

Digging In to Yelp Reviews

Text Mining Yelp Reviews for Sub Topic Ratings

BY SAM GUTENTAG

MARCH 15, 2018

Introduction

Yelp is the gold standard in the field of Restaurant Reviews and Ratings, and Yelp Stars are incredibly important. As an example of this importance, a 2011 report concluded that a Improving a Business's Yelp Star Rating by just a single Star, could lead to a 5-9% bump in revenue [1].

But as is the case with many user generated ranking systems, it can be hard for the average user or typical business owner to make sense of a single Star Rating for a business, with many locations being either heavily favored or frequently despised. Many users create accounts just to give poor reviews after a particularly bad experience, or on the other hand sign up to sing the praises of a local favorite.

How does a user distinguish between a pair of 4 Star Sushi Restaurants? Is the customer service at this 4 Star burrito place going to be a better choice than the 3 Star hole in the wall location where the servers are known to yell at you, but the chef is in a league all her own?

And what about a Restaurant Owner or Manager, when the Yelp Star Rating of your restaurant is the same as your rival across town, why do so many people go there instead?

But how does a user know to compare a four star sushi to a 3 star sushi restaurant to a 5 star roofing contractor? It may not be likely that a person would be looking at all of these just to find a good spot for date night, but it shows the downside of a single number rating system.

This project aims to use Reviews and Tips for restaurants in Arizona to build a system that will identify key positive and negative terms for each location. From here we will generate sub scores in the fields of Food Quality, Value, Cleanliness, Service, and possibly more!

This will help Yelp users make better, more informed decisions as well as provide Restaurant owners and managers helpful insight into areas in which their business is leading or trailing customer expectations and make improvements to drive sales.

Let's dig in, and uncover the details of user reviews and ratings.

The 9 Identified Review Sub Topics

To better identify Topics from user reviews, we are limiting the vocabulary of this analysis to be derived from Restaurants in the state of Arizona only. The source data set provided by Yelp [2] includes 501,250 Reviews from 7066 Unique Restaurants.

Using a Latent Dirichlet Allocation (LDA) model from the Natural Language Toolkit (nltk) Library, reviews are collected into 50 Topics with 10 Terms per Topic.

Topics highlighted in red are used for Sub Topic identification.

Topic ID	Topic Terms
0	['dessert', 'wine', 'ice', 'cream', 'cake', 'entree', 'meal', 'chocolate', 'course', 'appetizer']
1	['time', 'back', 'first', 'try', 'place', 'went', 'definitely', 'go', 'next', 'great']
2	['u', 'table', 'came', 'server', 'food', 'drink', 'waitress', 'asked', 'ordered', 'minute']
3	['dish', 'flavor', 'sauce', 'like', 'taste', 'menu', 'one', 'would', 'bit', 'meat']
4	['pizza', 'crust', 'slice', 'topping', 'cheese', 'pie', 'thin', 'good', 'sauce', 'pepperoni']
5	['crab', 'leg', 'shell', 'pound', 'coworker', 'saving', 'panini', 'hub', 'e', 'angry']
6	['chicken', 'rice', 'chinese', 'fried', 'food', 'beef', 'soup', 'egg', 'orange', 'sour']
7	['wait', 'minute', 'time', 'food', 'get', 'long', 'line', 'hour', 'waiting']
8	['great', 'nice', 'patio', 'atmosphere', 'outside', 'place', 'cool', 'fun', 'inside', 'food']
9	['order', 'ordered', 'called', 'delivery', 'extra', 'got', 'time', 'get', 'card', 'call']
10	['tempe', 'school', 'opening', 'mill', 'w', 'college', 'b', 'asu', 'high', 'grand']
11	['fresh', 'ingredient', 'olive', 'tomato', 'salad', 'oil', 'made', 'bianco', 'mozzarella', 'basil']
12	['best', 'year', 'place', 'restaurant', 'phoenix', 'one', 'ever', 'food', 'valley', 'family']
13	['table', 'dirty', 'clean', 'bathroom', 'floor', 'plate', 'hand', 'cup', 'paper', 'chair']
14	['chicken', 'fried', 'waffle', 'deep', 'breast', 'piece', 'cajun', 'tender', 'catfish', 'crispy']
15	['mouth', 'smell', 'taste', 'word', 'world', 'heaven', 'bud', 'holly', 'win', 'watering']
16	['always', 'love', 'place', 'get', 'time', 'food', 'go', 'favorite', 'great', 'never']
17	['food', 'experience', 'service', 'great', 'server', 'u', 'staff', 'friendly', 'restaurant', 'made']
18	['sushi', 'roll', 'tuna', 'fresh', 'chef', 'salmon', 'japanese', 'spicy', 'tempura', 'fish']
19	['option', 'menu', 'free', 'gyro', 'meat', 'vegetarian', 'veggie', 'choose', 'vegan', 'choice']
20	['food', 'like', 'place', 'ordered', 'tasted', 'bad', 'even', 'back', 'cold', 'taste']
21	['good', 'food', 'place', 'price', 'pretty', 'service', 'better', 'like', 'would', 'really']
22	['burger', 'fry', 'onion', 'bun', 'shake', 'ring', 'cheese', 'french', 'bacon', 'patty']
23	['bowl', 'pho', 'soup', 'noodle', 'roll', 'spring', 'broth', 'vietnamese', 'beef', 'rice']
24	['star', 'review', 'give', 'yelp', 'five', 'reason', 'rating', 'giving', 'based', 'shot']
25	['location', 'drive', 'employee', 'thru', 'window', 'car', 'fast', 'chipotle', 'dont', 'im']
26	['restaurant', 'parking', 'menu', 'lunch', 'small', 'lot', 'area', 'find', 'little', 'place']
27	['great', 'food', 'service', 'place', 'friendly', 'good', 'recommend', 'staff', 'delicious', 'price']
28	['wing', 'pasta', 'sauce', 'italian', 'garlic', 'bread', 'meatball', 'spaghetti', 'dish', 'ranch']
29	['steak', 'cooked', 'potato', 'perfectly', 'well', 'medium', 'rare', 'filet', 'mashed', 'perfection']
30	['bar', 'room', 'dining', 'bartender', 'area', 'sit', 'drink', 'airport', 'sat', 'table']
31	['breakfast', 'egg', 'coffee', 'pancake', 'bacon', 'morning', 'toast', 'brunch', 'biscuit', 'gravy']
32	['fish', 'shrimp', 'ramen', 'chip', 'fried', 'sea', 'ceviche', 'ahi', 'sauce', 'mahi']
33	['hot', 'dog', 'chicago', 'beef', 'chili', 'piping', 'italian', 'slider', 'wendy', 'pepper']
34	['place', 'like', 'get', 'know', 'go', 'one', 'say', 'want', 'make', 'really']
35	['good', 'really', 'got', 'ordered', 'little', 'nice', 'pretty', 'came', 'portion', 'also']
36	['sandwich', 'bread', 'sub', 'cheese', 'turkey', 'meat', 'subway', 'shop', 'lunch', 'philly']
37	['seafood', 'lobster', 'shrimp', 'pot', 'calamari', 'chop', 'sum', 'duck', 'dim', 'clam']
38	['taco', 'mexican', 'salsa', 'burrito', 'chip', 'bean', 'tortilla', 'food', 'carne', 'asada']
39	['bbq', 'pork', 'rib', 'meat', 'brisket', 'sauce', 'pulled', 'side', 'good', 'corn']
40	['salad', 'chicken', 'pita', 'dressing', 'greek', 'hummus', 'lettuce', 'wrap', 'fresh', 'lunch']
41	['kid', 'music', 'loud', 'wall', 'game', 'child', 'hole', 'playing', 'play', 'family']
42	['cheese', 'sweet', 'delicious', 'sauce', 'perfect', 'mac', 'onion', 'green', 'potato', 'pepper']
43	['happy', 'hour', 'beer', 'drink', 'selection', 'great', 'menu', 'special', 'price', 'deal']
44	['customer', 'manager', 'service', 'back', 'food', 'said', 'time', 'asked', 'one', 'would']
45	['buffet', 'indian', 'food', 'lamb', 'chicken', 'dish', 'naan', 'rice', 'masala', 'restaurant']
46	['son', 'sister', 'mother', 'brother', 'http', 'father', 'c', 'buck', 'law', 'wallet']
47	['night', 'dinner', 'last', 'party', 'friday', 'went', 'reservation', 'birthday', 'group', 'saturday']
48	['thai', 'tea', 'spicy', 'curry', 'pad', 'tofu', 'iced', 'dish', 'spice', 'rice']
49	['grill', 'cookie', 'hawaiian', 'cooky', 'mini', 'macaroni', 'island', 'paradise', 'scoop', 'nugget']

These 12 Topics are selected because they each pertain to a specific 'theme' or grouping of similar terms that can be used to identify Topics in review text. This collection of topics generated by the LDA model is ideal as it also separates out many topics that pertain to a cuisine of a restaurant and are going to be highly specific to a certain type of restaurant. Example Topic 22 ['burger', 'fry', 'onion',...] is not likely to be applicable to a restaurant focused on selling Breakfast or an Ice-cream Shop.

