

CS224W Final Project: Analyzing the Temporal Effects of Elite Membership on the Yelp Network

Fanny Chen - fcchen@stanford.edu
Laura Cruz-Albrecht - lcruzalb@stanford.edu
Kevin Khieu - kkhieu@stanford.edu

December 11, 2017

Abstract

The “Yelp Elite Squad” introduced in Yelp’s early years is a group of active Yelp users who are recognized for their contributions to Yelp’s mission to “connect people to great local businesses.” However, little research has been done on how Elites have made valuable impact in Yelp’s community. To examine impact, we look at changes over time in the Yelp graphs for two entities: the users and the businesses. We found that once a user becomes Elite, they write longer reviews more often that are received better by the community. We also found that, from the local business perspective, garnering more Elite reviews does not necessarily correlate with higher business rating or more reviews in general. Our paper concludes that Elite users make a more active and helpful Yelp community for other users, but may not necessarily directly benefit business growth.

1 Introduction

Yelp is an online service best known for publishing crowd-sourced customer reviews of local businesses. Since its launch in 2004, Yelp has maintained a “Yelp Elite Squad”, a group of active Yelp users whose membership is based on a mixture of nomination and non-publicized criteria. The group exists both to reward certain users for their contributions to Yelp and to encourage further positive behavior amongst non-Elite members. Although the program has been going strong for many years, there are not many obvious trends that show whether or not Elite membership has a valuable impact on Yelp’s community. Hence, this project seeks to study that. In particular, this project aims to investigate the effects of Elite Member status with respect to businesses as well as users, both before and after being assigned Elite Membership status. In our study, we ask the following questions:

- How do Elite Member reviews affect the growth of businesses?
- How does receiving Yelp Elite Membership status affect that user’s behavior in following years?

2 Related Works

In *An Evaluation of the Yelp Dataset* (Cui) [1], Cui attempted to study traits and differences between Elite members and non-Elite members. Cui found that Elite members on average own significantly more fans (10x), write many more reviews (7.6x), comment more than others (10.9x), and have more friends (7.8x) than non-Elite members. In other words, Elite users are much more active in the community. Overall, Cui’s analysis is based on basic metric and graphical analyses of only the user-user Yelp graph. Cui’s graph analysis only focuses on users, whereas our project seeks to study the graphical relations in both the user-user and user-business graphs in the Yelp dataset. We will also focus on the changes to the graph over time with respect to the effects on the graph after certain users are assigned “Elite” status.

In *An Analysis of the “Elite” Users on Yelp.com* (Crain, Heh and Winston) [2], Crain et. al. investigated the properties of Elite users in the Yelp network. In particular, they sought to investigate whether the properties of Elite users in the network align with the qualities Yelp claims characterize

Elite users, and also investigate which properties correlate most highly with Elite status. To investigate connectivity, they first looked at the Yelp social graph (edges between friends), investigating properties such as degree, betweenness centrality, PageRank score, the number of communities to which the user belongs to, and the effects of removing Elite/non-Elite nodes. They also constructed a “taste graph” connecting users who rated the same businesses similarly, and performed similar analysis. In both cases they found Elites have higher degree, betweenness centrality, PageRank, community count, and review count. Moreover, review count proved to be the strongest indicator of being an Elite user. This is confirmed in our own tests later in this paper.

In *Discovering Yelp Elites: Reifying Yelp Elite Selection Criterion* (Kim, Lin, Bang) [3], the authors use machine learning models to determine the most predictive features of an Elite user. Their results showed that user who wrote positive, cool, and useful reviews were more likely to be Elite. Elite Yelpers also tend to have more fans and Elite friends.

