

Starbucks

Social Media Analysis

3/5/2013

Introduction:

Starbucks Corporation is an American global coffee company and coffeehouse chain based in Seattle, Washington. Starbucks is the largest coffeehouse company in the world, with 18000 stores in 61 countries. Starbucks is known for its products like Macchiato, Latte, Frappuccino, Mocha, Espresso, Skinny beverages and the Americano. Starbucks is faced with serious competition from big companies such as Costa coffee, Dunkin' Donuts. Each of these has specialties which allow them to capture a certain portion of the market share.

In this analysis, we have tried to understand the public perception of Starbucks and its standing in the coffee chain market in comparison to its competitors.

In this project, we chose to analyze the following aspects:

- Public sentiments towards Starbucks
- Two of Starbucks' competitors-Dunkin Donuts and Costa Coffee
- Starbucks products: Macchiato, Latte, Frappuccino, Mocha, Espresso, Skinny beverages and the Americano
- Most popular coffee products are Latte and Mocha

Methods:

Data Collection

At first, we tried to use Twitter API to get tweets to do analysis. However, the GetSearch function in Python-Twitter Package did not work very well. Finally we find a Twitter-related website called Topsy (www.topsy.com) which has thousands of real-time tweets. So we selected the Topsy as a way to get tweets.

Programming Language

The programming language used in the project is Python

Packages

We used two main packages –BeautifulSoup and urllib2 to crawl the website of Topsy. We used matplotlib package to visualize our analysis. In the end, we got about 10,000 tweets which are much more than the tweets we could get from Twitter-API. The time period is from last 30 days to now.

Analysis

We used the sentimental analysis. In this analysis, we have tried to understand the positive and negative sentiment of the Starbucks as whole and its products. For this analysis, we made the exhaustive list of positive and negative words. We then calculated the ratio of positive words

over total words in a tweet and ratio of negative words over total words in a tweet. We categorized the tweets in the following manner:

- If the ratio of positive words is greater than the ratio of negative then the tweet is considered as positive
- If the ratio of positive words is lesser than the ratio of negative then the tweet is considered as negative
- If the ratio of positive words is equal to the ratio of negative then the tweet is considered as neutral.

For visualization purpose, we have ignored the neutral tweets and only plotted the percentages of positive and negative tweets

We also used the frequency analysis wherein we calculated the total number of tweets for a particular keyword.

Results:

Our analysis calculated the number of positive and negative tweets in 1000 tweets for Starbucks, Costa Coffee and Dunkin donuts. And our results show that the most positive tweets are in Costa Coffee and the least negative tweets appears in Costa Coffee. Starbucks has less number of positive tweets in comparison to its competitors. While the maximum number of negative tweets appears in Dunkin Donuts.

We also calculated the number of positive and negative tweets in 1000 tweets for each products of Starbucks (except for Americano (only 562) because it is seldom talked). The Mocha has the highest number of positive tweets which indicates that among all products of Starbucks, Mocha is famous among its customers. Latte got the second highest number of positive tweets. For Espresso, we got the highest number of negative tweets which indicates that Starbucks' Espresso is not liked by its customers.

When we compared the major coffee type i.e. latte and mocha among different coffee chains, we observed that Starbucks latte appears to be more talked about than its competitors' latte. While Costa Mocha appears to be more talked about in our analysis.

Conclusions:

Our analysis concludes that Starbucks does not have a positive perception in comparison to its competitors. We also concluded that Costa Coffee is more positively perceived among its customers. Even though Starbucks' Mocha got the highest number of positive tweets among all its products but it is not talked about more in social media in comparison to its competitors' mocha. Therefore, Starbucks needs to work on the social media aspects of its most popular products-Mocha. Our analysis also suggests that since Espresso has the highest number of negative tweets therefore Starbucks should try to understand the customer feedback for this product and should try to make changes.

Since it is the sample data set of only 10000 tweets, more representative data, if gathered would be more helpful in understanding the public sentiments towards Starbucks. Also, the classifier used in our sentimental analysis could be improved more by categorizing the words in a tweet in a more sophisticated way.

