group of customers with middle-of-the-road income. They want value and basic services.	\$50,121		247 Customers 5.06% of Total
<b>Frugal Follower</b>	An older, lower income customer, with a lower level of education. They are very loyal and only require the bare minimum to save on costs.	\$48,244		541 Customers 11.08% of Total
<b>Nomadic Navigator</b>	A younger, lower income customer group, with a lower level of education and a need for premium services.	\$43,442		475 Customers 9.73% of Total
<b>Platinum Patron</b>	An older, loyal customer with high income and is willing/able to spend on basics and premium services.	\$85,460		866 Customers 17.74% of Total
<b>Possibly Platinum</b>	A newer, younger, and highly educated group of customers that are willing to spend big on basic and premium services.	\$58,112		535 Customers 10.96% of Total
<b>Savvy Saver</b>	An older, loyal customer with the second highest avg. household income and a need for premium services.	\$63,583		235 Customers 4.81% of Total

Figure 1 displaying an overview of the profiles, including a brief description, average household income, and count of customers in each profile.

The distribution of customers amongst the profiles is fairly even, although there is thinner representation in the Savvy Saver and Essentials Hunter groups. The Platinum Patrons are the highest earning profile with an average household income of \$85,460. The Bargain Hunters represent the largest portion of the company's customer base, with an average household income of \$38,135.

## Bargain Hunter

Typically, younger customers with a low level of income and low level of education. They look mostly for the best deals on only the services they need to get by.

