Task 3

Yujia Wang

12/8/2021

Task 3: sentence-level analysis

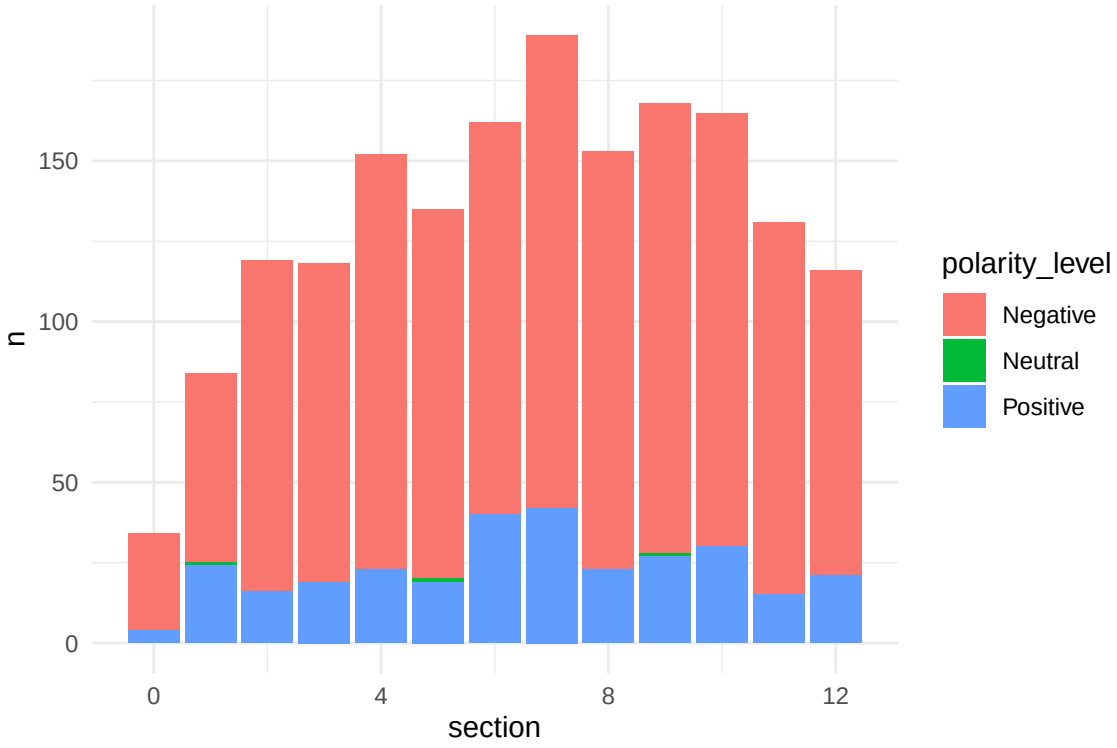
```
## [1] "test3"
```

input the book

Because test2 space is too busy, I use test3 space. And below is my "carroll/alice".

```
##  [1] ""                                "lewiscarrol/alice"
##  [3] "carrol/alice"                    "carroll/alice"
##  [5] "elisa/the_call_of_the_wild"      "handing/hw4"
##  [7] "zara/hw4"                        "handing/sea"
##  [9] "zara/A4"                         "zara/a4"
## [11] "elisa/wild"                      "dostoevsky/hw4"
## [13] "dostoevsky/crime_and_punishment" "handing/game1"
## [15] "handing/game2"                   "sisitzky/scarlet"
## [17] "william/test3"                   "zara/homework4"
## [19] "zara/submission4"
```