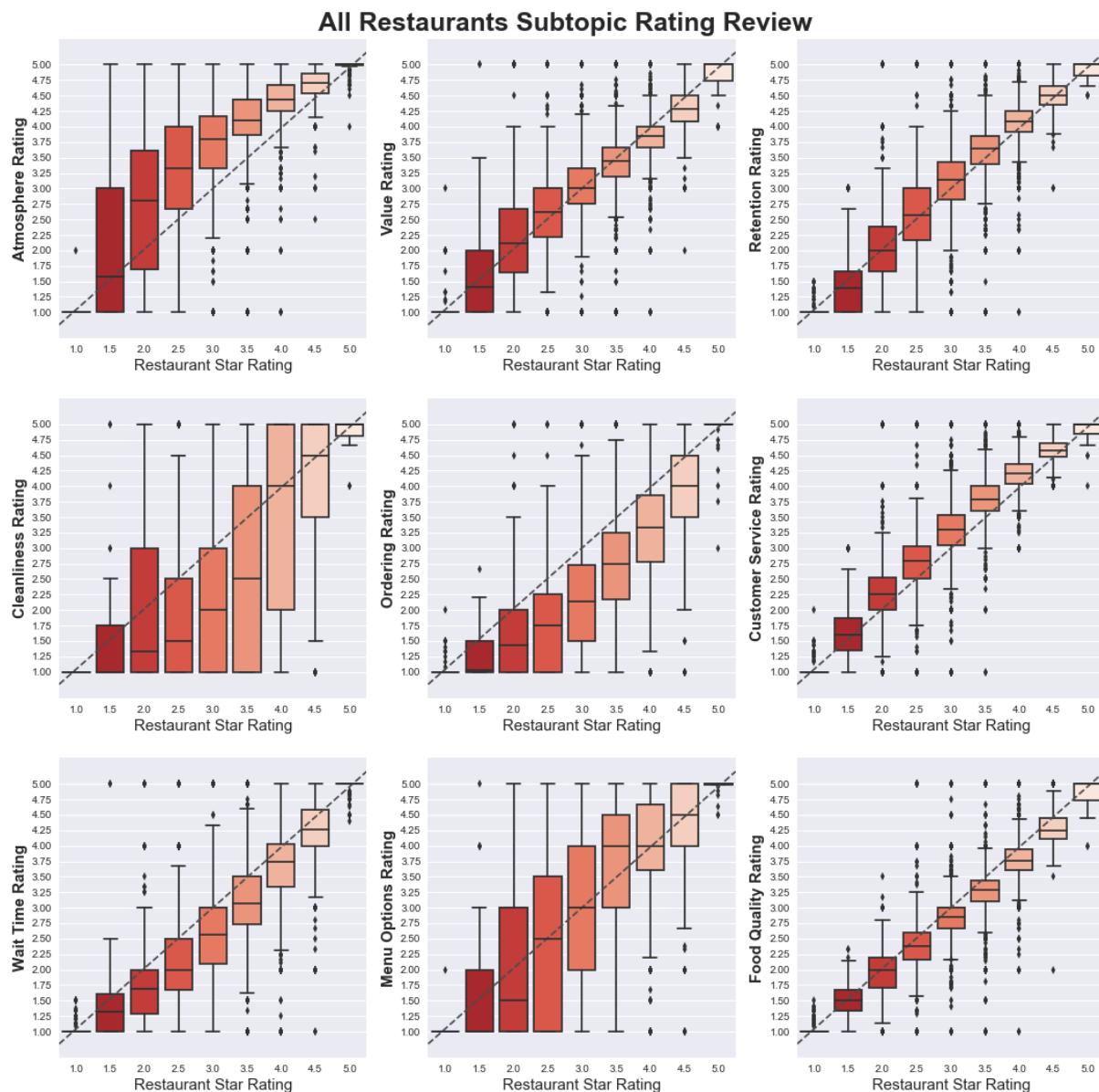
The 12 identified topics are given human readable labels, and a few are collected into groups, we are able to identify 9 Sub Topic areas in review texts that can be used to derive Sub Topic Ratings.

Retention	1 ['time', 'back', 'first', 'try', 'place', 'went', 'definitely', 'go', 'next', 'great']
Food Quality	3 ['dish', 'flavor', 'sauce', 'like', 'taste', 'menu', 'one', 'would', 'bit', 'meat'] 20 ['food', 'like', 'place', 'ordered', 'tasted', 'bad', 'even', 'back', 'cold', 'taste'] 21 ['good', 'food', 'place', 'price', 'pretty', 'service', 'better', 'like', 'would', 'really']
Wait Time	7 ['wait', 'minute', 'time', 'food', 'get', 'order', 'long', 'line', 'hour', 'waiting']
Atmosphere	8 ['great', 'nice', 'patio', 'atmosphere', 'outside', 'place', 'cool', 'fun', 'inside', 'food']
Ordering	9 ['order', 'ordered', 'called', 'delivery', 'extra', 'got', 'time', 'get', 'card', 'call']
Cleanliness	13 ['table', 'dirty', 'clean', 'bathroom', 'floor', 'plate', 'hand', 'cup', 'paper', 'chair']
Menu Options	19 ['option', 'menu', 'free', 'gyro', 'meat', 'vegetarian', 'veggie', 'choose', 'vegan', 'choice']
Customer Service	27 ['great', 'food', 'service', 'place', 'friendly', 'good', 'recommend', 'staff', 'delicious', 'price'] 44 ['customer', 'manager', 'service', 'back', 'food', 'said', 'time', 'asked', 'one', 'would']
Value	35 ['good', 'really', 'got', 'ordered', 'little', 'nice', 'pretty', 'came', 'portion', 'also']

Each Review is compared to the unique terms in these 9 Sub Topics to be labelled as pertaining to the Sub Topic or not. If a Review is heavily discussing Menu Options and Atmosphere, then the Review is given a Star Rating for these two Sub Topics equal to the original single Review Star rating given by the user. All other Sub Topics are set to null values, as not to interfere with business averaging functions later on.

Sub Topic Ratings Compared to Review Ratings

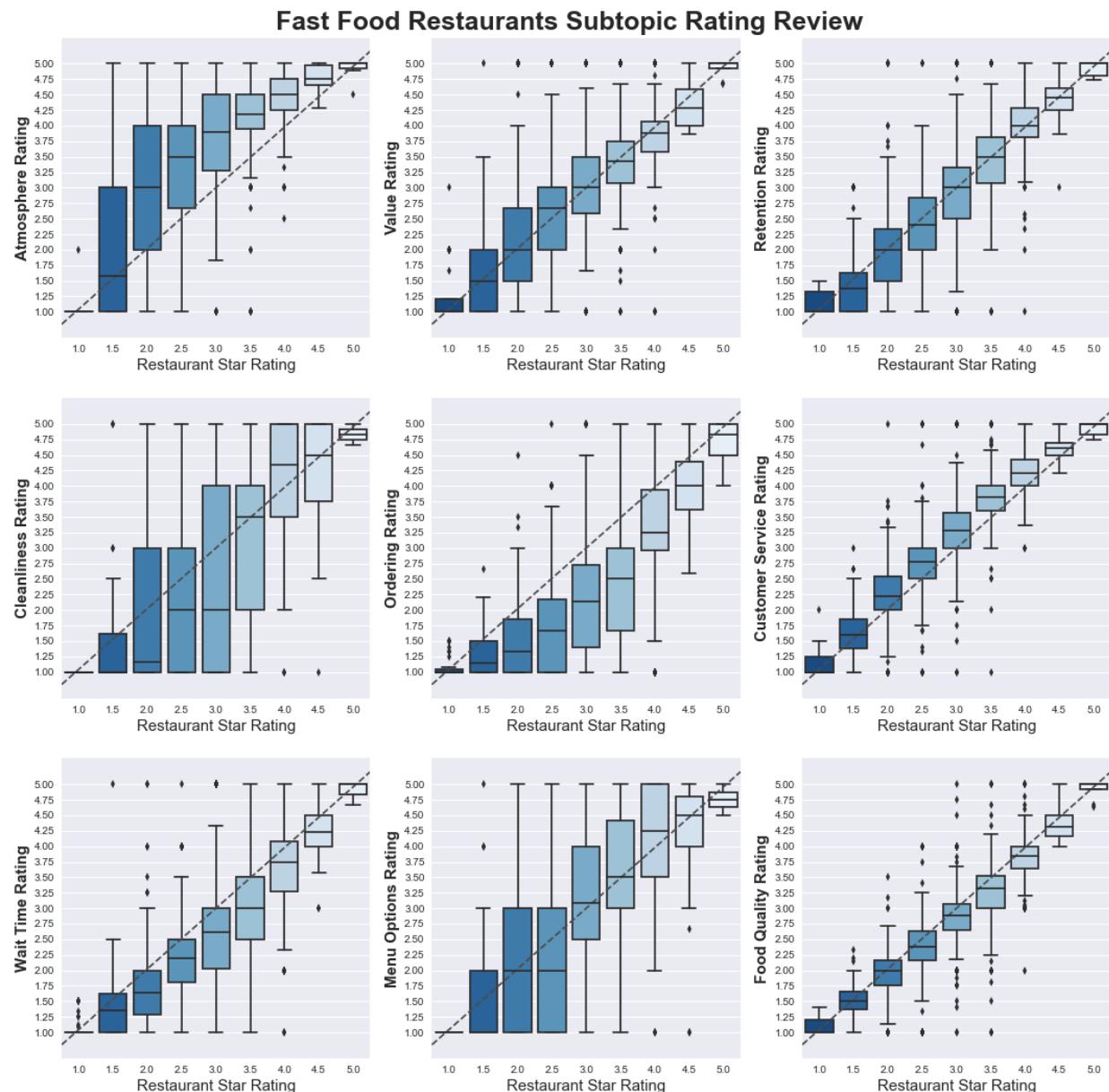
Looking first at the distribution of Review Stars as compared to the calculated Stars for each Sub Topic, a few trends become evident. First, Sub Topics like Value, Retention, Customer Service, and Food Quality are highly correlated to the original Review Stars. More skewed distributions such as Atmosphere and Wait Time indicate that a user is likely to oversell the Atmosphere and be a bit more critical of the Wait Time. Lastly, Sub Topics like Menu Options and Cleanliness are much more widely distributed, and are thus likely to vary greatly from restaurant to restaurant.



Fast Food Sub Topic Ratings Compared to Review Ratings

Fast Food Restaurants are a special subset, with user reviews being much more polarized than those given to non Fast Food Restaurants. The Sub Topic distributions are accordingly a bit different than those for all restaurant types.

The most noticeable differences are in Retention and Cleanliness. Customers are much more lenient on Cleanliness scores at Fast Food Locations than Non Fast Food Locations. That is to say, users are likely to comment on cleanliness frequently, but even when the cleanliness is lacking, still give a strong Review.



