**Module 8: Portfolio Project Option#1**

Joseph C. Hand

Colorado State University Global

MIS 445: Statistics in Business Analytics

Prof. Bateh, M.

8 August 2022

**Module 8: Portfolio Project**

In this assignment we are examining the data that has been collected for the company Telco Extra. Telco as the name implies is a telecommunications company that is looks to identify quality information from the data base on their customer characteristics. The company data we are looking at is a fairly large sample of the customer base, about 50% of the customer’s data has been collected. This is a large number for sample and should give enough data to help make the predictions on the customers on how they relate to Churn. Personal Identifiable Information (PII) has been removed from the data so as not expose information that could be used to harm the customers in the data.

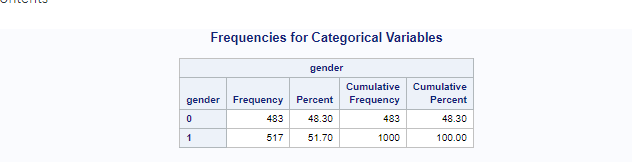
The company would like for us to examine the data to help us understanding of the customers who churn. Churn rate is a measure of the number of customers or employees who leave a company during a given period. (Techtarget,2020.). In this case it will be the customers we want to see who churn. The company suspects that churn rate may have something to do with income of the customer. We would like to predict the income of a customer/customer’s so that the company does not need to spendy money purchasing the income data from an outside vendor. This would help to reduce operating costs that could be put to better use for the customer base.

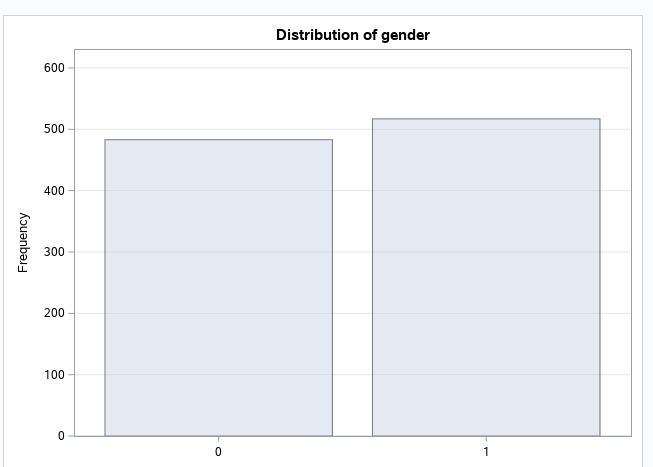
We are going to look at the variables in the data, however we will pay particular attention to the Age, year at current address, gender, level of education, income, marital status, region, customer category, and churn of the respondents in the data collected. We will be using SAS studio to help us with statistical calculations and graphing to give us a visual representation of the data and will help us to understand the numbers from the variables we collected.

