

The Change in the Relevance of the Animation Industry
Project Proposal

Drake Monfregola, David Vazquez, and Kimmy Navarro

University of Washington Bothell

BIS 412: Advanced Data Visualization

Prof. Caleb Trujillo

Sunday, February 4, 2024

The Change in the Relevance of the Animation Industry

Project Proposal

Solving the issue of how the animation industry's relevance has shifted in the last 30 years is our main focus for this data dashboard. The sudden decline in quality and general discourse in the social sphere has made the animation industry seem not as prominent as it was in the '90s and early 2000s. Though this might seem like fewer people are interested in these films, businesses like Disney and Illumination studios are continuing to produce films with increasingly higher budgets. So has the relevance of these films decreased? Our research plans to answer the question: is the animation industry as culturally relevant as it was in the past and what are the major factors contributing to this potential shift in relevance.

Figuring out the relevance of these animated films will have to be decided by focusing on multiple variables. What we will define as relevant for this study is any film with a high critic/user score, movies with a high-profit return, and movies with a large number of seats filled during a film's box office run.

Our intent for this dashboard is to see what is “wrong” with the animation industry or if there even is anything wrong with it at all, this project would be a valuable resource for animation studios like the aforementioned Disney and Illumination, to help them figure out exactly what needs to be done in order to get their animation industry back to the forefront of culture like how they were twenty years ago.

Questions this data dashboard is planned to answer are: Is animation undervalued? Is it still popular to this day? And Why has the quality of animation changed over time? Visuals are planned to accompany each of these questions and all will be used to help our team evaluate our main research question. So in total, we plan on making a minimum of 3 interactive graphs for our dashboard, however, as the project progresses our team might add more visualizations if necessary to further explain our findings.

When the data dashboard is complete, the dashboard should speak for itself and give a clear answer to our initial question of whether the animation industry is as culturally relevant as it was in the past and what are the major factors contributing to this potential shift in relevance.