Manger Reports

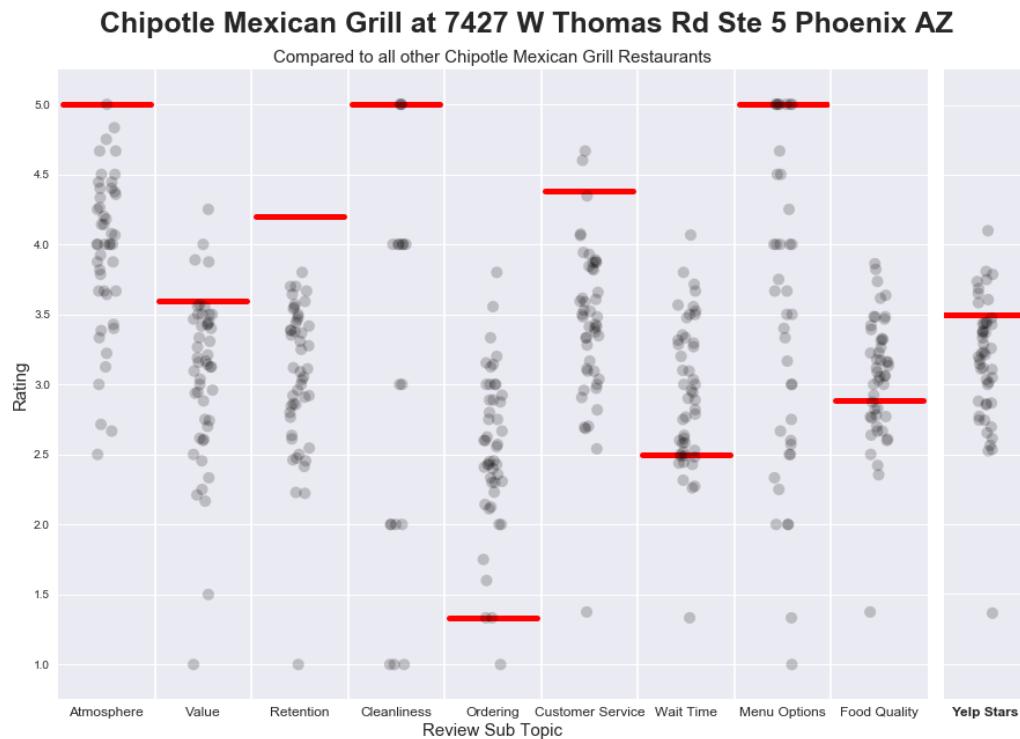
The last step of this analysis was to derive a set of comparisons that can be used by a restaurant owner or manager to better assess what customers are saying about their restaurant as well as how they compare to other restaurants and where they can improve.

A single Chipotle restaurant will be used as the sample business in this report. The **Red Lines** indicate the Stars in each Sub Topic as well as Overall Rating for the specified Subway restaurant. **Black Dots** are the Ratings of all other restaurants it is being compared to.

Report 1: Chipotle as compared to All Other Chipotle Locations

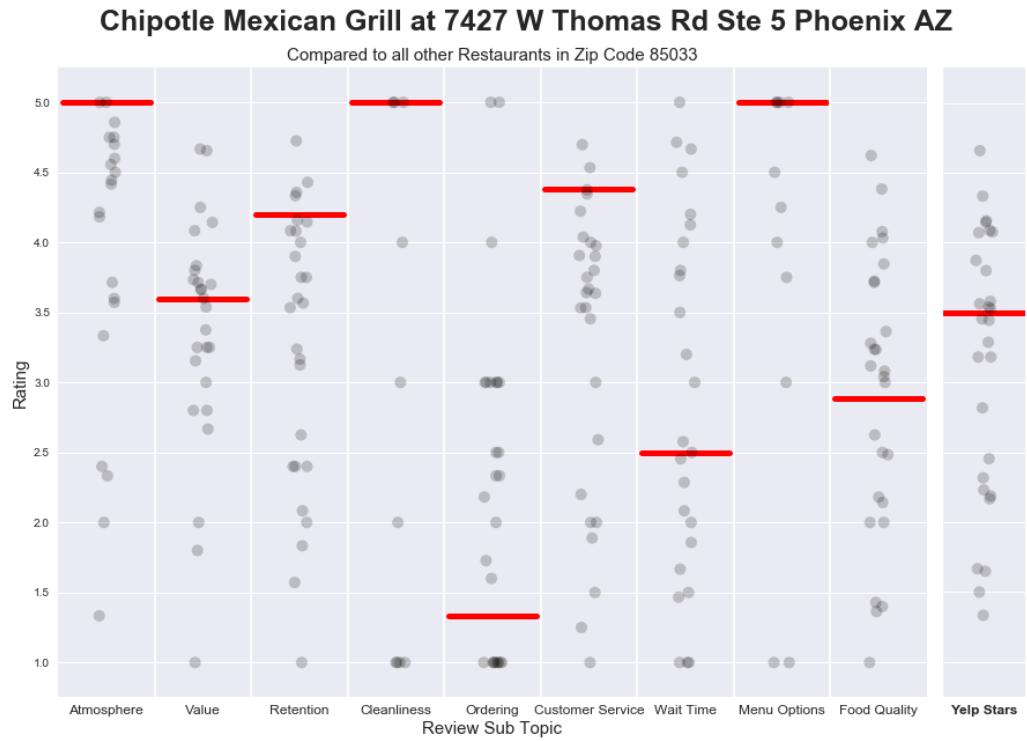
A 'Chain' restaurant is a collection of restaurants all bearing the same name. In this case, we are comparing our Chipotle, to all other Subway locations.

This Chipotle location is leading the pack in nearly all Sub Topics. However, compared to other Chipotle locations, Ordering and Wait Time are trailing far behind. The Manager of this location would be wise to investigate ways in which orders are being taken and the rate at which they are being prepared at this location can be improved.



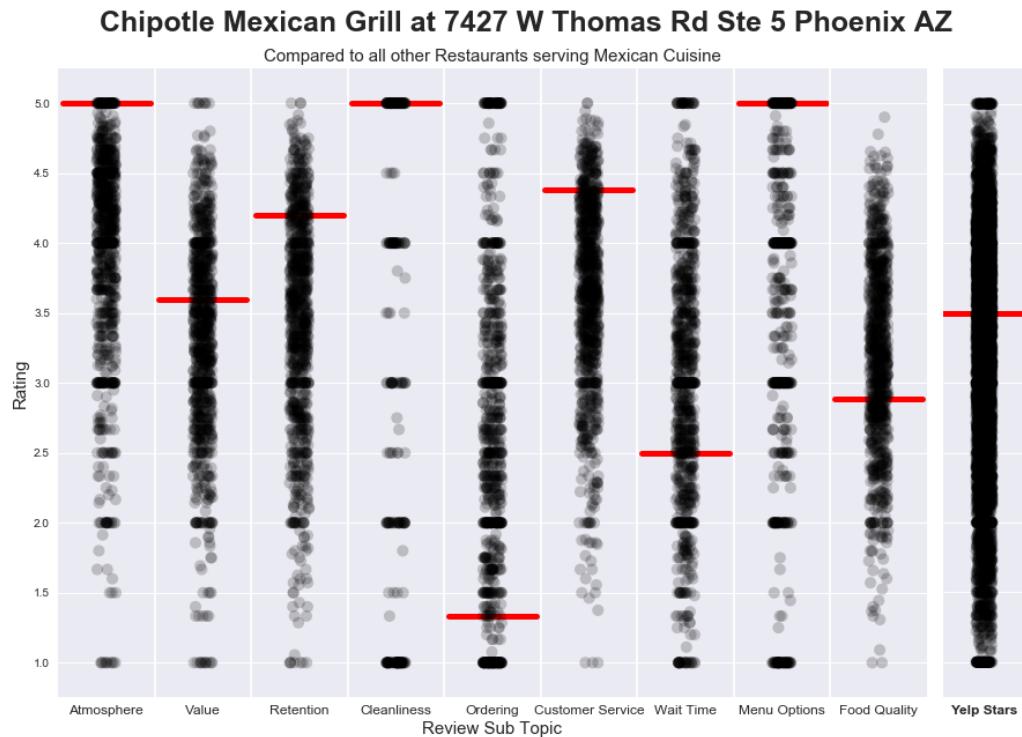
Report 2: Chipotle as compared to All Other Restaurants in the same Zip Code

Very similar trends in this comparison, this Chipotle is a local standout in all areas except for Ordering and Wait Time where it is trailing behind.



Report 3: Chipotle as compared to All Others Serving the Same Cuisine

This Chipotle is a leader among all Restaurants serving a Mexican Cuisine. Food Quality begins to be more of a concerning metric at this comparison. With the trend being towards 3.0-4.25 Stars while this location is just below the 3.0 Star mark.

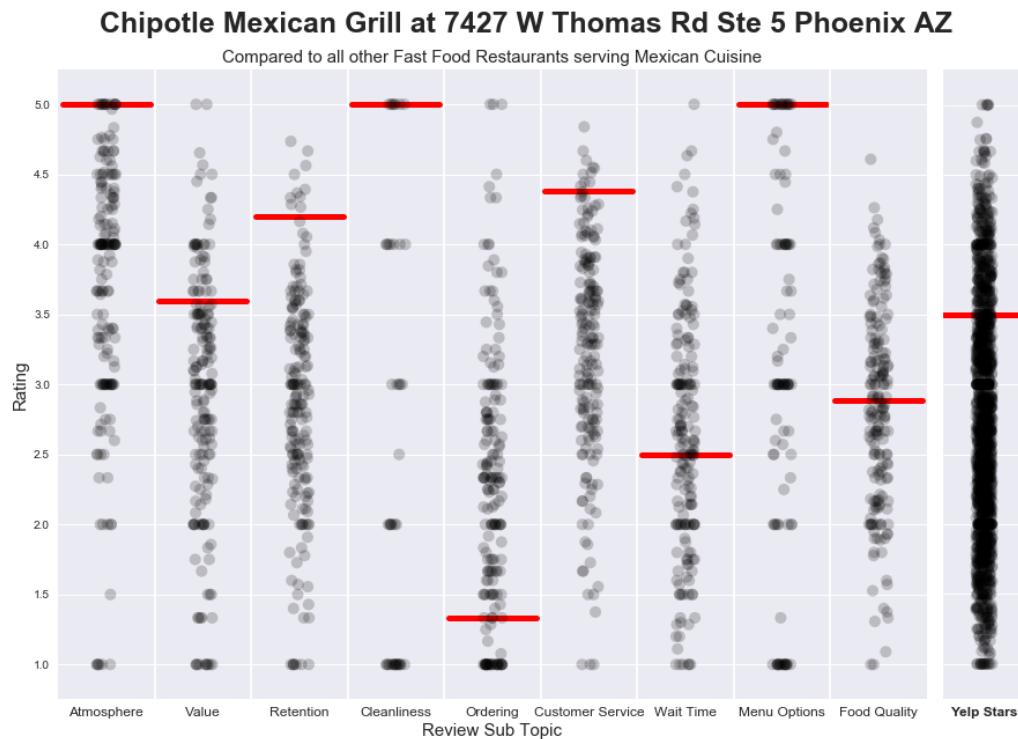


Report 4: Chipotle as compared to Fast Food Restaurants Serving the Same Cuisine

In Report 3 we saw how this Chipotle stacked up against all other restaurants serving a Mexican cuisine. But Chipotle is considered a 'Fast Food' option, and as such comparing this location to other Fast Food Mexican Restaurants gives us a better picture of where it stands against a more comparable competition.