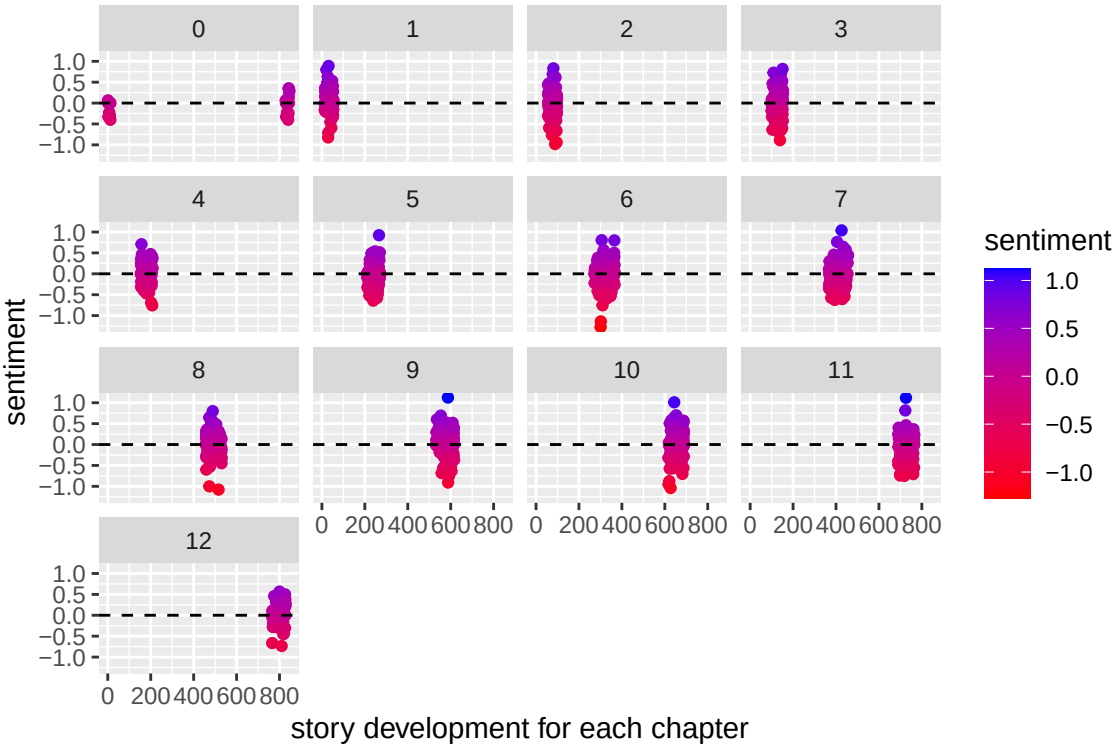
sentimentr



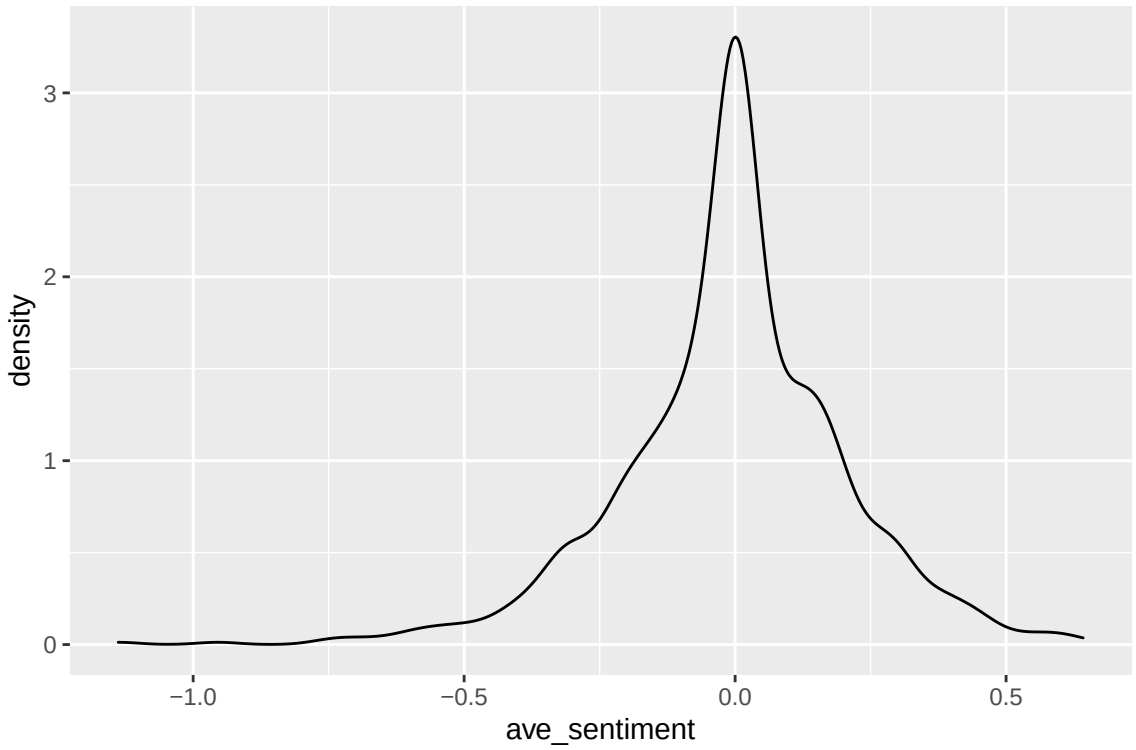
The graph above shows the percentage of emotions in each chapter.

The figure above shows the changes in positive and negative emotions as the storyline develops. The two lines are like two time series.

It can be seen that as the book goes forward, the positive emotions are less and the negative emotions appear more and more dense. This coincides with the middle and later part of the story, that is, the battle between Alice and the Red Queen. The adventure plot is becoming more and more exciting and thrilling, and more and more negative emotions.

