**<><><>PROBLEM 1<><><>**

Number of unique bigrams: 182246

Most frequent bigrams: [('with a', 4587), ('in the', 2595), ('of the', 2245), ('is a', 2056), ('on the', 2033)]

**<><><>PROBLEM 2<><><>**

Residual obtained: 1715.55282359

**<><><>PROBLEM 3<><><>**

Residual obtained: 1455.922293441

**<><><>PROBLEM 4<><><>**

Positive:

[(0.20595978853104524, 'not bad'), (0.20739457500788969, 'the best'), (0.21599260269442216, 'of these'), (0.23209416129721003, 'a bad'), (0.71053216725127544, 'sort')]

Negative:

[(-0.82557064428341109, 'sort of'), (-0.24323403776414548, 'water'), (-0.24034308329847279, 'corn'), (-0.2337507387037564, 'the background'), (-0.20574199313217217, 'straw')]

**<><><>PROBLEM 5<><><>**

Using the combination of unigrams and bigrams resulted in a better predictor. One, we can see how the negative/min predictors resulted with stop words, which seems to be true. In addition, we can see how the phrases correlated to the positivity and negativity more precisely.

Unigrams:

[(0.17953565223020301, 'stands'), (0.17968841874159203, 'okay'), (0.18160569739312082, 'murky'), (0.25065984633199245, 'modest'), (0.25226326129579446, 'juice')]

[(-0.38419429248466058, 'a'), (-0.3280567765234913, 'the'), (-0.32646097412740971, 'and'), (-0.29137707362423787, 'of'), (-0.25104187735801958, 'is')]

Bigrams:

[(0.25349606986092577, 'copper with'), (0.26023564312673175, 'even a'), (0.26523863019401184, 'dont think'), (0.3076108848520907, 'small head'), (0.34221493309409884, 'very much')]

[(-0.35482129365300835, 'with a'), (-0.34405888097763532, 'in the'), (-0.3402583067452094, 'of the'), (-0.31901941051531568, 'is a'), (-0.27198574026050126, 'on the')]

**<><><>PROBLEM 6<><><>**

Inverse document frequencies:

Foam = 0.9554602396075891

Smell = 0.5712174885030455

Banana = 1.6020599913279623

Lactic = 2.853871964321762

Tart = 1.0096611452123985

Tf-idf's in the first review:

Foam = 0.038998377126840365

Smell = 0.011657499765368276

Banana = 0.06539020372767193

Lactic = 0.1164845699723168

Tart = 0.02060532949413058

**<><><>PROBLEM 7<><><>**

Cosine similarity between first and second review: 0.91167782352553861

**<><><>PROBLEM 8<><><>**

Highest cosine similarity compared to first review: 1.0000000000000002

Text of the review:

"Poured into a Tulip glass and it developed a decent 1/16 inch head (could have been a bit bigger but it stuck around for a while). Blood red in color, smooth and very easy to drink. Slightly sweet and a little dry on the front end mixing with a nice caramel/CaraRed malt taste. The bittering hops show up in the middle with a nice slightly clean slightly sweet finish. Would definitely recommend! ~Jonathan Passow"