Here the Food Quality concerns fade away a bit, with Mexican Fast Food Restaurants having a Food Quality Rating more dispersed between 2.0 and 3.5. Suddenly this Chipotle's 3.0 is not so bad!

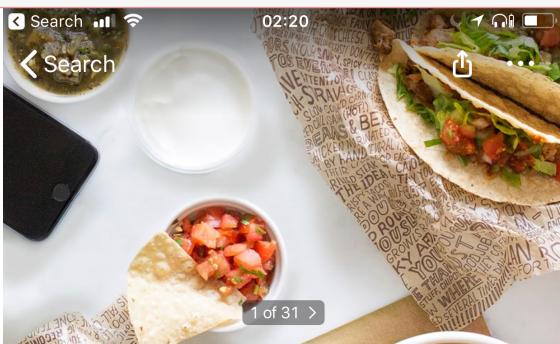
Another standout is the Retention Rating here. Retention is very highly rated at this location, which given the reportedly bad Ordering and Wait Time, is a head scratcher. But looking back at Report 2, there might just not be any other good options in this Zip Code, so this Chipotle is still the best option.



The New User View

With these new Sub Topic Stars for each location, the end Yelp users need a good way to see this information. These Sub Topics should be available for users who seek them, but given the wall of stars as displayed on the New Layout, would likely need to be presented only after a user interaction. Perhaps with a tap on the Star Rating bar in the 'Normal View' image would drop down to present a screen similar to the 'Sub Topic View'.

Normal View



Chipotle Mexican Grill Bookmark

★★★★★ 28 Reviews

\$\$\$\$ • Mexican, Fast Food

Closed 10:45 - 22:00

★ ★ ★ ★ ★

Add Photo Check In

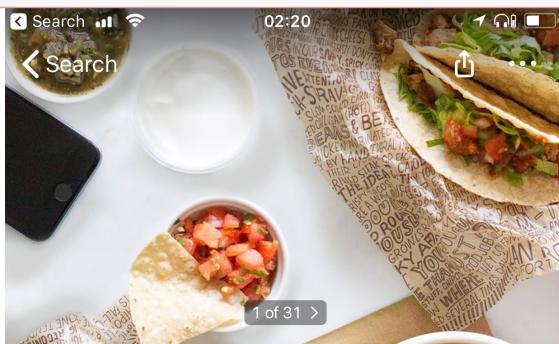


7427 W Thomas Rd, Ste 5, Phoenix, AZ 85033 >

Call (623) 240-1662 Get Directions 10 hr, 57 min

Nearby Search Me Delivery More

Sub Topic View



Chipotle Mexican Grill Bookmark

★★★★★ 28 Reviews

★★★★★ Atmosphere

★★★★★ Value

★★★★★ Retention

★★★★★ Cleanliness

★★★★★ Ordering

★★★★★ Customer Service

★★★★★ Wait Time

★★★★★ Menu Options

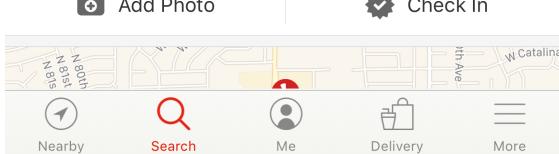
★★★★★ Food Quality

\$\$\$\$ • Mexican, Fast Food

Closed 10:45 - 22:00

★ ★ ★ ★ ★

Add Photo Check In



Nearby Search Me Delivery More

Findings Summary and Action Items

Mining Yelp Review Texts is a gold mine of information about restaurants and other businesses. While this analysis and investigation focuses on Restaurants in the state of Arizona. It would be worth while to next investigate for similar topics and trends in other states, or even other countries.

Continuing with this Arizona Restaurant subset, I would next move to a word2vec Continuous Bag of Words model for selecting tokens to pass off to an LDA model. This would help us determine more context in review texts, but would immensely increase the computational cost.

Additionally, Reviews occur over time, so it would be good to find trends over time in each Sub Review Topic and potentially identify the point of time in which a business booms or busts, ideally driving a means to deter businesses closing.

Methodology: Data Wrangling

The Yelp Open Dataset is provided directly by Yelp and consists of tips, reviews, users, businesses and check-ins data sets. Available for download in SQL or JSON format the dataset consists of more than 5.25 million reviews and over 1 million user tips from more than 1.1 million users. Included also is a collection of more than 156,000 business profiles.

The first big hurdle for analysis of the datasets is that the data file are large, and stored in JSON format. When loading business data only businesses located in Arizona are kept for processing, and from there filtering down to only Restaurants. This will provide us a still substantially large data set while limiting the domain of user reviews and tips to be more focused on restaurant related topics.

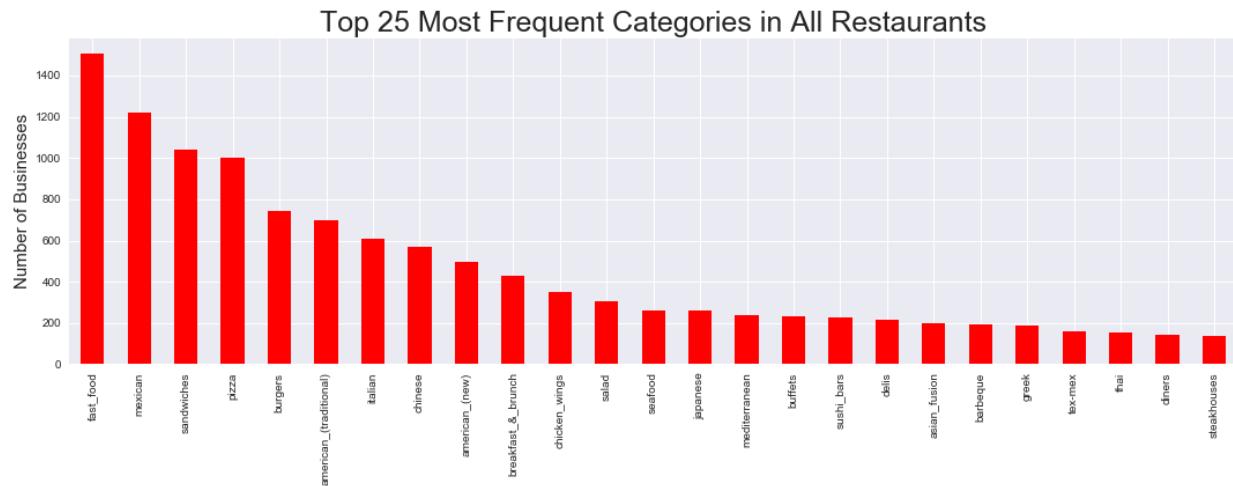
Business Dataset

The 'business.json' file consists of data for each business included in the datasets. Each business is assigned a unique 'business_id' attribute that can be used to cross reference reviews, tips, and checkins reported in the other data files. Business data contains information on the name, location, neighborhood, the total review count, the business' star rating, several attributes such as parking, validation, byob, and several other factors. These are largely ignored in the wrangling as they are all user reported and by default are set to nan, only updating to True or False when enough users report an update.

Business data also includes an array of cuisine categories for each business. This categories array and location are the largest filters used to limit the data set to only locations that are considered a 'restaurant' and are in the state of Arizona. A list of restaurant cuisines was manually collected from all possible reporting categories and a list of 221 possible options was retained. This list is used to cross

reference each business, removing any location that included a category value not in this list of restaurant categories.

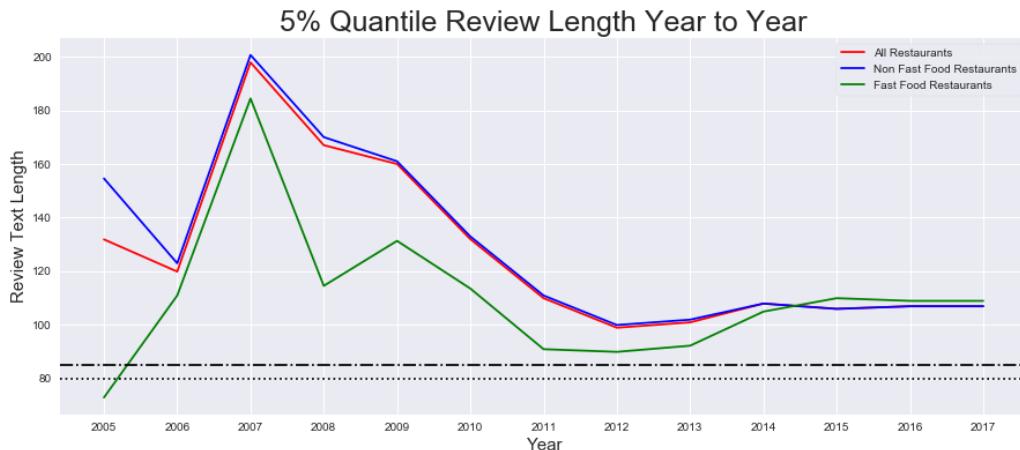
This reduced the initial set of 174,567 businesses to just 7066 total restaurants, 5561 non fast food restaurants and 1505 Fast Food restaurants.



