**Project Title**: Health Inspection Rating Impact on Yelp

**Team Members**: Ryan, Anjali, David, Albert

**Project Description/outline**:

This project seeks to answer the question: Does the restaurant health inspection grading impact the popularity of the restaurant?

**Research Questions to Answer**:

1. What is the correlation between restaurant inspection grades and popularity of the restaurant based on Yelp ratings?
2. At what point in the inspection grade scale does the Yelp rating get impacted?
3. Are there stronger correlations between inspection grades and Yelp ratings for certain ethnic restaurant groups?
4. Are there correlation between inspection grading and Yelp $ signs (or any other Yelp data attributes)?

**Data Sets to be Used**:

1. Environmental Health Inspection Results for Restaurants and Markets in Los Angeles County
   1. Source: <https://data.lacounty.gov/Health/LOS-ANGELES-COUNTY-RESTAURANT-AND-MARKET-INSPECTIO/6ni6-h5kp/data>
   2. Number of lines = 200,000
2. Yelp API Data
   1. Source: <https://www.yelp.com/developers>

**Rough Breakdown of Tasks**:

1. Pull in data from the two sources
2. Clean the data and extract the appropriate columns from each data set
3. Merge the data sets together by restaurant
4. Create a panda dataframe to group the restaurants by rating and ethnic food group
5. Run regression to see if there is a meaningful correlation for the different research questions we posed