Value-Based Figure 2

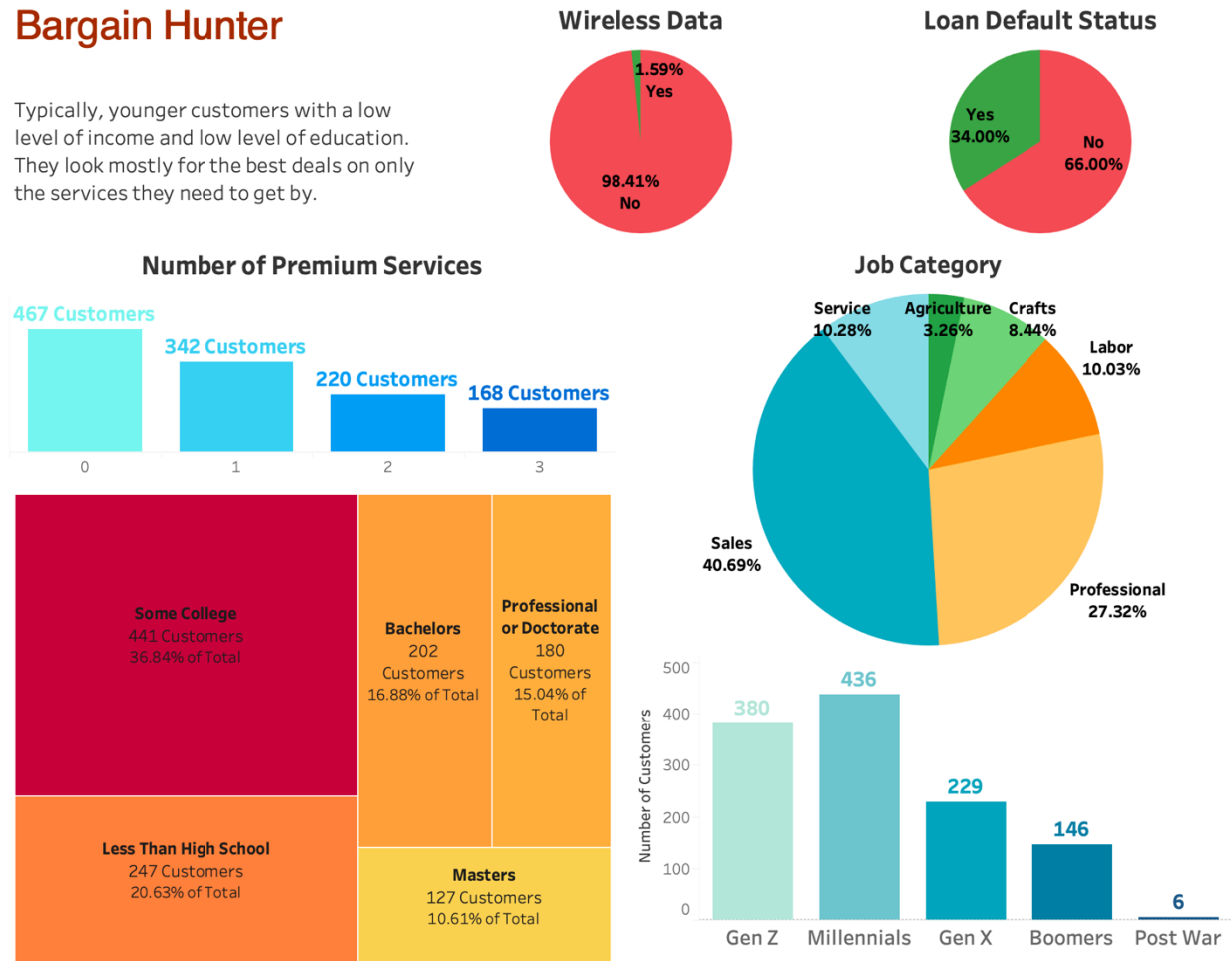


Figure 2 displaying the Bargain Hunter dashboard.

This group of customers is less financially stable with an average household income of \$38,135 and a loan default status of 34%. They tend not to subscribe to many premium services and don't use wireless data. They mainly work in sales. This group consists of mainly younger generations, primarily the millennials. This group also has the lowest percentage of bachelor's degrees, amongst the younger customer profiles, at 16.88%.

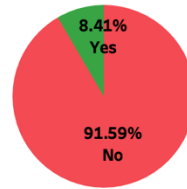


Value-Based Figure 3

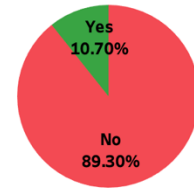
## Essentials Enthusiast

An older, loyal customer but does not have as much income to spend on premium services. However, has no issue spending a lot on the basic services.

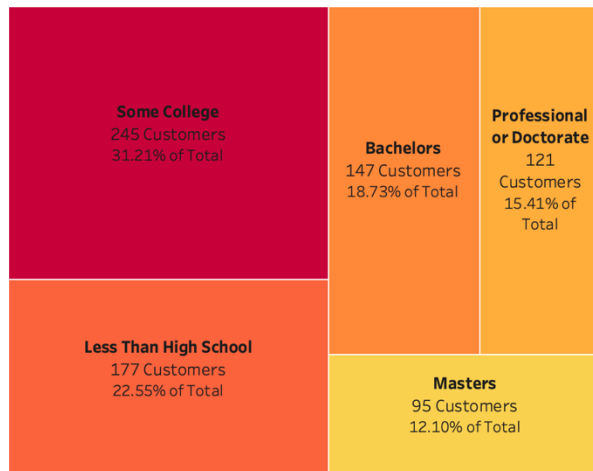
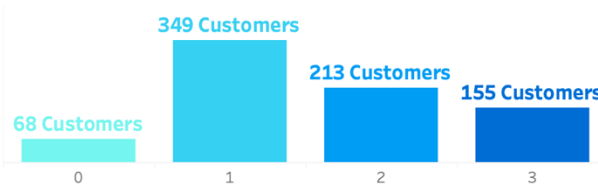
Wireless Data



Loan Default Status



Number of Premium Services



Job Category

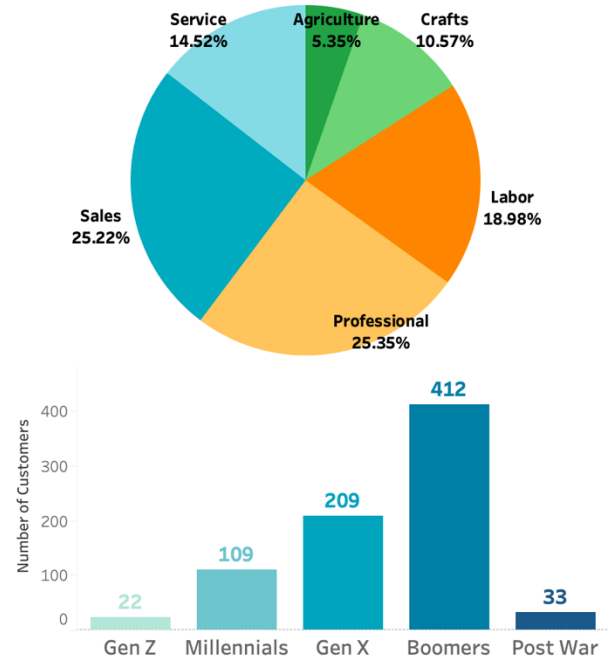


Figure 3 displaying the Essentials Enthusiast dashboard.