1. **Below is the Characterization data from the Telco Extra spread sheet:**

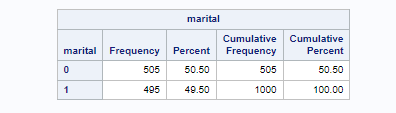
**Characterization TelcoExtra:**

**Gender Characterization:**



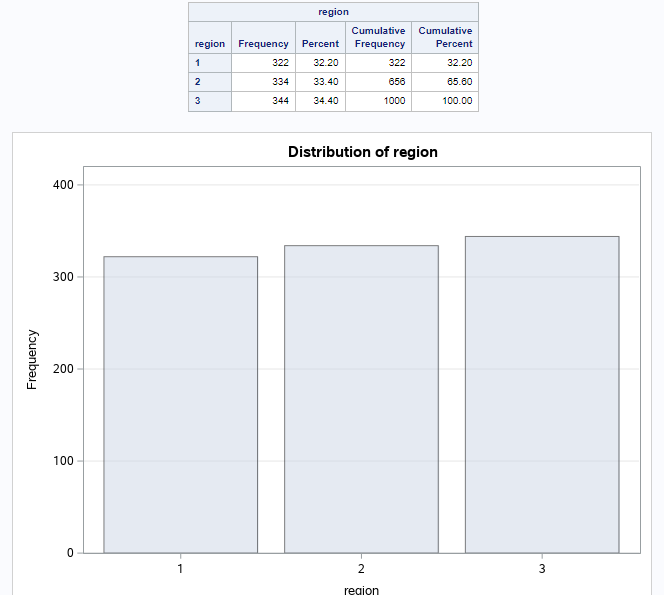


**Marital Characterizatioin Data.**

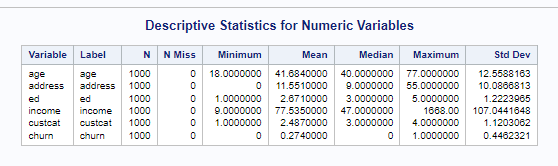




**Region Characterization Data:**

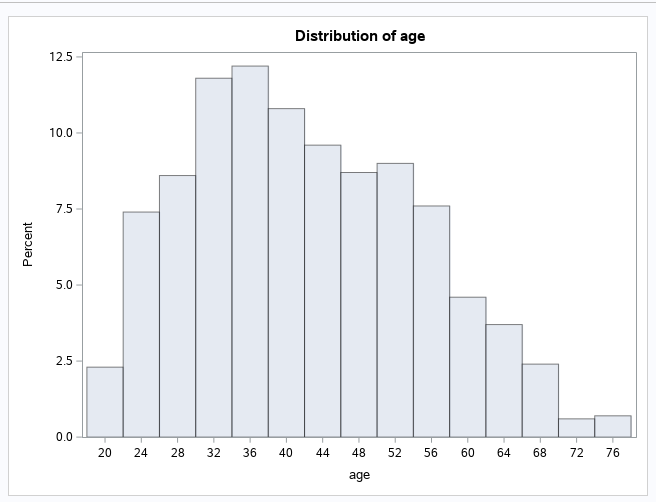


**Descriptive Data: Age, Address, Education, Income, Customer Category, Churn.**

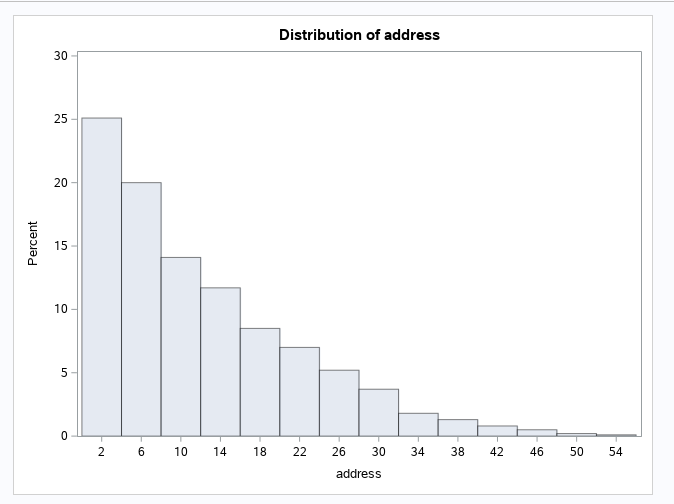


**Graph**

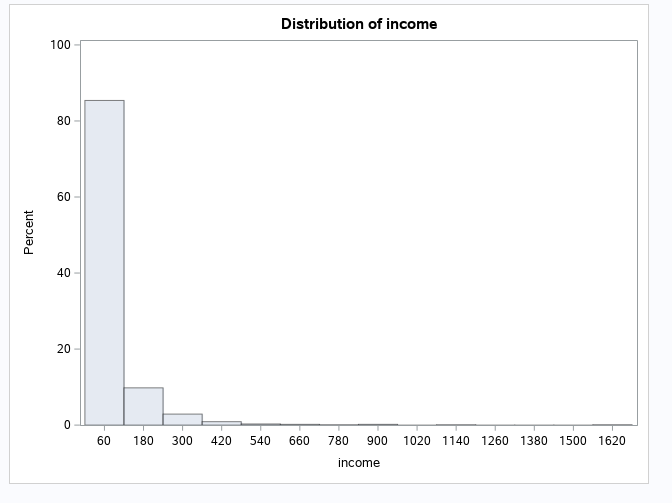