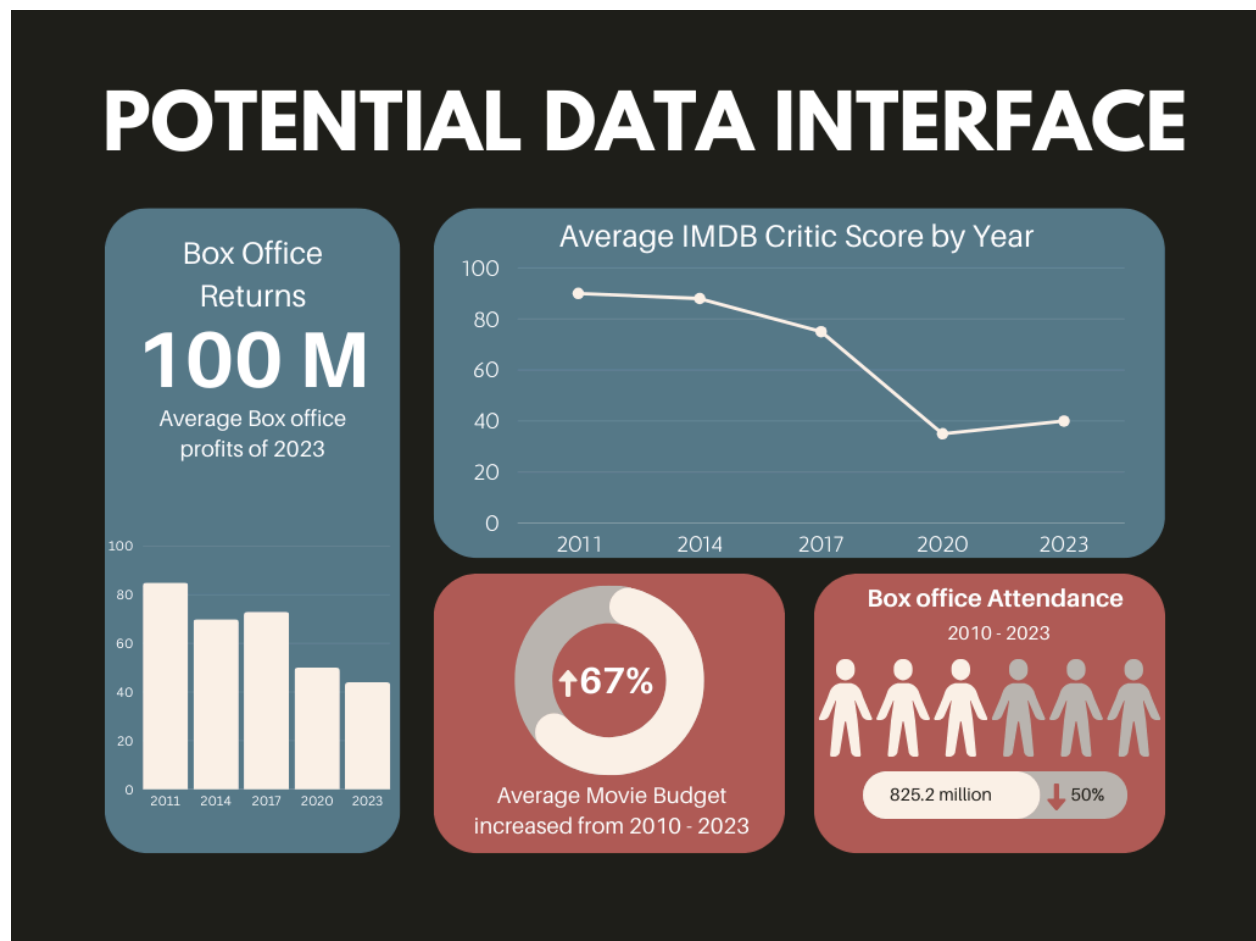


Figure 1: Mockup of potential data dashboard designed in Canva. Note that none of the numbers represented in this figure accurately represent any sort of data.

Timeline

- Week 1 Design:
 - Finding information and resources available to utilize for our research.
 - Importing the data sources that are necessary and important for our topic.
 - Drafting ideas for how our dashboard will be presented and what key ideas we want to use.
 - Stating our objectives of what we want to accomplish with our findings.
 - Potentially start a proposal repository on GitHub
- Week 2 Prototype:
 - Officially creating a repository on GitHub
 - Analyzing our gathered data by filtering through what variables we want to focus on.
 - Using these variables to create a table for our computed values such as averages, consensus, etc.
 - Utilizing the values from our table, we can create GGplot visualizations as well as using Tableau, for example.
- Week 3 Build:
 - Processing and finalizing our data and dashboards, making sure everything is well organized and aesthetically pleasing.
 - Summarizing our findings by grouping data sets, omitting anything unnecessary

- Adding textual information to our proposal for background information, context, and relevance—why it’s important.
- Week 4 Deploy:
 - Using feedback given by the professor as well as our peers, we will make keynotes on what we need to improve or change.
 - Utilizing our notes, we can make changes to our designs and finalize our proposal.
 - Testing our User Interface to make sure any changes need to be made; ensure it is running smoothly.
 - Finally, we will deploy our website containing the work we have done on this proposal.

Background

Our Original plan was to use the Statista Animation Industry dataset but after noticing its restrictions such as limited data and a paywall behind its sources, we decided to try out a few new datasets. The first was the “Top 1000 Movies” dataset published through Kaggle and the second being a large IMDb dataset that requires an API package as well as further technical skills to extract properly. We are still figuring out how to work with the packages on R Studio so for now, we will discuss the Kaggle dataset that we found.

The Kaggle dataset is a record of the Top 1000 movies up to the year 2020. Its original author is named Harshit Shankhdhar. The Metadata from the dataset is listed in columns and are summarized to have been brought from the IMDb site under a public domain license. There isn’t

any direct information from the author as to why the dataset was created however its inspirations are listed as the following: analyzing the gross and votes of a movie vs. its director and its cast, which actors prefer which film genres more, and finally which actors or combination of actors are bringing in the most votes and gross for a film.

The data is collected and organized in the following ways: Poster Link (Link or Image that IMDb is using), Series Title (Name of the movie), Release Year (Year the movie was released in theaters), Certificate (Certificate that the movie earned), Runtime (Total Runtime of the movie), Genre (Genre of the movie), IMDB Rating (Rating of the movie on the IMDb site), Overview (Short summary of the film), Meta score (Score the movie earned on metacritic.com), Director, Name of four Stars, Number of Votes, and the Gross. So far the only known limitations that it has is that there appear to be certain rows that are missing data in the certificate and meta-score columns and the dataset hasn't been updated since 2020.

On the Kaggle website that houses this dataset, the author has provided us a sample of what the data could potentially look like on a spreadsheet near the end of the page, giving us a general idea of what to expect from this dataset. The EDA also gives away how the data is shaped and how the code is built along with samples and images of what could result using certain code chunks. The only known use of the data is given on a page called "code next to the data card" which reveals the people who downloaded the data as well as what it was used for. It revealed that there are over 130 downloads thus far.

This dataset holds value for film studios because it provides insight into audience engagement as well as potential ways to increase profits by seeing where most of their pitfalls lie. These Animation studios will sometimes become out of touch with their current audience and even potential new audiences. Since our topic is on animated film, our plan with our own data

dashboard is to open the eyes of film studios, the Academy Awards as well as the American audience to keep the animation industry alive and profitable in the long run.

Works Cited

Statista Research Department. (2024, January 10). *Animation industry - statistics & facts*.

Statista.

<https://www-statista-com.offcampus.lib.washington.edu/topics/9725/animation-industry/#topicOverview>

Shankhdhar, H. (2021, February 1). *IMDB movies dataset*. Kaggle.

<https://www.kaggle.com/datasets/harshitshankhdhar/imdb-dataset-of-top-1000-movies-and-tv-shows?resource=download>