This customer group earns a moderate average household income of \$55,944 yearly and is responsible with their loans with a 10.7% default percentage. They tend to only be interested in the essential services and subscribe to few premium services and not to wireless data. The majority of this group has had at least some college education and either work in sales or professional careers. Interestingly, this profile also has one of the highest labor type career percentages. The group is comprised mostly of older generations, primarily the Boomers.

Value-Based Figure 4

## Essentials Hunter

A younger, highly educated group of customers with middle-of-the-road income. They place high value on basic services and most likely shop around for the best deals for these services.

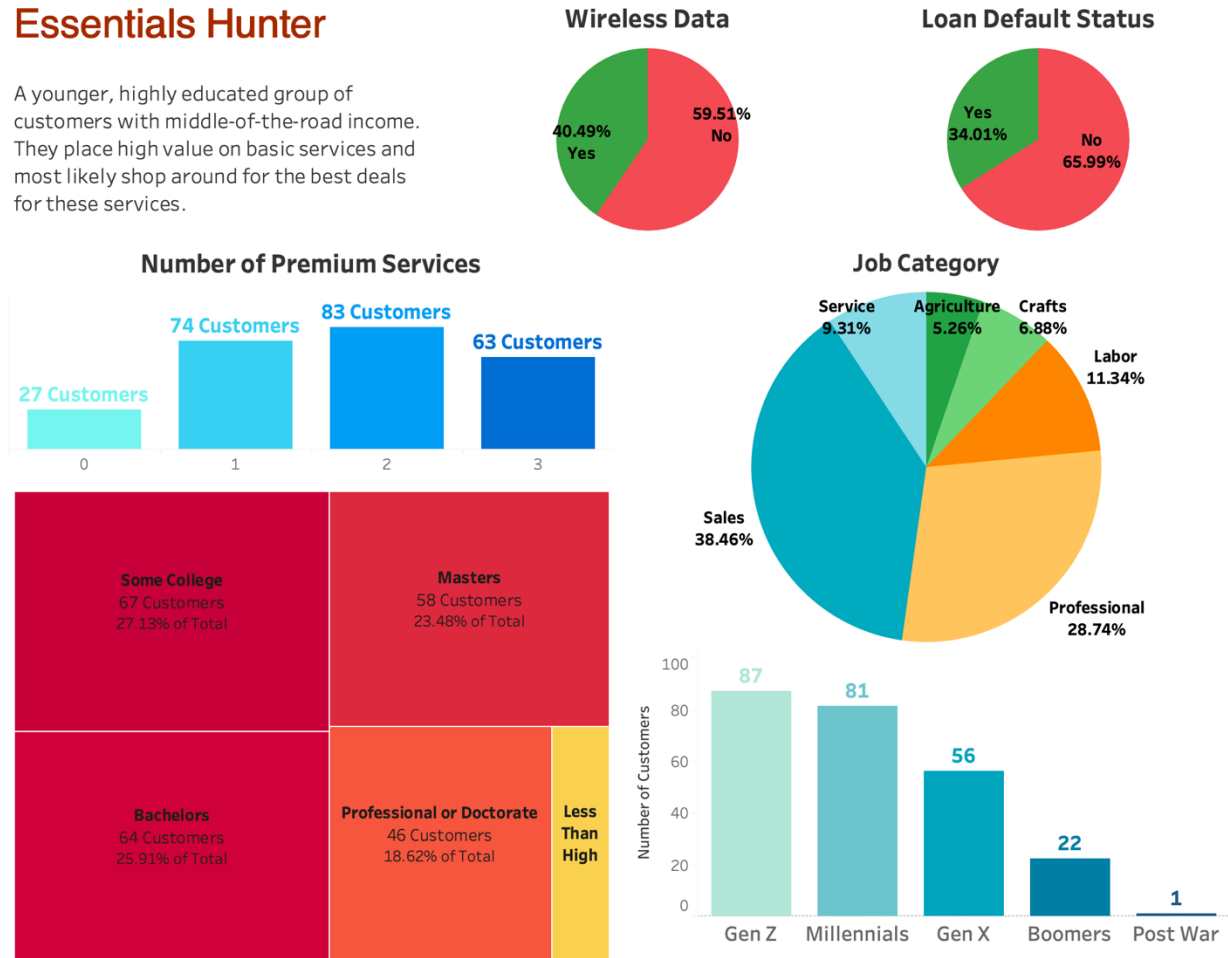


Figure 4 displaying the Essentials Hunter dashboard.

