Part 1: "The Shawshank Redemption" is the top-ranked movie in the "IMDb Top 250 Movies" list, with over 10,000 reviews to date. So, what factors contribute to its greatness according to these reviewers? Is it the cast, the director, the writer, or the film company behind it? Let's uncover the answer through our statistical analysis!

Analyzing the reviews of *The Shawshank Redemption* from Four dimensions	
1.Writing credits	• Stephen King((based on the short novel "Rita Hayworth and the Shawshank Redemption" by) • Frank Darabont (screenplay by)
2.Top 5 Cast and their characters in the movie (in credits order)	Tim Robbins ... Andy Dufresne Morgan Freeman ... Ellis Boyd 'Red' Redding Bob Gunton ... Warden Norton William Sadler. ... Heywood Clancy Brown ... Captain Hadley
3.Director	Frank Darabont
4. Company Credits	Castle Rock Entertainment

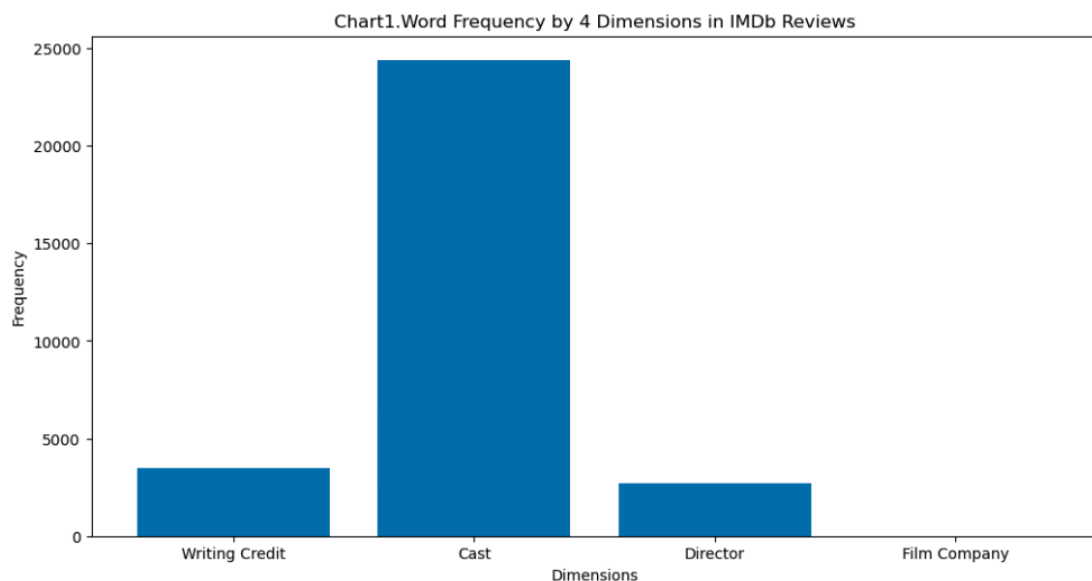Chart1.Word Frequency by 4 Dimensions in IMDb Reviews

Chart1.Word Frequency by 4 Dimensions in IMDb Reviews
From Chart 1, we can see that the most discussed aspect of the movie in reviews is the actors, followed by the authors and director. Interestingly, there were no mentions of the movie's production company among the over 10,000 reviews analyzed. The film company, Castle Rock Entertainment, is basically ignored by all reviewers which makes the incubator of this great film invisible to common audience.

