

Report Big Data project

Game of Thrones episodes data

Realised by Isabella Mi Hyun Kim, Vu Diep
5th year ISS - January, 2020

Introduction

The analysis of the data is a crucial part of data science to develop the essence of what information we can extract from a certain dataset and what type of questioning can this data raise.

From this investigation, we can decide which variables are relevant for the type of study we want to carry. It's important to say that every data analysis, as neutral as it may try to be, always carry an influence of the person who's exploiting the data. No data investigation is ever impartial.

In this project we chose to work with a dataset about a tv show called Game of Thrones that is openly accessible in the platform Kaggle. Game of Thrones is an American fantasy drama television series transmitted by HBO.

The dataset with the episodes' data of Game of Thrones movies can be found in: <https://www.kaggle.com/dasbootstrapping/game-of-thrones-episode-data>.

We divided the analysis of the data basically into 2 steps:

- 1. Defining questions:** Identifying relations between the variables that might be interesting or that we could make a correlation.
- 2. Visualization:** Trying to visually communicate through graphs the results we reached in the clearest way possible.

Analysis of the data

The number of viewers per episode and per season

We can see from the graph in Figure 1 that the number of US viewers grew almost linearly as the TV show progressed. That means that the show's audience was really stable.

However, from seasons 5 to 8, we can verify that the growth was not as big as in the first seasons and not as constant. The number of viewers started to vary drastically, for example, in season 8: episode 1 had nearly 12 million US viewers and the next episode had approximately 10 million viewers. This is the biggest decay of viewers between episodes in the whole series. The show lost almost 2 million viewers with the pilot of its last season.

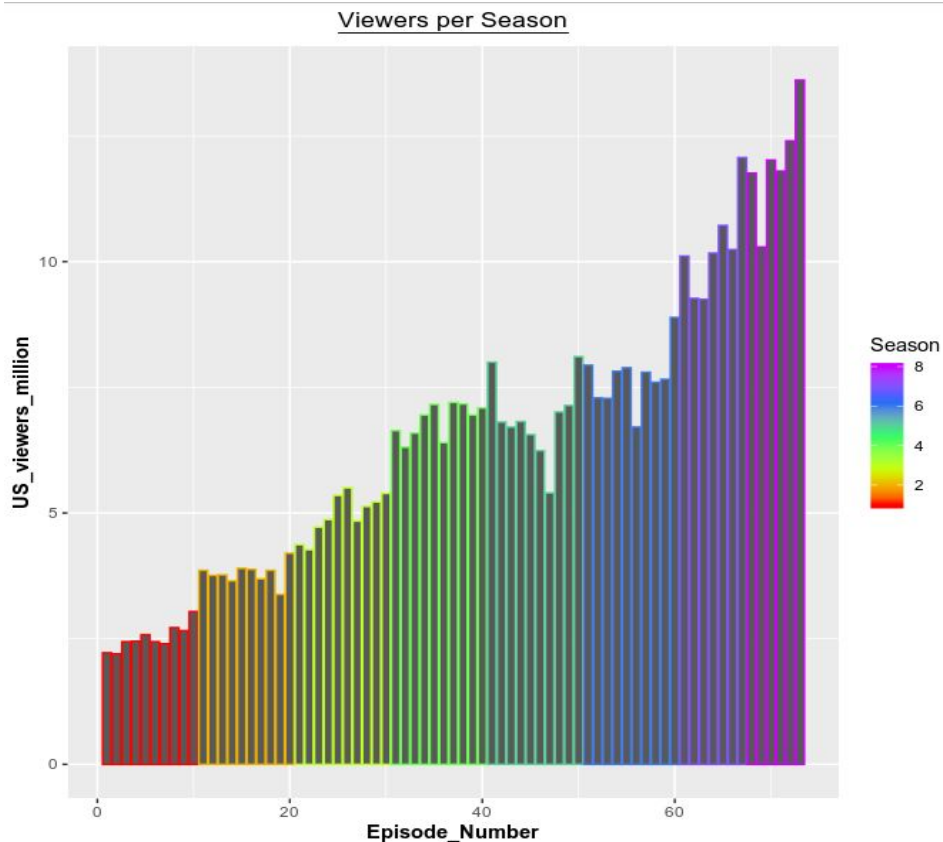


Figure 1: Number of US viewers per episode and per season.

