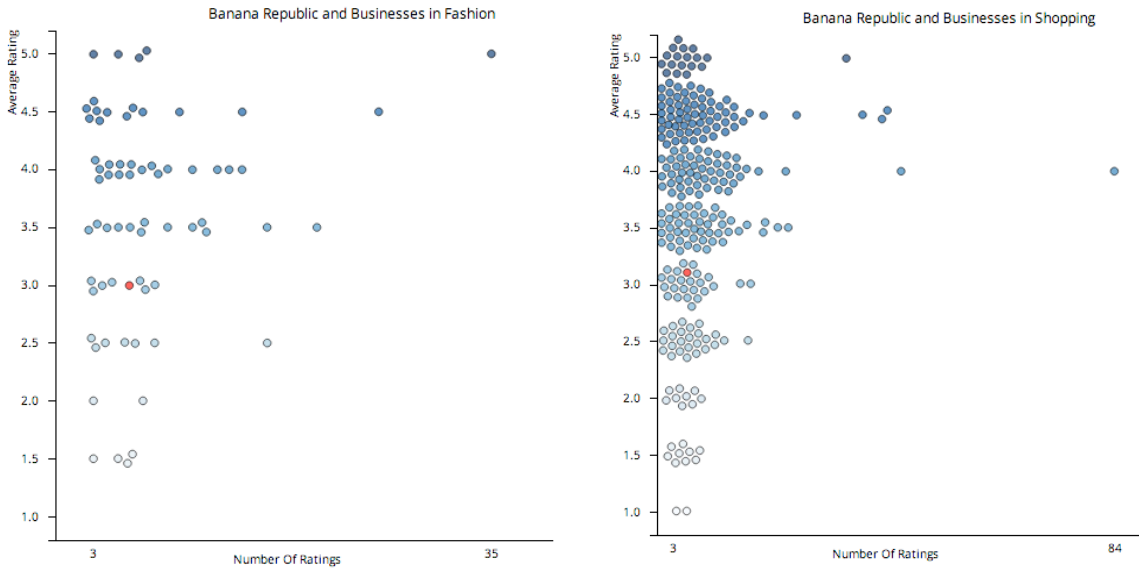# Visualizing Yelp Ratings: Interactive Analysis and Comparison of Businesses

Akhila Raju, Cecile Basnage, and Jimmy Yin

Abstract - In this paper, we present an interactive tool for visualizing customer reviews from the Yelp dataset. The goal was to provide a way for businesses to visualize their review data. These visualizations can aid business owners in understanding how their performance compares with other businesses of similar type category or location, to make more informed decisions that could improve their success. We used the open dataset from the Yelp Dataset Challenge and focused our analysis on one city to gain a more thorough understanding. We present the findings of the various methods we decided to filter the dataset and display the corresponding review information.

---

1 INTRODUCTION

Although Yelp provides consumers an efficient way to learn about businesses, there are fewer avenues for business owners themselves to explore data that is relevant to them. Owners do not always have a clear idea of the tradeoffs customers make when deciding where to shop. There are a multitude of factors at play – such as price, location, environment, ambiance, and a multitude of others that factor into the success of a business.

For problems with high-dimensional data, extracting relevant features is a challenge.[1] We therefore aim to gain more specific insights by filtering business review data across categories. It is known that Yelp ratings have a substantial effect on the success of a business. Simply an extra half-star rating "causes restaurants to sell out 19 percentage points more frequently." [2] However, there currently does not exist a tool for businesses to compare themselves to others in similar categories in order to determine why customers might be rating another business higher. Therefore, a deeper understanding of the various factors behind business success could prove valuable to them. Motivated by the lack of an existing interface for business owners to peruse pertinent data, we set out to create a tool that would help businesses understand their performance relative to others similar to them. This underutilized information has the potential to empower owners to make data-driven decisions in their business.