Chart2.Word Frequency by Four Dimensions in IMDb Reviews in more details

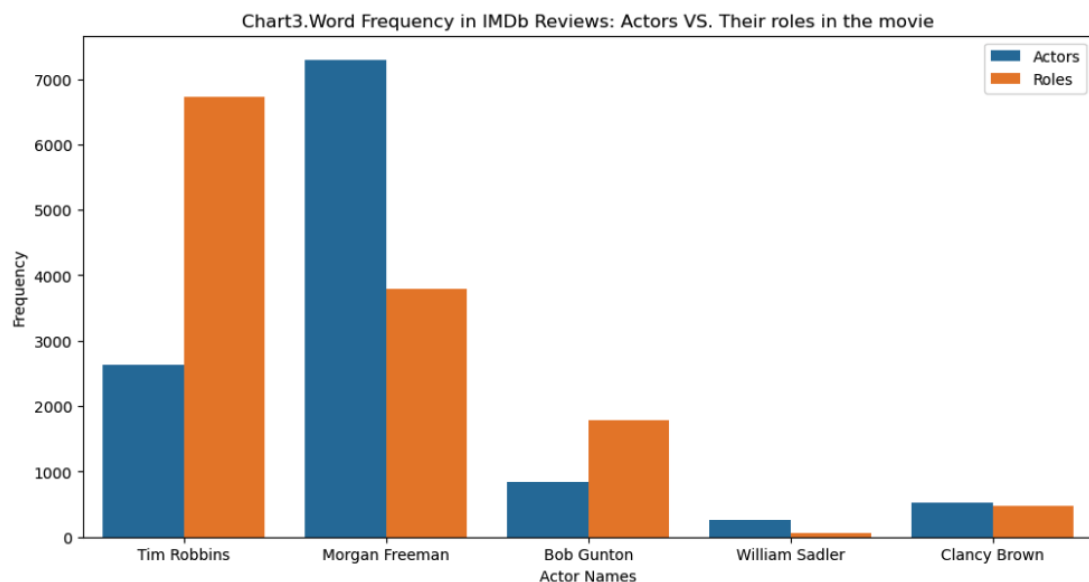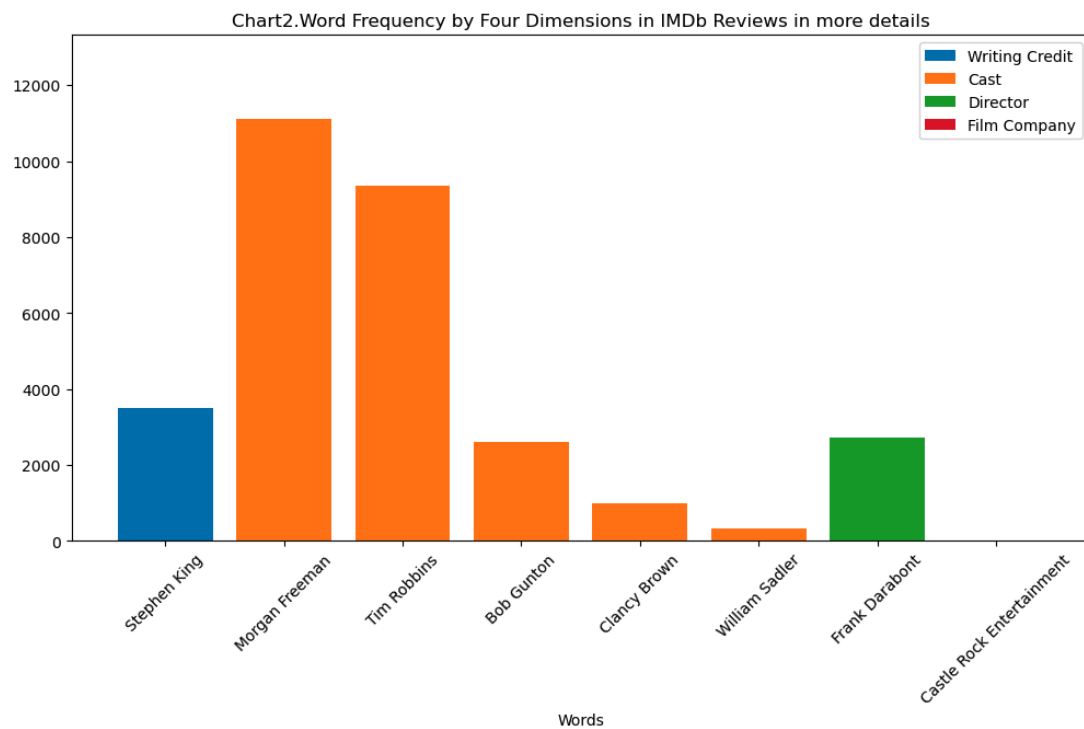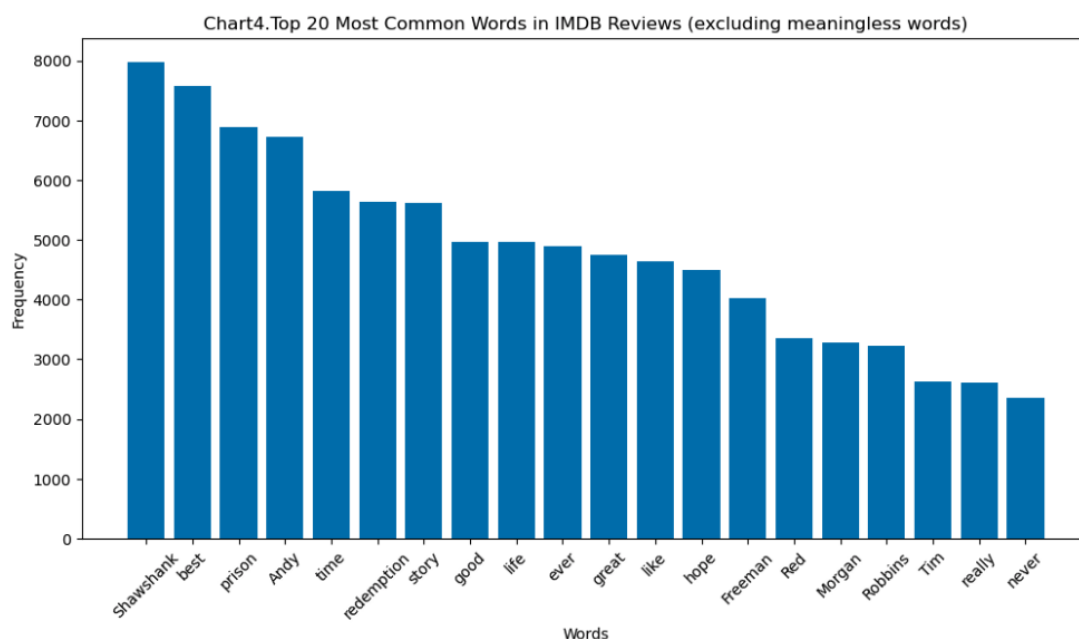Chart3.Word Frequency in IMDb Reviews: Actors VS. Their roles in the movie

