Netflix Dashboard:

Final Project Report

Aditya Pendyala

Michigan State University

CMSE 402: Data Visualizations Principles Techniques

Prof. Devin Silvia

Datasets

This project uses the Netflix dataset called "titles.csv" and an additional dataset called "countries.csv". While all the data comes from the "titles.csv" file, the "countries.csv" file helps map all the country codes to continents and names. The "titles.csv" file provides title names, movie or tv-show, age certification, IMDB scores, country of production, and year of release.

Goals and Arguments

Main goal of the project is to understand the content on Netflix using IMDB ratings. Some of the research questions are:

1. How have the IMDB ratings changed by genre over time?
2. How have the IMDB ratings for continents changed over time?
3. How do the IMDB ratings differ by continent?

There's a general trend of IMDB scores declining over the years, which could be due to the following reasons:

1. The quality of content has declined.
2. Maybe people just became more critical over time?
3. Probably averages are not a good metric.

Keeping those reasons in mind, it can be seen that the median IMDB score has an overall decline, regardless of genre and region. Continents Oceania, South America, and Africa have comparatively worse releases in 2022, as the IMDB averages don't even cross 7.0. Asia had the best average IMDB score all time, followed by North America. There's a slight jump in median IMDB scores during the years 2018 and 2019 which could be due to the pandemic.

Overall Experience

      Handling missing values was a challenge, the breaks in data over the years made the line plots initially non-comprehensible. Using interpolation of data, I was able to make the line plots continuous. Furthermore, the lack of support for HTML hosting in pptx, I was unable to make the geo-plot interactive. On the other hand, the clear trend in IMDB score decline made my project findings mostly conclusive!