Maybe this more irregular behavior of the telespectateurs came from the immense expectations created by the show. As the show progressed, people started creating more expectations. And with great promises can also come great disappointments, that we can see through the graph.

We can also verify that the biggest growth in the series came between seasons 3 and 4 (more than 2 million new US viewers). This fact could be related to one of the most famous episode from the whole series "The Red Wedding" (last episode of season 3). The show became viral on social media, many people recorded their reactions on Youtube for this episode and the show gained a lot of attention.

Number of character deaths per episode and per season

Game of Thrones started to gain popularity for its fame of killing its main characters ruthlessly and never before has a show done that. Therefore, the number of deaths per episode is a relevant data to be studied in our analysis.

Through the Figure 2, we can see that for the first 5 seasons the percentage of deaths stay quite linear (around 12%) these seasons match the fact that all the intrigues of the story are setting up. The season 6 marks a turning point of the main characters and changes the course of events.

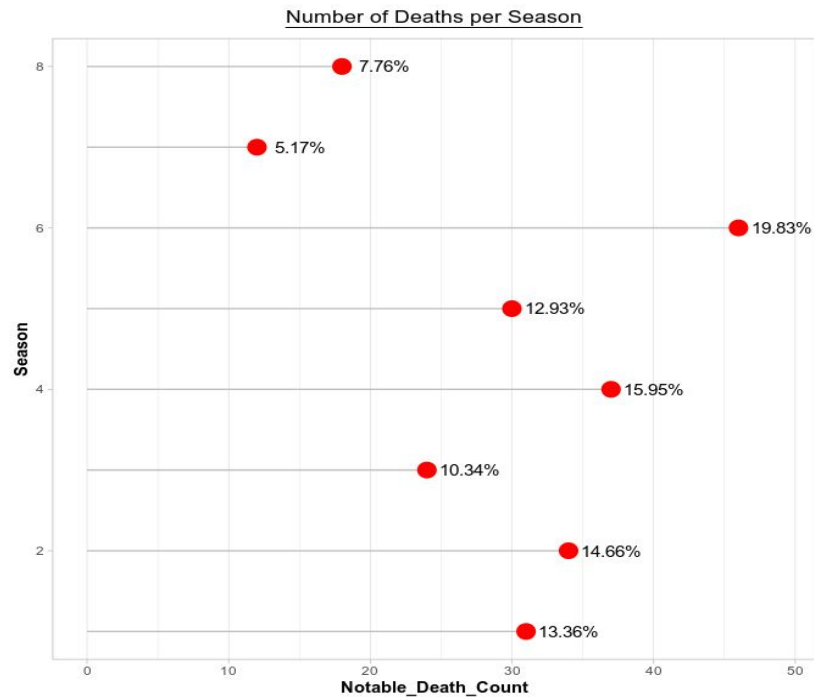


Figure 2: The average of deaths per season.

Also, we can verify by the graph in Figure 3 that each season has one death peak, this is a constant behaviour of the TV show makers to maintain its audience by breaking the expectation of their viewers.

