

Coke vs. Pepsi Social Media Analysis

Xin Qiu, Harshal Gogri

Coca-Cola and Pepsi are the most popular soda brands in the world. In this project, we are trying to perform a social media analysis on the two competing brands. We mainly used nltk package in Python3 and applied many of the concepts in natural language processing.

The first part of our report is the explanatory data analysis. We examined the original dataset and analyzed the source type. Also, we extracted and counted sentences that include one of the matching Brands: Coca-Cola or Pepsi.

The second part of our report is the sentimental analysis. In this part, we performed a sentiment analysis on both products and print out the overall relative sentiment. Then, for each Brand, we created lists for the main parts of speech and print Top n words and hashtags in each list using the Part Of Speech (POS) tagger. At last, we answered the three questions: Which brand has more positive sentiment? What attributes of the product are the reasons why the perception is more positive? What attributes of each product are the top contributors for their negative perception?

The last part of our report is summary and recommendations. Here, we summarized recommendations for Coke or Pepsi based on our findings of sentimental analysis of social media text.

Part I. Explanatory Data Analysis

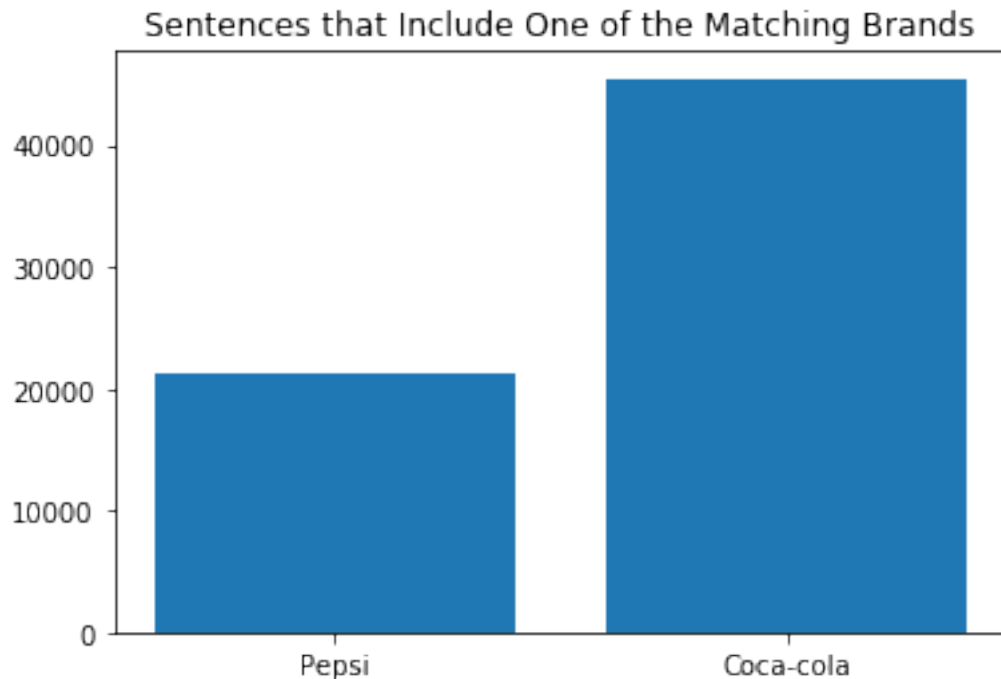
1) About the dataset

Our data is from <https://spaces.hightail.com/space/5yxHcCks2V>. It includes social media text data and their relevant information from 2017-09-01 to 2017-10-04 on Twitter, Instagram, etc. For our research purpose, we will only use the data from the first column, which is Sound Bite Text.

- Overview

From the word cloud graph of all Sound Bite Text (excludes from hashtags, mentions and links), we can see “Pepsi Gladiator” is a very hot topic, which demonstrates Pepsi’s effort in commercials. Also, “diet coke”, “slimming world” are mentioned a lot, indicating that people are paying more attention to health. In addition, we can see words like “music”, “photography” and “beauty” showing soda is now more related to lifestyle on social media.