This customer group earns a moderate average household income of \$50,121, however, struggle with financial responsibility with a loan default status percentage of 34.01%. This is a relatively educated group of customers, most with either bachelor's or master's degrees. They're less interested in premium services but do value wireless data. This group of customers typically works in either sales or in professional careers and is comprised mainly of Gen-Z and Millennials.

## Frugal Follower

An older, lower income customer, with a lower level of education. They are very loyal and only require the bare minimum to save on costs.

Value-Based Figure 5

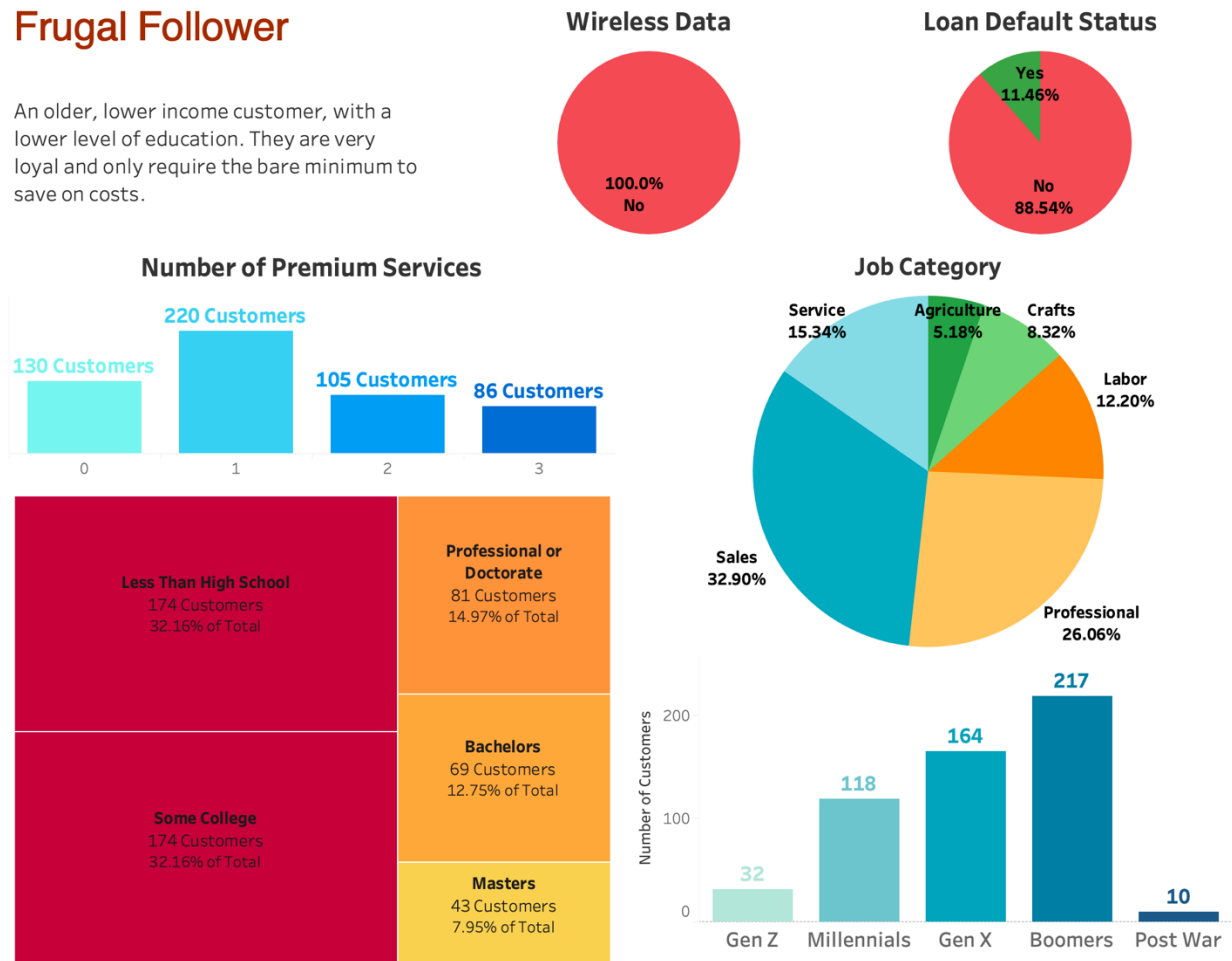


Figure 5 displaying the Frugal Follower dashboard.