3 Dataset Characterization

3.1 Overview of Dataset and Representation

The Yelp Dataset is provided as part of Yelp’s annual Dataset Challenge. The dataset is comprised of 4.7M Yelp reviews, along with associated users, businesses, and check-ins, across 12 metropolitan areas ranging in year from 2004 to 2017. We downloaded the dataset in JSON format. Then by using a combination of Python’s Pickle and JSON libraries, we read in the data into easy-to-access dictionaries that allowed us to look up reviews, users, and businesses in $O(1)$ time. We then used this data to construct the following graphs:

1. Aggregated User-Business Bipartite Graph: A bipartite directed graph from users to businesses indicating that a certain user reviewed a certain business. In the weighted edge case, the edges are weighted by number of stars in the review.
2. Aggregated User-User Graph: An undirected graph from users to users with edges existing between two users if they reviewed the same business.
3. Yearly User-User Graphs: An undirected graph from users to users with edges existing between two users if they reviewed the same business in the same year.

Due to the sheer size of the dataset, we needed to reduce our dataset for graphing purposes so that we could process efficiently in Snap.py. For graphing/analytics, we reduced our set to businesses in Nevada plus all related users and reviews. This resulted in 30K businesses, 680K reviews, and 198K users. Since these numbers were too large to process efficiently in Snap, we ended up taking a random subset of 2.5K businesses (+ related users and reviews). This resulted in 57K reviews and 38K users to test off of.

3.2 Dataset Statistics

3.2.1 Elite Member Presence

We investigated the overall presence of Elite members over the 10 year span from 2007 to 2016, both by count and by fraction of total users active each year, across the entire Yelp dataset. The results are plotted above. As expected, the empirical number of Elite users grew yearly. However, we found that the fraction of members that are Elite out of total active members (where we defined active as having written at least one review) in a year decreased over time, indicating that the general increase in active users in Yelp in recent years significantly outstrips growth in the Elite membership program.

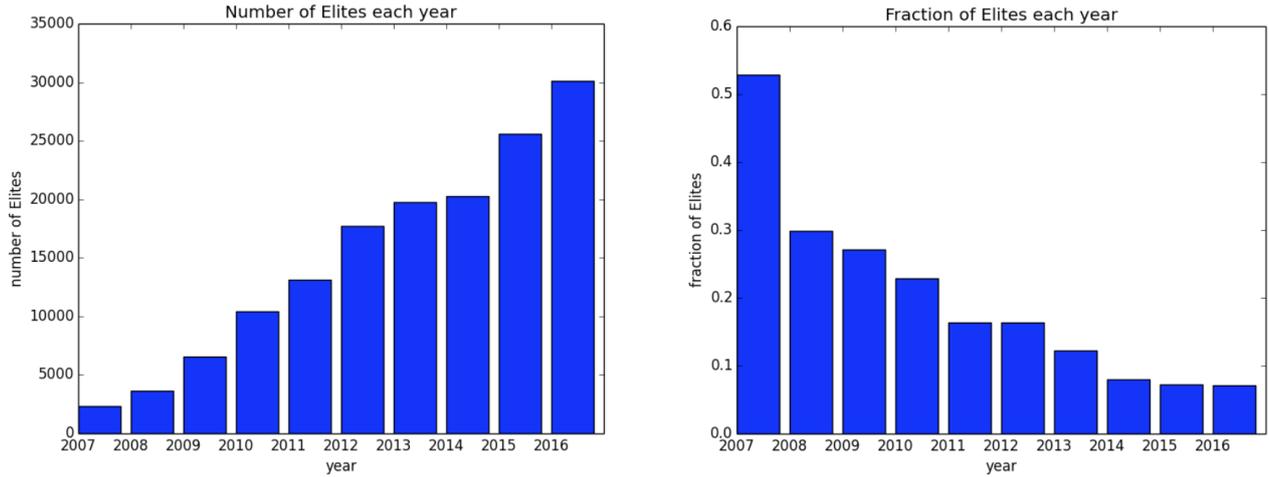


Figure 1: Elite member presence by year, by count (left) and fraction of total users (right)

3.2.2 User-User Graph Characterization

As described in Section 3, we created a User-user graph by connecting users who rated the same business (Graph 2). Here, we looked at 1) the degree distribution of the User-user graph derived from businesses in Nevada for the aggregated timespan of 2010-2014 inclusive, 2) the degree distribution for the pruned subset of (1); and 3) the degree distribution for the User-User graph derived from Wisconsin businesses in the same time period, as a comparison point for Nevada.

