Twitter Analysis of Lyft

Luying Cao

2017/12/15

For this Twitter data analysis project, I choose the keyword and brand "Lyft" to study. Lyft is a transportation network company to provide car transportations through mobile app. It was founded in 2012, and is now grow rapidly and becoming the largest competitor of Uber.

To start with, I used the "searchTwitter" function to generate 1,000 tweets that contain the keyword "#Lyft" in it. Second, I cleaned the texts in order to make sure that there is no special character or symbol in them.

After the text cleaning, I was ready to do some analyses. The first one I am going to find out is the word frequency in those texts. I used R to generate a word frequency table, which also identifies the nature of the words. From the table, it can be found out that the most frequent positive word "free" appears seven times of the most frequent negative word is "vex". It can be then inferred that customers using Lyft generally receive a cheap or even free ride and they are pretty satisfied. In addition, I did the sentiment analysis to have a better view of the general picture of both positive and negative words. According to the sentiment analysis, there are way more positive words than negative words, which can be concluded that the majority of customers who use Lyft have a generally satisfied experience.

Then, I generated two word clouds to have a more direct sight on the texts. In word clouds, words that appear more often will be shown bigger and bolder. Again, it can be easily seen that Lyft customers feel happy about their experiences.

Last but not least, I generated a map of United States showing where the tweets were published. Each location is represented by a red point in the

map. From the map, it can be found out that there are more tweets published either in New England area or California, which can be inferred that Lyft are more popular in these places. At the same time, only a few or none tweets have been published in the Middle and Northern part, which can be inferred that Lyft is not that popular or maybe do not exist in those places.

In conclusion, customers who use Lyft generally have a satisfied experience. It is essential for Lyft to keep its advantage of customer satisfaction. At the same time, there are some limitations of this data analysis project and there are a few things that I would like to address in the future. First, I could gather more tweets about Lyft to do more analyses. Second, I could gather tweets contain keyword "#Uber", which is Lyft's largest competitor, to generate comparisons. Third, I could find out which state or city has the highest amount of tweets or highest level of satisfaction of customers using Lyft.