

About that Tasty Hole-in-the-Wall ...

An Analysis of Restaurant Health Inspection
Grades/Scores Versus Yelp Reviews

Angelenos love love LOVE food

From the amazing produce to the mix of immigrant flavors, Los Angeles has a strong food-focused culture.

Project goals:

1. Relevant to where audience lives
2. Topic anyone can relate to
3. Compare 2 diverse data sets
4. Make the data visual



Do Health Inspections & Yelp Ratings Matter?

L.A.'s Department of Health checks everything from food temperatures to storage practices. Is there a correlation between inspection scores/grades and a restaurant's overall Yelp rating?



Our gut says “yes” ...

Dirty, dirty data ... an initial look

We combined 2 data sets and 19 yelp API calls (in sets of 5,000) to create a 30 column/ 52,000 row CSV.

```
raw_restaurants_df.head()
```

activity_date	employee_id	facility_address	facility_city	facility_id	facility_name	facility_state	facility_zip	food_type	...	program_name	program_status	r
11/17/2015 12:00:00 AM	EE0000726	7669 BEVERLY BLVD	LOS ANGELES	FA0025994	NAKKARA ON BEVERLY	CA	90036	['Thai', 'Asian Fusion']	...	NAKKARA ON BEVERLY	ACTIVE	
11/17/2015 12:00:00 AM	EE0000726	7669 BEVERLY BLVD	LOS ANGELES	FA0025994	NAKKARA ON BEVERLY	CA	90036	['Thai', 'Asian Fusion']	...	NAKKARA ON BEVERLY	ACTIVE	
11/06/2015 12:00:00 AM	EE0000726	7669 BEVERLY BLVD	LOS ANGELES	FA0025994	NAKKARA ON BEVERLY	CA	90036	['Thai', 'Asian Fusion']	...	NAKKARA ON BEVERLY	ACTIVE	
11/06/2015 12:00:00 AM	EE0000726	7669 BEVERLY BLVD	LOS ANGELES	FA0025994	NAKKARA ON BEVERLY	CA	90036	['Thai', 'Asian Fusion']	...	NAKKARA ON BEVERLY	ACTIVE	
11/17/2015 12:00:00 AM	EE0000950	6600 E OLYMPIC BLVD STE #106	LOS ANGELES	FA0055785	YUM YUM DONUTS	CA	90022	['Donuts', 'Coffee & Tea']	...	YUM YUM DONUTS	INACTIVE	

First, the numbers

Our main data points involve inspection scores/grades and Yelp ratings.

Turns out the average inspection ranks higher than most average restaurants.

Then we looked at each restaurant.

```
#Average Inspection Score
```

```
raw_restaurants_df["Score"].mean()
```

```
93.04407624633431
```

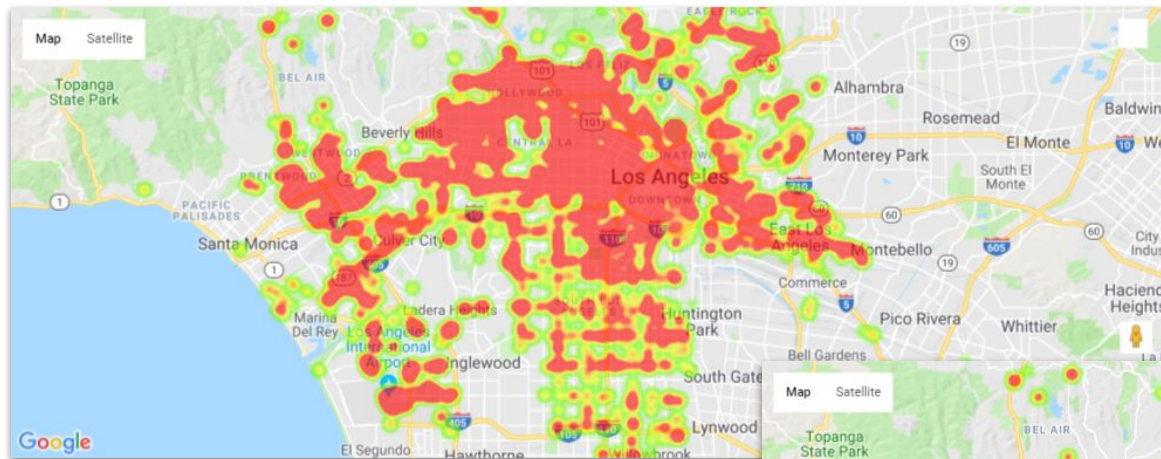
```
#Average Yelp Rating
```

```
raw_restaurants_df["Rating"].mean()
```

```
3.569516129032258
```