This customer group has a moderately lower average household income of \$48,244, but good loan default status percentage at 11.46%. This type of customer is clearly not interested in any premium services or wireless data. This profile has the highest percentage of “Less Than High School” education level and works primarily in sales. Other than the Savvy Savers, this group has the highest percentage of individuals working in service. Gen-x and. Primarily Boomers comprise this customer group.

## Nomadic Navigator

A younger, lower income customer group, with a lower level of education. This customer most likely requires premium features for work or lifestyle reasons. They likely frequently search for the best deals ..

Value-based Figure 6

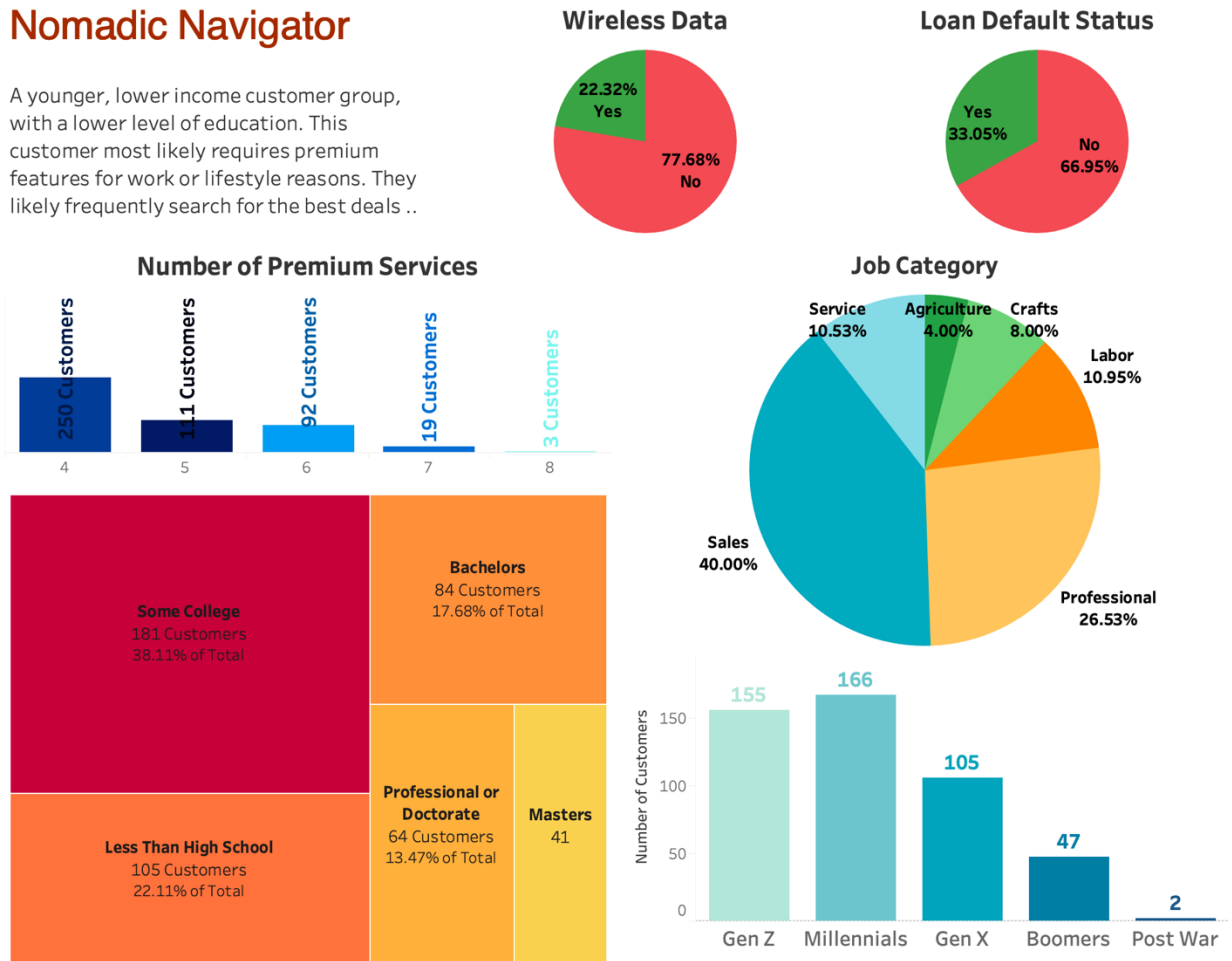


Figure 6 displaying the Nomadic Navigator Dashboard.