References:

<http://nealcaren.web.unc.edu>

<http://en.wikipedia.org/wiki/Starbucks>
www.topsy.com

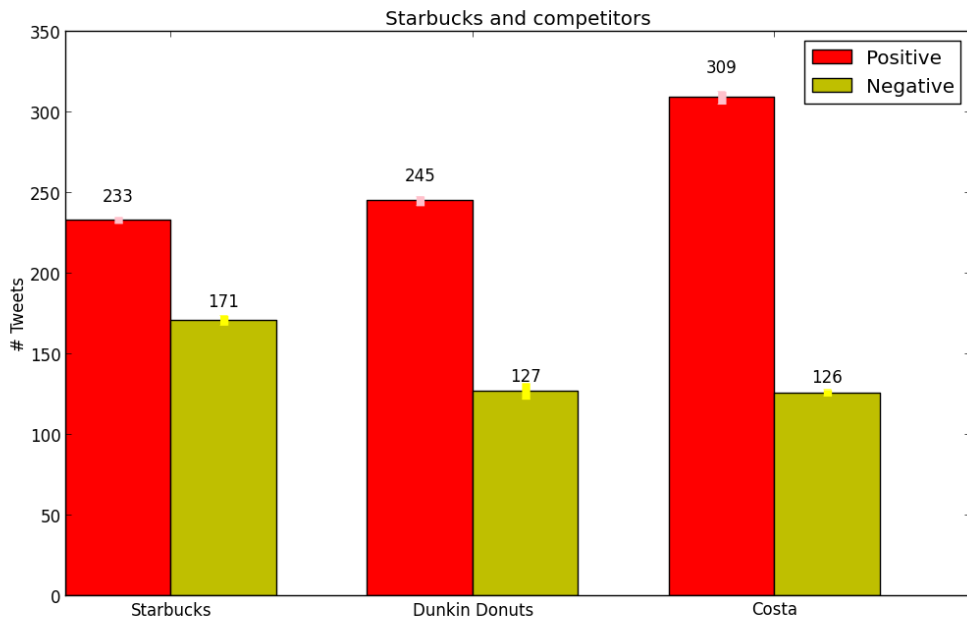
<http://www.starbucks.com/>

<http://www.dunkindonuts.com>

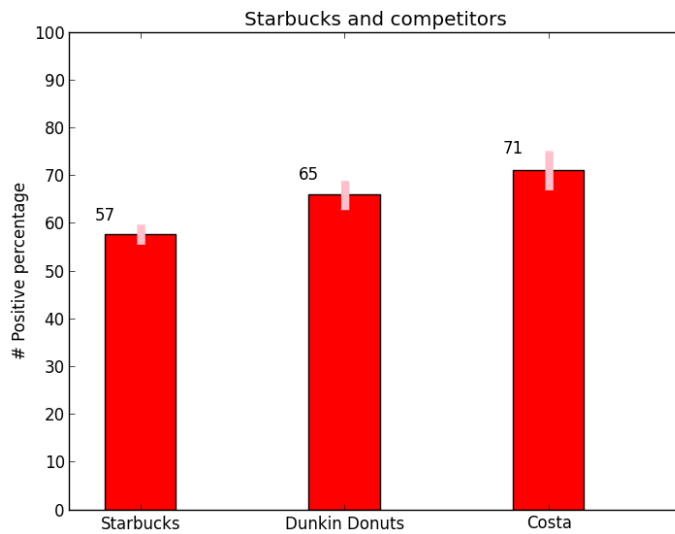
<http://www.costa.co.uk/>

<http://www.diveintopython.net/>

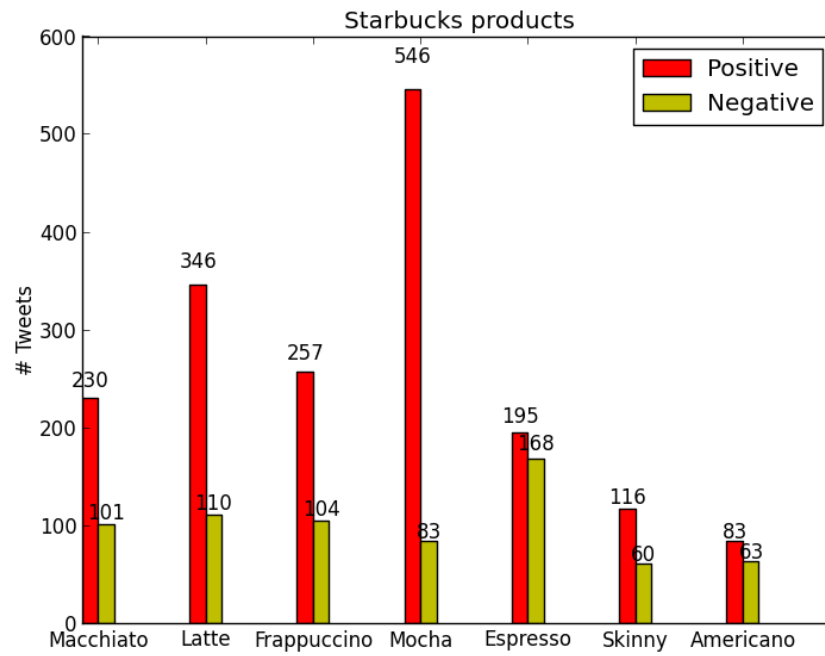
Appendices:



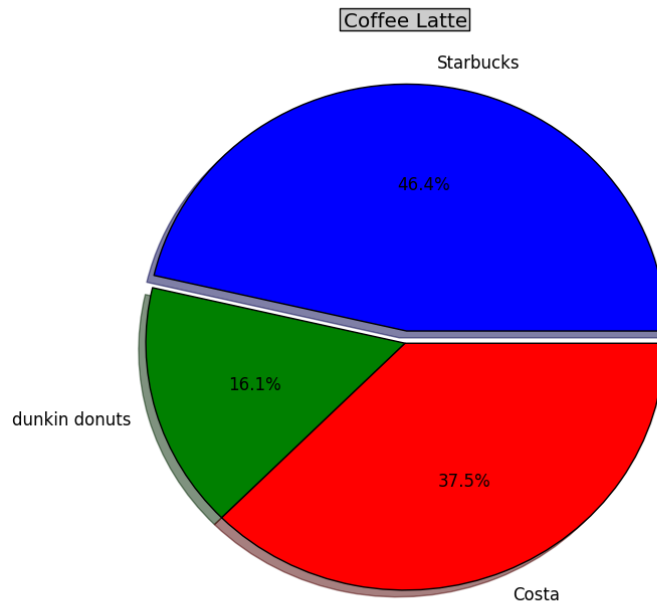