Reviews Dataset

The original ‘review.json’ contains over 5.25 million restaurant reviews, after we have pruned down our selection of businesses to include only Arizona Restaurants, we can filter these reviews to include only those reviews given to restaurants in this list. We end up with a set of Arizona Restaurant Reviews consisting of 501,250 reviews split into 40,349 Fast Food Reviews and 460901 Non Fast Food Reviews.

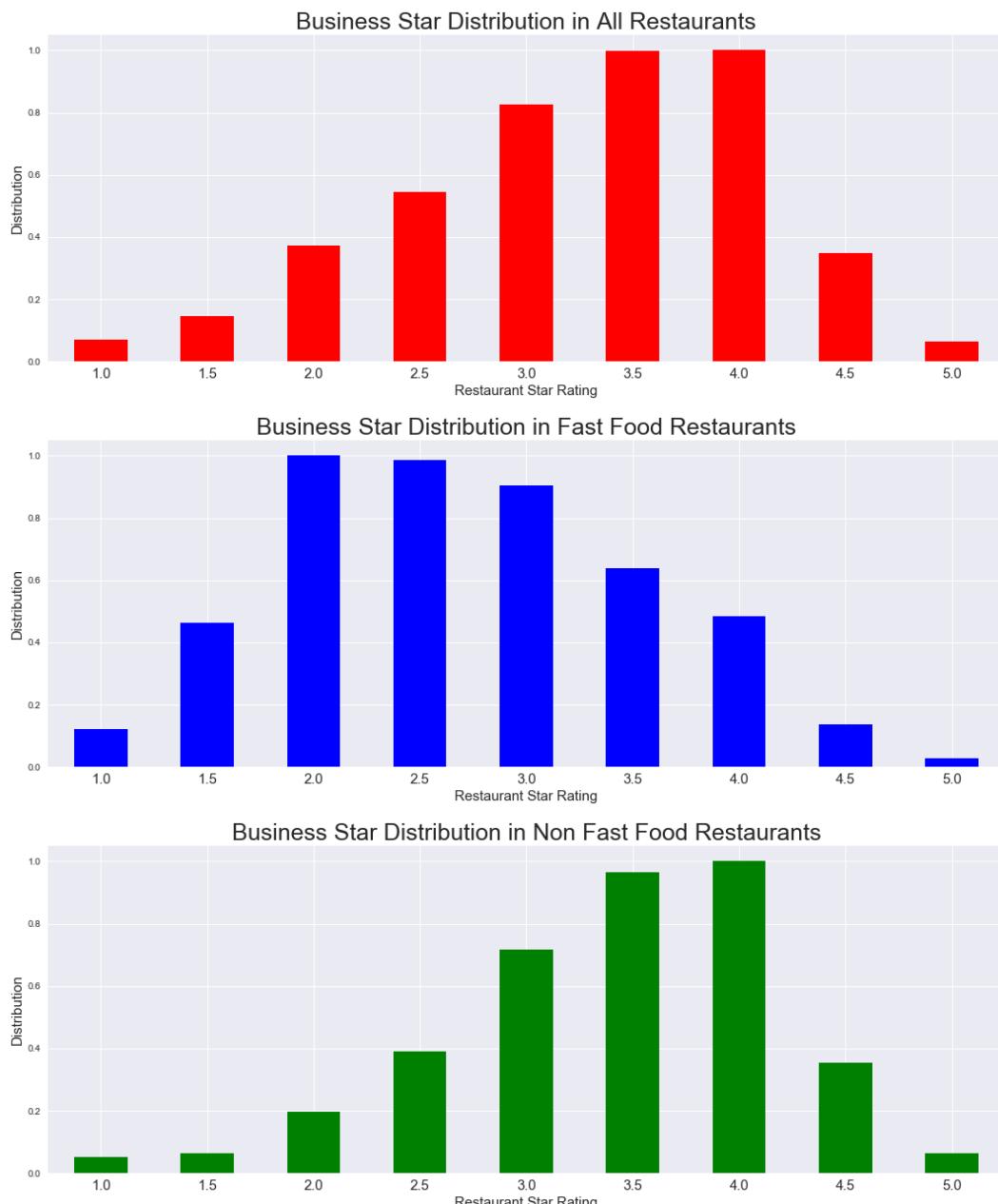
Next up, we clean up reviews by stripping out reviews with a text length less than 80 characters. The 80 character limit was reached by looking at the 5% Quantile of the review lengths year to year. Review lengths at this metric only dip below 80 characters in early 2005 so it is a safe line to draw in the sand.



Initial Findings - Restaurant Scores

Initial findings of the data set are interesting, and corroborate early beliefs that users are often creating accounts to write a few negative reviews or positive ones. Rarely is a review neutral (3 stars in this case). A wide disparity exists in the Business Star ratings of Fast Food restaurants and Non Fast Food restaurants.

Not quite perfect mirrors of each other, but it is immediately apparent that users are much more critical of fast food locations than they are of non fast food locations.



Initial Findings - Review Texts

After initial wrangling and prep of the review data, we tokenize the review text by removing stop words, removing punctuation, and lemmatizing (stemming) all reviews to clean tokens. Initial exploration of review text yields interesting results.

For each analysis, the tokens from all reviews are counted by frequency, and only tokens occurring more frequently than the 10,000th most frequent term are kept. Example if the 10,000th most frequent token appears 4 times in the corpus, only terms appearing 5 or more times are kept.

Given the difference between Fast Food and Non Fast Food review trends, I completed initial analysis of several subsets of the total reviews data set. Latent Dirichlet Allocation models are generated from 5 groups of reviews.

Several adjustments were made when running LDA models on the data set, on the next page, each run of the model and parameters I show each set of topics generated. Highlighted in Red are the the most easily summarized topic sets.

Model 1 Fast Food Review Categories:

4:Speed, 5:Drive Through, 12:Sanitary Status, 20:Order Accuracy, 24:Seating, 27:Price Value,
35:Customer Service, 41:Payment Options, 43:Order Wait Time, 46:Deals and Specials 48:Atmosphere

Model 2 Non Fast Food Review Categories:

1:Seating, 5:Loyalty, 8:Service, 15:Value, 16:Speed, 18:Customer Service

Model 3a All Restaurant Review Categories:

5:Loyalty, 8:Service, 11:Customer Service, 16:Bar Scene, 17:Value

Model 3b All Restaurant Review Categories:

1:Atmosphere, 8:Menu Option, 16:Wait Time, 17:Sanitary Status, 27:Service Time, 30:Customer Service,
33:Customer Service, 40:Bar Scene, 42:Food Quality, 44:Home Town Feel, 46:Menu Variety

Model 4 All Restaurant Review Categories, Nouns and Verbs Only :

3:Service, 10:Customer Experience, 13:Wait Time, 18:Problem Resolution, 19:Atmosphere, 24:Value,
27:Menu Variety

Model 5 All Restaurant Review Categories, all Tokens:

1:Retention, 3:Food Quality, 7:Wait Time, 8:Atmosphere, 9:Ordering, 13:Cleanliness, 17:Customer
Service, 19:Menu Options, 20:Food Quality, 21:Food Quality, 27:Customer Service, 35:Price Value,
44:Customer Service