2) Count sentences that include one of the matching Brands: Coca-Cola or Pepsi



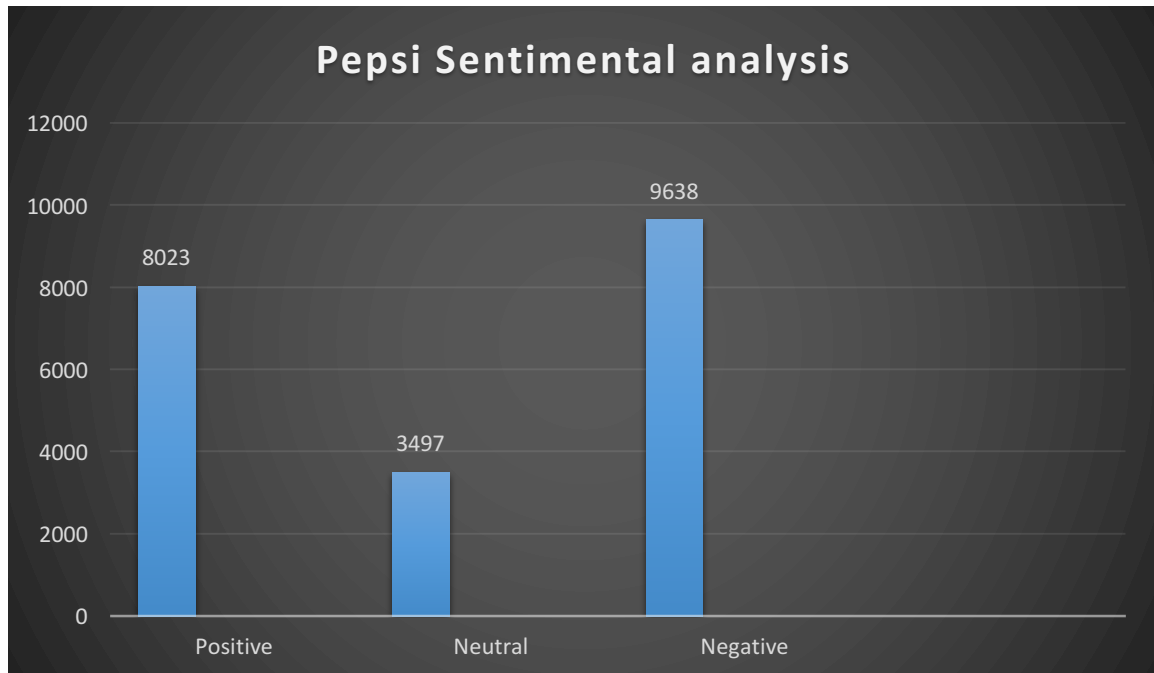
There are 21158 sentences that have mentioned Pepsi (identified by name 'Pepsi'), and 45459 sentences that have mentioned Coke (identified by name 'Coke' or 'Coca-Cola'). Since the number of sentences containing coke is more than twice of the number of Pepsi, it is obvious that Coca-Cola is more popular on social media.

Part II. Sentimental Analysis

The sentiments are captured with the help of compound rate obtained using Sentiment Intensity Analyzer in Python. The threshold chosen for positive sentiments is Compound rate > 0.2 . The threshold chosen for negative sentiments is Compound rate < -0.2 . The threshold chosen for neutral sentiments is Compound rate ≤ 0.2 and Compound rate ≥ -0.2 .

1) Comparative Sentimental Analysis of two brands

- Sentimental Analysis of Pepsi



From the above graph, the total number of sentiments captured is 21158.

The Positive sentiments percentage = $8023/21158 = 37.91\%$

The Negative sentiments percentage = $3497/21158 = 15.52\%$

Here are some of the example sentences associated with positive sentiment towards Pepsi:

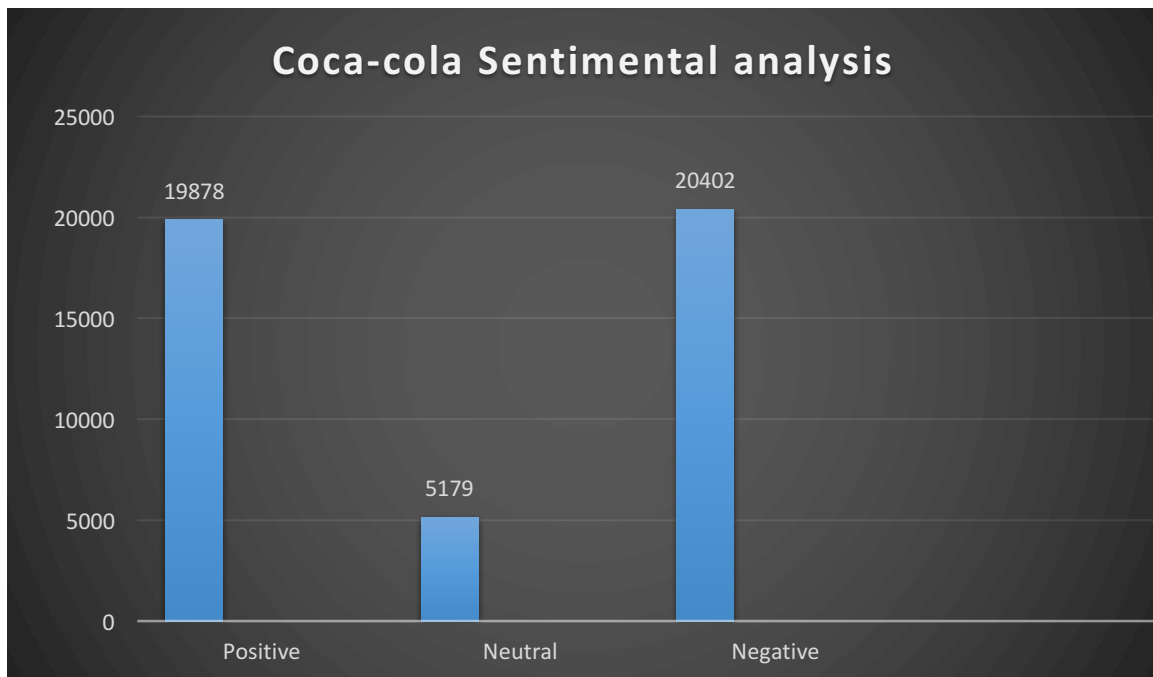
```
['Love the new Pepsi can ... #GoLions #Pepsi #PepsiandDetroitLions',  
 'RT @springvoice_boy: [PREVIEW] 170924 iKON Pepsi Fan Party #아이콘 #iKON  
#구준희 #JUNE pic.twitter.com/A38YfgXIr5',  
 "I just put the pepsi insert in - it's quite possibly a perfect watch.",  
 'Basically I think they should be like Pepsi and coke and have a vendor  
stock there own product.',  
 'Friend: Pepsi is fine.']
```

Also here are some of the example sentences associated with Negative sentiment towards Pepsi:

```
['The oncoming storm sucks water in think about opening a Pepsi bottle.',  
 "I've always hated Pepsi...bitch!",  
 'Damn these were simpler times so lemme break this down -lord caesar  
iglesias, who does not sing in this musical commercial, has captured  
britney spears, pink, and beyonce to battle it out gladiator style\n-our  
trio decides instead of fighting, they will instead rock so hard that the  
audience forgets about their battle to the death\n-their musical prowess  
is so damn powerful they rock the foundation of the earth and overthrow  
caesar iglesias along with his stash of ancient roman pepsi\n-beyonce,
```

britney spears, and pink drink pepsi while the audience cheers\n-enrique iglesias is eaten by a lion Tags: #amazing',
'First, there was the uproar over Kendall's ridiculous Pepsi commercial that attempted to use racial unrest and the historic Black Lives Matter movement as gimmicks to sell carbonated sugar water.']

- Sentimental Analysis of Coca-Cola



From the above graph, the total number of sentiments captured is 45459.

The Positive sentiments percentage = $19878/45459 = 43.72\%$

The Negative sentiments percentage = $5179/45459 = 11.39\%$

Here are some of the example sentences associated with positive sentiment towards Coca-Cola:

```
[ "Check out the Jordan's Cool Mix mix I created with the Coca-Cola freestyle® app.",
  "My friend Jim was married here, right where Mad Men's Don Draper, played by Jon Hamm, envisions the iconic Coca-Cola commercial while doing yoga.",
  'For those of you whose beverage of choice is Red Bull or Black Bull or any other artificial drink like Diet Coke or Mountain Dew, let me explain.',
  'A friend of mine calls #Pepsi "Blue Coke"']
```

Also some of the example sentences associated with Negative sentiment towards Coca-Cola:

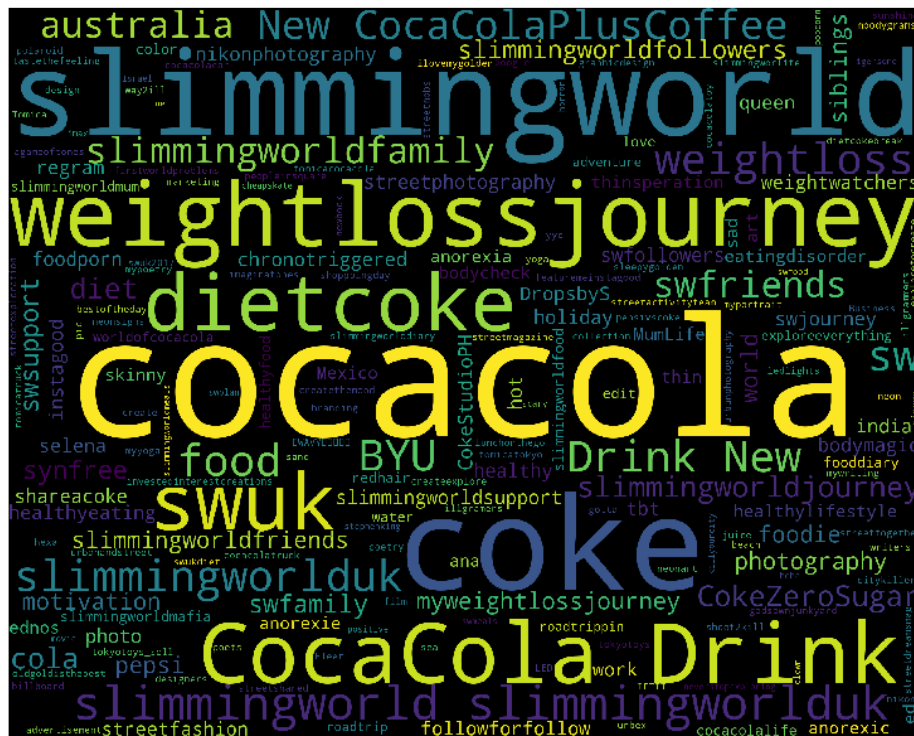
- [illegible]

- Among the top 50 hashtags associated with positive sentiments of Coca-Cola, we can notice that many of them are related to health and fitness, like #slimmingworld, #weightloss, #fitness and #healthyliving. Also, #instagood, #yummy and #foodporn are mentioned a lot, which shows that Coca-Cola does have an appealing taste. #dietcoke and #cokezerosugar are popular choices for people who want to drink soda but keep fit as well.

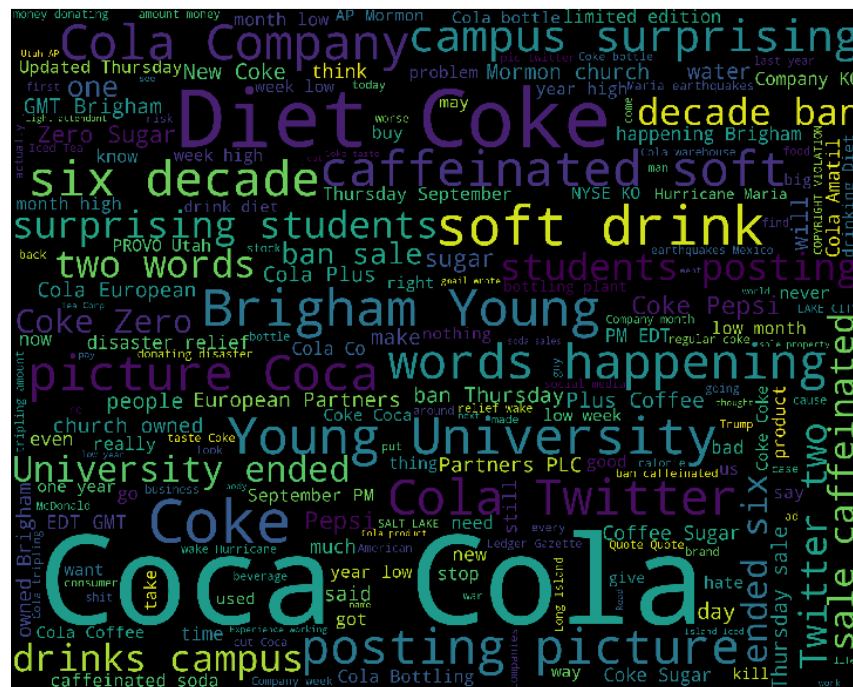
Among the top 100 words associated with positive sentiments of Coca-Cola, there is also “diet” which corresponds to the hashtags. Also, customers like its bottling. In addition, Coca-Cola has a good reputation among customers as we can see words like “rating”, “brand” and “original”. While the taste is “sweet” and “American”, Coca-Cola leaves customers an impression of “global”. Plus, Coca-Cola performs well in shares.