Interestingly, we see that the aggregated graph's degree distribution seems to mostly follow the power law, a segment of the graph diverges from the power law. This suggests that there are some users with fairly high degree that comprise a higher proportion of nodes in the graph than we would expect. This may arise from the fact that users are connected to all other users that have reviewed a same business; so, by reviewing a business that has already been reviewed by many users can cause the user's degree to jump significantly. In the random sample of the User-user graph, we see a distribution that echoes that of the non-pruned graph, but with higher variance between the points.

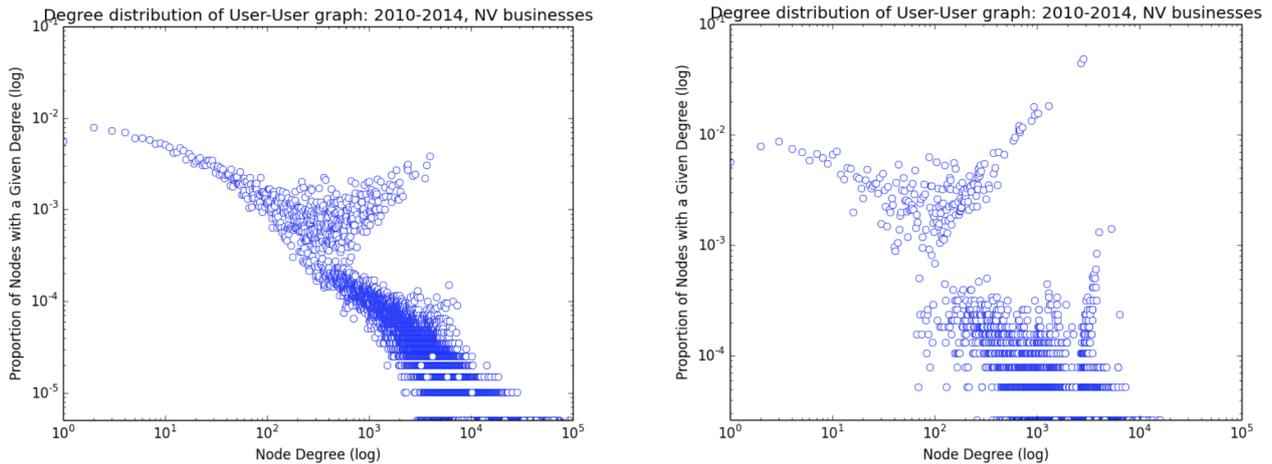


Figure 2: Plots of the User-user graph degree distribution derived from businesses in Nevada 2010-2014 (left); and the User-user graph degree distribution for the pruned version (right).

For Wisconsin, we see distribution seems to follow the power law as well. It also possesses a few node degrees with higher proportion than expected for a power-law distribution, but to a lesser extent than in the Nevada graph.

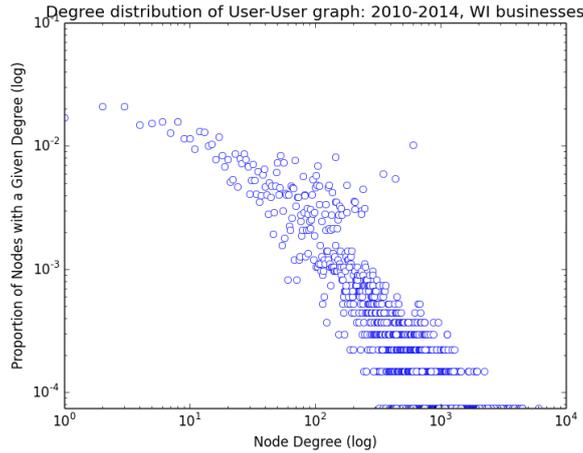


Figure 3: Plot of the User-user graph degree distribution derived from businesses in Wisconsin 2010-2014.