2 RELATED WORK

Review visualization may be categorized as either quantitative feature based or content based.[3] The former are used to understand quantitative data such as rating and price. In contrast, content reviews incorporate the content of the text into the visualization itself. Wang et al use this approach to create a word cloud that allows users to visualize the text that occurs the most in different reviews.[3] Other research implements the latter visualization method and uses natural language processing to filter reviews and visually display their opinion.[4]

2.1 Extracting Subtopics from Reviews

In order to visually display insights from a group of documents, there are several methods for extracting topics from text. A common choice is the Latent Dirichlet Algorithm, a Bayesian generative model for text. In "Improving Restaurants by extracting Subtopics from Yelp Reviews," Huang et al use Latent Dirichlet Algorithm to uncover latent subtopics in Yelp reviews and visualize the results. LDA "is used to discover underlying topics in a text document. [it] assumes that a corpus of text documents cover a collection of K topics. Each topic is defines as a multinomial distribution over a word dictionary |V| words."[1] The result was predictions of star ratings within a certain sub-topic discovered by applying the algorithm. These predictions were then displayed visually. Although our focus is not in applying an algorithm, we similarly hope to provide insights by filtering data with subtopics.

2.2 Semantic Topic Modeling

Other research has attempted to tweak LDA. Linshi proposes an approximation of LDA that "conditions topics' term distributions not only on the Dirichlet parameter, but also on star ratings".[5] Linshi proposes that ratings are a function of words with positive and negative connotations. The paper displays graphical representations in the form of pie charts to demonstrate that a modified LDA is more successful in identifying the human interpretation of review texts compared to the traditional LDA. Linshi presents the introduction of code words as an effective means to better approximate the subtopics of review text. Since a tweaking of the algorithm drastically changes the results of LDA, we can see the limitations of using simple mathematical models without human understanding to interpret large datasets.

2.3 Understanding Ratings with Review Text

Similarly, McAuley and Leskovec's attempt to merge latent review topics with latent ratings with the HFT model.

They note that "in spite of the wealth of research on modeling ratings, other forms of feedback present in online review websites--namely, the reviews themselves-- is typically ignored…. [this is] a major shortcoming of existing work on recommender systems."[6] HFT thus aligns factors in product ratings with factors in product reviews. In our future work, we hope to incorporate text reviews into our tool as well.

3 METHODS

Yelp is an informative tool that allows users to explore and compare businesses through customer reviews. Yelp provides transparency and detail about businesses that it may be difficult for users to otherwise get access to without physically going the restaurant - such as noise level, ambience, and attire. Crowdsourcing reviews allows users to get a more rounded analysis of the business they are exploring.

Yelp provides useful insight to the consumer about businesses. However, Yelp lacks an interface made for business owners. Given how much data Yelp has access to, we found that a business owner-facing application is an area largely unexplored. Such a tool could provide business owners with the insights they need to improve their business in accordance with how customers rate their business.

3.1 User Interviews

Initially, we assumed a business comparison tool would be helpful to business owners. We decided to take a user-centered design approach and interview business owners in the cities of Berkeley and San Francisco. We created a set of questions which would allow us to understand how business owners interact with Yelp, and how they compare themselves to competition based on similarity and proximity.

Interview Questions

1. How do you identify your competition?
2. How do you use Yelp to see how your business is doing?
3. How does the information you find about your business on Yelp affect your business?
4. How do reviews of your competition affect your business?
5. If you could compare your store to businesses around you, what would you want to know?

6. If you could compare your store to business that are similar to you, what would you want to know?

7. Tell me a time when you saw something on Yelp that impacted your business.

We spoke to fifteen businesses, three of which were in Berkeley. After speaking to a few business owners in Berkeley, we found that competition wasn't something they worried about. These business owners cared much more about providing good products and services. This may be because college students are more concerned with the price and convenience of a business than its quality. Thus, we interviewed business owners in San Francisco, where the consumer base is highly economically and socially diverse.

We spoke to a businesses in San Francisco's Embarcadero and Mission districts. We spoke to businesses varying in age and category. Some had been open for over three decades, whereas others had been started within the last year. Among these businesses were restaurants, clothing stores, an apothecary, and a juice bar.

Consolidation of Interview Results

We gained very useful insights from speaking to businesses that guided us in our design process and implementation. We found that the businesses we spoke to often had a negative view of Yelp; yet all businesses utilize Yelp and check the reviews left by customers at least once a week. Most business owners expressed distaste towards the presence of anonymous "hate" reviews which oftentimes do not provide much insight or suggestions for how a business can improve. When asked about comparison, These businesses expressed in interest in wanting to see general trends in their Yelp reviews, such as how the business is doing over time.

