February 14, 2022

Emerson College Bootcamp, Data Analytics

**GitHub:** https://github.com/sthurstone/Project-IV---T-S-Team.git

**Group Name:** T&S Group

**Members**: Tanya Adams and Stephanie Thurstone

**Project IV:** Tracking Consumer Trends – Walmart Bakery

**Hypothesis:** Bakery sales vary based on specific factors such as time of day, week, year, etc. however, there is always a surplus being disposed of at the end of the day. What consumers are most likely to stop by the bakery on any given day?

**Data Sources:**

* Walmart API - <https://developer.walmart.com/>
* <https://www.kaggle.com/rutuspatel/walmart-dataset-retail>
* <https://www.kaggle.com/yamqwe/walmart-product-listing-data-2020e>
* https://www.kaggle.com/namithadeshpande027/walmart-stock-historical-data

**Procedures:**

1. Determine exact datasets to be utilized and merge and clean the data. Determine any unnecessary elements and drop them from the data set.
2. Explore and study the datasets
3. Use API Key to pull JSON files and utilize data
4. Determine the what items are most likely to be purchased by a specific consumer thru segmentation/etc.

**Anticipated Technological Library:**

* **Python Pandas**
* **Python MatPlotLib**
* **Scikit-Learn**