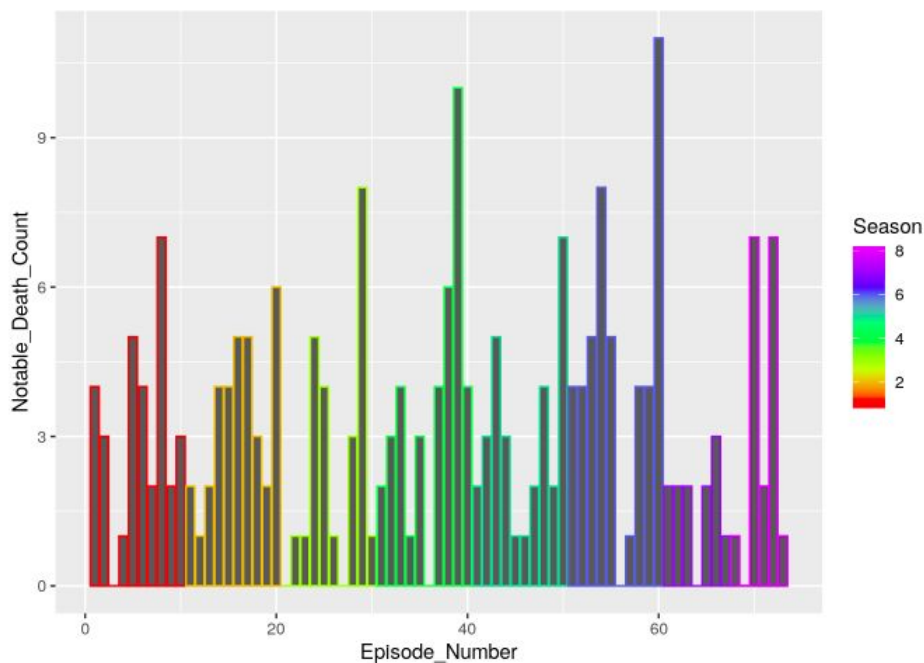


Figure 3: Number of character deaths per episode and per season.

Another thing we can notice is that these peaks on the number of deaths occur always at the end (last 2 episodes) of the seasons. This is a well-known plot strategy, called “cliffhanger”, to keep up the audiences for the next season to come.

Number of minutes per season

As it's possible to see from the graph in Figure 4, the number of minutes of each season of Game of Thrones was maintained almost constant (around 550 minutes per season)

until season 6. For the last 2 seasons, the runtime was reduced drastically. This could be related to the fact that for the last 2 seasons the level of the production got really high as well as the paycheck of the cast. By the end of the series, there were actor getting paid over \$500.000 per episode.

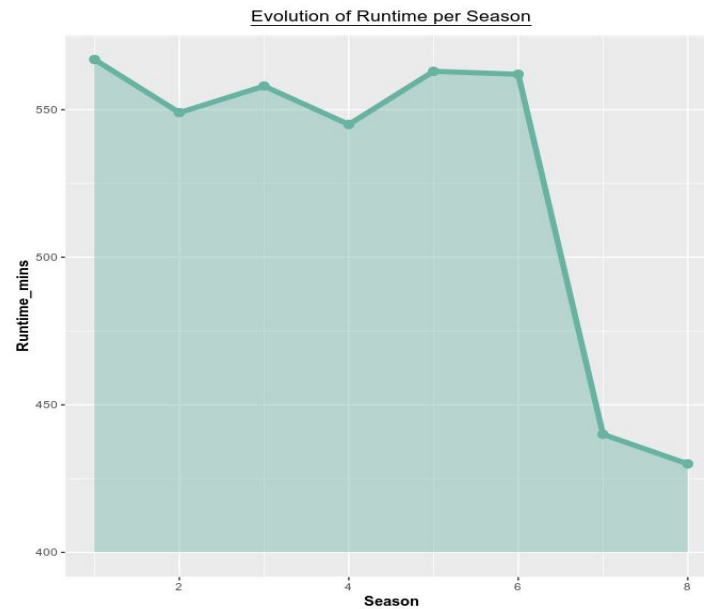


Figure 4: Runtime per season in minutes.

The opinion of the public

As we can see from the spyper graph in Figure 5, the TV show was really well rated in IMDB. The mean of the rating of all seasons was 8,89 (out of 10). All the seasons were rated between 8 and 10, except for the last season. This indicates that GoT was able to maintain a constant and high popularity throughout its duration.

The last season was the one rated the worst with 4 out of 10. And the best rated season was the 6th one, even though this season lost audience presenting more viewers at the first episode than the last. It's curious to notice also that season 6 contains the episode with the highest number of deaths.

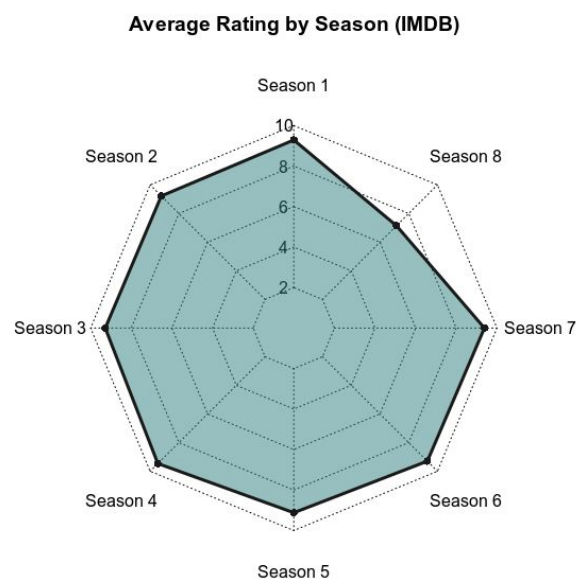


Figure 5: Average IMDB rating made per season.