Although most businesses were interested primarily in their own business's trends, some businesses expressed interest in a simple business comparison tool, such as one that can compare their business to those across their category (such as wanting to compare a juice store to all juice stores in San Francisco). These business owners were generally younger-generation entrepreneurs who regularly used Yelp to get insights on the performance of their own business as well as their competitors'.

The problem we wanted to approach is to that business owners cannot easily extract useful information and about their business and see trends in the progress of their business. Keeping the human-centered design process in mind, we designed a tool that would allow users to explore businesses and view visualizations of ratings and review content.

We initially designed the tool to allow for exploration; users can enter a category for the type of business, search for a keyword, or change the location. We iterated on this initial design to be more focused towards use by a business owner in comparison to a consumer. In the final design phase, we chose to allow users to enter their business name and enter a location-based filter for which they wanted to compare their businesses. Because this design has one use-case, it is more focused and requires minimal cognitive load on the user.

3.2   Data Wrangling

The dataset provided by Yelp was formatted in JSON (JavaScript Object Notation) files, with each file containing one JSON object per line. For this project, we used the business and review data, which contained the variables of interest to us, including business attributes, star ratings, and review metadata.

The first step of software implementation involved manipulating the data in a way that could be handled readily by JavaScript and the D3 library. This initial step was done in a Python script. Given the sheer size of the review file (1.05GB), we tested out various JSON decoders and found that UltraJSON, a library written in pure C, gave us the best results. Ultimately, we chose to filter the dataset by a single city (Madison, WI) and store each JSON object in a large array to support decoding by the D3 JSON function. By focusing our implementation on the data of a single city, we reduced the size of the dataset from 42,153 businesses and 1,125,458 reviews to 1,630 businesses and 31,305 reviews. Future implementations of our tool will likely require a way to handle the data of multiple cities, especially in order to support the comparison of data across different areas.

```
{
    'type': 'business',
    'business_id': (encrypted business id),
    'name': (business name),
    'neighborhoods': [(hood names)],
    'full_address': (localized address),
    'city': (city),
    'state': (state),
    'latitude': latitude,
    'longitude': longitude,
    'stars': (star rating, rounded to half-stars),
    'review_count': review count,
    'categories': [(localized category names)]
    'open': True / False (corresponds to closed, not business hours),
    'hours': {
        (day_of_week): {
            'open': (HH:MM),
            'close': (HH:MM)
        },
        ...
    },
    'attributes': {
        (attribute_name): (attribute_value),
        ...
    },
}
```

Figure 1. Example JSON object structure for Yelp business data. This can be manipulated easily in JavaScript using object syntax (e.g. myobject.name *or* myobject['name']*)*

3.3 Software Implementation

Our implementation primarily utilized D3, a JavaScript library designed for the creation of interactive data visualizations. The library allows for easy linking of data attributes to visual encodings within a relatively flexible framework. Much of our code was drawn from examples by Mike Bostock, a creator of D3. In addition to D3, we also used the Google Maps API to create an interactive map. Styling and website object properties were manipulated via CSS and HTML files.

The first tool that we successfully implemented was a map of the city of Madison showing the location of each business in our dataset. In order to support the filtering of businesses by distance, we added the ability to manipulate the businesses shown by means of an interactive slider. Later on, this would become a simple box in which a user could specify a number of miles in which they wish to filter businesses. One difficulty we discovered while filtering businesses was that we would have to keep multiple copies of the data in order to keep track of 1) the businesses within the radius of interest and 2) the businesses within that radius that exist within some category. However, the size of our data was not so large as to make this a problem in terms of efficiency.

Next, we used the data to generate samples of the types of graphs we would like to have in our final product. We

created functions to generate the following visualizations for a single business:

1. *Distribution of ratings*
    This was implemented as a bar chart showing the frequency of each star rating (1 through 5) for a single business.

2. *Number of useful votes vs. star rating*
    This scatter plot showed every review for a given business, with its x-position representing how many "useful" votes it received and its y-position representing what rating it gave for the business.