This customer group earns on the lower end of the average household income spectrum at \$43,442 and has a high loan default status percentage. This group is interested in premium features, having at enrolled in at least 4 but are not interested in wireless data, at only 22.32%. The majority of this profile has at least some college education and work primarily in sales. This group consists mainly of younger generations, Gen-Z and Millennials.

Value-Based Fig. 7

## Platinum Patron

An older, loyal customer with high income and is willing/able to spend on basics and premium services.

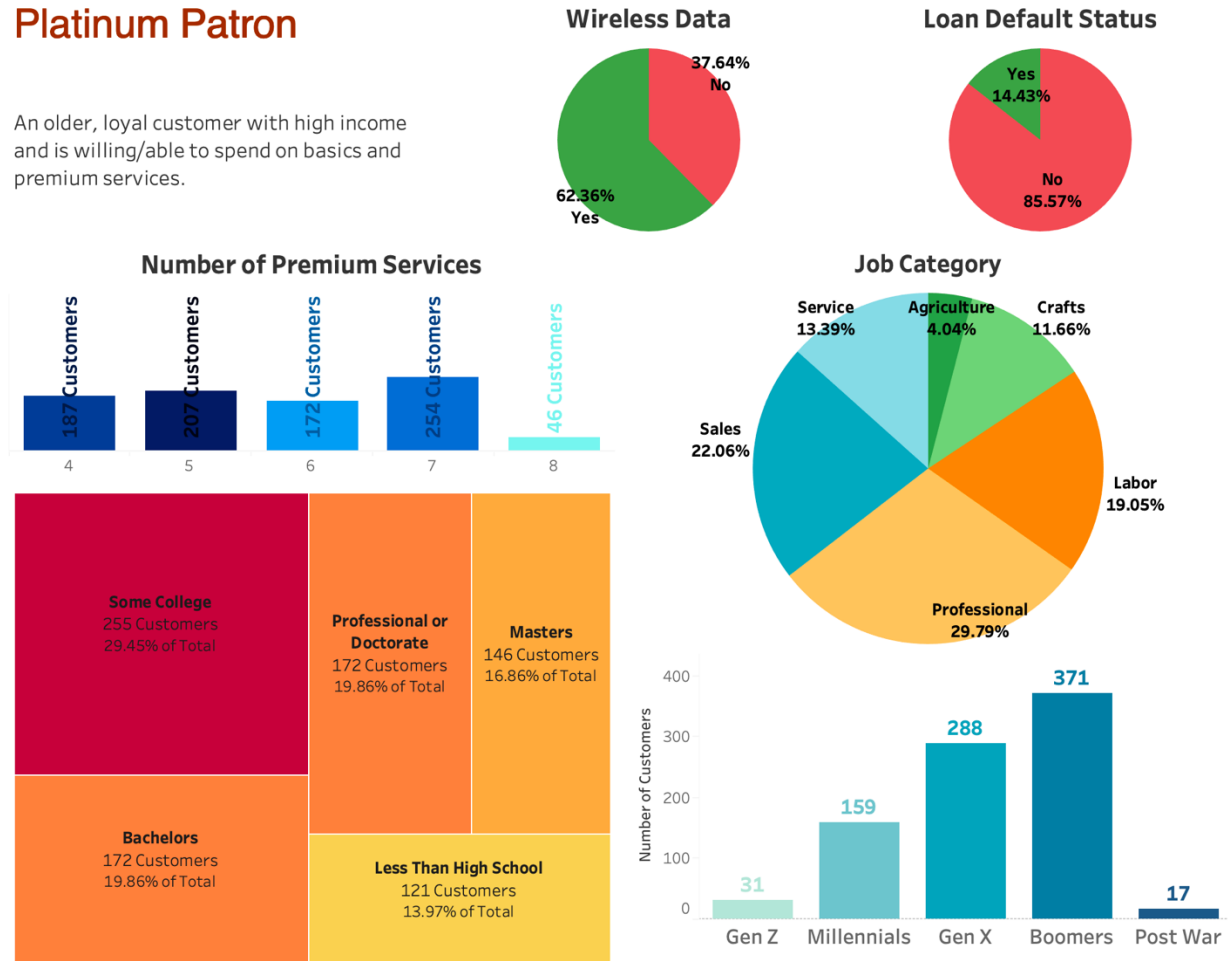


Figure 7 displaying the Platinum Patron dashboard.