Correlation between votes and ratings in IMDB

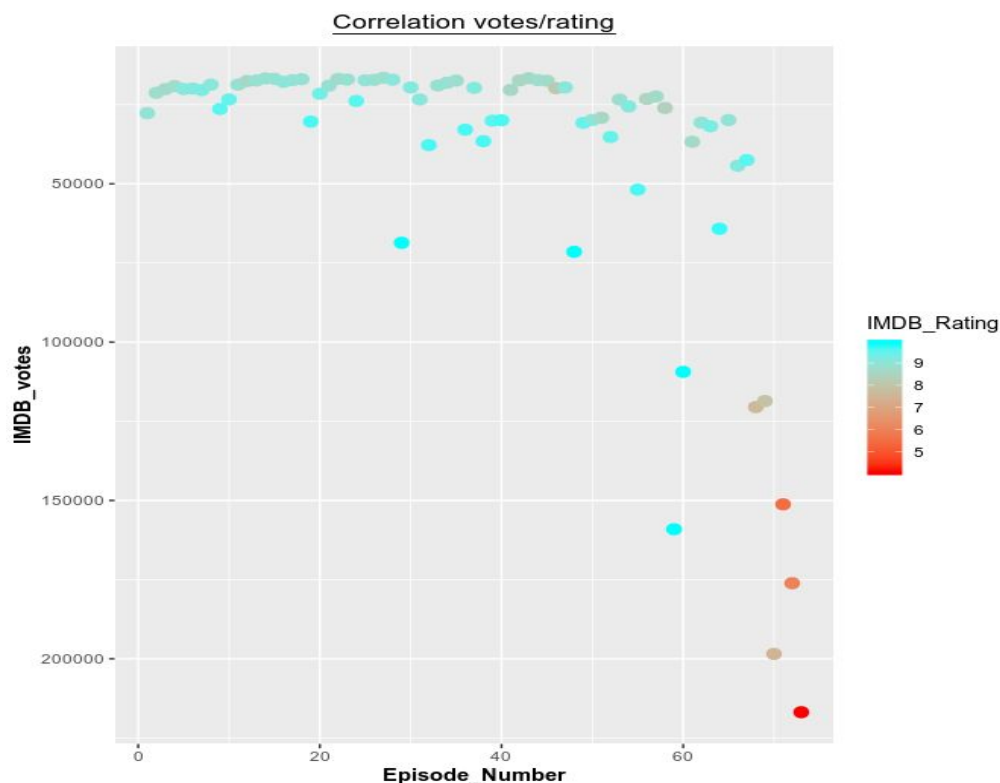


Figure 6: Correlation between IMDB votes and rating per season.

Also from the graph in Figure 6 we are able to verify that all the seasons received a rate greater than 8, for exception of the last season (season 8). This was the unique season that presented ratings lower than 8, which indicates a general disapproval of the audience on this last season. All the episodes of the last season were rated with less than 8, and also these are the only 6 episodes from all the 73 in the whole series that got a rate lower than 8.

It is also relevant to notice that the last season (8) was the most badly rated season of all and at the same time it was the one that received the biggest number of votes (216.879 users voted). This could indicate a tendency of the viewers to enter IMDB's website mostly when they are disappointed with an episode. When episodes are well rated, between 8 and 10, the number of people who entered the IMDB website to vote was distinguishably lower.

Conclusion

Through this report we are able to have a global vision on the progress of one of the most popular series in the world, we have identified TV show makers' strategies regarding the audience in order to increase it, and figured out the crucial points that make Game of Thrones one of the best series. Even if the last episode, which put an end for the whole series, was the most deceptive one.

We would like to thank Roberto PASCUA for his advice during the TDs.