4 Characterizing Non-Temporal Elite Differences

Here, we look at key differences between graph structure and attributes of Elite users and reviews.

4.1 Characterizing differences in the User-Business Bipartite Graph

As an initial analyses, we looked at differences in degree distribution for users in the user-business bipartite graph. In other words, how do the number of reviews written by a user differ depending on their Elite status? Below, we present degree distributions for several different graphs under different conditions.

4.1.1 Full user-business graph

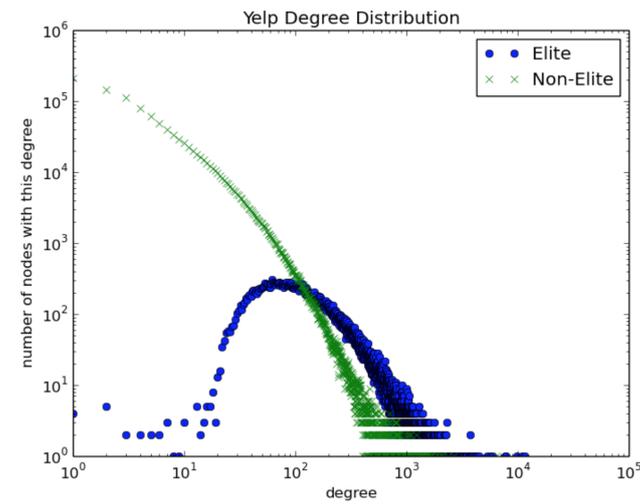


Figure 4: The degree distribution for users in the full user-business bipartite graph.

This graph shows the degree distribution for users in the full user-business bipartite graph. Here, we consider a user Elite if they have been a Yelp Elite for at least 1 year. While non-Elite users have a degree distribution that follows the power law, Elite users do not show this same kind of distribution. Instead of the majority of users having low degree like the power law would predict, the majority of Elite users have very high degree, peaking just under 100 reviews per Elite user. There is also a higher

concentration of Elite users who have very high degrees of 1000. This confirms existing research (see Related Works) that shows that Elite members generally write more reviews than regular users.

4.1.2 Nevada User-business graph

The graphs below were generated with a subset of the full user-business bipartite graph where businesses are located in Nevada. On the left, we filter the original full-dataset degree distribution to display users only if they have at least one edge to a business in Nevada. As a result, each user node retains its original degree from the full graph, but the number of these user nodes we display is limited to users who reviewed a business in Nevada. On the right, we again filter the nodes to those connected to businesses in Nevada, but this time we removed all other business nodes completely from the graph. Thus, the degree of each user node is limited to only reviews written for businesses in Nevada.

