Telco Extra Portfolio

Joshua Cray

MIS445 – Statistics in Business Analytics

Colorado State University – Global Campus

Alin Tomoiaga

May 8, 2020

Telco Extra Portfolio

According to Frankenfield, J (2019) churn rate is also known as the rate of attrition, which is also the rate that customers stop doing business with a company. Frankenfield also points out churn rate is typically expressed as a percentage of service subscribers who cancel their subscription within a predetermined time frame. Another example Frankenfield provides in terms of churn rate is the rate at which employees leave their respective jobs in a given time frame. This brings us to the point of this paper which is to recognize the variables which impact the overall churn rate. The point of identifying the variables impacting the churn rate is to reduce the churn rate of the customer base and extend repeat business. The data provided to us makes up approximately half the Telco customers which were selected randomly. Each of the cases is from separate customers varying demographics and service usage information which was purchased from an outside vendor. If Telco can better understand their customer base and why they churn then they have a better chance of determining a way to deter customers from ending their subscription to the services from the company.

**Data Introduction**

Customer churn presents difficulties for companies, especially in the telecommunications field companies are continually looking for ways to foresee customer churn. Given how many telecommunications companies it is a competitive industry. With the industry being so competitive companies have to figure a way to stay ahead of their competitors. Reducing customer churn rate plays a role in this if the company can identify the factors which contribute to customers canceling the services they receive through the company. If these factors can be determined by Telco then they can reduce the customer churn rate and increase the number of returning customers.Telco is trying to determine the variables in their customer base which are playing a role in their customer churn. This is why Telco Extra bought the collected data to in turn enable them to predict the income of their customers. Predicting the customers would allow the company to not have to pay to buy the income data from another outside vendor. The variables contained in the data consist of: age, level of education, agen, income, years at current address, and income. I would expect as age increases the churn rate decreases, as the level of education increases the churn rate should decrease. I would also expect a positive correlation between the age and income variables. There should also be a positive correlation between the variables of income and time at the current address.

**The Data**

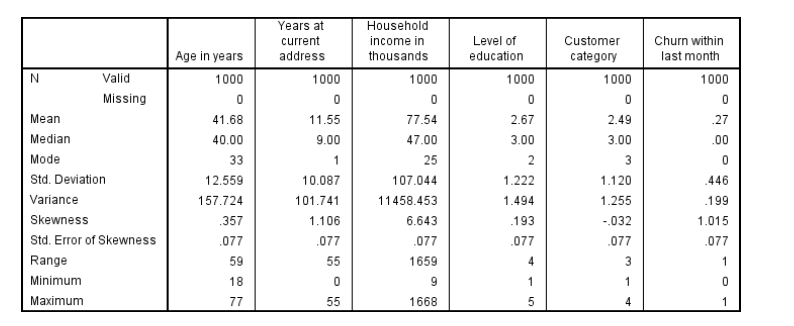
Since churn rate describes the rate at which customers stop subscribing to a company’s services Telco Extra is concerned with slowing this rate to retain customer loyalty to their business. The information contained in this report will aid in Telco’s efforts to reduce the churn rate in their customer base. Through the research this report contains Telco will be able to have a more thorough understanding of the customers who are cancelling their subscriptions to Telco’s services and will also be able to more accurately predict the income of the customers.

In this report population refers to the total of all Telco’s customer base, however the data contained in the provided file is approximately half Telco’s customers who were selected at random.

**Sample Characteristics**

The sample characteristics include age, time at current address, income of home (in 1000’s) educational level, customer category, and last month's churn. These characteristics are shown in the table below.

*Characteristics*



From looking at the characteristics table we can see the company’s customers average age is 42, and the standard deviation is 12.6 units. For age the skewness comes in at 0.357 which is skewed to the right which tells us there are some customers which are 80 years old or older.

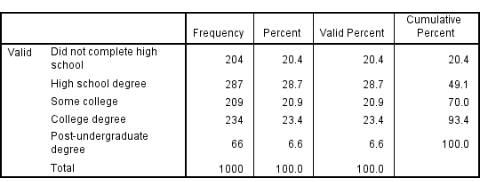
We can also see from this table the average time the customers have lived at their current address is 11.55 years andthe standard deviation for this is 10.08 with a skewness of 1.106, skewed to the right. This indicates there are customers who have lived at their current address for 55 years or more.

The table shows the average income for the household of the customers is 77.54 and since this is in thousands that becomes $77,540, the standard deviation for this shows to be 107.04. For average home income the skewness is 6.643, again skewed to the right which means there are customers with even higher incomes for their household.

Educational level in this data set is broken down to: didn’t complete highschool, highschool degree, some college, college degree, and post-undergraduate degree.

The second table shows the level of education of the Telco Extra customers. Of the sample provided in this data the majority of the customers involved either obtained their high school diploma, had some amount of college or followed through to getting their college degree.