G1- Number of positive and negative tweets in 1000 tweets for each Coffee Company



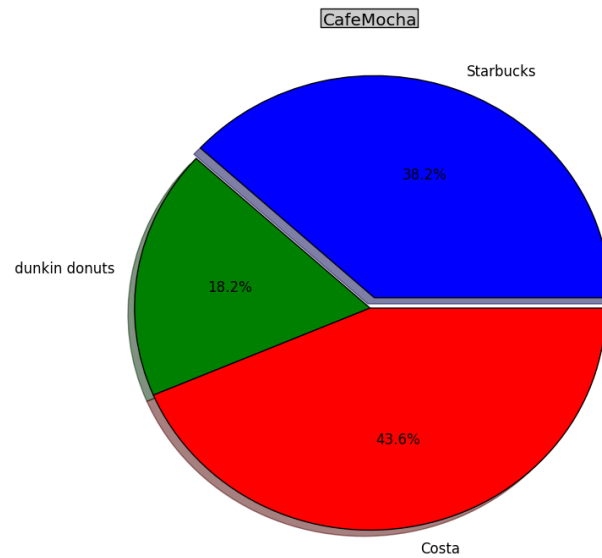
G2- Percentage of positive tweets for each Coffee Company



G3- Number of positive and negative tweets in 1000 tweets for each products of Starbucks



G4- Percentage of each coffee company mentioned in the tweets for latte



G5- Percentage of each coffee company mentioned in the tweets for Mocha