Chart2.Word Frequency by Four Dimensions in IMDb Reviews in more details
As we know from Chart 1 that the Cast takes the most credits for the reviewers. Let's take a look at which actor do they like most. Chart 2 gives us the answer: Morgan Freeman and Tim Robbins and their roles in the movie are mentioned way more than the rest three actors.

Chart3.Word Frequency in IMDb Reviews: Actors VS. Their characters in the movie
Chart 3 takes a closer look at the comparison between the 5 main actors and their roles in the movie. It can be seen that some actors are more popular than their roles while others aren't. To be specific, Morgan Freeman, William Sadler, and Clancy Brown themselves are more discussed by the reviewers than their roles in the movie, especially Morgan Freeman who is even more popular than the first character Tim Robbins. For the rest of actors, their roles are more discussed than themselves, especially for Tim Robbins.

Part 2. In part one, we assumed four factors to check the word frequency. Let's ask the statistics to give us the full picture of the word frequency in the review contents and titles. What are the most frequent words, both positive and negative, that people used to describe this movie?

Chart4.Top 20 Most Common Words in IMDB Reviews (excluding meaningless words)

High-Rate Movie Titles Word Cloud

Low-Rate Movie Titles Word Cloud

Chart4.Most Common Words in IMDB Reviews

After calculating the frequency of all words in all contents, we find those top 20 most common words. As we can see, those words can be divided into 2 categories. One is about the people and plots relating to the movie itself, i.e., Shawshank, prison, Andy, redemption, story, life, Freeman, Red, Morgan, Robbins, Tim..., the other is about the reviewers' praise of this movie, i.e., best, time, good, ever, great, like, hope, really, never...And this distribution is very similar to the distribution in High-rate Movie Word Cloud of all titles.

World Clouds

Upon comparing the two word clouds, it is clear that positive words dominate the high-rating reviews. *The Shawshank Redemption* is a timeless classic that has left a lasting impact on multiple generations, evident from the prominence of the word "best" in the cloud. However, upon examining the low-rating reviews, we can see negative words like "overrated," "average", "mediocre", "long", "boring," "nothing" and "overblown" emerging in the word cloud. The reason behind those negative reviews is that some people think it's too long and boring and shouldn't be rated so high. This indicates that even a universally acclaimed film like The Shawshank Redemption cannot cater to every individual's cinematic preference.