	Restaurant	Score	Rating	# of Reviews
13750	SUSHI AI	90	4.5	36
13792	HOMESTATE	96	4.0	91
13743	K&A CATON CHINESE REST.	90	3.0	143
13720	KOBUNGA KOREAN GRILL	97	4.0	50
13887	WALL STREET PIZZA	91	3.0	44

If ya can't stand the heat ...

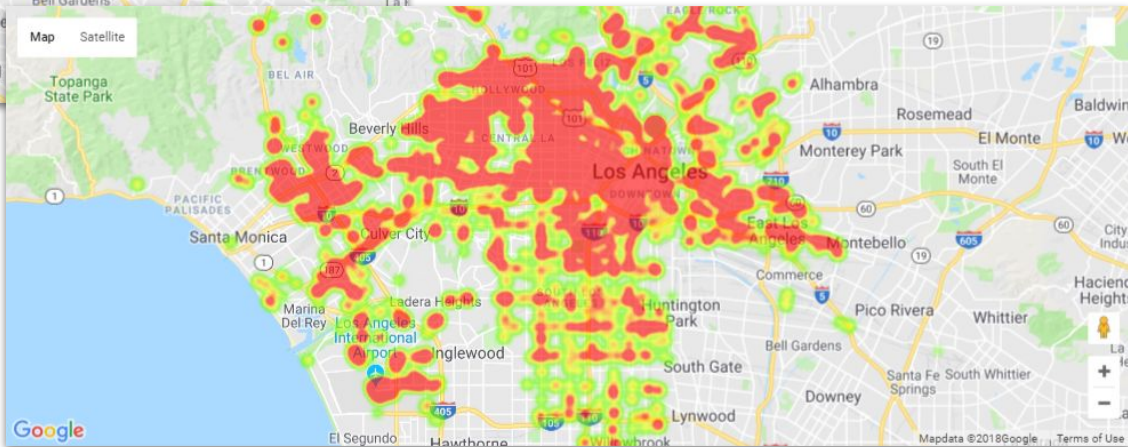


<<<

This heat map shows Health Score (with higher score in warmer colors).

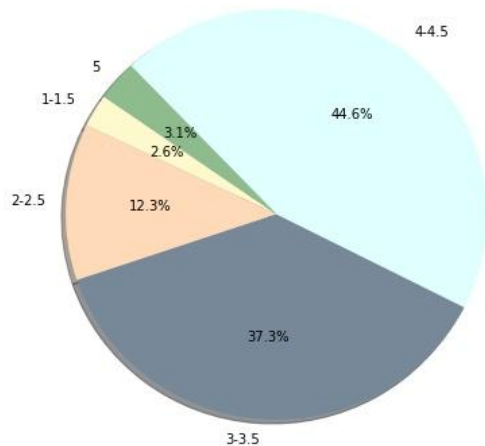
>>>

This heat map shows Yelp Grades (with higher grades in warmer colors)

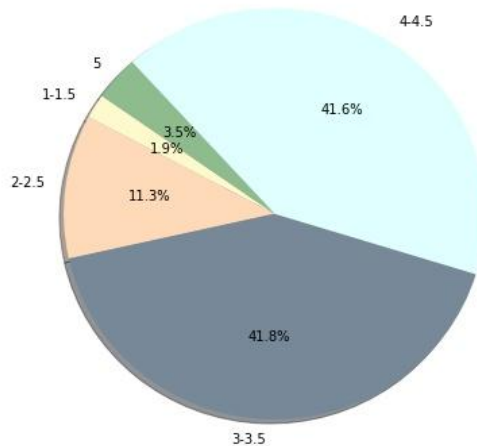


Now we know our A, B, Cs ...

