The project that was chosen was option 2, which was “I am the CEO of a telecom firm, what is causing the churn in our customer base?” Upon trying to understand the real problem, I asked the client Professor Droescher about his concerns toward the issue. We concluded that his concerns were based on the care about the overall customer base and wants it to continue growing. So, what we’re trying to find out is what opportunities do we have for reducing customer loss while increasing customer accounts?

After this was discovered, the next thing that was focused on was the things needed to know to solve this problem. Some of the things that were needed to know was the amount of people coming and going, duration, services offered, price, address or zip code, payment method, reviews, reason for leaving, etc. the data was then checked and many of the things that were mentioned was in the data set. If it wasn’t the available data would have been manipulated to become information that can be used. Good thing everything was there.

So, the first thing I did was connect to the server on azure data studio. I first looked at the columns that were there and seen how much data there is. I then wrote a query to extract it and combine the tables needed in to one big table. I saved it as a CSV file and imported it to python. Once I imported the csv file and imported any necessary libraries, I first began cleaning the data. I checked for any duplicated information, their data types, null values, and I changed the name of the values to help me perform analysis. My Cleaning was complete, and I began data manipulation. I used mostly scatterplots to see trends and correlations. Most of the columns were numerical. I used a group by to count the number of values in each row. I used a histogram to see the relationship between the ages and the reason for leaving.

It was discovered that the biggest reason that customers are leaving is because of competitors. I found out the ages that were mostly leaving from the competitors. I also discovered why they chose the competitors and did not stay, such as made better offer, better devices. It was discovered that the attitude of customer support was another big reason why people were leaving. Offers were also a cause for churn. Offer E was the worst offer that was being offered. The services that were also a cause to churn was internet phone service along with unlimited data. Fiber Optic internet turned out to be the worst type of internet showing a big cause of churn. Month to Month contracts were the lease effective and were causing a tremendous amount of churn. This was all discovered in python, and I saved the data set and brung it to tableau.

I was able to get more information from tableau such as the customer base and specific numbers from the data. Most of the customers are from California. The cities that are churning the most are San Diego, Los Angeles, San Jose, Escondido, and Fallbrook all due to competitors. The dashboard was created and that concluded the work done for the project aside from the presentation deck.