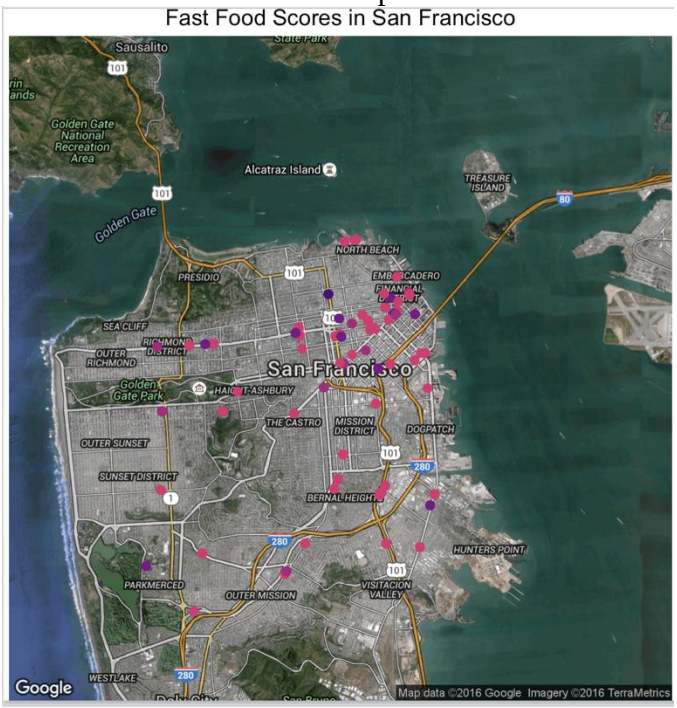
Homework 5: Team 17

In this report, we looked at the health rating information for multiple California counties. These data frames consist of health inspection information for several California counties, including San Francisco, Berkeley, Alameda, and Yolo. Each county has similar information such as business names, respective ID's assigned to each, and longitude and latitude of the businesses which allows us to map them. The datasets provided to us focus on health safety information. We decided to investigate whether the likelihood of lower health code rating is higher for fast food restaurants than for regular food establishments, which we defined as everything except for fast food restaurants. For our investigation, we looked at ten of the top fast food chains in California to see whether they showed a trend of lower health inspection ratings compared to the rest of the food establishments contained in the data sets. This question is of interest because there is a widespread belief that fast food restaurants tend to be more prone to health safety issues. To investigate this, we utilized the scores of each restaurant for each county and the coordinates of each business to create maps and compare.

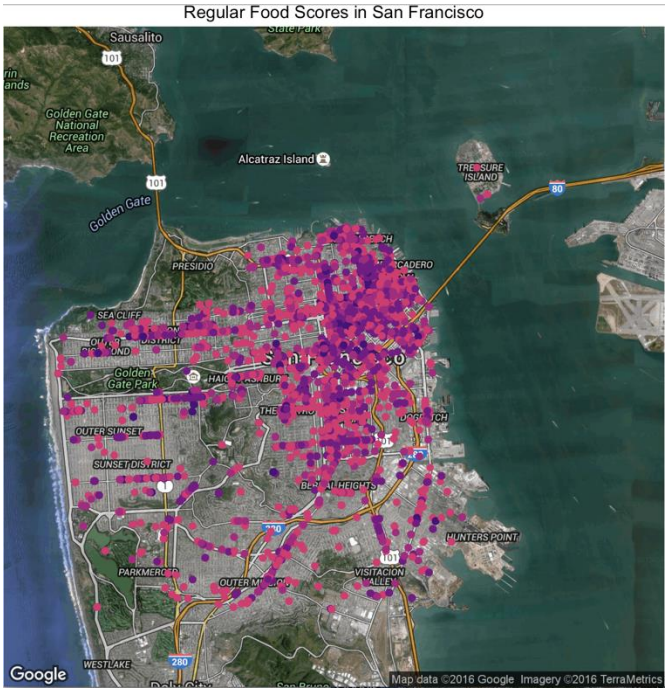
Using McDonald's, Subway, Burger King, In-n-Out, Wendy's, Taco Bell, Pizza Hut, Arby's, KFC, and Jack in the Box as our sample of fast food chains we looked at the distribution and spread of health inspection scores for each county. San Francisco utilizes a numeric score in order to assign businesses operating condition categories. These categories include 'Good', 'Adequate', 'Needs Improvement' and 'Poor'. Below we compare these fast-food chains with other establishments. We computed scores and assigned ratings based on violation frequency to mimic San Francisco's four tiered system.