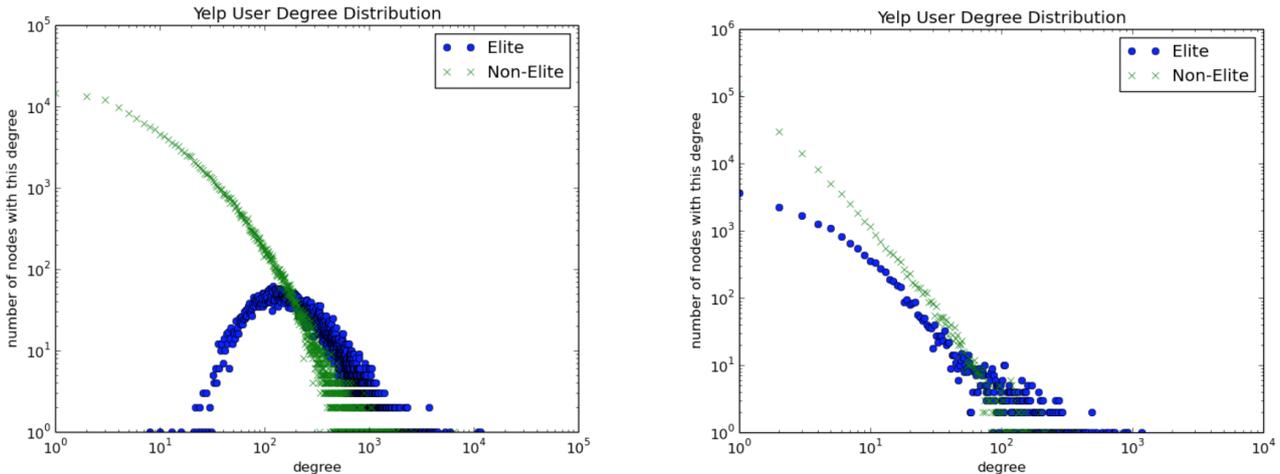


Figure 5: Left: the degree distribution for users in the full user-business bipartite graph, filtered to show only users who have reviewed at least one Nevada business. Right: the degree distribution for users in the Nevada user-business bipartite graph, where only businesses in Nevada and users who review those businesses exist as nodes.

These graphs merit many interesting observations. First off, we notice that in the degree distribution with removed business nodes (on the right), the Elite degree distribution looks a lot more like a power-law. Specifically, most of the Elite users who reviewed a Nevada business only reviewed 1 business. This is quite different from the other two graphs, where almost no Elite users only reviewed a single business. In addition, the peak in the graph on the left tells us that Elite reviewers for Nevada businesses are even more active than the typical Elite user, with the peak happening just over 100 reviews per Elite user. This tells us that the more active Elite users are tourists to Nevada (perhaps to Las Vegas, which is one of the largest attractor of tourists in Nevada), because although they have reviewed many businesses, only a few of those businesses are in Nevada.

4.2 Characterizing Differences between Elite and non-Elite Reviews

We investigated several characteristics to learn more about the differences between Elite and non-Elite users with respect to the reviews that they write for businesses. The reason for this was to see if there were any noticeable differences in the way Elite users write reviews versus non-Elite users; discovering this would show that Elite and non-Elite users have differing qualities in the reviews they write, thus suggesting an important difference between the two groups. For each of the following subsections, we calculated differences by taking the average of each metric across all Elite/non-Elite members who wrote reviews in Nevada between 2010 and 2014.

Review Word Count: We first investigate the average number of words per review for both groups. The statistic can be seen as indicative of review quality with regards to how much effort a

user is putting into his/her reviews. The average word count for Elite members' reviews was found to be 189.0 words, whereas non-Elite members wrote an average of 112.4 words per review. This means that there is a considerable difference between the length of reviews for Elite and non-Elite members, with roughly 77 more words written in Elite reviews.

Number of Reviews: In addition, we also looked into the number of reviews written by Elite members versus non-Elite users. Intuition would tell us that Elite users write more reviews, and our calculations on the Nevada graph confirmed this. We calculated an average of 3.48 reviews written by Elite users in a given year, in comparison to an average of 0.55 reviews written by non-Elite users. These numbers indicate that Elite users contribute more reviews to Yelp than the average reviewer and are one of our strongest indicators of difference between the two groups.

Polarity: Polarity is a measure of how positive or negative a piece of text is, ranging from -1 to 1. We ran each review's text through TextBlob, a python package for NLP analysis/feature extraction, resulting in an average Elite review polarity score of 0.2088 and an average non-Elite review polarity score of 0.2417. When we take the absolute value of each review's polarity before incorporating it into our average metrics, the resulting average Elite review polarity score becomes 0.2192, and the average non-Elite polarity score is 0.2752. These values indicate that there is a very slight decrease in the polarity of reviews written by Elite members versus non-Elite members, perhaps indicative of the idea that when users become Elite, they tend to write less extreme reviews.