Model 1 - Fast Food Reviews

Review Subset: Reviews Given to Fast Food businesses, most frequent tokens only
Number of Topics: 50
Number of Terms: 10
Number of Passes: 50
Results: 4:Speed, 5:Drive Through, 12:Sanitary Status, 20:Order Accuracy, 24:Seating,
27:Price Value, 35:Customer Service, 41:Payment Options, 43:Order Wait Time, 46:Deals and Specials
48:Atmosphere

```
0  ['chipotle', 'bowl', 'burrito', 'rice', 'bean', 'meat', 'chicken', 'steak', 'guacamole', 'get']  
1  ['taco', 'bell', 'burrito', 'mexican', 'salsa', 'carne', 'asada', 'good', 'food', 'tortilla']  
2  ['counter', 'one', 'behind', 'people', 'employee', 'order', 'line', 'guy', 'working', 'front']  
3  ['cold', 'fry', 'food', 'old', 'never', 'hot', 'warm', 'ever', 'even', 'fresh']  
4  ['lunch', 'quick', 'rush', 'dinner', 'work', 'stopped', 'get', 'today', 'busy', 'grab']  
5  ['drive', 'thru', 'window', 'car', 'minute', 'get', 'order', 'food', 'inside', 'wait']  
6  ['panda', 'express', 'job', 'team', 'professional', 'thanks', 'thank', 'polite', 'guy', 'done']  
7  ['asked', 'said', 'told', 'back', 'ordered', 'wanted', 'got', 'went', 'would', 'could']  
8  ['box', 'jack', 'server', 'surprised', 'waitress', 'rocket', 'johnny', 'mall', 'pleasantly', 'la']  
9  ['chicken', 'wei', 'pei', 'rice', 'food', 'dish', 'roll', 'orange', 'chinese', 'teriyaki']  
10  ['place', 'try', 'back', 'best', 'great', 'good', 'go', 'time', 'owner', 'definitely']  
11  ['location', 'one', 'time', 'staff', 'new', 'visit', 'particular', 'store', 'employee', 'need']  
12  ['dirty', 'clean', 'place', 'food', 'dining', 'area', 'kid', 'room', 'bathroom', 'floor']  
13  ['fish', 'chip', 'bar', 'priced', 'outstanding', 'beer', 'game', 'reasonably', 'menu', 'selection']  
14  ['lot', 'parking', 'across', 'street', 'wall', 'hole', 'space', 'sooo', 'spot', 'basic']  
15  ['manager', 'called', 'back', 'call', 'order', 'store', 'phone', 'would', 'number', 'could']  
16  ['place', 'would', 'like', 'good', 'really', 'better', 'much', 'get', 'little', 'pretty']  
17  ['hubby', 'lemon', 'sugar', 'malt', 'mild', 'fruit', 'cinnamon', 'btw', 'gf', 'content']  
18  ['review', 'star', 'yelp', 'write', 'read', 'based', 'negative', 'writing', 'reading', 'http']  
19  ['drink', 'soda', 'ice', 'cup', 'machine', 'tea', 'water', 'coke', 'cream', 'iced']  
20  ['order', 'time', 'get', 'wrong', 'go', 'every', 'right', 'location', 'never', 'correct']  
21  ['food', 'fast', 'good', 'service', 'place', 'pretty', 'really', 'friendly', 'restaurant', 'like']  
22  ['breakfast', 'mcdonald', 'egg', 'morning', 'coffee', 'mcdonalds', 'sausage', 'bacon', 'whataburger', 'brown']  
23  ['sub', 'subway', 'firehouse', 'meatball', 'shop', 'best', 'meat', 'bread', 'way', 'veggie']  
24  ['table', 'inside', 'restaurant', 'outside', 'area', 'seating', 'sit', 'clean', 'station', 'nice']  
25  ['sandwich', 'bread', 'cheese', 'meat', 'good', 'subway', 'turkey', 'like', 'mayo', 'tomato']  
26  ['pizza', 'topping', 'crust', 'place', 'pie', 'good', 'mod', 'great', 'like', 'sauce']  
27  ['price', 'small', 'size', 'special', 'portion', 'large', 'nothing', 'money', 'full', 'average']  
28  ['like', 'place', 'get', 'know', 'food', 'go', 'want', 'people', 'say', 'one']  
29  ['chicken', 'wing', 'piece', 'sauce', 'kfc', 'got', 'spicy', 'meal', 'ordered', 'tender']  
30  ['potato', 'wendy', 'gravy', 'mashed', 'popeye', 'sweet', 'airport', 'whopper', 'cheese', 'quarter']  
31  ['u', 'man', 'young', 'lady', 'hair', 'took', 'came', 'gave', 'went', 'guy']  
32  ['time', 'first', 'ordered', 'got', 'back', 'like', 'one', 'went', 'tasted', 'would']  
33  ['salad', 'fresh', 'mike', 'healthy', 'jersey', 'dressing', 'option', 'love', 'pita', 'ingredient']  
34  ['school', 'phoenix', 'carl', 'high', 'jr', 'bk', 'north', 'downtown', 'mcd', 'american']  
35  ['customer', 'service', 'employee', 'rude', 'manager', 'ever', 'worst', 'food', 'bad', 'attitude']  
36  ['night', 'open', 'late', 'close', 'hour', 'closed', 'door', 'sign', 'friday', 'closing']  
37  ['year', 'last', 'week', 'time', 'ago', 'food', 'eat', 'day', 'month', 'went']  
38  ['container', 'grill', 'heat', 'fork', 'plastic', 'put', 'throw', 'spoon', 'set', 'standard']  
39  ['sauce', 'flavor', 'green', 'red', 'hot', 'good', 'fresh', 'menu', 'delicious', 'also']  
40  ['yes', 'mac', 'cheese', 'big', 'drop', 'yeah', 'everyday', 'heart', 'blue', 'denny']  
41  ['card', 'extra', 'charge', 'pay', 'cash', 'cent', 'use', 'credit', 'money', 'dollar']  
42  ['always', 'love', 'get', 'location', 'time', 'go', 'friendly', 'never', 'staff', 'one']  
43  ['order', 'minute', 'wait', 'food', 'time', 'line', 'ready', 'long', 'waiting', 'waited']  
44  ['shake', 'burger', 'style', 'n', 'fry', 'animal', 'double', 'menu', 'love', 'east']  
45  ['sonic', 'drink', 'tot', 'happy', 'hour', 'frosty', 'slush', 'cherry', 'w', 'car']  
46  ['free', 'coupon', 'lettuce', 'wrap', 'buy', 'use', 'deal', 'birthday', 'one', 'learned']  
47  ['dog', 'hot', 'beef', 'chicago', 'chili', 'italian', 'fry', 'cheese', 'good', 'portillo']  
48  ['great', 'food', 'friendly', 'service', 'staff', 'place', 'delicious', 'fresh', 'clean', 'awesome']  
49  ['burger', 'fry', 'onion', 'cheese', 'guy', 'five', 'good', 'cheeseburger', 'bun', 'bacon']
```

Model 2 - Non Fast Food Reviews

Review Subset: Reviews Given to Non Fast Food businesses, most frequent tokens only

Number of Topics: 20

Number of Terms: 15

Number of Passes: 50

Results: 1:Seating, 5:Loyalty, 8:Service, 15:Value, 16:Speed, 18:Customer Service

```
0 ['place', 'back', 'try', 'got', 'time', 'good', 'really', 'went', 'first', 'ordered']
1 ['great', 'bar', 'highly', 'patio', 'beer', 'atmosphere', 'place', 'nice', 'outside', 'area']
2 ['salad', 'dessert', 'dish', 'bread', 'wine', 'pasta', 'meal', 'sauce', 'delicious', 'cheese']
3 ['like', 'food', 'place', 'would', 'could', 'one', 'taste', 'bad', 'even', 'really']
4 ['steak', 'shrimp', 'fish', 'fried', 'cooked', 'seafood', 'ordered', 'crab', 'potato', 'good']
5 ['place', 'always', 'love', 'best', 'food', 'get', 'go', 'great', 'time', 'favorite']
6 ['chicken', 'rice', 'dish', 'thai', 'soup', 'chinese', 'food', 'noodle', 'pho', 'sauce']
7 ['breakfast', 'egg', 'coffee', 'pancake', 'bacon', 'morning', 'toast', 'waffle', 'potato', 'brunch']
8 ['great', 'food', 'service', 'friendly', 'good', 'staff', 'place', 'back', 'delicious', 'definitely']
9 ['bbq', 'pork', 'rib', 'meat', 'sauce', 'side', 'brisket', 'mac', 'cheese', 'pulled']
10 ['burger', 'sandwich', 'fry', 'cheese', 'onion', 'bread', 'good', 'vegan', 'bun', 'sub']
11 ['sushi', 'roll', 'hour', 'happy', 'tuna', 'fish', 'fresh', 'chef', 'spicy', 'salmon']
12 ['time', 'location', 'year', 'last', 'restaurant', 'visit', 'since', 'food', 'first', 'one']
13 ['restaurant', 'make', 'menu', 'know', 'business', 'owner', 'review', 'one', 'staff', 'experience']
14 ['pizza', 'wing', 'crust', 'cheese', 'good', 'sauce', 'italian', 'slice', 'topping', 'thin']
15 ['good', 'food', 'place', 'price', 'pretty', 'really', 'get', 'like', 'would', 'little']
16 ['table', 'u', 'server', 'minute', 'drink', 'food', 'came', 'wait', 'time', 'service']
17 ['chicken', 'salad', 'pita', 'gyro', 'dog', 'greek', 'hummus', 'meat', 'tea', 'lunch']
18 ['order', 'food', 'time', 'said', 'back', 'would', 'service', 'asked', 'get', 'customer']
19 ['taco', 'mexican', 'salsa', 'burrito', 'chip', 'bean', 'good', 'cheese', 'green', 'tortilla']
```

Model 3a - All Restaurant Reviews, Nouns Only

Review Subset: All reviews, most frequent noun tokens only

Number of Topics: 20

Number of Terms: 15

Number of Passes: 50

Results: 5:Loyalty, 8:Service, 11:Customer Service, 16:Bar Scene, 17:Value

```
0 ['sushi', 'fish', 'shrimp', 'crab', 'place', 'seafood', 'tuna', 'chef', 'salmon', 'time']
1 ['breakfast', 'egg', 'coffee', 'bacon', 'waffle', 'morning', 'place', 'toast', 'potato', 'gravy']
2 ['dessert', 'cream', 'dish', 'menu', 'meal', 'dinner', 'restaurant', 'cake', 'appetizer', 'chocolate']
3 ['chicken', 'sauce', 'meat', 'bbq', 'pork', 'rib', 'side', 'flavor', 'potato', 'tender']
4 ['taco', 'street', 'fish', 'bell', 'tuesday', 'shrimp', 'shop', 'nacho', 'pastor', 'shell']
5 ['love', 'place', 'food', 'family', 'time', 'service', 'kid', 'staff', 'year', 'owner']
6 ['location', 'restaurant', 'phoenix', 'place', 'area', 'parking', 'lot', 'wall', 'spot', 'menu']
7 ['steak', 'pita', 'gyro', 'chicken', 'salad', 'hummus', 'meat', 'wrap', 'plate', 'lamb']
8 ['table', 'food', 'server', 'service', 'restaurant', 'time', 'drink', 'minute', 'waiter', 'meal']
9 ['salsa', 'chip', 'bean', 'food', 'cheese', 'rice', 'tortilla', 'margarita', 'sauce', 'chicken']
10 ['place', 'review', 'food', 'star', 'thing', 'people', 'make', 'way', 'something', 'day']
11 ['order', 'time', 'food', 'customer', 'service', 'minute', 'manager', 'location', 'employee', 'people']
12 ['burger', 'fry', 'dog', 'onion', 'cheese', 'place', 'time', 'order', 'soda', 'shake']
13 ['chicken', 'airport', 'jungle', 'masala', 'phx', 'sample', 'jerk', 'tikka', 'flight', 'island']
14 ['dish', 'chicken', 'rice', 'soup', 'spicy', 'noodle', 'food', 'sauce', 'beef', 'curry']
15 ['pizza', 'wing', 'place', 'cheese', 'sauce', 'slice', 'time', 'pie', 'try', 'order']
16 ['bar', 'hour', 'night', 'drink', 'beer', 'place', 'music', 'friday', 'atmosphere', 'fun']
17 ['food', 'place', 'service', 'price', 'restaurant', 'lunch', 'time', 'quality', 'menu', 'staff']
18 ['salad', 'sandwich', 'bread', 'lunch', 'cheese', 'chicken', 'pasta', 'sub', 'tomato', 'meat']
19 ['bowl', 'pho', 'tea', 'spring', 'water', 'ice', 'cup', 'broth', 'pot', 'express']
```