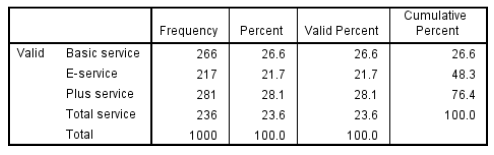
*Level of Education*



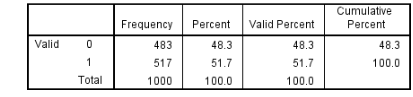
As you can see by looking at the table the majority of the Telco customers have attained their college degree.

The next table is the customer category table.

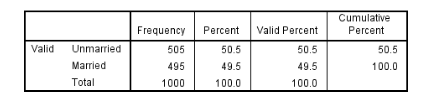
*Customer Category*



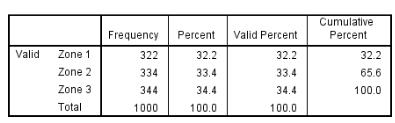
In the gender table male gender is denoted as (0) and female as (1).

*Gender*

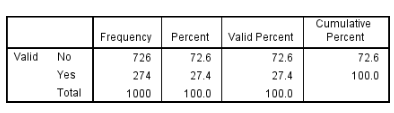
As you can see in the prior table there are marginally more female customers than males which were included in the data pool.

*Marital Status*

There was a slight difference in marital status in the given customer base, slightly less customers included were married.

*Location*

*Last Months Churn*

The percentage of the churn for Telco was only 27.4% in the last month, which leaves the company with a 72.6% of no churn of customers.

Churn Variable

Null Hypothesis: There is not a direct connection between last month's churn and the customer's age.

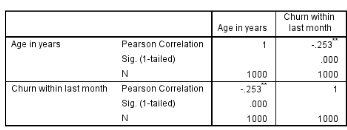
Alternative Hypothesis: There is a negative correlation between the last month's churn and customers' age.

Null Hypothesis: There is not a correlation between last month's churn and educational level of the customer base.

Alternative Hypothesis: There is a positive correlation between educational level and the churn for last month.

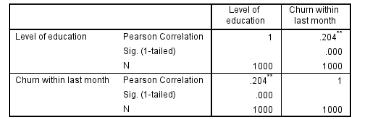
According to Bock, T. correlation is the strength of a linear relationship between two given variables. The correlation between the customer age and last month's churn variables is significant r=-.253, with a p-value of <0.05. This shows a weak negative relationship, so as age increases the churn rate increases. This supports the hypothesis. This information is shown in the table below.

*Education level & Churn*



The analysis of the correlation between customer educational level and last month’s churn rate is provided in the table below. This correlation shows to be significant with a r value of .204 and a p-value of <0.05. This shows a weak positive relationship of the two variables. This shows us as the average customer's education level increases so does the churn rate probability as well. The hypothesis of a negative correlation between the two variables isn’t supported by this.

*Educational Level & Churn*



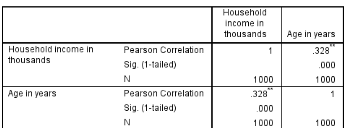
Income-Variable

Null Hypothesis: there is not a significant correlation between the variables of total income of the home and age.

Alternative Hypothesis: There is a significant positive correlation between the home income and customers age variables.

The correlation between the variables of total income of the home and age of the customer is significant. It has an R-value of -.328, a p-value of <0.05. This indicates a weak positive relationship between the two variables. Overall, this tells us the total income of the home increases as the age of the customer increases.

*Total Income & Age*

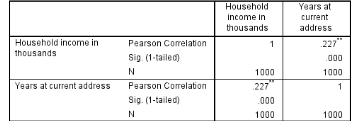


Null Hypothesis: There is not a correlation of significance between the variables of the income of the home and time lived at current address.

Alternative Hypothesis: There is a significant correlation between the variables of overall income of the home and time lived at current address.

The table below shows the correlation of the overall income of the home to the time spent at the customer’s current address. The correlation is significant, the r value is .227, p-value is <0.05, this shows a weak positive relationship between the two variables. So this shows when the overall income of the home increases the time lived as the current address increases as well.

*Total Income & Time at Address*



**Conclusion**

After interpreting the data and comparing the correlations of each variable Telco can better determine why they were experiencing the customer churn they were. From here Telco can see what attributes the customers who are unsubscribing from Telco’s services have to give them a better idea of how to reduce this churn and keep their customers. There were a few variables which we discovered to be associated with the churn rate. These variables with an association included: educational level, overall income of the home, and time spent at current address. Educational level and churn showed customers with higher educational levels had a higher churn. Telco customers with a higher overall income in their home correlated to older customers. The overall time lived at the current address also rose as the income of the household increased. Telco should be able to take this analysis and make further improvements to reduce their customer churn.

Reference

Frankenfield, J. (September 2, 2019). Churn Rate. Retrieved from

<https://www.investopedia.com/terms/c/churnrate.asp>

Bock, T. (N.D.) What is Correlation? Retrieved from

<https://www.displayr.com/what-is-correlation/>