Subjectivity: Subjectivity scores range from 0 to 1, where 0 indicates greater objectiveness and 1 indicates greater subjectiveness. These scores were also calculated by using TextBlob. When calculating subjectivity, Elite reviewers averaged 0.5475 in comparison to 0.5616 for non-Elite reviewers. These scores indicate that there is little difference in how the subjectivity of user changes upon becoming Elite.

Compliments: Each Yelp review comes with fields for three kinds of compliments: "funny", "cool", and "useful". For each review, we totaled these compliments and found the average number of compliments per review for Elite users versus non-Elite users. Elite users received an average of 6.0221 compliments, whereas non-Elite members received 1.508 compliments per review, hereby showing that Elite users receive nearly 4x more compliments on their reviews than non-Elite members.

Spelling: Our last metric for review quality was to see how many spelling errors were made in users' reviews based on their Elite status. To find this, we used TextBlob's spellcheck features, which check for misspelled words in sentences. From this, we found that both non-Elite and Elite reviews had an average of 0.11 spelling errors per review, which means we can infer that there is little difference in the spelling quality of reviews between Elite and non-Elite members.

5 Characterizing Temporal Elite Differences

We now investigate the two key temporal questions posed in the introduction.

5.1 How do Elite Member reviews affect the growth of businesses?

The goal in asking this question was to see if the presence/actions of Elite members affected businesses positively in any way. Finding a such pattern would help lead us to believe that Elite members play an influential role in boosting (or diminishing) the success of businesses and thus provide a reason for businesses to prioritize their treatment of Elite members (which, we may add, is already being done via businesses hosting special "Yelp Elite Only" events).

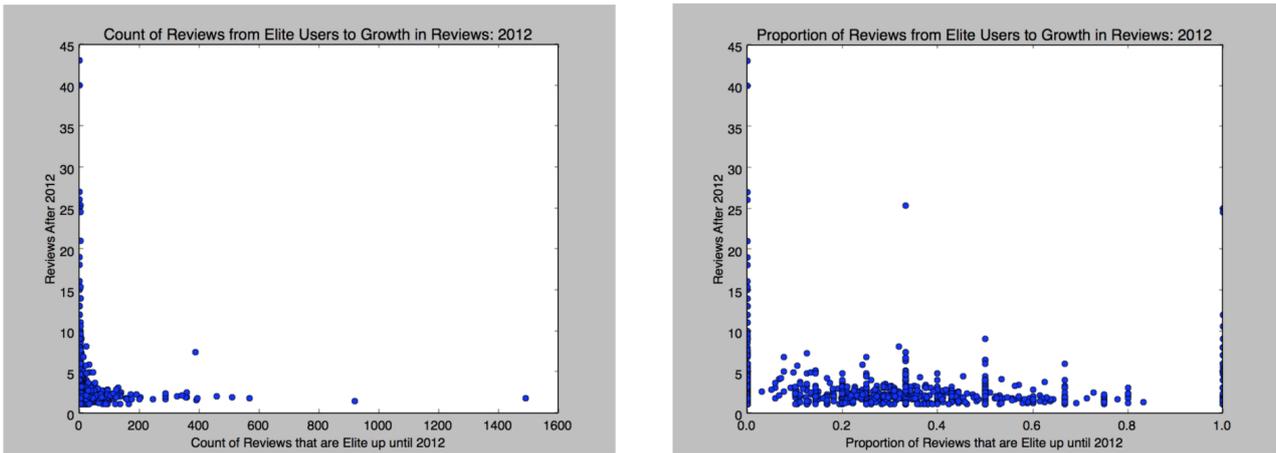


Figure 6: Plots comparing both count and proportion of Elite reviews a business has up until 2012 to its growth in terms of number of reviews gained after 2012.