Model 3b - All Restaurant Reviews, Nouns Only

Review Subset: All reviews, most frequent noun tokens only
Number of Topics: 50
Number of Terms: 10
Number of Passes: 30
Results: 1:Atmosphere, 8:Menu Option, 16:Wait Time, 17:Sanitary Status, 27:Service Time, 30:Customer Service, 33:Customer Service, 40:Bar Scene, 42:Food Quality, 44:Home Town Feel, 46:Menu Variety

```
0 ['food', 'service', 'price', 'place', 'quality', 'restaurant', 'portion', 'time', 'atmosphere', 'size']  
1 ['kid', 'coupon', 'tuesday', 'deal', 'dad', 'adult', 'regret', 'meal', 'def', 'b']  
2 ['burger', 'fry', 'onion', 'place', 'cheese', 'shake', 'ring', 'bacon', 'well', 'time']  
3 ['star', 'review', 'yelp', 'reason', 'read', 'rating', 'rate', 'update', 'living', 'state']  
4 ['lunch', 'day', 'today', 'time', 'work', 'week', 'dinner', 'spot', 'service', 'yesterday']  
5 ['spicy', 'curry', 'food', 'spice', 'dish', 'chicken', 'pad', 'rice', 'flavor', 'place']  
6 ['bowl', 'chicken', 'rice', 'noodle', 'egg', 'food', 'soup', 'ramen', 'sauce', 'place']  
7 ['chicken', 'strip', 'mall', 'waffle', 'piece', 'breast', 'tender', 'crispy', 'shawarma', 'meal']  
8 ['dish', 'restaurant', 'menu', 'course', 'meal', 'chef', 'well', 'cuisine', 'flavor', 'try']  
9 ['sauce', 'onion', 'pepper', 'tomato', 'flavor', 'mushroom', 'oil', 'cheese', 'taste', 'lemon']  
10 ['box', 'bomb', 'chop', 'stew', 'claim', 'alot', 'vacation', 'bento', 'earth', 'freshness']  
11 ['hand', 'daughter', 'bathroom', 'child', 'hair', 'parent', 'restroom', 'paper', 'book', 'cover']  
12 ['plate', 'lamb', 'vegetable', 'platter', 'rice', 'leaf', 'kabob', 'meat', 'sampler', 'combination']  
13 ['order', 'time', 'line', 'minute', 'food', 'drive', 'counter', 'guy', 'girl', 'store']  
14 ['salad', 'pita', 'gyro', 'hummus', 'lettuce', 'wrap', 'side', 'meat', 'bread', 'jungle']  
15 ['place', 'food', 'thing', 'time', 'make', 'something', 'way', 'taste', 'try', 'cook']  
16 ['table', 'minute', 'waiter', 'food', 'water', 'restaurant', 'waiting', 'place', 'service', 'people']  
17 ['picture', 'groupon', 'step', 'baby', 'crap', 'floor', 'foot', 'photo', 'fly', 'hell']  
18 ['night', 'dinner', 'friday', 'date', 'saturday', 'evening', 'friend', 'girlfriend', 'thursday', 'movie']  
19 ['shrimp', 'crab', 'seafood', 'lobster', 'fish', 'scallop', 'puff', 'leg', 'oyster', 'sea']  
20 ['area', 'phoenix', 'trip', 'restaurant', 'city', 'arizona', 'hotel', 'finding', 'town', 'summer']  
21 ['soup', 'pho', 'noodle', 'spring', 'broth', 'place', 'beef', 'bowl', 'restaurant', 'service']  
22 ['pie', 'joe', 'apple', 'market', 'goat', 'lasagna', 'farm', 'partner', 'bruschetta', 'pecan']  
23 ['country', 'fruit', 'refreshing', 'lemonade', 'sugar', 'cinnamon', 'juice', 'pineapple', 'air', 'temperature']  
24 ['cream', 'dog', 'ice', 'chipotle', 'chocolate', 'cookie', 'cheese', 'cooky', 'scoop', 'vanilla']  
25 ['chip', 'salsa', 'fish', 'margarita', 'food', 'guacamole', 'place', 'restaurant', 'service', 'dip']  
26 ['tea', 'drink', 'soda', 'cup', 'bag', 'tip', 'charge', 'coke', 'machine', 'container']  
27 ['server', 'drink', 'food', 'hour', 'time', 'service', 'order', 'min', 'half', 'friend']  
28 ['location', 'tempe', 'downtown', 'chandler', 'chain', 'mesa', 'drive', 'school', 'mile', 'year']  
29 ['group', 'parking', 'lot', 'party', 'people', 'door', 'reservation', 'birthday', 'sign', 'everyone']  
30 ['staff', 'food', 'family', 'year', 'time', 'visit', 'place', 'restaurant', 'service', 'son']  
31 ['breakfast', 'egg', 'coffee', 'bacon', 'place', 'morning', 'toast', 'potato', 'gravy', 'brunch']  
32 ['style', 'mouth', 'tasteless', 'word', 'suck', 'walking', 'bf', 'express', 'edamame', 'calzone']  
33 ['customer', 'service', 'manager', 'time', 'food', 'employee', 'business', 'owner', 'care', 'restaurant']  
34 ['pasta', 'bread', 'sausage', 'spaghetti', 'garden', 'olive', 'catfish', 'alfredo', 'lil', 'square']  
35 ['beef', 'bean', 'rice', 'combo', 'cheese', 'chile', 'corn', 'tamale', 'side', 'ground']  
36 ['taco', 'tortilla', 'street', 'place', 'salsa', 'quesadilla', 'bell', 'meat', 'nacho', 'food']  
37 ['love', 'place', 'food', 'owner', 'time', 'service', 'try', 'everything', 'make', 'everyone']  
38 ['pork', 'meat', 'bbq', 'rib', 'wing', 'sauce', 'side', 'place', 'flavor', 'potato']  
39 ['pizza', 'place', 'cheese', 'slice', 'wing', 'time', 'try', 'pie', 'pepperoni', 'sauce']  
40 ['bar', 'dining', 'drink', 'area', 'sit', 'room', 'seating', 'bartender', 'hour', 'beer']  
41 ['wife', 'valley', 'gem', 'west', 'sister', 'disappoint', 'crepe', 'sooo', 'coast', 'scratch']  
42 ['bland', 'money', 'salt', 'greasy', 'smell', 'waste', 'hate', 'disappointment', 'save', 'soggy']  
43 ['wall', 'boyfriend', 'dollar', 'hole', 'mark', 'crave', 'college', 'father', 'spending', 'christmas']  
44 ['beer', 'fun', 'music', 'atmosphere', 'game', 'watch', 'place', 'tv', 'folk', 'vibe']  
45 ['option', 'delivery', 'mom', 'vegan', 'gluten', 'pop', 'driver', 'app', 'eater', 'jack']  
46 ['menu', 'item', 'selection', 'buffet', 'variety', 'try', 'choice', 'everything', 'thing', 'lot']  
47 ['sandwich', 'cheese', 'bread', 'sub', 'mac', 'meat', 'place', 'shop', 'subway', 'try']  
48 ['steak', 'dessert', 'meal', 'dinner', 'cake', 'potato', 'entree', 'appetizer', 'bread', 'salad']  
49 ['sushi', 'place', 'tuna', 'chef', 'salmon', 'fish', 'hour', 'california', 'service', 'price']
```

Model 4 - All Restaurant Reviews, Nouns and Verbs Only

Review Subset: All reviews, most frequent noun and verb tokens only
Number of Topics: 30
Number of Terms: 15
Number of Passes: 50
Results: 3:Service, 10:Customer Experience, 13:Wait Time, 18:Problem Resolution, 19:Atmosphere, 24:Value, 27:Menu Variety

```
0  ['shrimp', 'fish', 'crab', 'seafood', 'lobster', 'fried', 'leg', 'sauce', 'clam', 'oyster']
1  ['parking', 'door', 'lot', 'closed', 'drive', 'car', 'sign', 'street', 'find', 'mile']
2  ['food', 'eat', 'place', 'could', 'got', 'tasted', 'would', 'go', 'eating', 'looked']
3  ['u', 'server', 'food', 'drink', 'service', 'came', 'time', 'waiter', 'table', 'ordered']
4  ['flavor', 'sauce', 'dish', 'spicy', 'taste', 'bland', 'spice', 'bit', 'ordered', 'would']
5  ['ice', 'cream', 'bowl', 'soda', 'cup', 'drink', 'water', 'coke', 'machine', 'fountain']
6  ['taco', 'mexican', 'salsa', 'burrito', 'chip', 'bean', 'food', 'cheese', 'tortilla', 'margarita']
7  ['lunch', 'salad', 'tea', 'curry', 'today', 'pad', 'soup', 'dinner', 'lettuce', 'went']
8  ['phoenix', 'year', 'arizona', 'know', 'lived', 'find', 'life', 'city', 'moved', 'mouth']
9  ['star', 'place', 'review', 'give', 'would', 'try', 'say', 'yelp', 'think', 'go']
10  ['customer', 'service', 'food', 'employee', 'owner', 'people', 'make', 'need', 'guy', 'staff']
11  ['salad', 'dessert', 'pasta', 'wine', 'bread', 'cheese', 'chocolate', 'dish', 'cake', 'appetizer']
12  ['place', 'get', 'time', 'go', 'love', 'food', 'ca', 'order', 'wait', 'come']
13  ['table', 'minute', 'wait', 'u', 'food', 'order', 'waiting', 'seated', 'waited', 'people']
14  ['order', 'said', 'asked', 'told', 'time', 'would', 'ordered', 'called', 'manager', 'got']
15  ['pizza', 'wing', 'crust', 'cheese', 'place', 'sauce', 'topping', 'slice', 'pie', 'salad']
16  ['burger', 'fry', 'dog', 'onion', 'cheese', 'place', 'shake', 'got', 'ring', 'bacon']
17  ['sandwich', 'bread', 'sub', 'cheese', 'meat', 'beef', 'shop', 'subway', 'store', 'got']
18  ['night', 'dinner', 'went', 'u', 'time', 'husband', 'friend', 'meal', 'food', 'restaurant']
19  ['food', 'place', 'service', 'restaurant', 'love', 'staff', 'price', 'recommend', 'family', 'atmosphere']
20  ['bbq', 'pork', 'rib', 'meat', 'side', 'mac', 'sauce', 'brisket', 'cheese', 'pulled']
21  ['location', 'tempe', 'strip', 'mall', 'chandler', 'downtown', 'west', 'mesa', 'valley', 'corner']
22  ['steak', 'cooked', 'potato', 'well', 'medium', 'perfection', 'side', 'fried', 'filet', 'mashed']
23  ['chicken', 'pita', 'gyro', 'salad', 'hummus', 'meat', 'rice', 'lamb', 'plate', 'ordered']
24  ['money', 'tip', 'price', 'pay', 'card', 'dollar', 'charge', 'buck', 'coupon', 'cost']
25  ['chicken', 'rice', 'soup', 'fried', 'noodle', 'pho', 'beef', 'roll', 'food', 'dish']
26  ['breakfast', 'egg', 'coffee', 'bacon', 'pancake', 'waffle', 'morning', 'toast', 'brunch', 'biscuit']
27  ['menu', 'option', 'food', 'item', 'price', 'restaurant', 'buffet', 'portion', 'choice', 'variety']
28  ['bar', 'beer', 'place', 'drink', 'hour', 'music', 'atmosphere', 'fun', 'night', 'area']
29  ['sushi', 'roll', 'hour', 'place', 'tuna', 'ramen', 'chef', 'salmon', 'fish', 'spicy']
```