3. *Ratings over time*
    This chart was the same as Chart 2, but with the date of the review on the x-axis.

We also added a function to visualize a group of businesses within a certain category. For this, we chose to plot the number of reviews for each business against its average star rating. Furthermore, this visualization could be filtered by distance from a specified business. Thus, a user could see where their business stands in relation to similar ones in the area. Naturally, if a business owner was interested in seeing its greatest competition, they would look toward the upper right of the chart to see those with the highest average rating and highest number of reviews (as those metrics determine how well a business is perceived by Yelp users).

Once we implemented basic versions of the above visualizations, we noticed that oftentimes there were multiple businesses occupying a single point. For reasons of clarity, we needed a way to expose each of these points to better represent all of the businesses at once. Thus, we decided to use D3's force directed layout functionality, which, by adding a repulsive force to each point, ensured that no collisions occurred between them. We chose to use this approach instead of jittering, which creates randomness across multiple instances of a single graph. Furthermore, we added the ability to toggle this collision detection feature and left it off by default, in order to avoid misleading users regarding the absolute position of the points on each axis.

To support selection of a single business, we used a jQuery autocomplete widget to list possible businesses as the user types into a search box. One problem we ran into was duplicate names, as in the case of franchises with multiple locations in Madison (such as Walgreens or Target). Thus, if there were multiple franchises within our dataset, we attached its address alongside it to differentiate between each business. This would allow business owners, who are

familiar with the address of their business, to quickly select their business.

3.4   Graph Design Decision

Due to the fact that "using color effectively is surprisingly difficult,"[7] we made very specific choices when choosing the scheme for our visualizations. We decided to use a d3 library of ColorBrewer palettes. ColorBrewer is an online tool that helps users identify appropiate color schemes. These pre-determined palettes are more perceptually effective because they ensure the viewer can distinguish each color as distinct. [8]

4   RESULTS
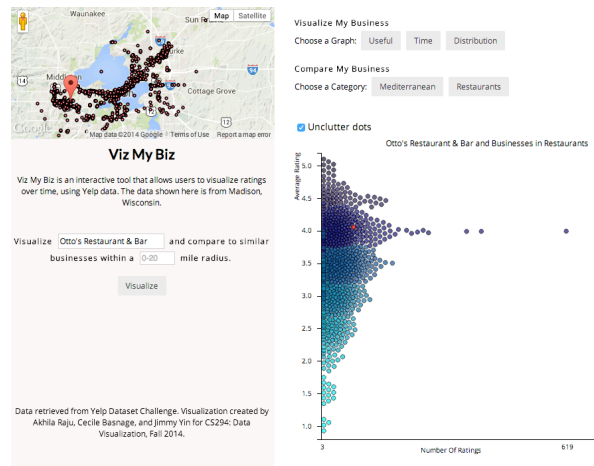
Figure 2. The final design of our application.

Our final product is designed as an interactive display, allowing users to first select a business of interest and indicate the businesses they want to compare themselves to based on the radius in miles from their business of interest. We chose on having the available mileage in a range of 0-20, which

Our final product is designed as an interactive display, allowing users to first select a business of interest and indicate the radius in miles which filters the businesses they can compare themselves to. We chose the available mileage to be in the range of 0-20, which covers all other businesses on the map when a business is chosen (including the businesses that are on the corners of the map). We implemented error handling for this feature; in the case that a user inputs a number greater than 20, all of the data is

selected, disregarding the radius entered. In the case that a user doesn't input a number, all data is selected as well.

Running time of our application and the time it takes to generate a visualization is instantaneous.

5   DISCUSSION

We can discuss the insights of our tool from its 4 primary visualizations:

1. Rating over Number of Reviews
2. Reviews over Time
3. Rating over Number Useful votes
4. Average Rating over Number of Reviews by Category

