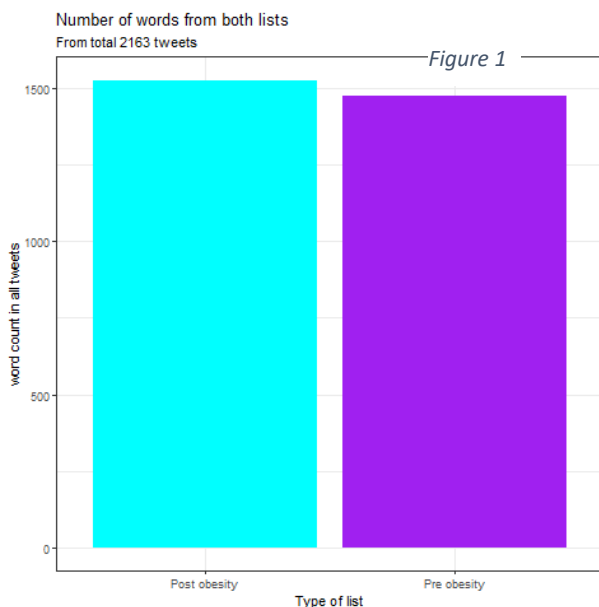Shekhar Karmarkar

DS 710 Project

Comparison of preventive and disease related words in obesity tweets

Introduction:

Obesity is very common disorder in society. Obesity has two main aspects. One is preventive aspect or a facet before obesity happens. Another is disease related aspect or when something goes wrong or some disease occurs due to obesity. Prevention is always better than cure.
In this study of obesity related trends, I have created two lists of words for both aspects. The words in preventive list would be like fiber, diet, run, calorie, etc. The words in disease related list would be like heart, cancer, arthritis, kidney, etc. this study wants to compare which group of words is more common in tweets. It will reflect view of society towards obesity. As obesity is very large commercial market, this study will help to companies to know which aspect of obesity is more common in society.

Data collection:

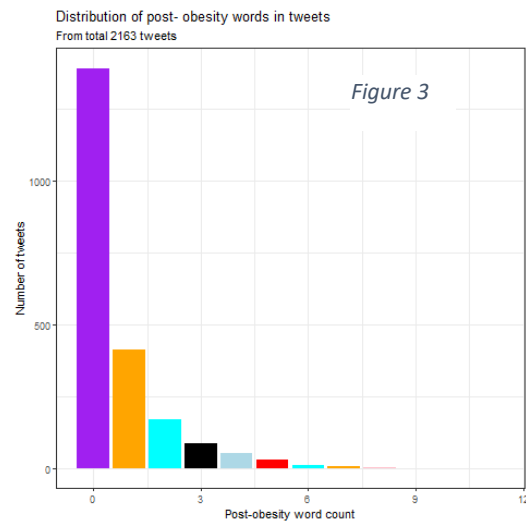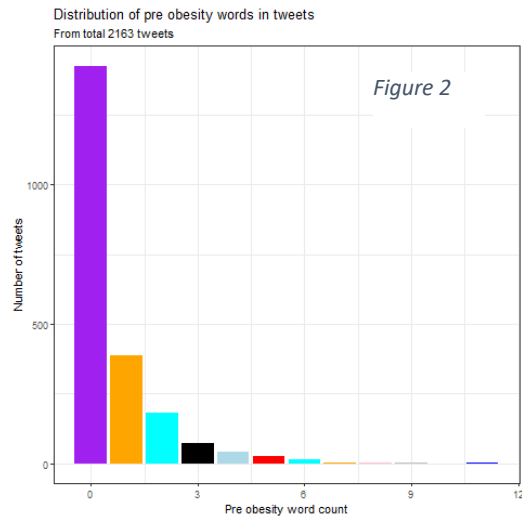Python is used to collect tweets with REST API. Tweets were collected over period of one week. I searched tweets for #Obesity tag so that the study is restricted to obesity. Duplicate tweets are deleted. Each tweet is used to count total number of words from both lists. A tweet can contain words from both lists, and they are counted separately.

Figure 1

Analysis:

As shown in Figure 1, total number of words in both groups are almost same. I have performed one sample proportion test to compare number of words. The hypothesis is that words from both groups will be same and proportion would be

50%. From the test, it can be concluded that the number of words from both groups are same.

As shown in figure 2 and 3, the number of words from both groups are shown per tweet. Many tweets do not contain any word from any list of words. But the distribution from both groups is almost same.

Figure 2

Figure 3

Conclusion:

From this study, it can be concluded that both preventive aspect and disease aspect are equally represented in tweets. We can say that the society is interested in both facets of obesity. A company which do business in the field of obesity should focus on both aspects as both are equally important.