

A top-down view of various fresh ingredients on a wooden surface. In the center is a large piece of raw salmon with a sprig of dill. To the left is a whole green onion. Above the salmon is a small white bowl filled with yellow oil. To the right is a halved avocado with its pit, a lemon wedge, and a head of purple garlic. Below the salmon are several cherry tomatoes and a lemon half. In the bottom right corner is a white bowl filled with a mix of nuts like almonds, cashews, and walnuts.

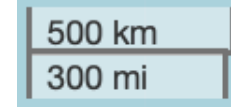
Analyzing Fad Diet Tweets

Vinh Tran
Capstone #2
7/3/2019

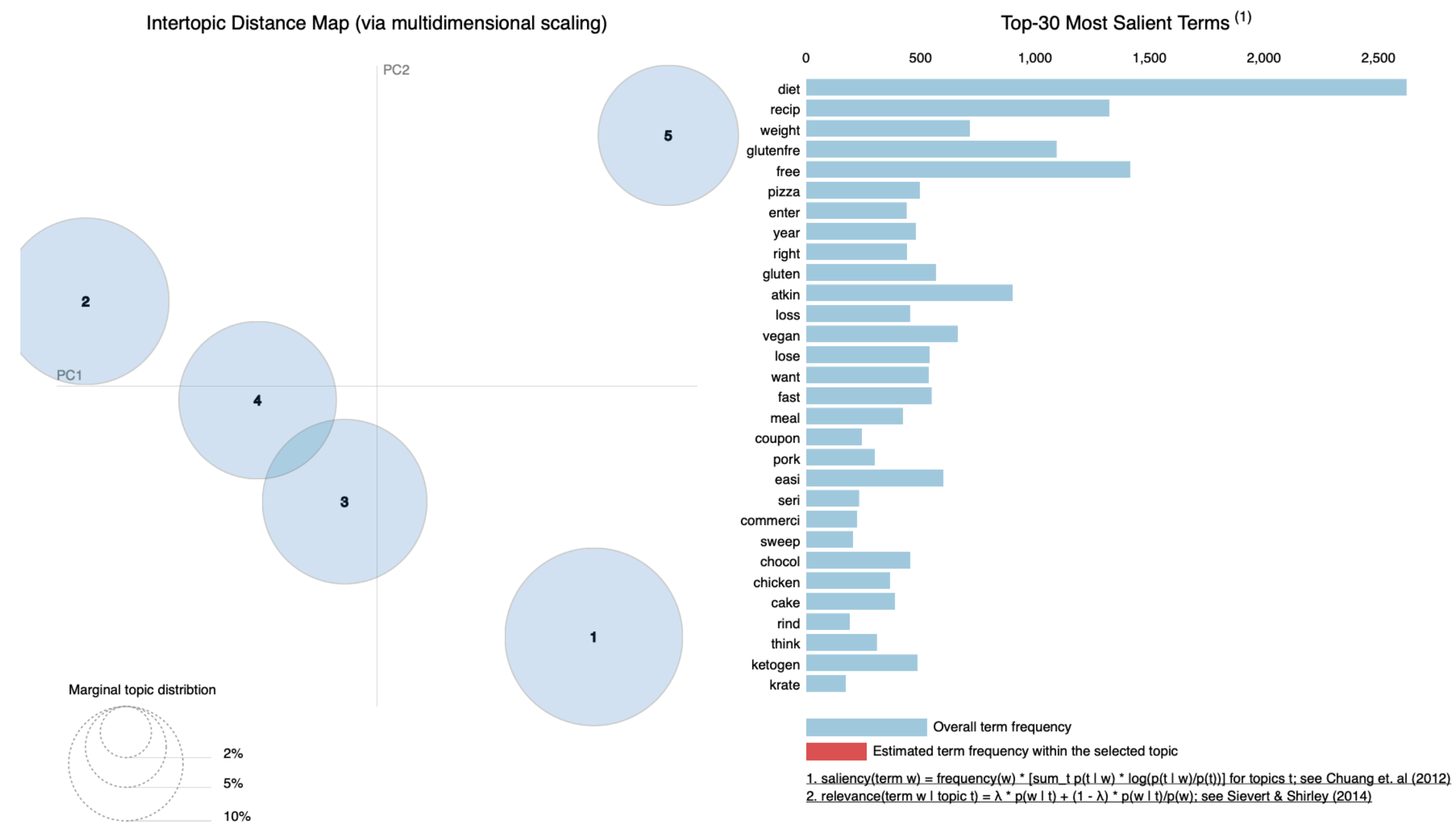
https://github.com/vinhtranh/fad_diets_ml



