Nhi Le

D207 Exploratory Data Analysis

Western Governor University

March 15, 2023

**A1.  Question for analysis**

Are churn and Item 1, Item 2, and Item 3 on survey dependent? I want to know if the timely manners from the eight-question survey (timely response, timely fixes, and timely replacements) affect churn.

**A2.  Benefit from analysis**

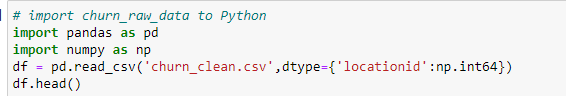
To prevent or reduce churn, the company needs to know what issues they have. By identifying which variables have statistically significant relationship with churn, the stakeholders can understand why customers stopped using the services. Through analyzing the importance of timely factors to customers, the company can decide whether they affect churn and if it is worth it to focus on improving our timely response, timely fixes, and timely replacements.

**A3.  Data Identification**

I will use the variable ‘Churn’, which is a categorical variable including two values ‘Yes’ and ‘No’. The categories ‘Item1’, ‘Item2’, and ‘Item3’ will be used. They are categorical variables having values from 1 to 8. They are the responses from customers on how they think these factors are important to them. Since I am trying to find the dependence among Churn and timely factors, ‘Item1’, ‘Item2’, and Item3’ representing timely response, timely fixes, and timely replacements. I will also utilize variables ‘Bandwidth\_GB\_Year’ and ‘MonthlyCharge’. They are continuous variables for univariate and bivariate analysis.

**B1.  Code**

I will use the chi-square technique.



Graphical user interface, text

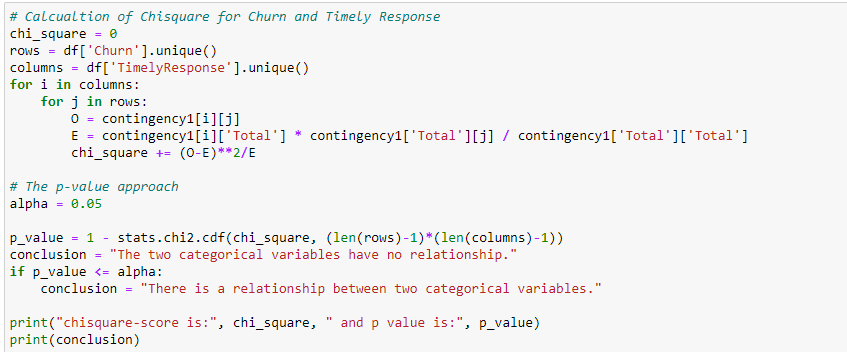
Description automatically generated

Text

Description automatically generated

Graphical user interface, text, application

Description automatically generated with medium confidence



Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

A picture containing application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

**B2.  Output**

The results for Churn and Timely Response with chi-square analysis:



The results for Churn and Timely Fixes with chi-square analysis:



The results for Churn and Timely Replacements with chi-square analysis:



**B3.  Justification**

Since ‘Churn’, ‘TimelyResponse’, ‘TimelyFixes’, and ‘TimelyReplacements’ are categorical variables, I decided to use chi-square test. It is a statistical analysis to identify the relationship between categorical variables. This analysis method helps to determine the difference between the observed and the expected date. Therefore, it will show the correlation between 2 categorical variables (Shinde, 2021). Chi-square analysis will answer the question if Churn and the 3 timely factors are dependent.

**C.  Univariate Statistics**

Continuous variables: MonthlyCharge, Bandwidth\_GB\_Year

Categorical variables: TimelyResponse, TimelyFixes

**C1.  Visual of Findings**

Table

Description automatically generated

Table

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Chart, histogram

Description automatically generated

Chart, histogram

Description automatically generated

Chart

Description automatically generated

Chart, box and whisker chart

Description automatically generated

Chart, box and whisker chart

Description automatically generated

**D.  Bivariate Statistics**

Continuous variables: MonthlyCharge, Bandwidth\_GB\_Year

Categorical variables: TimelyResponse, Churn

**D1.  Visual of Findings**

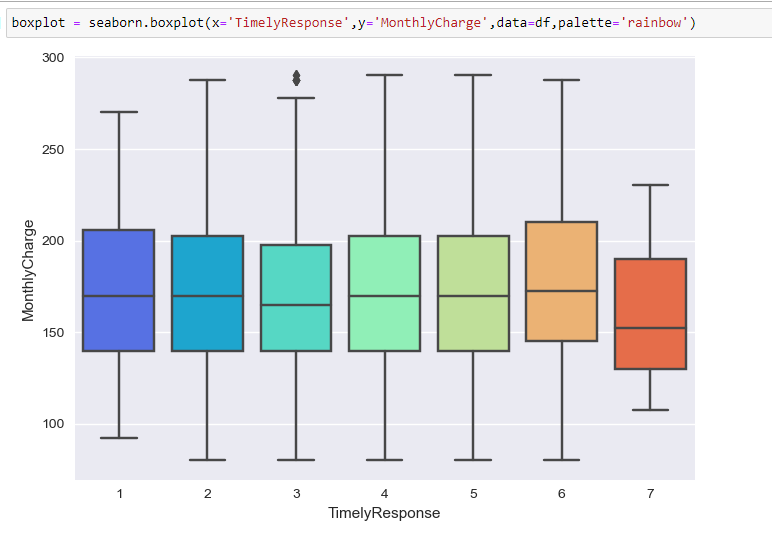
Churn and Timely Response:

MonthlyCharge and Bandwidth\_GB\_Year:

Chart, scatter chart

Description automatically generated

TimelyResponse and MonthlyCharge:



**E1.  Results of Analysis**

The p-value of the chi square test between Churn and Timely Response is 0.6318335816054494. The p-value of the chi square test between Churn and Timely Fixes is 0.5093789499498207. The p-value of the chi square test between Churn and Timely Replacements is 0.6148391285975547. As p-values are higher than 0.05, I cannot reject the null hypothesis. Churn and the 3 timely factors have independent relationship. Therefore, the customers’ responses to Timely Response, Timely Fixes, and Timely Replacements are not statistically significant to Churn and cannot determine if they impact churn.

**E2.  Limitations of Analysis**

The variables used for this test are categorical, which is a great condition for the chi square test. However, it limits to only 2 variables for the analysis at a time. Another limitation is that the chi square analysis can only identify if two variables are related to one another but cannot imply whether that one variable has any causal effect on the other (The University of Utah, n.d.).

**E3.  Recommended course of action**

Since Timely Response, Timely Fixes, and Timely Replacements do not have dependent relationships with Churn, we do not need to prioritize these factors to reduce churn. However, the churn rate was close to 30% and I believe any small action can possibly help to reduce churn. It is still essential to improve customer satisfaction. I would recommend that the company track the time customers need to wait to get our response, and how long we need to spend to fix or replace issues with our services. Then we will have a better observation of the timely factors’ performance.

**F.  Video**

<https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=4e08438e-77f8-42d0-ade4-afc6013e6455>

**G.  Sources for third-party code**

De-Yu, C. (2023, January 5). Chi-square test, with python. Medium. Retrieved March 15, 2023, from https://towardsdatascience.com/chi-square-test-with-python-d8ba98117626

**H.  Sources**

Shinde, Y. (2021, June 25). Chi-square test - use, implementation and visualization. Analytics Vidhya. Retrieved March 14, 2023.

*The chi-square test for Independence*. Chi-Square - Sociology 3112 - Department of Sociology - The University of Utah. (n.d.). Retrieved March 14, 2023, from https://soc.utah.edu/sociology3112/chi-square.php.