Our main way of testing this question was to see, for starters, if there is any correlation between the number of Elite reviews a business has and its future success. For this test, we compared both the count and proportion of Elite reviews a business has up until 2012 to the amount of reviews they get after 2012. From these plots, there is little to no correlation that having more Elite users, both with regards to proportion of reviews and count, has a noticeable effect on the future number of reviews for a business. This would initially lead us to believe that businesses aren't affected by the presence of Elite members reviewing them.

To confirm this, we also checked to see if a percent change in the number of Elite reviews in a given year correlated with a relatable change in average rating. To plot this, we looked at the correlation between the percent change in average rating of a business and the percent change in the number of Elite reviews from years 2010 - 2011, for all businesses with at least 40 reviews in 2010 (all of which have at least 1 Elite reviewer).

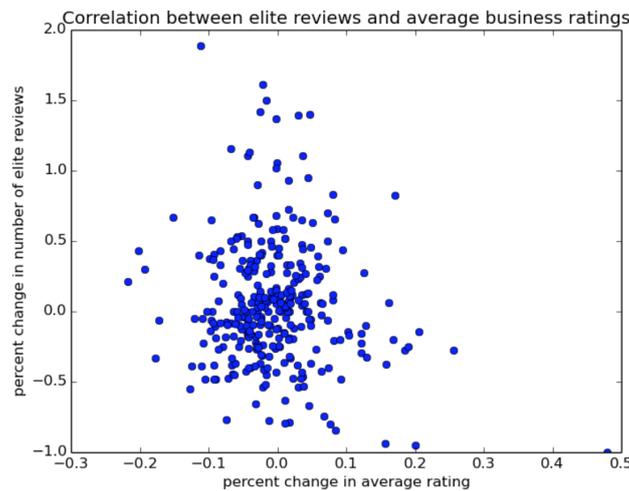


Figure 7: Correlation plot between Elite reviews and average business rating.

From the graph, it is clear that there is little to no correlation between number of Elite reviews and the ratings of businesses. Pearson correlation coefficient is -0.12422924 . When combined with our findings from our previous plots, there seems to be no pattern between Elite members' reviewing behavior and a corresponding business' growth, thus allowing us to infer that Elite members have little to no considerable impact on businesses.

5.2 How does receiving Elite Membership status affect that user’s behavior in following years?

This question seeks to understand the immediate impact of Elite membership assignment on a Yelp user’s behavior. Combined with our findings in section 4.2, the findings in this section help to show whether or not users act differently after being assigned “Elite”, or if they’ve always acted the way they have (and “Elite” membership serves as just a label/trophy for doing what they do). We studied the following:

Review Word Count: The average change in word length of an Elite member’s reviews from the last year they were Elite to the first year they became “Elite” was found to be 38.47 words, compared to 1.006 for users who were non-Elite in 2 consecutive years. This was computed by finding the average word length of a user’s reviews the year before becoming Elite, the average word length of that same user’s reviews in the year they were Elite, and then finding the average difference between those two across all users. Combined with our earlier statistic from section 4.2 on aggregated word count, this finding reveals that just the assignment of “Elite” led to people having longer reviews, indicating that this behavior is correlated with the event of a user gaining Elite status.

Number of Reviews: We looked at the average number of reviews written by an Elite user in their first year after gaining Elite status, and compared this number to the average number of reviews written by a non-Elite user in a given year. When studying this, we found that Elite users wrote an average of 3.29 reviews after just becoming Elite, compared to 0.236 reviews per year for non-Elite users, for an average difference of 3.054 reviews. This number indicates that upon being given “Elite” status, users seem to visit more business and thus write more reviews. Because this was highly correlated with the event of just becoming an Elite member, we have strong reason to assume that users wouldn’t have written this many reviews without being named Elite.

Polarity and Subjectivity: Using similar methods, we also conducted basic NLP on the change in polarity and subjectivity of a user’s reviews upon just becoming Elite. Through this we found that upon just becoming Elite, users had an average change of -0.0055 in review polarity and -0.0072 in review subjectivity, indicating little to no significant difference to review sentiment upon just becoming Elite.