**AGE**



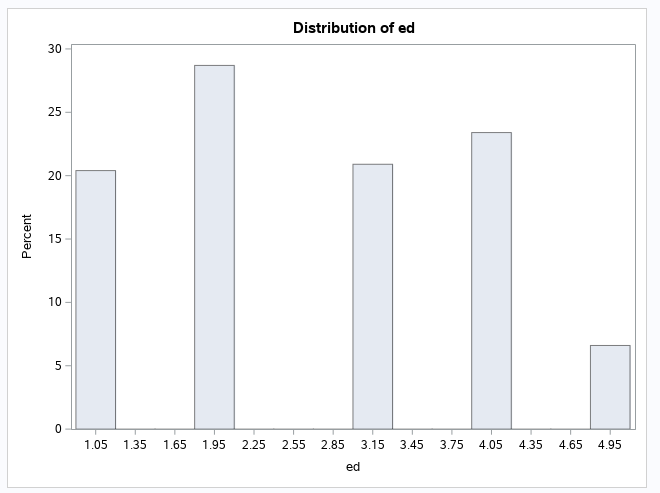
**ADDRESS**

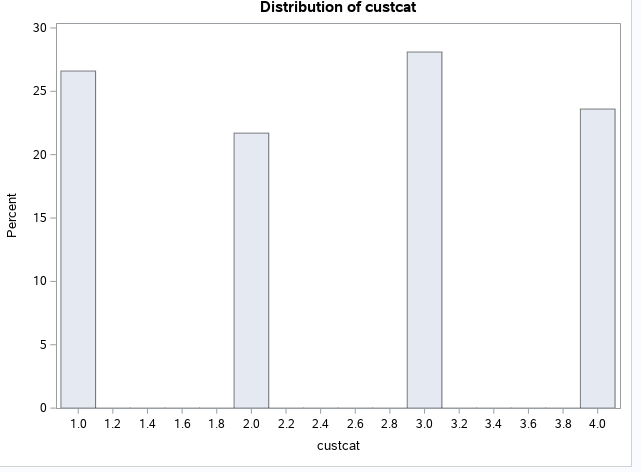


**INCOME:**

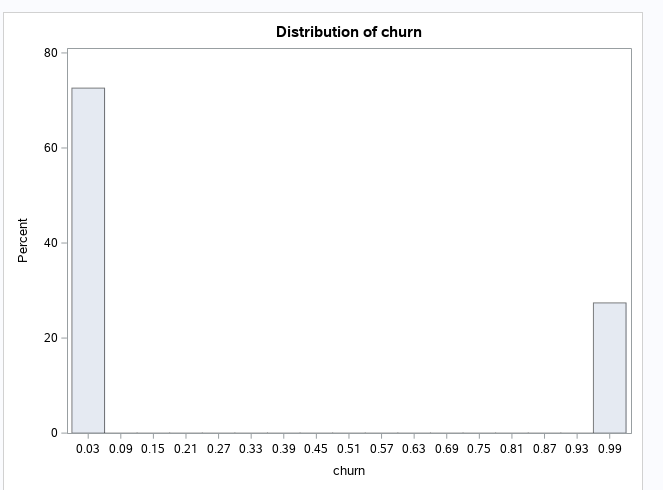


**LEVEL OF EDUCATION:**



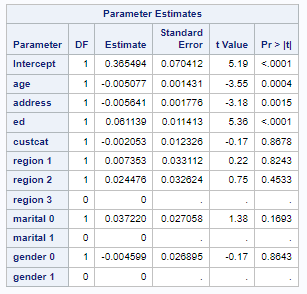


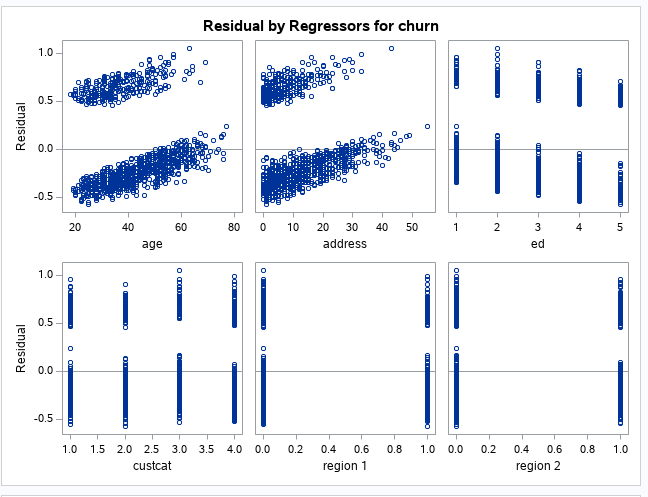
**CHURN**



**Multiple Linear Regression Analysis**

**CHURN AS DEPENDENT VARIABLE:**





**Hypothesis:**

*H0: Age, Gender, Region, Level of Education, Marital Status and Address are not related to CHURN outcome.*

*Ha: Age, Gender, Region, Level of Education, Marital Status and Address are related to CHURN outcome.*