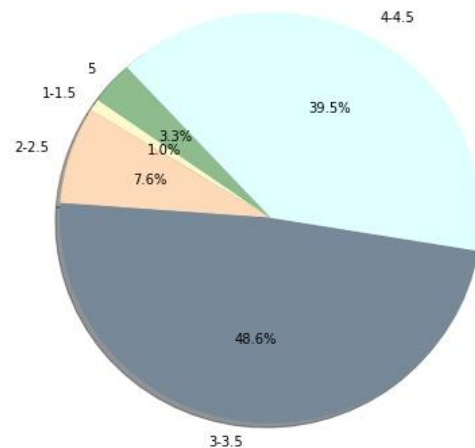
% Total Ratings for A Health Score Restaurants



% Total Ratings for B Health Score Restaurants



% Total Ratings for C Health Score Restaurants



Even after breaking down by health inspection grade, no trend was apparent on initial review.

Ratings count, so we counted the ratings

Here's a look at the percentages of ratings by each star level:

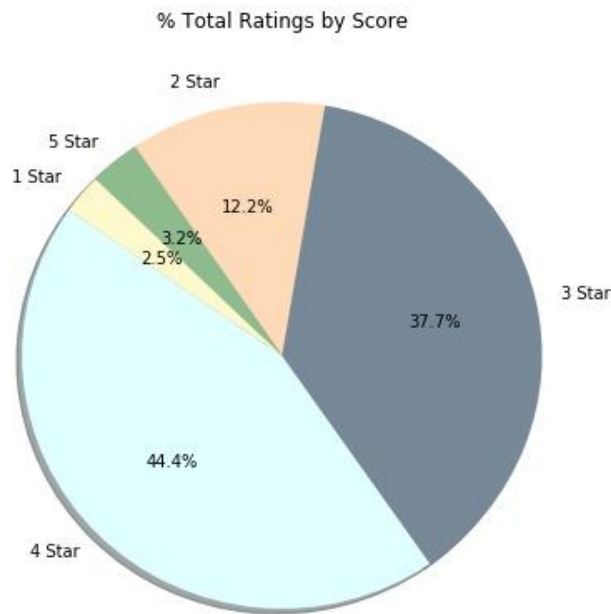
1-star = 863 ratings

2-star = 4,169 ratings

3-star = 12,860 ratings

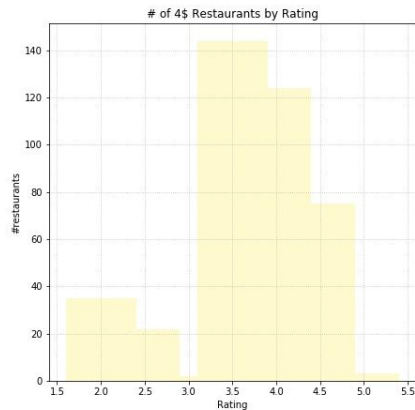
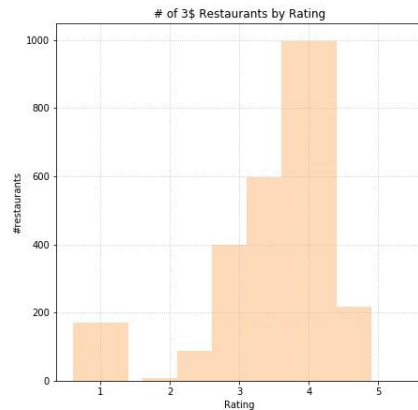
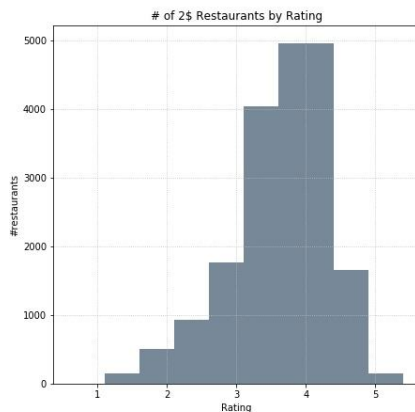
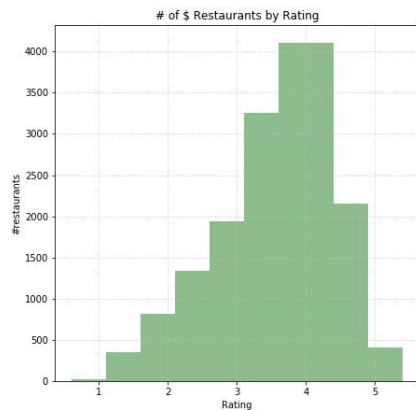
4-star = 15,429 ratings