Rating over Number Reviews

This simple visualization lends itself as a summary of how many of each rating a business received. This tool is interesting for businesses as it helps them to visualize the distribution of the ratings used to calculate the average. For example, we can compare the two businesses with similar ratings: "Otto's bar and Grill" and "MadCity Music exchange." The businesses have a similar average rating (4 and 4.5 stars respectively) but their distribution curves look very different. This tool could be helpful for businesses to get an intuitive sense of where their rating average comes from.
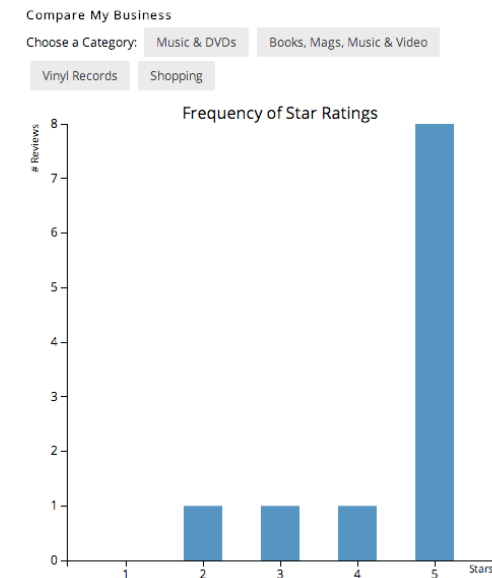
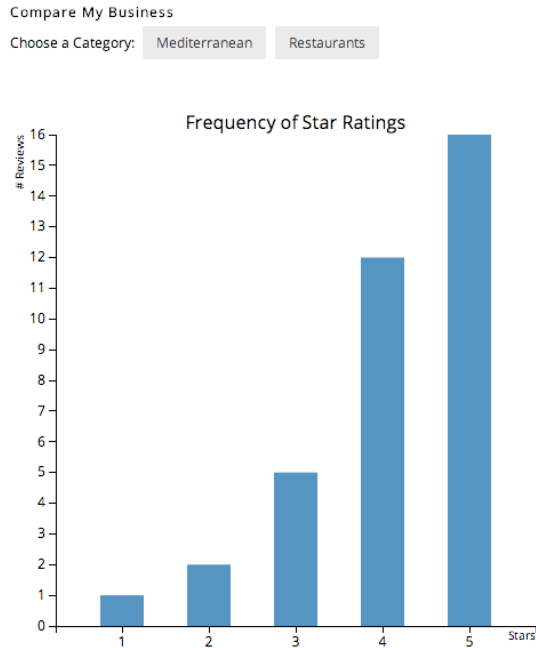Figure 3. The distribution of ratings. This business has eight five-star reviews.

Figure 4. Distribution of ratings. This business has

Reviews over Time

In Figure 3, it is clear that there is an outlier. The corresponding business is the most consistently highly rated business in all of Madison. Scrolling over the point, a user could see the name of this business and search it again for more insights. In this case, searching the restaurant name, "The Old Fashioned," we might endeavor to determine how it has aggregated so many reviews. Looking at the graph "Reviews over Time," we can gain interesting insights. "Reviews over time" illustrates that the restaurant has been open for nearly a decade.

Secondly, Figure 3 illustrates that there has been a trend of an increasing number of high-rated reviews over time. This tool could allow businesses to compare the number of high and low rated reviews over time as a possible metric of service improvement.
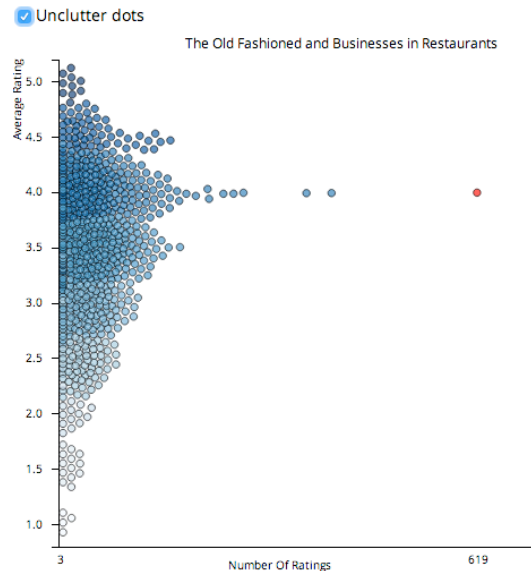