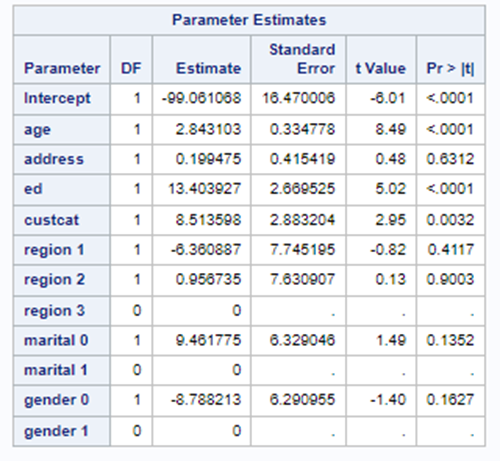
**SAS Output (CHURN DEPENDENT VARIABLE)**

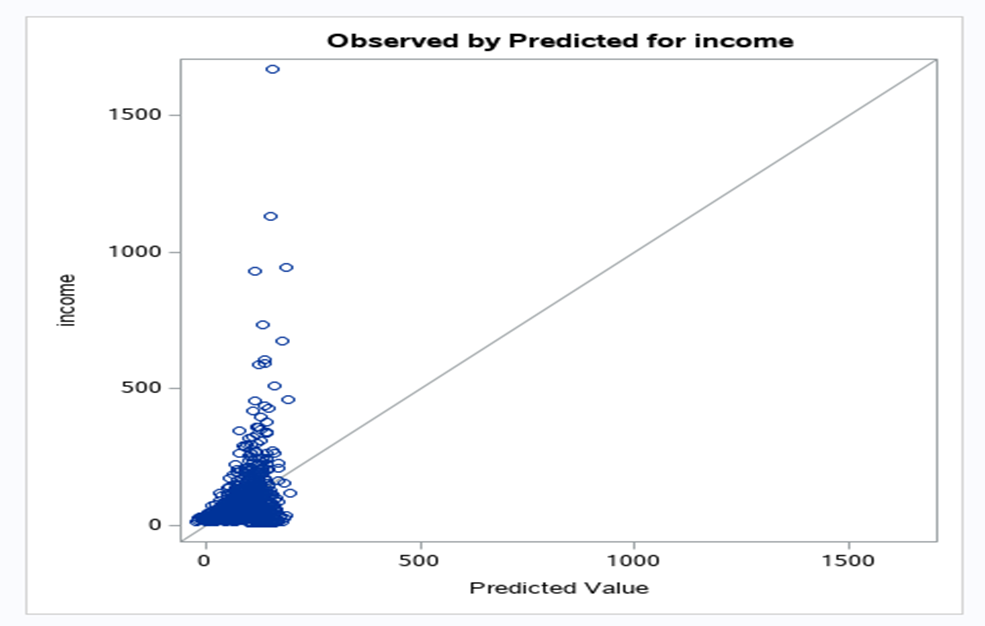
Multiple regression analysis allows researchers to assess the strength of the relationship between an outcome (the dependent variable) and several predictor variables as well as the importance of each of the predictors to the relationship, often with the effect of other predictors statistically eliminated.