First, we begin with a visualization of our top ten fast food restaurants in San Francisco.



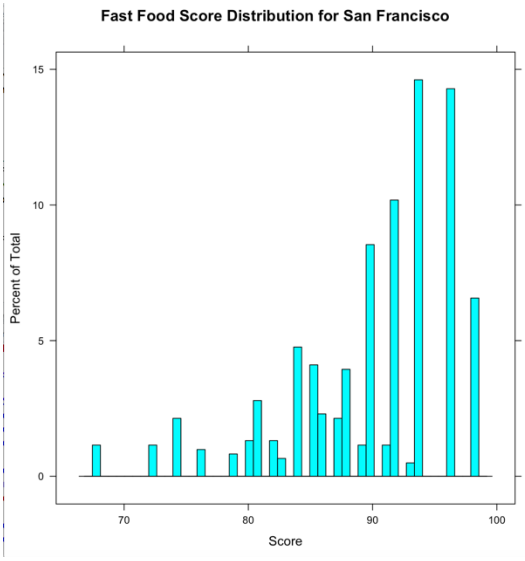
In the maps of scores in San Francisco above and below, dark purple colors signify bad health code ratings and light pink colors signify good health code ratings. The colors in between are ratings that fall in between the two extremes. When we compare the map of our top 10 pick of fast food restaurants (above) to all the other restaurants in the area (below) we see that the fast

food restaurants have mostly high scores with a few low scores dispersed throughout. When we look at regular food scores below, we see a strong variety of scores in very close geographical proximity to each other.

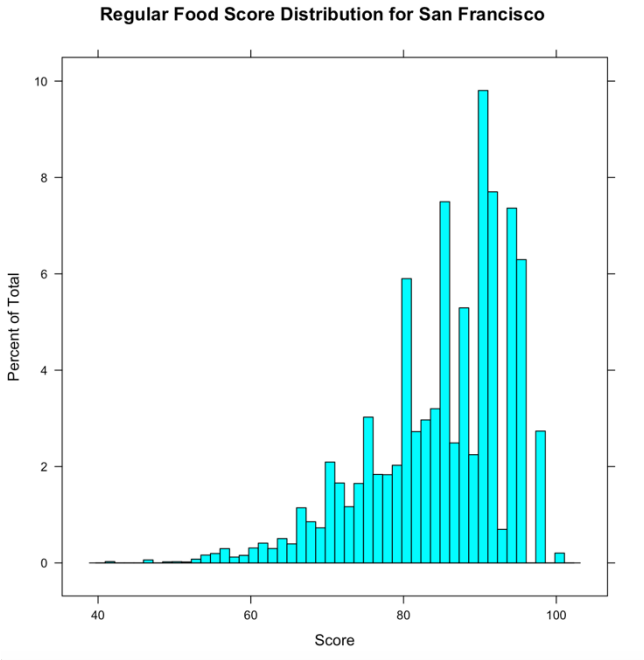


This could be due to the larger volume of restaurants in the regular food scores map, but the northeastern section of San Francisco pictured below features many more low scored restaurants when restaurants not in our top ten pick of fast food restaurants are included. If anything, we can safely say that the top ten fast food restaurants are not more likely to have health code violations than other restaurants.