Part3. So, even though The Shawshank Redemption is generally regarded as a great movie, people's views towards it do vary. Does its rating also vary over time?
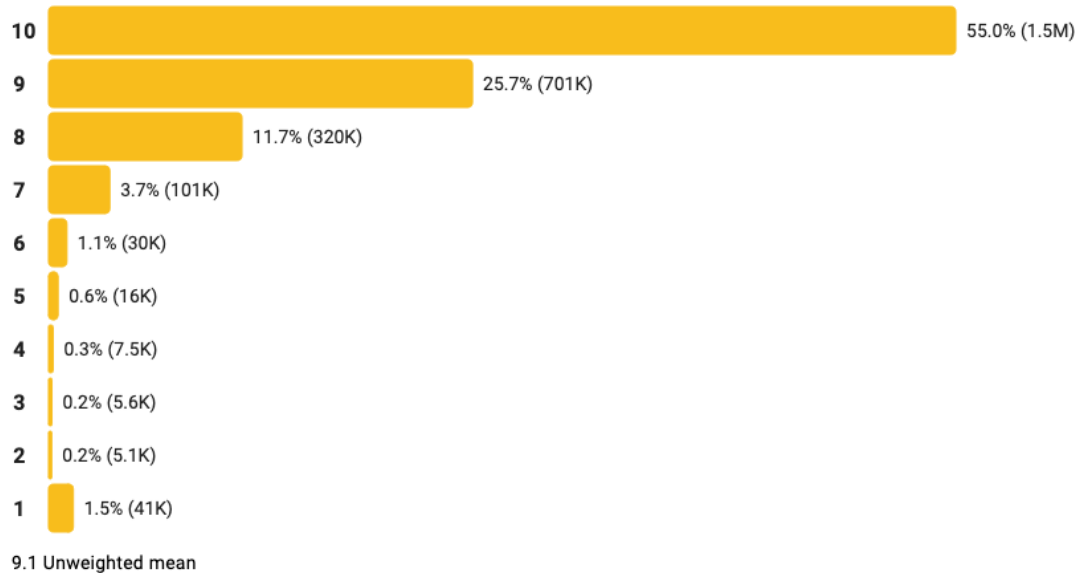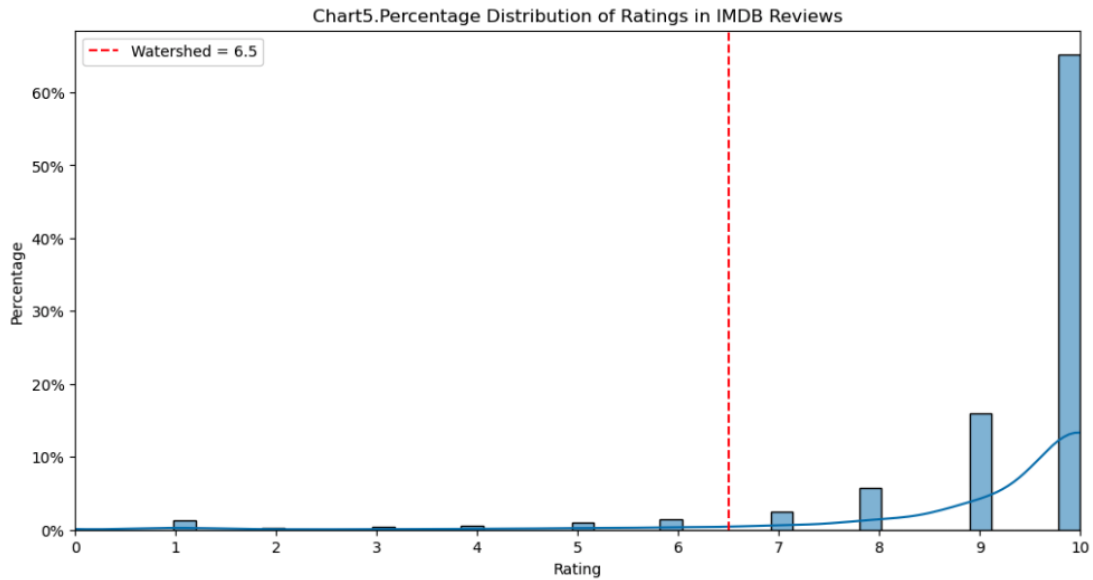
Chart5. Percentage Distribution of Ratings in IMDB Reviews