7

- Top hashtags associated with negative sentiments



- Top words associated with negative sentiments



- Among the top 50 hashtags associated with negative sentiments of Coca-Cola, we can notice that many of them are related to eating disorder problems, such as #eatingdisorder and #anorexie. Though words about fitness and weight losing are included in positive sentiment sentences of Coca-Cola, they also appear a lot in negative sentiment sentences, such as #bodycheck, #synfree and #slimmingworldjourney. Besides, #cocacolapluscoffee shows up perhaps because people do not like the product of “Coca-Cola Coffee Plus”.

4) What contributes to Pepsi's positive perception?

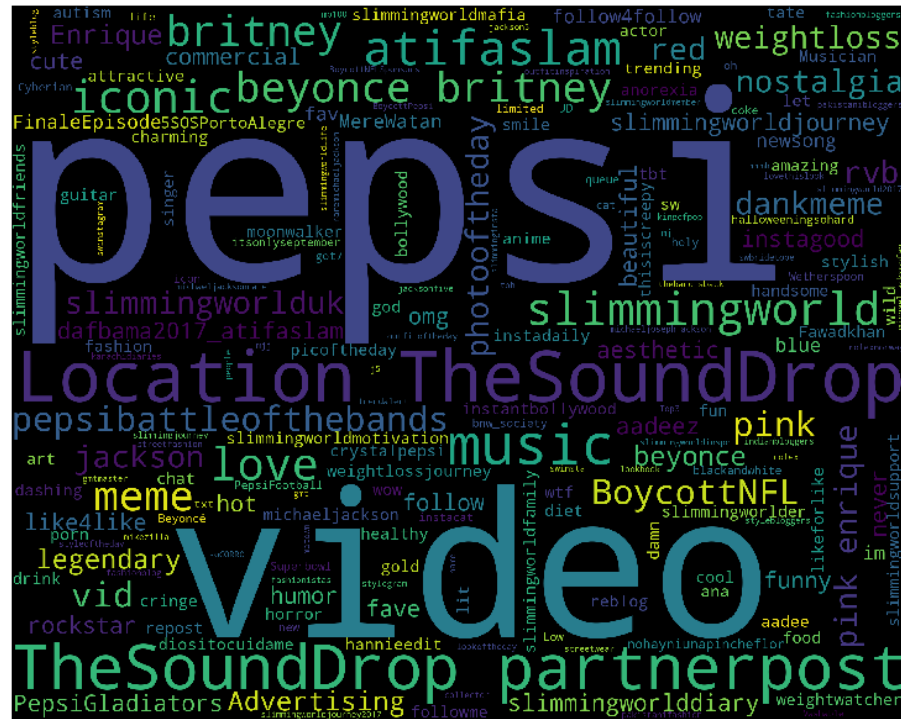
- [illegible]

-
- A word cloud visualization of tweets mentioning brands. The words are arranged in a circular pattern, with larger words indicating higher frequency. Brands like Pepsi, Coca-Cola, Frito-Lay, and Max are prominent.

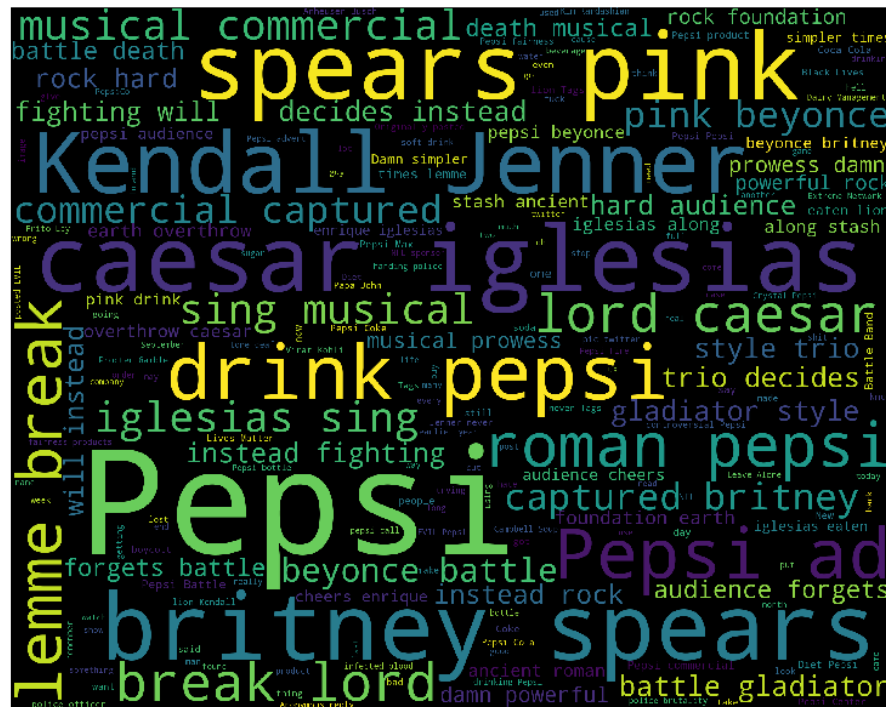
- Among the top hashtags associated with positive sentiments of Pepsi, we can notice that many of them are related to the #iKON commercial in South Korea. #JINHWAN leading actor in it and the words which are not shown properly are also Korean characters. We can say Pepsi had a success in marketing in South Korea. In addition, tags like #WithPepsiCommunity shows the company's caring for the community.