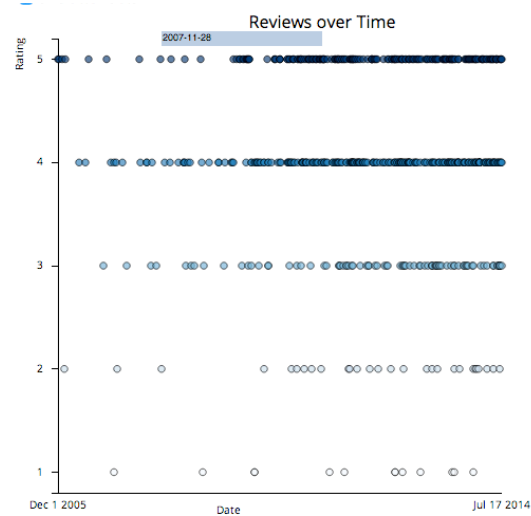Figure 5. The business being compared is indicated by the red dot.

Figure 6. This business has many 3+ ratings halfway between 2005 and 2014.

Rating over Number of Reviews: by Category

We provide the user with analytics to see how their business is reviewed relative to other businesses in that category. For example, viewing "Otto's" in the category "Mediterranean" allows the user to see the statistics in a text box : "This business has an average rating of 4 stars.

60 % of businesses in this category have an equal or lower rating."

Visualizing data of businesses that pertain to multiple categories provide valuable insights about their performance across different sectors. For example, clicking "Banana Republic", we can see 4 different categories: Shopping, Fashion, Men's clothing, and Women's Clothing. The audience gains insight into the business both from the visualization and the analytics provided. For example, compared to other businesses in the category "Shopping," "18 % of businesses in this category have an equal or lower rating" than Banana Republic. Comparing the store's performance in the category of Men's Clothing versus Women's Clothing provides valuable insight. In Women's Clothing, the user can see that, "14 % of businesses in this category have an equal or lower rating." In contrast, in Men's Clothing, the user can see that " 27 % of businesses in this category have an equal or lower rating." Thus, a Banana Republic business owner could conclude that their store was rated nearly twice as high as a Men's Clothing store than a Women's Clothing store with respect to those categories.

Another intriguing insight that can be gained from visualizing business review data across categories is looking into how many competitors a given business faces within a certain category. Again, the restaurant La Taguara can be viewed as a case study. La Taguara falls under 3 categories: "Restaurants," "Latin American," and "Venezuelan." La Taguara's average rating of 4.5 stars is greater than 90% of all restaurants and 86 % of all Latin American restaurants in Madison. However, viewing the graph of La Taguara across Venezuelan restaurants, we can see that La Taguara is the only restaurant of this category in the entire city. It therefore faces no competition as a Venezuelan restaurant. This is in contrast to other Latin American restaurants within the radius.

Exploratory Data Analysis

Lastly, our tool lends itself easily as a tool for exploratory data analysis. For example, suppose a business owner wants to know who is the most highly rated restaurant in a certain category. The owner of "El Rincon Tico," a Latin American restaurant with an average rating of 3.5 stars, could determine who are the higher-rated restaurants in their category. They might click on the 4-star restaurant Inka Heritage, an outlier on the graph. They can then reset the tool and search "Inka Heritage" to find out more data. Clicking through the categories that "Inka Heritage" pertains to: Latin American, Restaurants, Peruvian, and Seafood. The user can see that the restaurant performs well across categories but is an especially high-rated outlier in