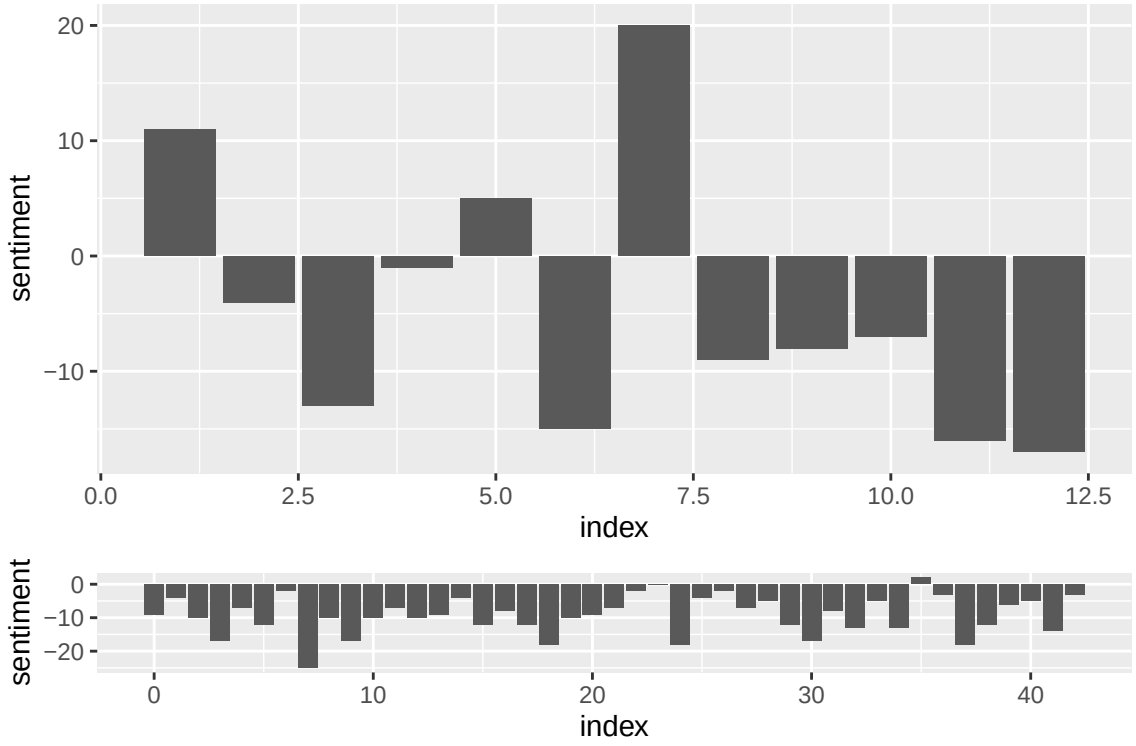


The above plot is a point plot of the emotions of each chapter.



The above plot is the average sentiment density plot. The curve shows a left-skewed distribution, that is, the density of negative emotions is greater.

