Steven Garrett

DATA1201 Final Exam

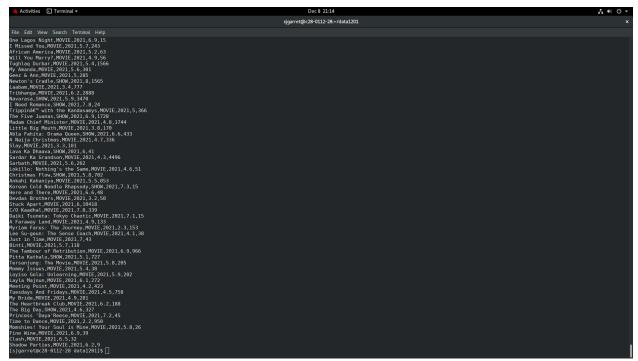
12/8/2023

Has Netflix Grown Oversaturated Over Time?

Question: My question stemmed from my own personal experience, and my perception of the media landscape around me. For the past few years I have noticed a general air of overwhelming indecision when using modern streaming platforms. From this inconvenience, I began to wonder if there was anything that could be causing this difference in experience. I then remembered a phenomenon that was experienced with a different type of media: Shovelware. This is the phenomenon in which low quality media is "shoveled" onto a platform to increase revenue. This was mainly an issue with Nintendo products from the 2010s, but Netflix could be suffering the same ailment. I wished to see if there was a general trend in the quality of programming on the site, and/or there was a drastic change in the amount of content put on Netflix over the years. After finding this question, it did not evolve significantly over the course of my research.

Getting the Data:





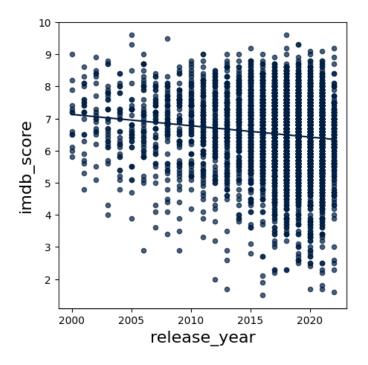
I received my data from Kaggle, which had information regarding items in Netflix's library and their associated IMDb scores. This data also held internal ids for these items, a very long description of the show/movie, age-rating, and runtime. All were

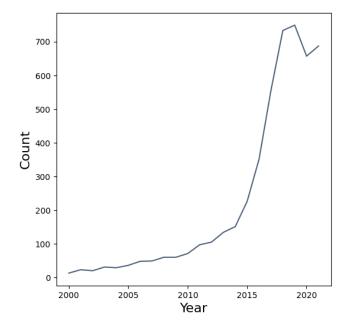
unnecessary for my purposes, so I cut them from my dataset using the Bash command line. The above picture is my data pre-simplified, and the lower post-simplified. The main visual noise being the description, removing it created a much more usable dataset. There are some ethical concerns from utilizing this database, as I am unsure if the internal ids were scraped from Netflix's internal database. However the data of which I am using is all publicly available, so there is limited real concern about the data within itself. However there is a concern with the lack of globalization within my dataset. As the data scraped is from American netflix and IMDb is primarily an American site, the data may not be replicable in areas outside of the United States. This is to be taken into consideration while coming to any conclusions.

Exploring the Data

From the data there seems to be a general downward trend in quality within Netflix's library overtime, and a drastic increase in content being flooded onto the platform. These are shown in these visualizations. The first graph regards the IMBd rating of movies versus the release year. As we can see the content on Netflix has had a slight drop in quality over time, and a drastic increase in the outliers with scores of 3 and below. Movies that have absolutely abysmal ratings. The next graph shows the amount of content on Netflix that was released each year from 2000 -

2022. This shows a drastic increase in new content being added to the site





especially after the year 2015. This shows that there is data to support the perception that Netflix has been flooded by content over the past few years. And

that they may be suffering from "shovelware" in the same vein as the media types that have suffered from it before.