Compliments: The last metric we used to test users’ behavior right after becoming Elite was the number of compliments that a user received upon becoming Elite. When comparing the reviews of a member just before gaining Elite status with reviews right afterwards, we discovered an average gain of 0.5998 more compliments for these users. For non-Elite to non-Elite, this number was 0.1791. This is indicative of two things: first that these users might be writing higher quality reviews, and/or second that other users in Yelp are seeing these “Elite” reviews and are upvoting more due to the fact that they are from Elite people. When combined with the rest of the findings in this paper, our hypothesis (which we lacked the time to test) was that the prior is true, especially given the finding that users write significantly longer reviews right after becoming Elite.

Overall, these findings show that there is a clear immediate change in behavior correlated with a user becoming Elite, specifically with regards to these users writing longer reviews, reviewing a greater number of businesses per year, and writing reviews that garner more praise. This shows that at the very least, Yelp’s decisions to assign Elite status is helping to incentivize interaction in its community.

6 Conclusion and Directions for Future Work

This research sheds interesting light upon the nature, behavior, and impact of Elite users within the Yelp network. Through our analysis, we were able to characterize facets of Elite users on both a non-temporal and temporal level. On a non-temporal level, we have two main findings: 1) while Elite users have high degree, many of those users review very few businesses in Nevada, indicating that Elite users are more likely to perhaps be tourists to the area than non-Elite users; and 2) Elite reviewers and reviews vary from their non-Elite counterparts across various attributes: number of reviews, length of reviews, and compliments received; but not along others, such as review subjectivity, polarity, or spelling quality.

On a temporal level, we sought to answer the two key questions posed in the introduction: 1) how do Elite member reviews affect the growth of businesses; and 2) how does receiving Elite membership status affect that user’s behavior in following years.

Our analysis on the first question yielded that in fact Elite member reviews do not seem to correlate with the growth of businesses according to the metrics we used. However, our research into the second driving question uncovered that receiving Elite Membership does indeed affect that user’s behavior in the following years. Through our analysis, we found that users do indeed change their behavior upon becoming Elite: namely, they write more reviews (+3.29 more reviews than per year, compared to +0.236 for non-Elite users), longer reviews (+38.47 more words per review, compared to 1.006), and receive more compliments per review (+0.6 more compliments per review, compared to 0.18). Taken together, these findings indicate that transitioning to being an Elite member does fundamentally impact the user’s behavior in the network.

In the future, it would be interesting to further investigate different ways in which Elite users may change in behavior upon becoming Elite, as well as further ways in which they may influence the behavior of other users in the network. In particular, it would be interesting to analyze more deeply the correlation between Elite reviews and growth of businesses, to see if there is a correlation along a different metric than what we considered, or whether there fundamentally is none. Another area for future research would be investigating how Elite users may or may not influence the behavior of other users in the network.

7 References

1. Cui, Yan. 2015. An Evaluation of the Yelp Dataset. Columbia University arXiv: 1512.06915v2 [cs.SI] 24 Dec 2015
2. Kevin Crain, Kevin Heh, and Johnny Winston. 2014. An Analysis of the “Elite” Users on Yelp.com. Stanford University. snap.stanford.edu. CS224W Course Project
3. Caroline Kim, Gordon Lin, and Honam Bang. 2015. Discovering Yelp Elites: Reifying Yelp Elite Selection Criterion. University of California - San Diego

8 Contributions

- Fanny: Created graphs for user-business degree distributions, dataset parsing, preliminary data analysis, problem formulation, writing up the report.
- Laura: Dataset parsing, created graphs for User-user graph degree distributions and for Elite user presence over time, data analysis, writing up the report.
- Kevin: Created graphs for count/proportion of Elite reviews to business growth, dataset parsing, NLP for analyzing Yelp reviews, data analysis, poster.