5) What contributes to Pepsi's negative perception?

- Top hashtags associated with negative sentiments



- Top words associated with negative sentiments



- Analysis of selected hashtags and words

Among the top 50 hashtags associated with negative sentiments of Pepsi, we can notice that many of them are related to celebrities who starred in Pepsi's commercials such as #britney, #beyonce, and #enrique. This is also revealed in top words associated with negative sentiments of Pepsi when we see names like "Kendall Jenner".

The controversial, radical and virtue-signaling commercials bring Pepsi to an unfavorable position to some degree. We came to the conclusion as we see top hashtags like #pepsibattleofthebands and #boycott nfl, and top words like "commercial", "political", "protest", "virtue-signaling" and "controversial".

Part III. Summary and Recommendations

1) Summary of Sentimental Analysis Result

From our sentiment analysis of the two brands, we can conclude that Coca-Cola has more positive sentiments. For Pepsi, the positive sentiments percentage is 37.91%, and the negative sentiments percentage is 15.52 %; for Coca-Cola, the positive sentiments percentage is 43.72 %, and the negative sentiments percentage is 11.39 %.

There are three main attributes that makes Coca-Cola more positive. First, Coca-Cola has a good taste that customers think is "original" and "American". Second, products like Diet Coke and Coke Zero Sugar are welcomed by consumers because they cater to the fashion of staying fit and weight losing towards a healthy lifestyle. Last, Coca-Cola has a successful global brand building strategy which earned the brand good reputation among consumers across the world.

But Coca-Cola also gets some negative perceptions. Those complaints are mainly about health issues. Some are worried about the ingredients of sugar and caffeine, and some may suffer from eating disorder problems because of the product. Also, it seems the special edition of "Coca-Cola Coffee Plus" is not a very successful product.

Pepsi is very active in social events and maintains a high exposure rate. Also, Pepsi is in close relationship with other food brands, which would promote the consumption of the product.

Pepsi achieved great success in some of its ads, but it also gets a lot of negative perceptions mainly because of its controversial commercials. Some of them may be too radical and virtue-signaling, thus not liked by the mass customer. Because of this complex social impact, customers may spend too much time debating about irrelevant things rather than focus on the product itself.

2) Recommendations for the Two Brands

Based on the results of sentimental analysis, we would like to give the following recommendations for the two brands.

For Coca-Cola, the company may want to keep their original taste (as people like the taste) and be very careful to trying to release new flavors. In addition, since Diet Coke and Coke Zero Sugar has already proved their success on market, the company can make an effort to further reducing the sugar level or caffeine because fitness and healthy lifestyle is what people are pursuing these days. Lastly, Coca-Cola can consolidate its successful global branding by connecting more to health and fitness.

For Pepsi, we would recommend the company to be more conservative in commercials as controversial advertisements may distract customers' attention from the actual product. Since Pepsi is often mentioned with other food brands, the company can try more bundle sales. Also, Pepsi can learn from Coca-Cola and try to develop a healthier product.

Appendix

1) Split of work

Explanatory Data Analysis	Harshal Gogri, Xin Qiu
Sentimental Analysis	Harshal Gogri
Visualization and Recommendations	Xin Qiu

2) Codes (see Jupyter Notebook File)