

Computer Science Major | Statistics Minor

<https://github.com/johann017>

- Part of the Data modeling and monitoring team focused on designing the UI.
- Created user application using React.js allowing users to interact directly with backend.
- Utilized AWS Lambda functions in Python to help execute jobs, thus increasing user functionality.
- Participated in daily Scrums and collaborated extensively with team members.
- Presented to a department wide audience given the user functionality of the application.
- Currently working on deploying the code to Test.

- Used Tweepy to access Twitter's API to collect tweets and stored them as a CSV file.
- Cleaned, tokenized, stemmed, and lemmatized tweets for Machine Learning models.
- Used Natural Language Toolkit (NLTK) and Sentiment Polarity for data analysis.
- Utilized Linear SVC, Bernoulli Naive Bayes, and Linear Regression models to analyze which model best fits the data.

- Utilized the DC Metro System's API to collect information about stations.
- Provided users with lines, transfer station details, and total travel time along with fares based on peak/off-peak hours.
- For scalability, accounted for future additions of stations.

- Extracted dataset from Kaggle, cleaned and performed exploratory analysis by inspecting each variable.
- Manipulated data to help with analysis using methods that standardized data and ran Machine Learning models like K-Nearest Neighbors, Decision Tree, and Random Forest to identify which best fits the data.
- Used a Confusion Matrix in addition to the accuracy score to help determine the performance of each model.

- Introduction to Data Science

- MySQL