Fad Diet Twitter Data



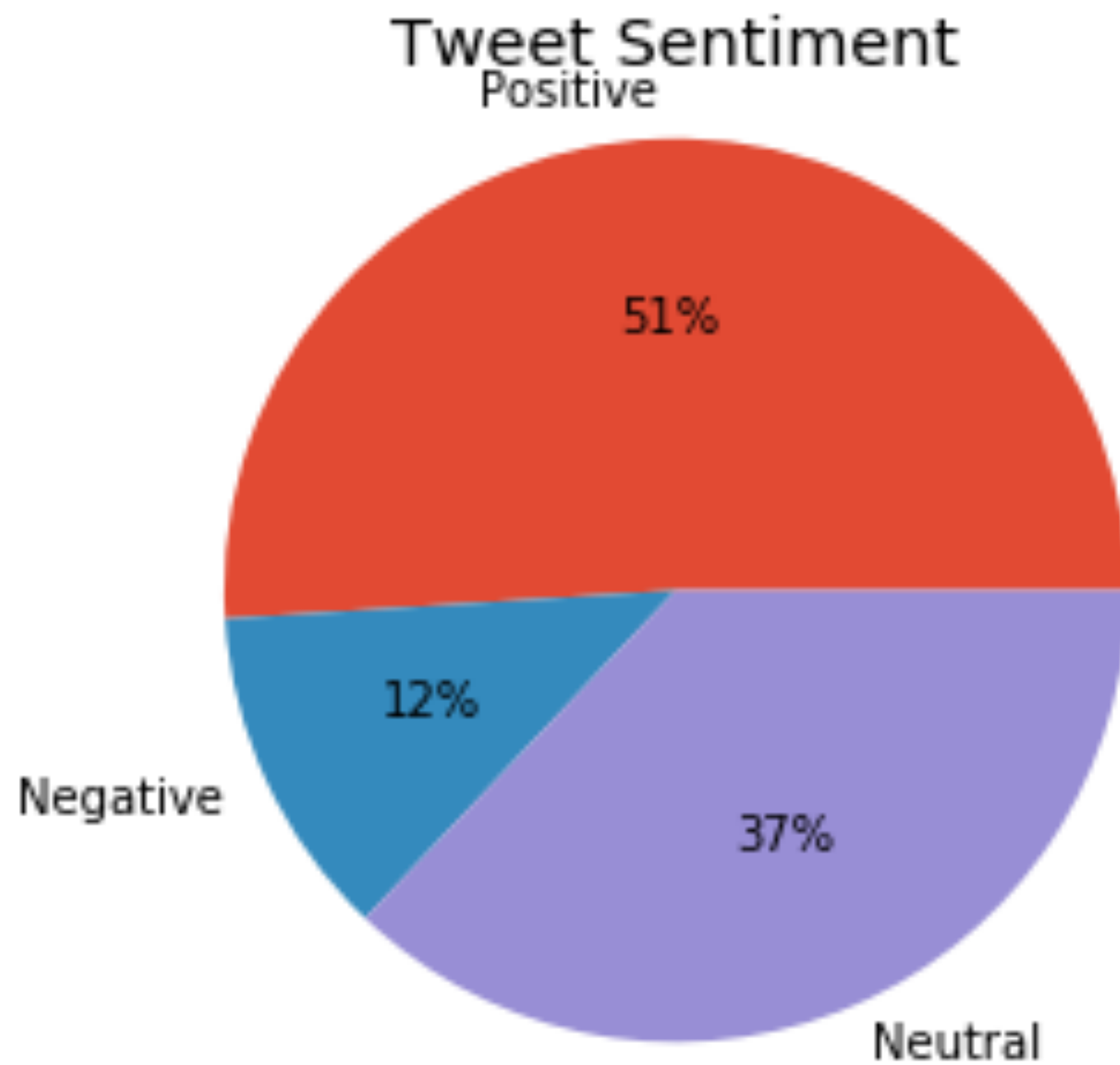
Topic Modeling using LDA



Perplexity tfidf: -7.57

	Keywords	Topic
1	go, chicken, dinner, fast, think, tri, recip, diet, right, best	Cooking
2	diet, paleo, know, thank, great, day, vegan, time, feel, meal	Feel-good diets
3	good, weight, idea, loss, diet, fee, like, love glutenfre, help	Weight-loss
4	food, come, friendli, happi, health, thing, father, ketodiet, work, ketogen	Health
5	atkin, glutenfre, week, carb, bread, today, coffe, lose, delici, chocol	Indulgences