Model 5 - All Restaurant Reviews, All Tokens

Review Subset: All reviews, most frequent tokens only
Number of Topics: 50
Number of Terms: 10
Number of Passes: 50
Results: 1:Retention, 3:Food Quality, 7:Wait Time, 8:Atmosphere, 9:Ordering, 13:Cleanliness, 17:Customer Service, 19:Menu Options, 20:Food Quality, 21:Food Quality, 27:Customer Service, 35:Price Value, 44:Customer Service

```
0  ['dessert', 'wine', 'ice', 'cream', 'cake', 'entree', 'meal', 'chocolate', 'course', 'appetizer']
1  ['time', 'back', 'first', 'try', 'place', 'went', 'definitely', 'go', 'next', 'great']
2  ['u', 'table', 'came', 'server', 'food', 'drink', 'waitress', 'asked', 'ordered', 'minute']
3  ['dish', 'flavor', 'sauce', 'like', 'taste', 'menu', 'one', 'would', 'bit', 'meat']
4  ['pizza', 'crust', 'slice', 'topping', 'cheese', 'pie', 'thin', 'good', 'sauce', 'pepperoni']
5  ['crab', 'leg', 'shell', 'pound', 'coworker', 'saving', 'panini', 'hub', 'e', 'angry']
6  ['chicken', 'rice', 'chinese', 'fried', 'food', 'beef', 'soup', 'egg', 'orange', 'sour']
7  ['wait', 'minute', 'time', 'food', 'get', 'order', 'long', 'line', 'hour', 'waiting']
8  ['great', 'nice', 'patio', 'atmosphere', 'outside', 'place', 'cool', 'fun', 'inside', 'food']
9  ['order', 'ordered', 'called', 'delivery', 'extra', 'got', 'time', 'get', 'card', 'call']
10 ['tempe', 'school', 'opening', 'mill', 'w', 'college', 'b', 'asu', 'high', 'grand']
11 ['fresh', 'ingredient', 'olive', 'tomato', 'salad', 'oil', 'made', 'bianco', 'mozzarella', 'basil']
12 ['best', 'year', 'place', 'restaurant', 'phoenix', 'one', 'ever', 'food', 'valley', 'family']
13 ['table', 'dirty', 'clean', 'bathroom', 'floor', 'plate', 'hand', 'cup', 'paper', 'chair']
14 ['chicken', 'fried', 'waffle', 'deep', 'breast', 'piece', 'cajun', 'tender', 'catfish', 'crispy']
15 ['mouth', 'smell', 'taste', 'word', 'world', 'heaven', 'bud', 'holy', 'win', 'watering']
16 ['always', 'love', 'place', 'get', 'time', 'food', 'go', 'favorite', 'great', 'never']
17 ['food', 'experience', 'service', 'great', 'server', 'u', 'staff', 'friendly', 'restaurant', 'made']
18 ['sushi', 'roll', 'tuna', 'fresh', 'chef', 'salmon', 'japanese', 'spicy', 'tempura', 'fish']
19 ['option', 'menu', 'free', 'gyro', 'meat', 'vegetarian', 'veggie', 'choose', 'vegan', 'choice']
20 ['food', 'like', 'place', 'ordered', 'tasted', 'bad', 'even', 'back', 'cold', 'taste']
21 ['good', 'food', 'place', 'price', 'pretty', 'service', 'better', 'like', 'would', 'really']
22 ['burger', 'fry', 'onion', 'bun', 'shake', 'ring', 'cheese', 'french', 'bacon', 'patty']
23 ['bowl', 'pho', 'soup', 'noodle', 'roll', 'spring', 'broth', 'vietnamese', 'beef', 'rice']
24 ['star', 'review', 'give', 'yelp', 'five', 'reason', 'rating', 'giving', 'based', 'shot']
25 ['location', 'drive', 'employee', 'thru', 'window', 'car', 'fast', 'chipotle', 'dont', 'im']
26 ['restaurant', 'parking', 'menu', 'lunch', 'small', 'lot', 'area', 'find', 'little', 'place']
27 ['great', 'food', 'service', 'place', 'friendly', 'good', 'recommend', 'staff', 'delicious', 'price']
28 ['wing', 'pasta', 'sauce', 'italian', 'garlic', 'bread', 'meatball', 'spaghetti', 'dish', 'ranch']
29 ['steak', 'cooked', 'potato', 'perfectly', 'well', 'medium', 'rare', 'filet', 'mashed', 'perfection']
30 ['bar', 'room', 'dining', 'bartender', 'area', 'sit', 'drink', 'airport', 'sat', 'table']
31 ['breakfast', 'egg', 'coffee', 'pancake', 'bacon', 'morning', 'toast', 'brunch', 'biscuit', 'gravy']
32 ['fish', 'shrimp', 'ramen', 'chip', 'fried', 'sea', 'ceviche', 'ahi', 'sauce', 'mahi']
33 ['hot', 'dog', 'chicago', 'beef', 'chili', 'piping', 'italian', 'slider', 'wendy', 'pepper']
34 ['place', 'like', 'get', 'know', 'go', 'one', 'say', 'want', 'make', 'really']
35 ['good', 'really', 'got', 'ordered', 'little', 'nice', 'pretty', 'came', 'portion', 'also']
36 ['sandwich', 'bread', 'sub', 'cheese', 'turkey', 'meat', 'subway', 'shop', 'lunch', 'philly']
37 ['seafood', 'lobster', 'shrimp', 'pot', 'calamari', 'chop', 'sum', 'duck', 'dim', 'clam']
38 ['taco', 'mexican', 'salsa', 'burrito', 'chip', 'bean', 'tortilla', 'food', 'carne', 'asada']
39 ['bbq', 'pork', 'rib', 'meat', 'brisket', 'sauce', 'pulled', 'side', 'good', 'corn']
40 ['salad', 'chicken', 'pita', 'dressing', 'greek', 'hummus', 'lettuce', 'wrap', 'fresh', 'lunch']
41 ['kid', 'music', 'loud', 'wall', 'game', 'child', 'hole', 'playing', 'play', 'family']
42 ['cheese', 'sweet', 'delicious', 'sauce', 'perfect', 'mac', 'onion', 'green', 'potato', 'pepper']
43 ['happy', 'hour', 'beer', 'drink', 'selection', 'great', 'menu', 'special', 'price', 'deal']
44 ['customer', 'manager', 'service', 'back', 'food', 'said', 'time', 'asked', 'one', 'would']
45 ['buffet', 'indian', 'food', 'lamb', 'chicken', 'dish', 'naan', 'rice', 'masala', 'restaurant']
46 ['son', 'sister', 'mother', 'brother', 'http', 'father', 'c', 'buck', 'law', 'wallet']
47 ['night', 'dinner', 'last', 'party', 'friday', 'went', 'reservation', 'birthday', 'group', 'saturday']
48 ['thai', 'tea', 'spicy', 'curry', 'pad', 'tofu', 'iced', 'dish', 'spice', 'rice']
49 ['grill', 'cookie', 'hawaiian', 'cooky', 'mini', 'macaroni', 'island', 'paradise', 'scoop', 'nugget']
```

Applying Sub Topic Stars

To assign a Sub Topic to a review, we count the number of tokens in each review that appear in each Sub Topic. Note that some Sub Topics include more than one set of tokens.

Assigning Topics to Reviews

Each review can be tagged with more than one Sub Topic, being tagged with a subtopic was determined by passing a min token count to the function assigning Sub Topics.

Each Review is given a Sub Topic Rating equal to the original Review Stars the yelper gave their review. Example, a Yelper write a 4 Star Review and write in depth about the value and cleanliness of the Restaurant, the value and cleanliness Sub Topic Ratings will both be set to 4 Stars for that Review, all other Sub Topic Rating are left (null) so as not to skew averages for that Restaurant.

Calculating Restaurant Sub Topic Stars

Calculating the Sub Topic Stars for each Restaurant is easy, and performed the same way Yelp calculates Review Stars. A simple average is applied to each Sub Topic Review Star count, grouped by Restaurant business_id.

Sources

[1] Reviews, Reputation, and Revenue: The Case of Yelp.com by Michael Luca [https://www.hbs.edu/faculty/Publication%20Files/12-016_a7e4a5a2-03f9-490d-b093-8f951238dba2.pdf]

[2] The Yelp Open Dataset [<https://www.yelp.com/dataset>]

Code for these Analysis can be found at on my [GitHub Page](#)