Overall, we can see that the rate distribution of all reviewers is very much similar to all ratings distribution provided by the IMDb website.
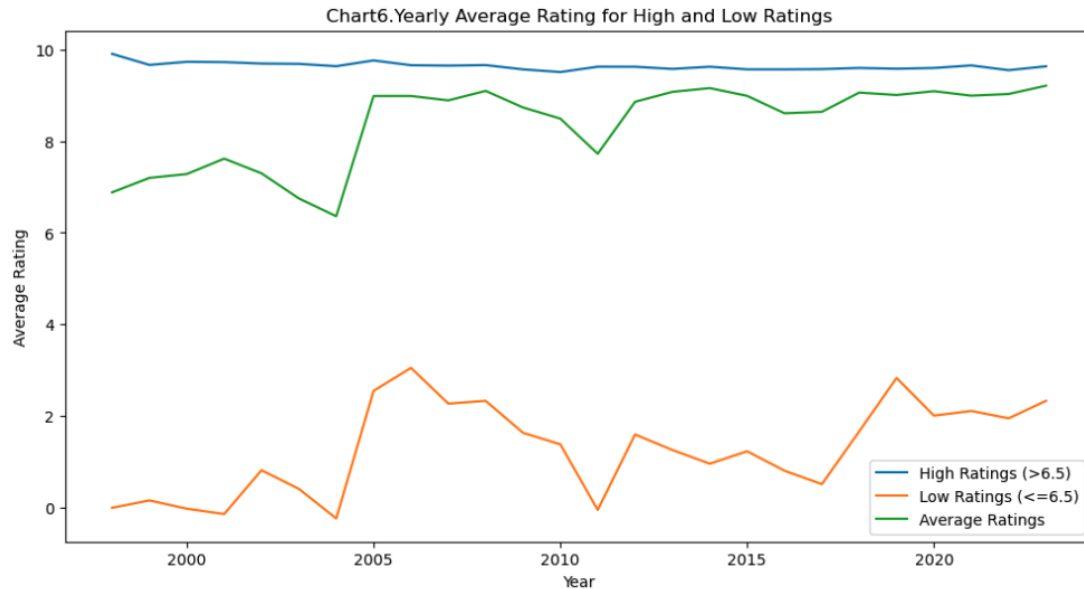
Chart 6: Yearly Average Rating for High and Low Ratings
The graph presents an intriguing phenomenon whereby the audience's ratings are predominantly concentrated in two extremes - either near perfect scores or very low scores - with minimal middle scores. This trend remains consistent across the years, except for a slight increase in low scores between '05 and '07. The polarization of audience ratings for this movie indicates that most individuals ardently fond of the genre while a few strongly disapproving of it.
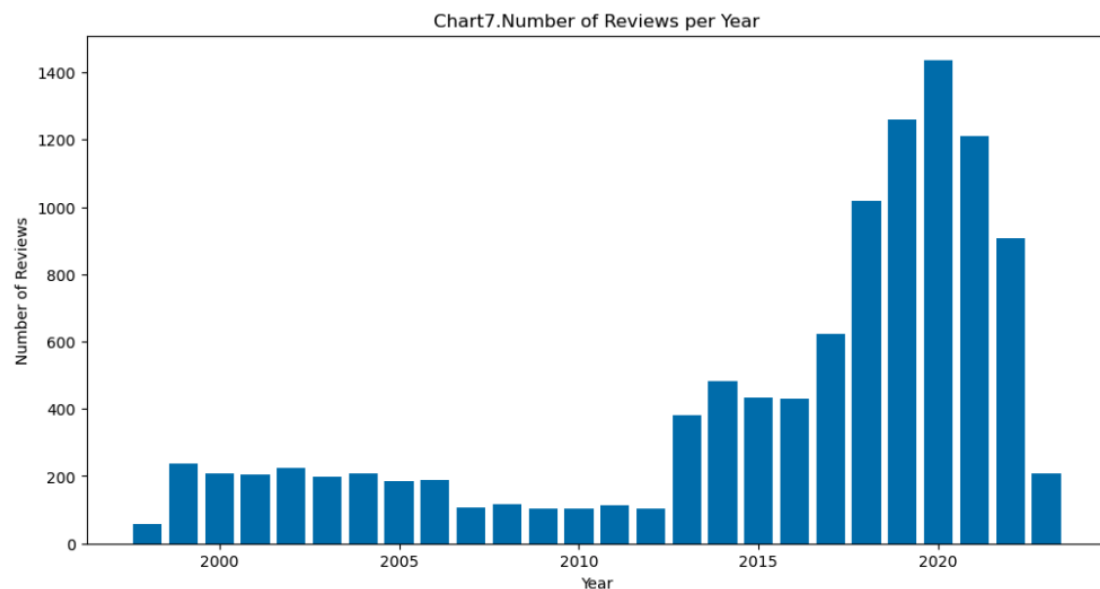
Chart7: Number of Reviews per Year
This graph reveals a fascinating trend in the number of movie reviews over time. The quantity of reviews remained relatively consistent during the early years, with only minor fluctuations. However, a significant increase occurred around 2013, which I