5-star = 1,079 ratings



Is the price right?

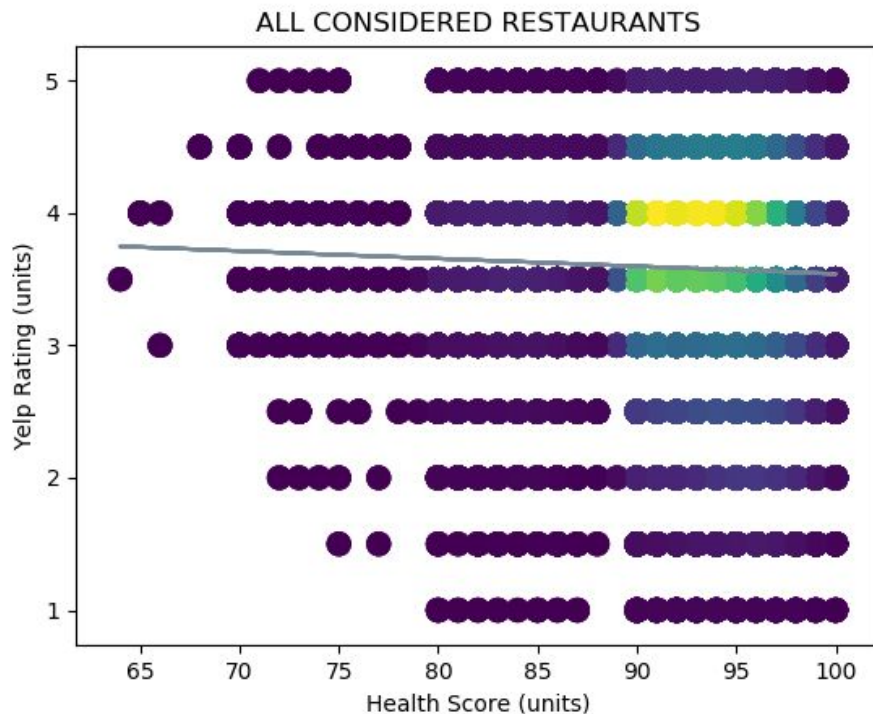
Distribution of restaurants in each \$ group shows similar pattern in that majority are categorized in 3–4 ratings regardless of the price point.



Looking at each restaurant ...

The densest part shows most health scores are in the low 90s (“A”) and most Yelp ratings are 4-stars.

The line of best fit is, in fact, not a good fit as the data is spread everywhere.



Statistically speaking ...

Along with some freedom fries ...

Degrees of Freedom	29,432
Significance Level	5%
Critical Value	29,832.2066026104033
Chi Squared Value	5,204.175393281761
Result	Hypothesis Rejected

If time were no issue ...

Further analyses that can be done with our raw & clean data:

1. Regression analysis of Yelp ratings and price:
 - Can a restaurant increase/decrease price point to increase rating?
2. Sentiment analysis on Yelp reviews per star-rating
 - Do the comments sentiments reflects the star rating?
3. Correlation analysis of health scores vs. types of restaurants
 - Do local cuisine get better scores vs. international cuisines?