

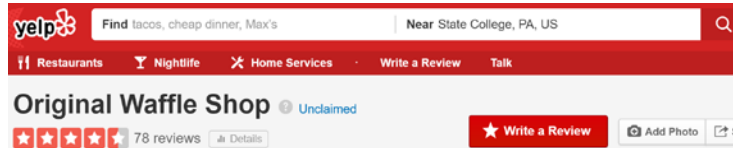
Why Statistics is Not Data Science

Chris Malone | Tisha Hooks

www.statsclass.org/uscots2017

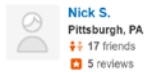
Goal: What is the relationship between the Yelp rating of one’s experience and the positivity / negativity of their review for a restaurant in State College.

Consider an investigation of Yelp customer reviews on restaurants in State College PA. The goal of this investigation will be centered around the positivity or negativity of the customer reviews for the Original Waffle Shop – a popular breakfast location in State College.



Example reviews left by customers of the Original Waffle Shop.

Review #1



The good:
The food is great. The portions are perfect.
Pecan waffles are delicious. The sausage is real.
They have different flavors of syrup. The bathroom door can be opened with your foot.

The Bad:
It's hard to get refills on drinks.

Review #2



It is hard to screw up breakfast food. At the same time, it is also hard to impress with breakfast food. This place didn't screw up, nor did it impress; just your regular breakfast food. Maybe I just don't enjoy stuffing myself full of greasy or butter-heavy things early in the morning.

A list of positive /negative words is needed to quantify the reviews. One such list, AFINN-111, contains about 2500 words with each word given a value between -5 and 5. A snippet of the AFINN-111 word list is provided here.

Word	Value
catastrophic	-4
hell	-4
badly	-3
screw up	-3
imperfect	-2
unhappy	-2
unhealthy	-2
hard	-1
broken	-1
cancel	-1
:	:

:	:
:	:
accept	1
agree	1
appreciation	2
approval	2
great	3
impress	3
perfect	3
terrific	4
wonderful	4
outstanding	5

Data Sources

- Original Waffle Shop Yelp Reviews: <https://www.yelp.com/biz/original-waffle-shop-state-college>
- A-FINN Word Lists: http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=6010

Why Statistics is Not Data Science

Chris Malone | Tisha Hooks

www.statsclass.org/uscots2017

Identify the procedural steps in providing an *analytics* solution for the goal stated above.

Step 1	
Step 2	
Step 3	
Step 4	
Step 5	
Step 6	
Step 7	
Step 8	
Step 9	
Step 10	