

The experiments that I ran were text analysis on the two groups classname and rating. Found the most common word for each classname and rating, the average length of review measured in words, and the average number of good words and bad words. The highest average number of good words in the classname is the unknown class name. And the highest average number of bad words in the classname is 'trend'. In the Rating group, the good words has a direct correlation with the rating number. The higher the rating the higher the good word average. Where as the bad words has a negative correlation. The lower the rating the higher the bad word average. Initially, I looked at the 500 most common words in the higher ratings for good words and in the lower ratings for the bad words. Then I looked online for a txt with a list of bad words and good words and opened it in my program.

Classname:

Class Name Unknown
Most Common word: [('soft', 7)]
Average length of review: 30.785714285714285
Average number of good words: 0.1136890951276102
Average number of bad words: 0.016241299303944315

Fine gauge
Most Common word: [('sweater', 714)]
Average length of review: 54.49818181818182
Average number of good words: 0.07471475278574764
Average number of bad words: 0.01906652432107827

Skirts
Most Common word: [('skirt', 1228)]
Average length of review: 58.44973544973545
Average number of good words: 0.06696840771250114
Average number of bad words: 0.015950031682809812

Jeans
Most Common word: [('jeans', 930)]
Average length of review: 56.516129032258064
Average number of good words: 0.06695051215599161
Average number of bad words: 0.018372824879674194

Chemises
Most Common word: [('dress', 1)]
Average length of review: 19.0
Average number of good words: 0.10526315789473684

Average number of bad words: 0.0

Sweaters

Most Common word: [('sweater', 1131)]

Average length of review: 57.13655462184874

Average number of good words: 0.07066955914255249

Average number of bad words: 0.02062727506710299

Trend

Most Common word: [('dress', 83)]

Average length of review: 67.89915966386555

Average number of good words: 0.05804455445544555

Average number of bad words: 0.021905940594059405

Knits

Most Common word: [('top', 2915)]

Average length of review: 51.447449927730744

Average number of good words: 0.08176673623374539

Average number of bad words: 0.01842992454647616

Dresses

Most Common word: [('dress', 9201)]

Average length of review: 62.21743946827029

Average number of good words: 0.06512748250040697

Average number of bad words: 0.01808206495197786

Intimates

Most Common word: [('bra', 67)]

Average length of review: 50.253246753246756

Average number of good words: 0.08024292544256365

Average number of bad words: 0.018090192531334797

Lounge

Most Common word: [('dress', 227)]

Average length of review: 53.25470332850941

Average number of good words: 0.07513791135628685

Average number of bad words: 0.01595152042175059

Outerwear

Most Common word: [('coat', 260)]

Average length of review: 68.7560975609756

Average number of good words: 0.06247782901738205

Average number of bad words: 0.018401915572898192

Shorts

Most Common word: [('shorts', 224)]

Average length of review: 54.95268138801262

Average number of good words: 0.06986222732491389

Average number of bad words: 0.01584385763490241

Pants

Most Common word: [('pants', 908)]

Average length of review: 59.65057636887608

Average number of good words: 0.06769732471767619

Average number of bad words: 0.016631439096563803

Layering

Most Common word: [('wear', 47)]

Average length of review: 45.04109589041096

Average number of good words: 0.0739051094890511

Average number of bad words: 0.019920924574209246

Legwear

Most Common word: [('leggings', 72)]

Average length of review: 45.90909090909091

Average number of good words: 0.07326732673267326

Average number of bad words: 0.016897689768976897

Sleep

Most Common word: [('wear', 67)]

Average length of review: 44.79385964912281

Average number of good words: 0.08410848918045628

Average number of bad words: 0.014687163419171643

Blouses

Most Common word: [('top', 2317)]

Average length of review: 57.73329028091702

Average number of good words: 0.0766890380313199

Average number of bad words: 0.019110738255033555

Swim

Most Common word: [('suit', 261)]

Average length of review: 59.63428571428572

Average number of good words: 0.07172288233039478

Average number of bad words: 0.015331544653123802

Jackets

Most Common word: [('jacket', 487)]

Average length of review: 59.13494318181818

Average number of good words: 0.06687324349643295

Average number of bad words: 0.01818356513175278

Casual bottoms

Most Common word: [('pants', 2)]

Average length of review: 20.0

Average number of good words: 0.05

Average number of bad words: 0.0

Rating:

4

Most Common word: [('dress', 2248)]

Average length of review: 60.475083710852864

Average number of good words: 0.06635139008311837

Average number of bad words: 0.01784830245707288

2

Most Common word: [('dress', 728)]

Average length of review: 60.630670926517574

Average number of good words: 0.05410646347761021

Average number of bad words: 0.027221853362420563

1

Most Common word: [('dress', 363)]

Average length of review: 56.44536817102138

Average number of good words: 0.047741283901782144

Average number of bad words: 0.033033854440633746

5

Most Common word: [('dress', 5539)]

Average length of review: 54.59096793846623

Average number of good words: 0.08107595342854831

Average number of bad words: 0.014845835995502446

3

Most Common word: [('dress', 1372)]

Average length of review: 62.401950539881575

Average number of good words: 0.05796624171113443

Average number of bad words: 0.02342651097367657

