**Obesity, Fast Food Prevalence, and Fast Food Ratings by State**

David Brehm

DSC 540

Final Project: Milestone 1

04/11/2021

**Three Data Sources:**

* **CSV Data**

Nutrition, Physical Activity, and Obesity from the CDC. This includes data on adult’s weight status by state and year. It appears to track between 2011-2019.

**Source:**

<https://catalog.data.gov/dataset/nutrition-physical-activity-and-obesity-behavioral-risk-factor-surveillance-system>

* **HTML Data**

Fast food restaurants per 100,000 people. This data includes overall fast food restaurants per 100,000 people by state, as well as a breakdown by popular fast food chains.

**Source:**

<https://www.nicerx.com/fast-food-capitals/>

* **API Data**

Yelp user reviews. This data contains reviews, ratings, restaurant type (filtered by fast food for this project), and location.

**Source:**

<https://www.yelp.com/developers/v3/manage_app>

**Data Relationship:**

The relationship between these data sources will be US states.

**Goals:**

Through these data sources, I hope to observe the relationships between obesity, fast food restaurant concentration, and fast food user ratings. I would have liked to view the yearly data on fast food restaurants per 100,000 people, but I have been unable to find that in a single data source or an HTML data source. The CSV data will need some manipulation as it looks like for each year, the obesity rates are broken out by demographics such as age, race, and education. These different rows of data will need to be looked at by their rates and sample sizes to get an overall obesity rate by state. An assumption on this topic could be that obesity rates have a positive correlation with fast food density since those are commonly associated together, so I wanted to see how true that is by state. I am not sure how either of these would influence restaurant ratings, but I thought it would be an interesting aspect to look into as well. Maybe states with higher obesity have a higher average rating for fast food restaurants, or more ratings. The main data used from the Yelp API will probably be state, rating, and number of ratings.