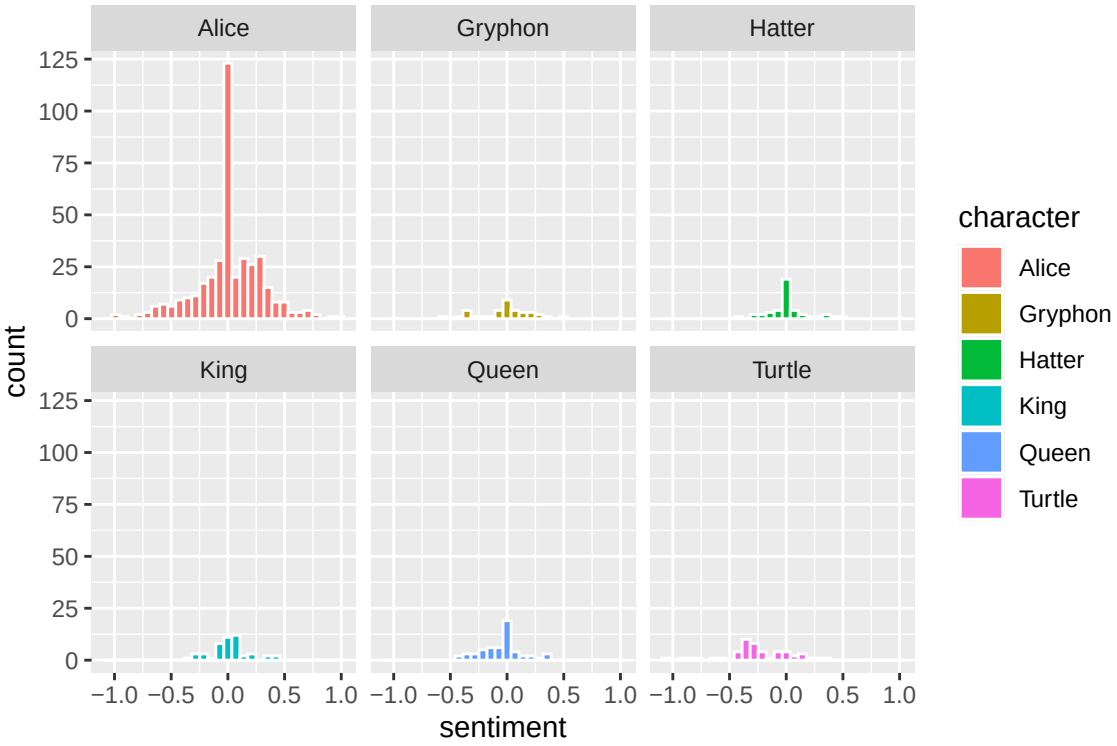
Compare task 3 with task 2



The above figure shows that task 2 and task 3 use the same method-Bing method to calculate emotional scores. From a trend point of view, the result of task 3 revealed more positive emotions than the result of task 2. The positive emotions in Chapters 1, 4 and 6 do not match the results of Task 2.

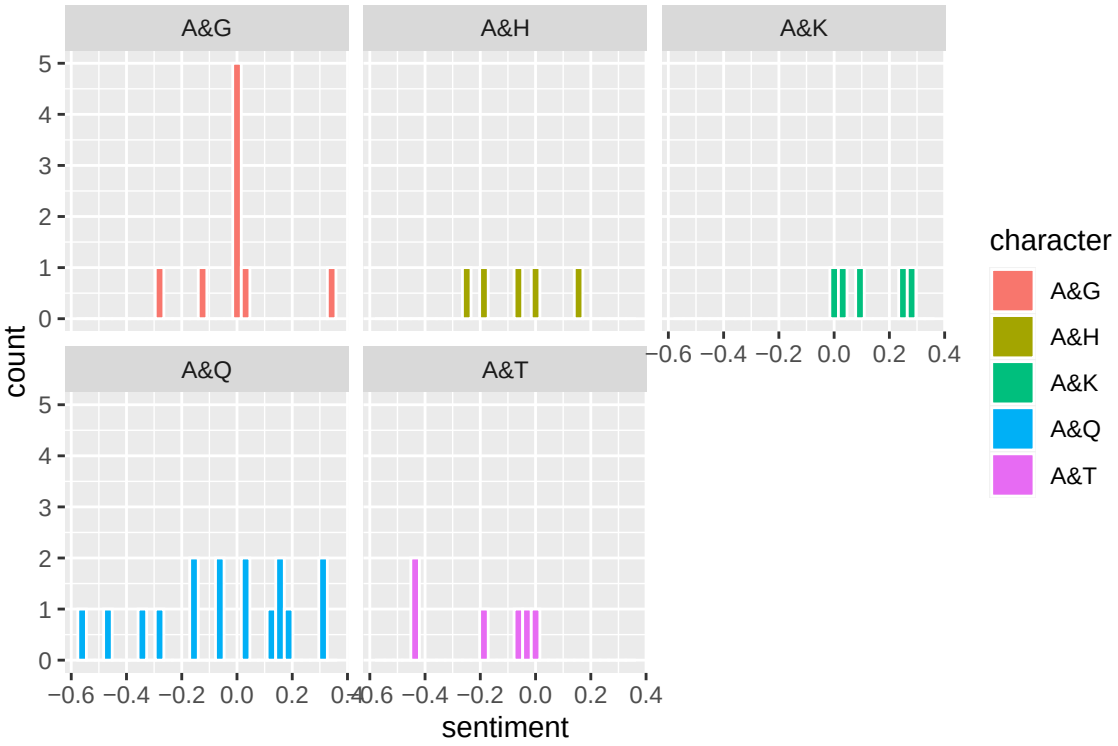
Extra credit: character-level analysis

According to Task 2, I selected six characters with the highest frequency. They are Alice, Queen, King, Turtle, Hatter and Gryphon.



Then, I filtered them from each chapter, each paragraph, and each sentence. Draw a histogram of their emotional scores.

It can be seen that the emotional scores of Gryphon and Hatter are positive, Turtle and Queen are negative, and Alice and King are relatively neutral.



Since Alice is the absolute protagonist of the book, I also calculated the emotional scores of the other 5 characters when they met her.

It can be seen that Alice and King had a very positive mood when they met. But her encounter with Turtle was negative. It can be inferred that when Alice is with different characters, the direction of the storyline is good or bad.

Reference

Haviland Wright, tnum - instructions and examples - v5.Rmd https://www.gutenberg.org/ebooks/11 https://www.r-bloggers.com/2020/04/sentiment-analysis-in-r-with-sentimentr-that-handles-negation-valence-shifters/ https://www.rdocumentation.org/packages/stringr/versions/1.4.0/topics/str_detect