Sentiment Analysis

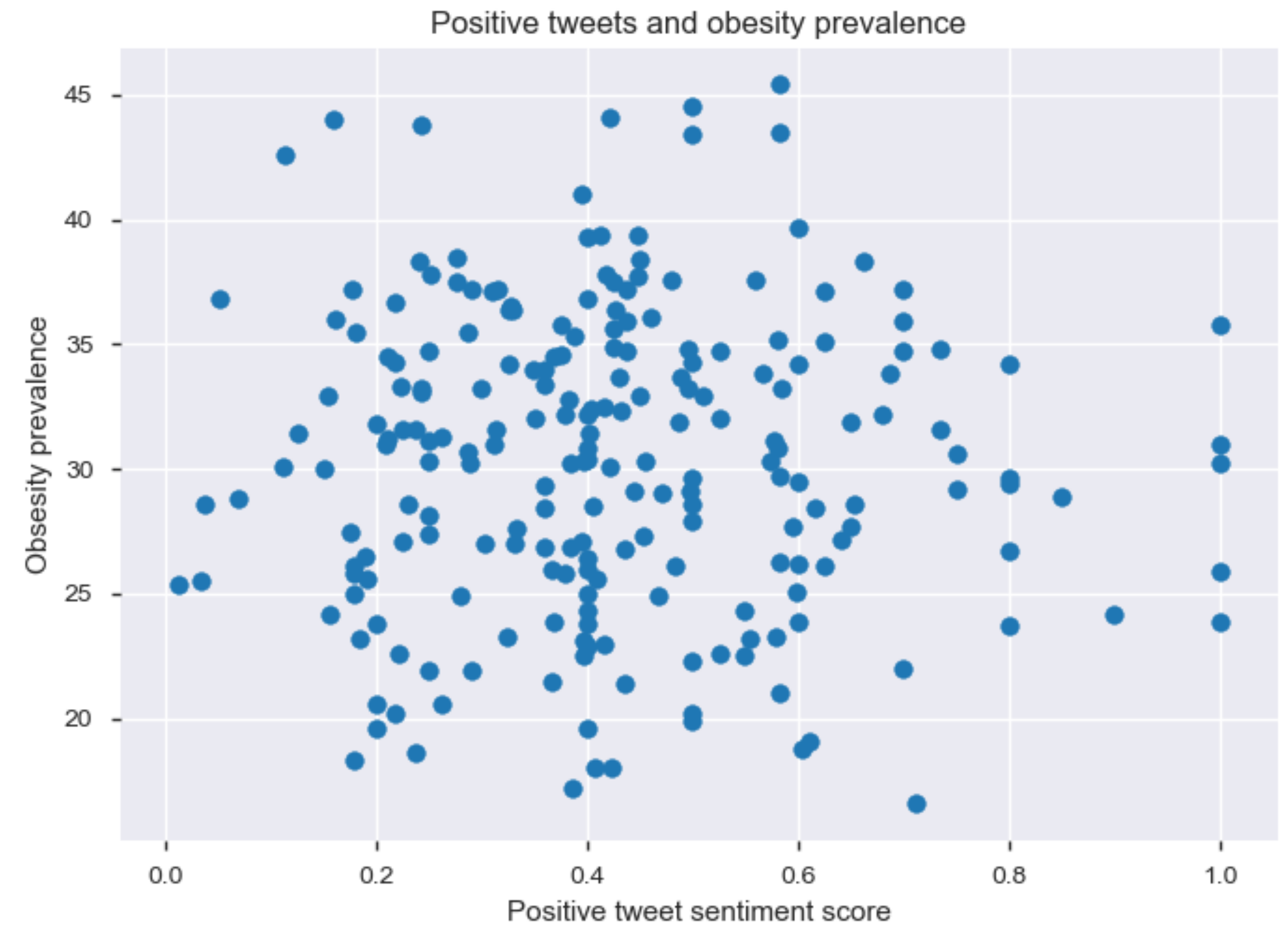


Naive Bayes Classifier accuracy
0.774869109947644

Linear Regression

Are positive tweets associated with obesity?

...NO, but...



Conclusions and Next Steps

- Emoji's
- More data!!!!
 - Continuously collect more tweets on AWS, utilize Spark.
- Capstone 3 goal is to make a prediction fad diet positive sentiment score in a city. This can be generalized to understand how certain markets respond to fad diets.
 - Combine with city level demographic data.