More people likely became aware of the movie review website during this period and began contributing reviews. The graph also shows a sharp increase in reviews in 2019, which I believe was a result of the COVID-19 pandemic. With more people spending time indoors due to quarantine measures, there was an increase in online activity, including more interest in older movies. This surge in attention ultimately led to a significant increase in the number of reviews for this movie in 2019.

Part 4. Alright, we have examined enough of people's attitudes towards this movie. Now let's explore their other preferences. Do they prefer reviews without spoilers? Do they tend to write longer reviews or shorter ones?
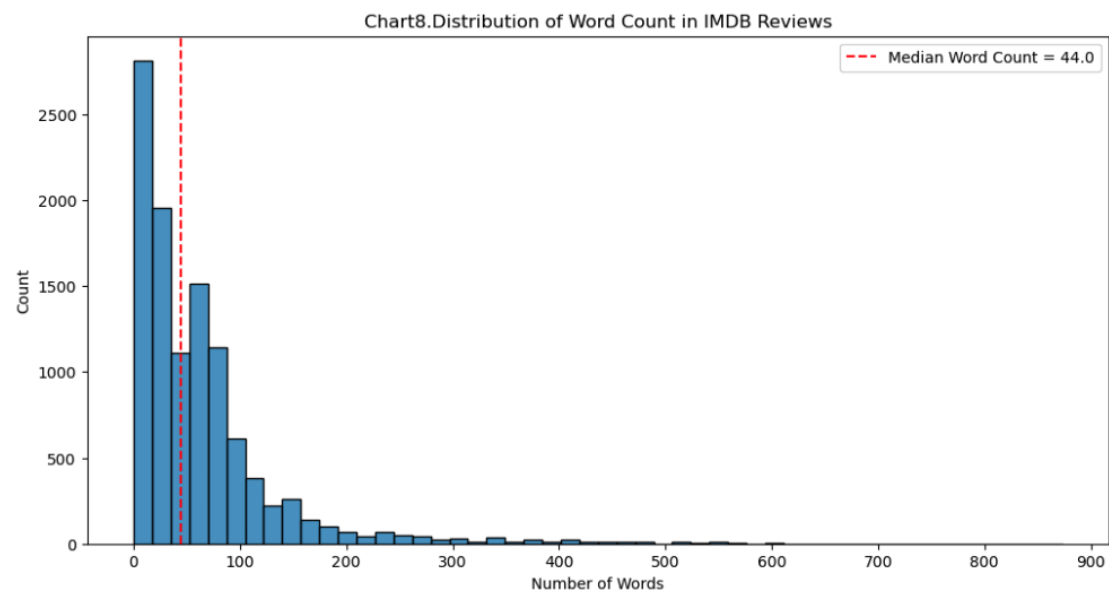