The Platinum Patron is a financially stable group, as can be seen with the low loan default status and the \$85,460 average household income. These customers also tend to have at least some college education and work in professional careers. The Platinum Patrons enjoy premium services, having subscribed to at least four of our premium services and values wireless data. Interestingly, this profile has the highest percentage of labor workers. This profile also tends to be on the older side, with the Boomers making up the majority of this profile.

Value-Based Fig. 8

## Possibly Platinum

A newer, younger, and highly educated group of customers that are willing to spend big on basic and premium services. They are frequently searching for the best upgrades and new technologies.

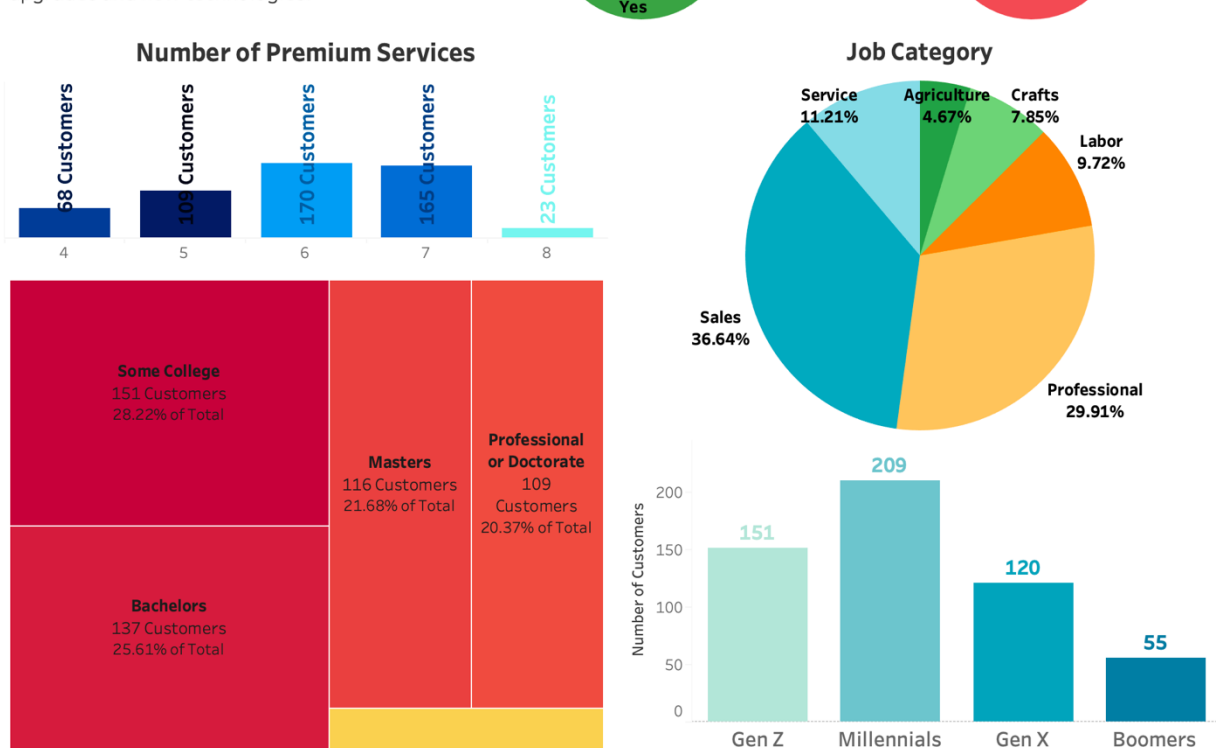


Figure 8 displaying the Possibly Platinum dashboard.

This is moderate earning average household income group at \$58,112 with a high loan default status percentage at 34.95%. This group is interested in premium services, enrolling in at least four premium services and highly values wireless data at 89.16%. This is a highly educated group that works primarily in sales and professional careers. These customers tend to be in the younger generations, Gen-Z, Gen-X, and primarily the Millennials.

## Savvy Saver

An older, loyal customer with the second highest avg. household income. They most likely have a need for premium services due to work or family needs, but they save on costs by cutting back on basic services.