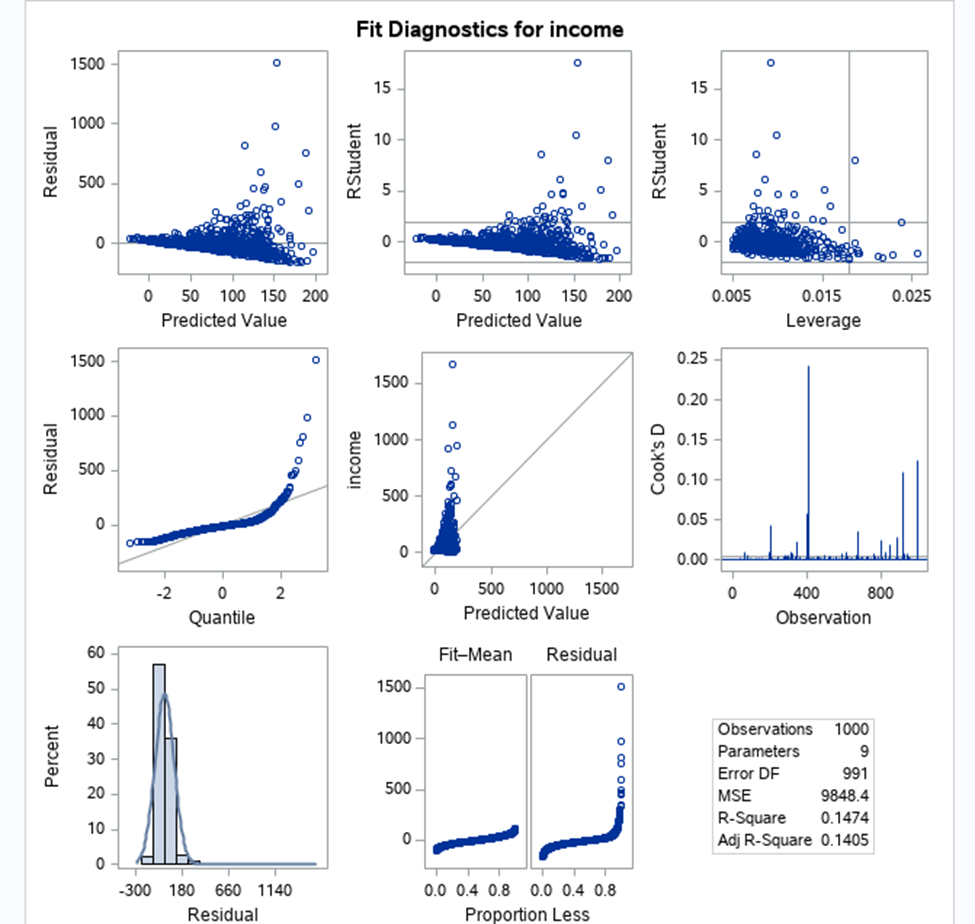
Our data in the first linear regression model that we used, we made Churn the dependent variable and compared the remaining variables as the independent. What we are able to tell from our graphs that were generated is that we see when churn.

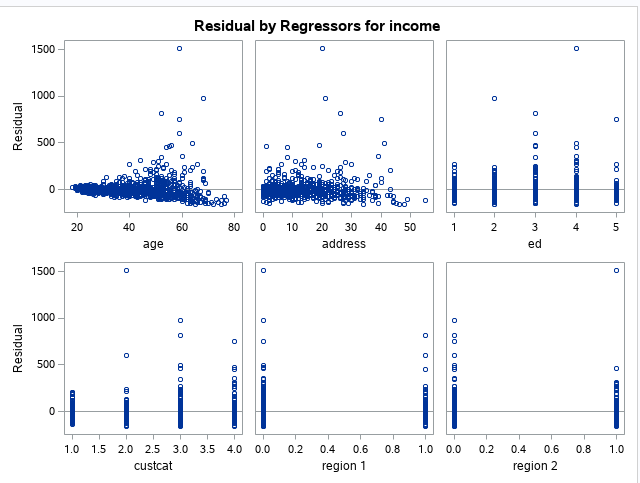
The Parameter Estimates chart above shows that Age, Address, and education have P values les that .05. The churn rates are very close and shows that relationship is strongly correlated with one another. The longer a customer has been at the area they are living it seems they tend to stay with the company.

**INCOME AS DEPENDENT VARIABLE**

****







**Hypothesis:**

*H0: Age, Gender, Region, Level of Education, Marital Status and Address are not related to INCOME.*

*Ha: Age, Gender, Region, Level of Education, Marital Status and Address are related to INCOME.*

**SAS Output (INCOME DEPENDENT VARIABLE)**

The output parameters chart above shows us that the age and education have p-values lower than .05. The graph strength shows there is no strong correlation to income amongst the other variables. The relationship with the variables also shows a negative direction in age and address.

**Conclusion:**

Our data has shown that age and income have a positive correlation. The amount of time someone has been living at the same address shows us that these factors have an influence on churn rates. People who have settled at their address the longest and that have also gained income through time have a lower churn rate. This information may seem intuitive but in order to stay ahead of possible competition, Telco Extra should manage or market to help create newer customers and find ways to maintain them over time.

**References**

Biscobing, J. (n.d.). *What is churn rate? - Definition from WhatIs.com*. SearchCustomerExperience. Retrieved August 5, 2022, from https://www.techtarget.com/searchcustomerexperience/definition/churn-rate#:~:text=Churn%20rate%20is%20a%20measure.

Elliott, A. C., & Woodward, W. A. (2016). *SAS essentials: mastering SAS for data analytics*. Wiley.

Petchko, K. (2018). *Multiple Regression Analysis - an overview | ScienceDirect Topics*. Www.sciencedirect.com. https://www.sciencedirect.com/topics/economics-econometrics-and-finance/multiple-regression-analysis