Chart8.Distribution of Word Count in IMDB Reviews
From this chart, we can safely say that people prefer writing shorter reviews(<=200words) with only a few like writing really long ones.
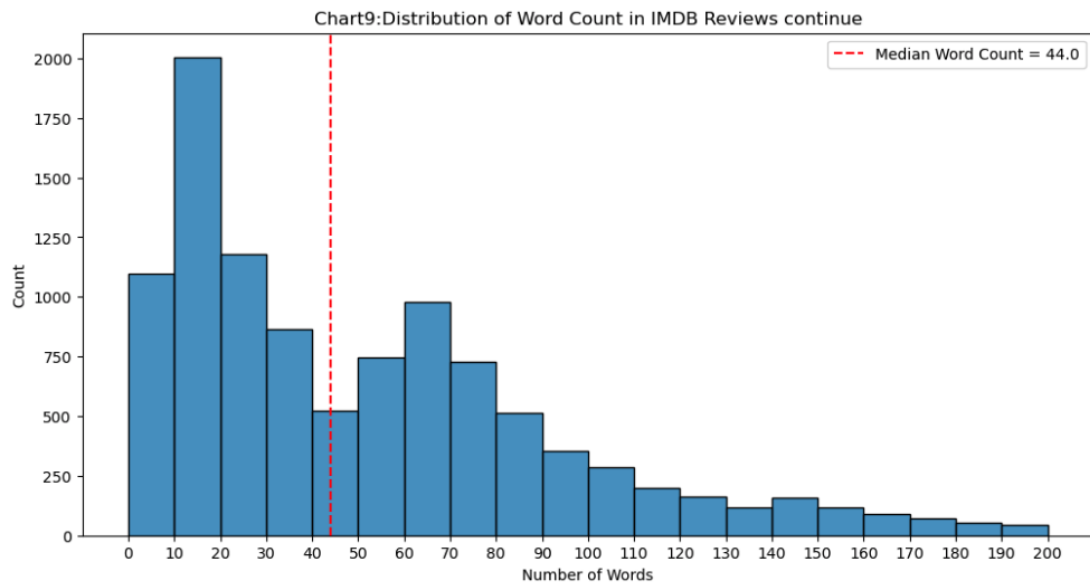
Chart9.Distribution of Word Count in IMDB Reviews continue
When we take a closer look at reviews less than and equals to 200 words, we can see the trend that the number of reviews decreases as the word amount increases. Particularly, 10-20 words are the most popular range of words that reviewer choose to go.
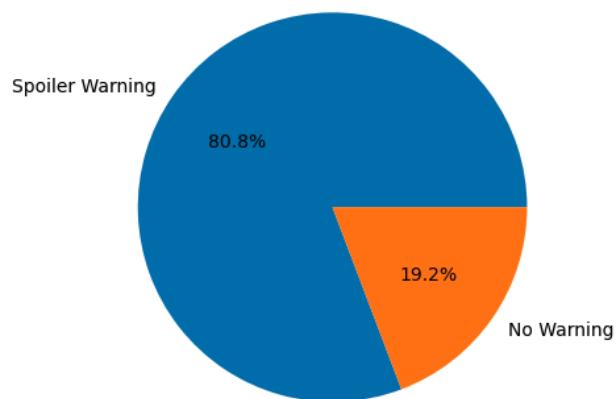

Chart10.Proportion of Spoiler Warning
From the spoiler warning chart, it can be seen that only a small part of the reviews (19.2%) has spoiler warning tags. We can assume that most reviewers probably tend to protect the first-hand experience of a great movie like Shawshank so that they avoid giving spoiler reviews.

Chart11.Proportion of Votes by Spoiler Warning

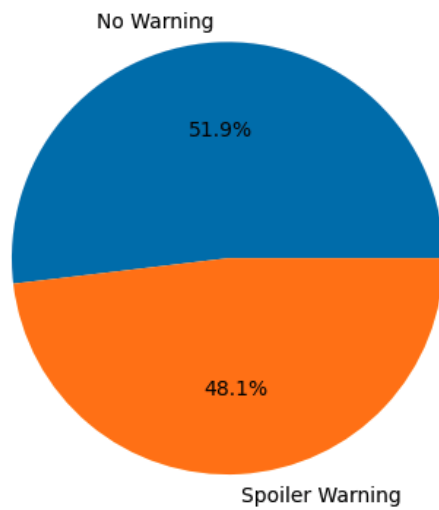No Warning

51.9%

48.1%

Spoiler Warning

Chart11.Proportion of Votes by Spoiler Warning

According to the percentage of accumulated votes of reviews both with and without spoiler warning, 51.9% of the votes goes to non-spoiler reviews while the rest 48.1% goes to spoiler warning reviews. There is no significant difference between them which makes it hard to tell which kind of reviews are more popular. As the spoiler warning reviews only take 19% of whole reviews, we may predict that audiences tend to like reading spoiler warning reviews.