Customer Support Analysis

Business Understanding

- **What problem are you trying to solve, or what question are you trying to answer?**

- I will be analyzing the top 5 companies with complaints on twitter, I will help see the average response time for each of them as well as the change of sentiment after they've interacted with the companies.

- **What industry/realm/domain does this apply to?**

- Customer Support

- **What is the motivation behind your project? (Saying you needed to do a capstone project for flatiron is not an appropriate motivation)**

- I believe we've all had issues with companies, so I'm trying to see which companies have the most complaint and can get a positive response after helping the customer.

Data Understanding

- **What data will you collect?**

- I will be using a customer support tweets dataset.

- **Is there a plan for how to get the data (API request, direct download, etc.)?**

- Direct download.

- **What are the features you'll be using in your model?**

- Tweet id, Author id, Inbound, Created at, Tweet, Response Tweet id, In response tweet id.

Data Preparation

- **What kind of preprocessing steps do you foresee (encoding, matrix transformations, etc.)?**

- Text cleaning and normalization, tokenization, feature extraction, vectorization, data splitting and data augmentation.

- **What are some of the cleaning/pre-processing challenges for this data?**

- Handling imbalances, emojis, unstructured text, multi topic tweets, sarcasm in tweets.

Modeling

- **What modeling techniques are most appropriate for your problem?**

- Time series analysis, clustering analysis and classification analysis.

- **What is your target variable? (remember - we require that you answer/solve a supervised problem for the capstone, thus you will need a target)**

- For the response time it will be calculated in hours, for the change in sentiment it will be a numerical measure.

- **Is this a regression or classification problem?**

- A regression problem, although it might involve both.

Evaluation

- **What metrics will you use to determine success (MAE, RMSE, etc.)?**

- MAE, RMSE, Pearson Correlation Coefficient, MAPE.

Tools/Methodologies

- **What modeling algorithms are you planning to use (i.e., decision trees, random forests, etc.)?**

- Decision Trees, SVM, RNNs.