the category "Seafood". Knowing the categories a competitor performs the best in could potentially help a business to make decisions that would drive in more customers.
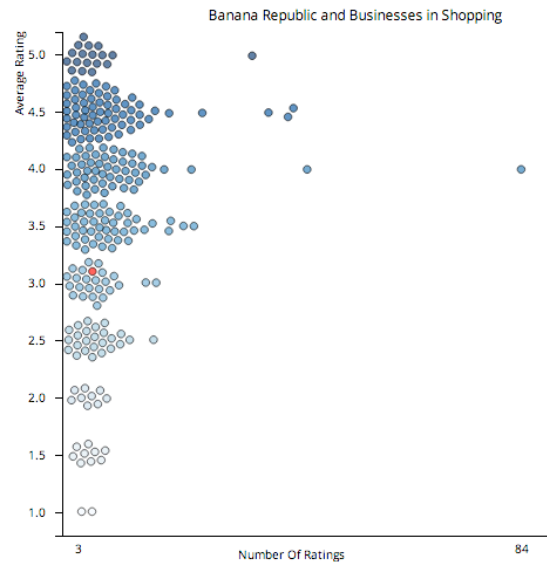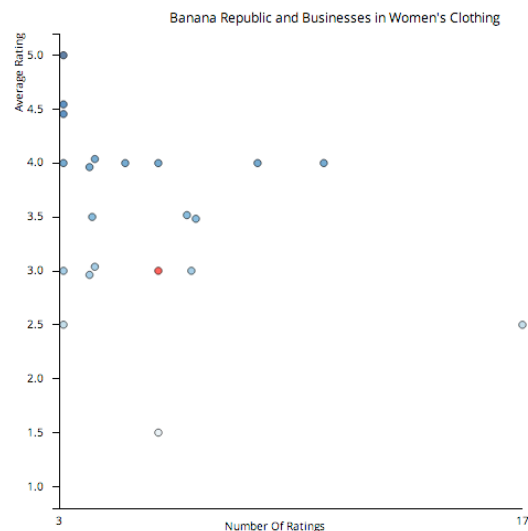

Figure 7. Most businesses have few reviews.

Figure 8. The business, indicated in red, is one of five businesses with 3.0 stars.

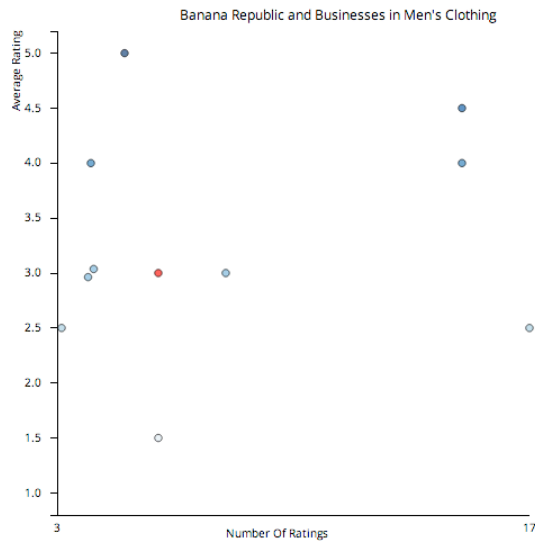Banana Republic and Businesses in Men's Clothing

Figure 9.

6 FUTURE WORK AND CONCLUSION

Since our dataset belongs to Madison, Wisconsin, we weren't able to get business owners to test our interactive display. Ideally, we would like to conduct user tests with business owners from Madison and see how they interact with the display and what features could be improved or changed.

Quite a few business owners whom we interviewed stated that they actively compare their business to businesses in other cities. One business owner said he reached out to business owners that aren't close in geographic proximity because it removes them as competition, so he can gain useful insights about what works for a business and what doesn't. Allowing for cross-city comparison would be a next step to explore.

We would like to implement sentiment analysis of reviews, which would provide more insight into a review than the current rating system. This would allow for more in-depth analysis and comparison between businesses with the same rating. We would also like to implement a popular keywords feature, which allows for users to see what the popular words for a business and across categories.

ACKNOWLEDGEMENTS

REFERENCES

[1]   J. Huang, E. Joo, and S. Rogers. " Improving Restaurants by Extracting Subtopics from Yelp Reviews. De partment of Computer Science, University of California, Berkeley. 2013.

[2]   M. Anderson and J. Magruder. "Learning from the Crowd." *The Economic Journal*. 2011.

[3]   J . Wang, J. Zhao, S. Guo, C, North. "Clustered Layout Word Cloud for User Generated Review." Virginia Tech and University of Toronto. 2014.

[4]   Liu, B., and Street, S. M. Opinion Observer: Analysizing and Comparing Opinions on the Web. WWW '05 Proceedings of the 14th International Conference on the World Wide Web (2005).

[5]   J.Leskovecand, J.McAuley."Hidden Factors and Hidden Topics: Understanding Rating Dimensions with Review Text." Department of Computer Science, Stanford University. 2013.

[6]   J . Linshi. "Personalizing Yelp Star Ratings: a Semantic Topic Modeling Approach." Yale University. 2014.