Value-Based Fig. 9

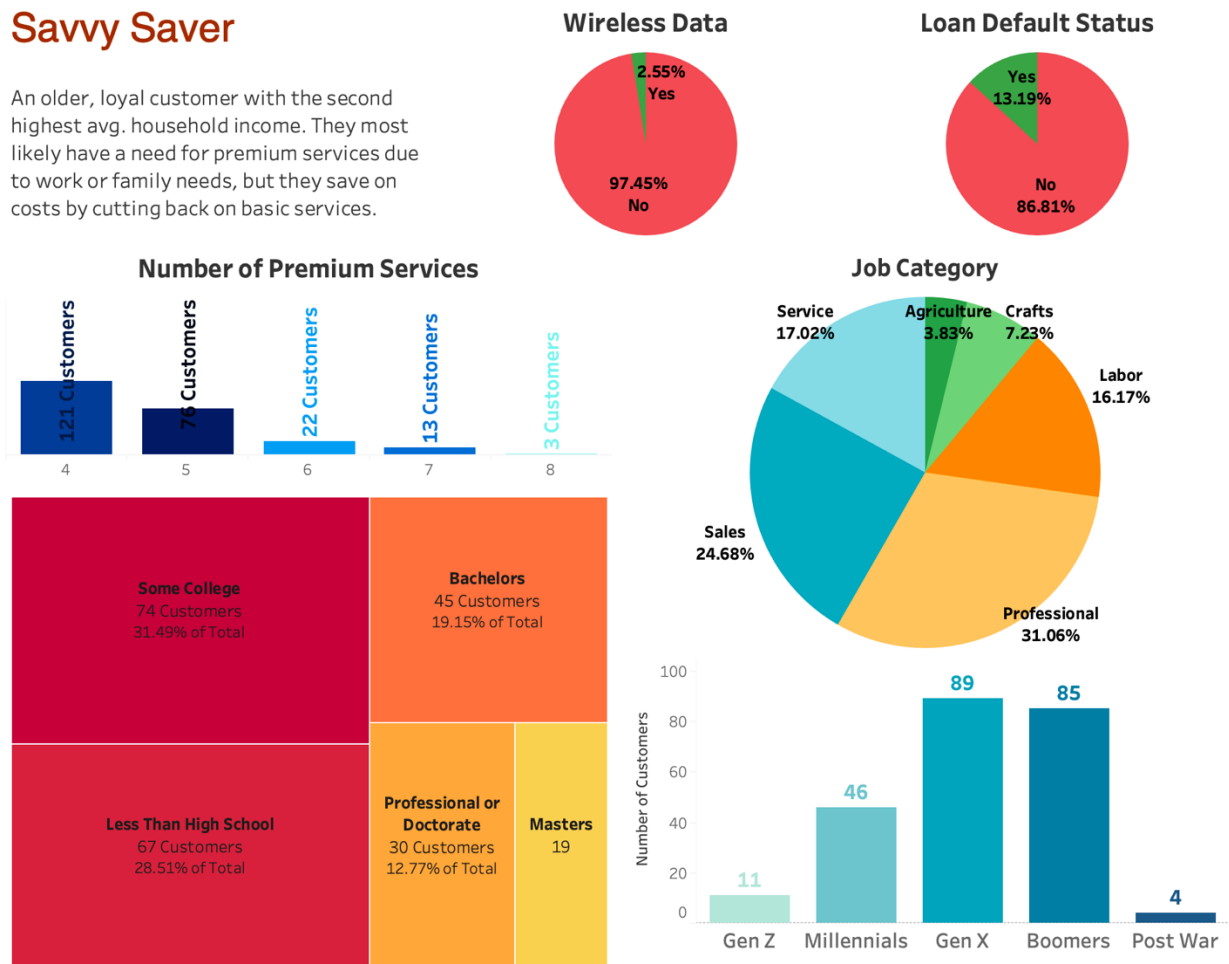


Figure 9 displaying the Savvy Saver dashboard.

The Savvy Savers are on the higher end of average household income at \$63,583 and on the lower end of loan default status at 13.19%. This group does value premium services, likely out of necessity, and are not interested in wireless data services. This group is on the slightly lower end of education level and work primarily in professional settings. They also have the highest percentage of service job representation. This group is mostly middle-aged and older, comprised mainly of Gen-X and Boomers.