Kyle Dunn

Text Mining Final

**PART 1 – Word Cloud**

Word cloud 1 - positive

**A close up of words

AI-generated content may be incorrect.**

In the positive word cloud more positive words appear like good, great, best, pos, and funny

Word cloud 2 – negative

A close up of words

AI-generated content may be incorrect.

Opposed to the positive word cloud the negative word cloud has more negative words like neg, least and bad.

**PART 2 – Jittered Scattered Plot**

A graph showing a number of words

AI-generated content may be incorrect.

**Part 3 – Bing**

**A screen shot of a graph

AI-generated content may be incorrect.**

Part 3 chart makes sense the negative and positive review in their respective places, but something I noticed when creating the word clouds reviews for films. Reviews often use the same words across many reviews. So, the plot being in the center makes sense.

**PART 4 - NCR sentiment lexicon**

**A red and blue squares

AI-generated content may be incorrect.**

**PART 5 – Topic Model**

**A group of colorful bars

AI-generated content may be incorrect.**

With the 20 topics in all of them have similar top words, the main top word between them in all categories is ‘film’ and the second is ‘movie’. Typical movie review terms: plot story character/characters and story are common in each topic. In peculiar topic 10 is about Lucas films and the star wars movie phantom menace.

**PART 6 - 75%/25% text training**

1000 positive and negative reviews.

neg 208 58

pos 42 192

Negative reviews

208 is negative reviews that are negative.

58 are negative reviews but are positive so the 58 are incorrect.

Positive Reviews

42 reviews are positive and are actually positive.

192 reviews are thought to be positive and are correctly positive.

A black numbers and a white background

AI-generated content may be incorrect.A black text on a white background

AI-generated content may be incorrect.