San Francisco Restaurant Data Analysis

We initially set out to answer the following questions:

1. **How many restaurants open in San Francisco per year and how many close?**
   1. Based on our calculations, of the restaurant data we were able to pull, 8.6% of the restaurants closed. We feel the data set we ended up isn’t a good representation of all restaurants in San Francisco. Instead, it is a representation of the restaurants that had complete information on Yelp. We would have ideally liked to include timestamps to when the restaurants closed to see if there were any years where turnover was more volatile.
2. **What neighborhoods see the most turnover?**
   1. When starting this project, we were really ambitious and aimed to generate a Heat Map to show what areas in San Francisco saw the most openings and closures.
   2. Due to not being able to capture neighborhoods in the government data source or Yelp API data, we had to adjust our calculations to see how many restaurants opened and closed in each zip code. Please note, a zip code can contain multiple neighborhoods, and a neighborhood can contain multiple zip codes. Since our bar chart makes it hard to visualize neighborhoods, the best solution was to compare the most popular zip codes to a map on San Francisco. Our data shows that the Mission district contains the most restaurants, and subsequently has the most closures as well. This makes sense to us since the Mission is considered a popular neighborhood going through a lot of gentrification.
3. **What type of restaurants are most populous?**
   1. Next, we set out to find what type of food services were most populous and most popular, which the top three categories were Coffee& Tea, Chinese and Mexican cuisine. This makes sense to us due to San Francisco’s strong coffee culture as well as concentrated Chinese and Mexican populations. Due to time constraints, we did not group certain categories like including Sushi Bars and Ramen with Japanese cuisine, or Coffee & Tea with Cafes.
4. **What type of restaurants get better ratings, i.e. do pricey restaurants receive better ratings?**
   1. We were disappointed to find that the data constraints couldn’t yield any meaningful graphs to demonstrate price versus rating, due to the type of data we pulled. Yelp provided a pricing scale of $, $$, $$$, $$$$, and a rating scale of 1-5 stars. When we tried to plot this on a scatterplot, we basically ended up with a grid.
   2. Instead, we delved deeper into looking what categories of restaurants were most popular. We found this by comparing the total number of reviews per category, average rating, and average rating. It turns out that San Francisco LOVE brunch, each brunch spot received an average of 1064 reviews, with an average rating of 4 stars.