Outcome of your EDA

We discovered that M1 communications plan to roll out charging for international calls would have a significant impact on their customers. Initially the data favored making the change as the ratio of calls made internationally was statistically irrelevant. Once we looked at the duration of the calls, it became apparent that if M1 was to charge by minute versus per call the impact to the customer would be dramatic. The average international call duration was three times the length when compared to the average domestic call. There is reason to believe that if a per minute charge plan was implemented, the international call duration may decrease which would negate the potential profit.

What do you feel was missed during the analysis?

I would have liked to use the number of times the customer contacted customer service to make some sort of significant connection, but the variable did not prove to have impact on the independent variables used for most of the analysis. In reflection, I think I could have used the customer service contact data to warn the M1 stakeholders of markets that may be resistant to the new fee structure. I also considered pulling more than a month’s worth of data to do an analysis over time. I decided not to take that route as the results didn’t change or document usage behaviors that would be impactful for the stakeholders to be aware of.

Were there any variables you felt could have helped in the analysis?

Knowing the country, the customer was calling would have really helped to better organize the data for stakeholders. Knowing the duration of international calling minutes by country and domestic origin would have been exciting to see.

Were there any assumptions you felt were incorrect?

When souring the dataset, I assumed the ‘charges’ column was the total charge per international call the customer made. In doing the summary work and cleaning the data set, I discovered it was the rate per minute it cost the company. My hypothesis assumed there would be no significant impact to the customer base by passing off this operational cost, but without knowing the operational cost per minute for domestic calls my assumption was unfounded.

What challenges did you face, what did you not fully understand?

The data set was more incomplete than I originally thought and only included ‘Active’ customers who made international calls. This made it difficult to predict and compare the two profiles of customers, ‘those who made international calls’ and ‘those who do not make international calls. This meant that I had to pivot my analysis to the impact of just one segment of the customer base. There was an opportunity to do some predictive analysis for ‘New Customers’, but there was not enough information available about the markets to know how competitive international calling would be.