**Project 1 Analysis/Outline – Team 4**

**Credit Card Customer Churn Analysis  
Team Members:** Will Cutrone, Edyson Katuuo, Subhashree Mallick, Jessica Hunter

**Introduction:**

Our objective is to evaluate data to discover trends that may or may not cause people to leave their credit card service. Customer Churn is one of the most important and challenging problems for businesses such as Credit Card companies and service providers worldwide. Customer churn metrics can help businesses improve customer retention. Some common reasoning of customer churn are; lack of usage of the product, poor service and better price somewhere else. Regardless of the reasoning, it costs more to acquire new customers than it does to retain existing ones. This has a direct impact on operating costs and marketing budgets within the company. Because of the significant importance of customer churn within a business, people are investing more time and effort in finding out the reasoning within their organizations, how they can accurately predict the type of existing customers that can stop doing business with them and what they can do to minimize the customer churn.

We will go through some consumer data and see how we can leverage data insights and predictive modeling in order to improve customer retention.

**Dataset:**

In our analysis, we use Python, Jupyter Notebook and Matplotlib. We analyzed a CSV file (bank\_churners.csv) found on Kaggle.com. We cleaned the data by removing all customers with a zero utilization ratio and any other extraneous information. We created two new datasets for existing customers and attritted customers.

**Questions:**

* What factors cause customers to churn?
* Does utilization ratio have an effect on churning? - based on distribution of the box plot and p-value this category is not statistically significant
* What does age have an effect on churning, if any?
* Does income have an effect on churning?
* What are the statistically significant areas?
  + total transaction amount
  + total transaction count
  + months inactive
* What is the profile of someone who is likely to churn?

**Observations:**

Summarize your conclusions. This should include a numerical summary (i.e., what data did your analysis yield), as well as visualizations of that summary (plots of the final analysis data

Discuss the implications of your findings. This is where you get to have an open-ended discussion about what your findings "mean".

Tell a good story! Storytelling through data analysis is no different than in literature. Find your narrative and use your analysis and visualization skills to highlight conflict and resolution in your data.

## Data Cleanup & Analysis

With data in hand, it's time to tackle development and analysis. This is where the fun starts!

Inevitably, the analysis process can be broken into two broad phases: **Exploration & Cleanup** and **Analysis** proper.

As you've learned, you'll need to explore, clean, and reformat your data before you can begin to answer your research questions. We recommend keeping track of these exploration and cleanup steps in a dedicated Jupyter Notebook, both for organization's sake and to make it easier to present your work later.

Similarly, after you've massaged your data and are ready to start crunching numbers, you should keep track of your work in a Jupyter Notebook dedicated specifically to analysis.

During both phases, **don't forget to include plots**! Don't make the mistake of waiting to build figures until you're preparing your presentation. Creating them along the way can reveal insights and interesting trends in the data that you might not notice otherwise.

We recommend focusing your analysis on techniques such as aggregation, correlation, comparison, summary statistics, sentiment analysis, and time series analysis.

Finally, be sure that your projects meet the [technical requirements](https://github.com/RutgersCodingBootcamp/RUT-SOM-DATA-PT-11-2020-U-C/blob/master/Class_files/07-Project-1/ProjectGuidelines/TechnicalRequirements.md).

## Presentation

After you've analyzed your data to your satisfaction, you'll put together a presentation to show off your work, explain your process, and discuss your conclusions.

This presentation will be delivered as a slideshow, and should give your classmates and instructional staff an overview of your work. PowerPoint, Keynote, and Google Slides are all acceptable for building slides.

As long as your slides meet the [presentation requirements](https://github.com/RutgersCodingBootcamp/RUT-SOM-DATA-PT-11-2020-U-C/blob/master/Class_files/07-Project-1/ProjectGuidelines/PresentationRequirements.md), you are free to structure the presentation however you wish, but students are often successful with the format laid out in the [presentation guidelines](https://github.com/RutgersCodingBootcamp/RUT-SOM-DATA-PT-11-2020-U-C/blob/master/Class_files/07-Project-1/ProjectGuidelines/PresentationGuidelines.md).