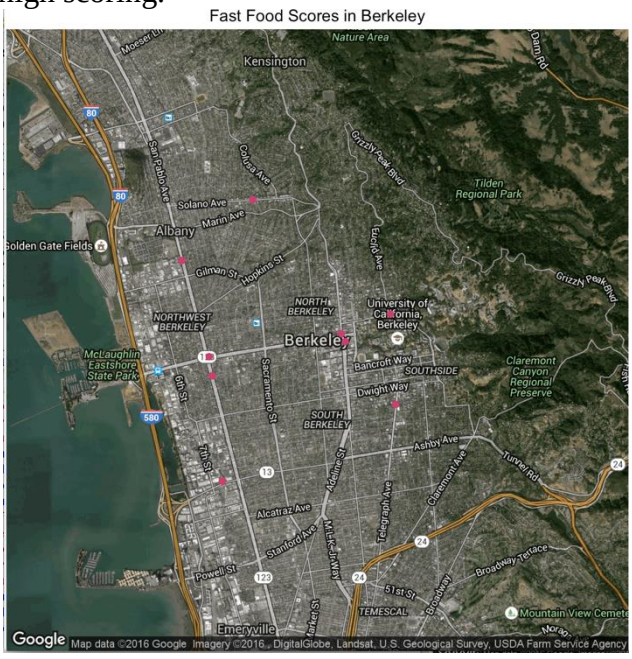
This is further evident in the following histograms of the fast food chains versus the rest of the sample.



The only striking difference between the two is that the fast food score distribution has fewer observations than the other food distribution. Both distributions are heavy with observations in the high score range, scores ranging 80 and up. This means that most of the restaurants in both subsets of the data have decent health safety scores. Again, we see it corroborated that the top ten fast food restaurants do not differ greatly in terms of rating distribution.



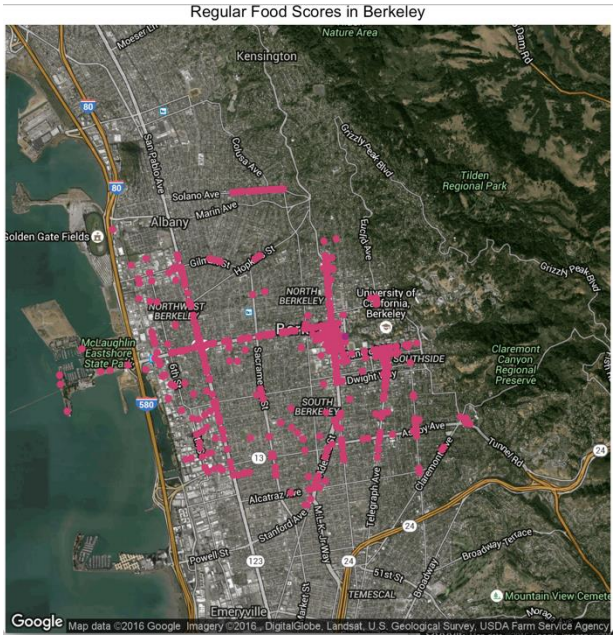
When it comes to Berkeley, the scores of both fast food chains along with the rest of the sample are consistently high scoring.



An interesting thing to note is the low volume of fast food locations in Berkeley, as compared to San Francisco. It appears that Berkeley is, in that respect, a more nutritious place to

eat. The same goes for sanitation and health code ratings. For the "regular" food scores in Berkeley map below, the homogenous high scoring color of the vast majority of points may seem not to offer much information, but when compared to San Francisco's "regular" food scores map, we see that Berkeley outcompetes San Francisco greatly in terms of health safety ratings. The color schemes for both San Francisco and Berkeley maps are set to the same breaks due to their similar numeric scoring method.

Regular Food Scores in Berkeley



While Yolo county does not use a numerical scoring system in their health inspections, the type variable allowed us to make note of failures by counting the follow up inspections. As described on Yolo county's organizational website follow up inspections are made when the business fails the initial inspection with two or more major violations. Among Yolo county's inspections only 706 of the 16129 required follow up inspections.

Across these datasets instances of fast food chains having higher health code violations are not outside the expected spread that is found in the rest of the data points. This isn't necessarily what one would expect from fast food restaurants which often have the stigma of being unclean. However, the results of our plots and maps did not find this expected outcome. Instead we found that they fit into the standard range.

References

Health Inspection Scoring Information:
San Fancisco
https://www.sfdph.org/dph/EH/Food/Score/
Alameda
https://www.acgov.org/aceh/documents/NewsRelease-Grading2012-06-28.pdf
Yolo County http://www.yolocounty.org/health-human-services/health-department/restaurant-search/violations