Teesside University

Data Visualization

A Discovery to Your Best Movie Streaming Channel

By Group 2 Assessment Team (Friday Lab)

11/01/2023

1.0 Project Brief

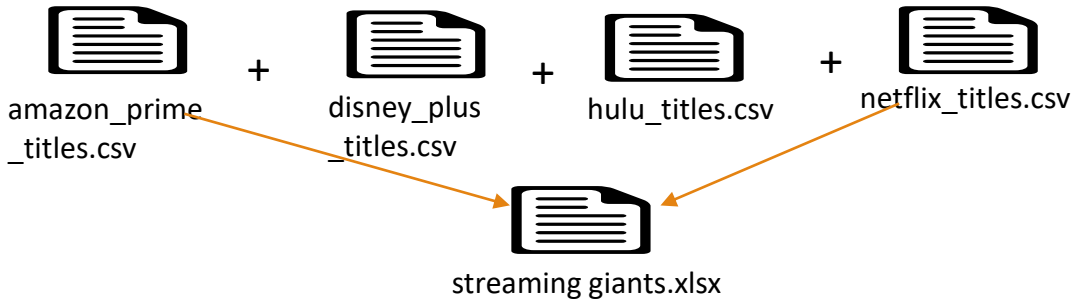
Over the past few years now, the world has evolved from traditional movie suggestions by word of mouth, advert posts on billboards, radio jingles, to a more advanced and modern way of communicating quality information that tends to inform, educate, entertain, or even compel, through digital means for the good interest of either its targeted audience, publishers, or both.

This could not have been possible without the dependency of historical datasets (data) to make powerful assumptions - as we now have in Machine Learning and AI-enabled applications; or to tell stories of what has happened in a business transaction of a certain period for better decision-making in future; or to make compelling discoveries and suggestions based on the personal choices we make in our daily lives – shopping, movies, vacations, etc.

Audience: Movie Lovers (General Public)

In this project, we have used infographic designs and analytics to create a pathway of discovering to the movie lovers, the best streaming channel between Amazon, Disney Plus, Hulu, and Netflix - that suits the most for different reasons. We make the discoveries while you decide what is best for you.

2.0 Description of Datasets

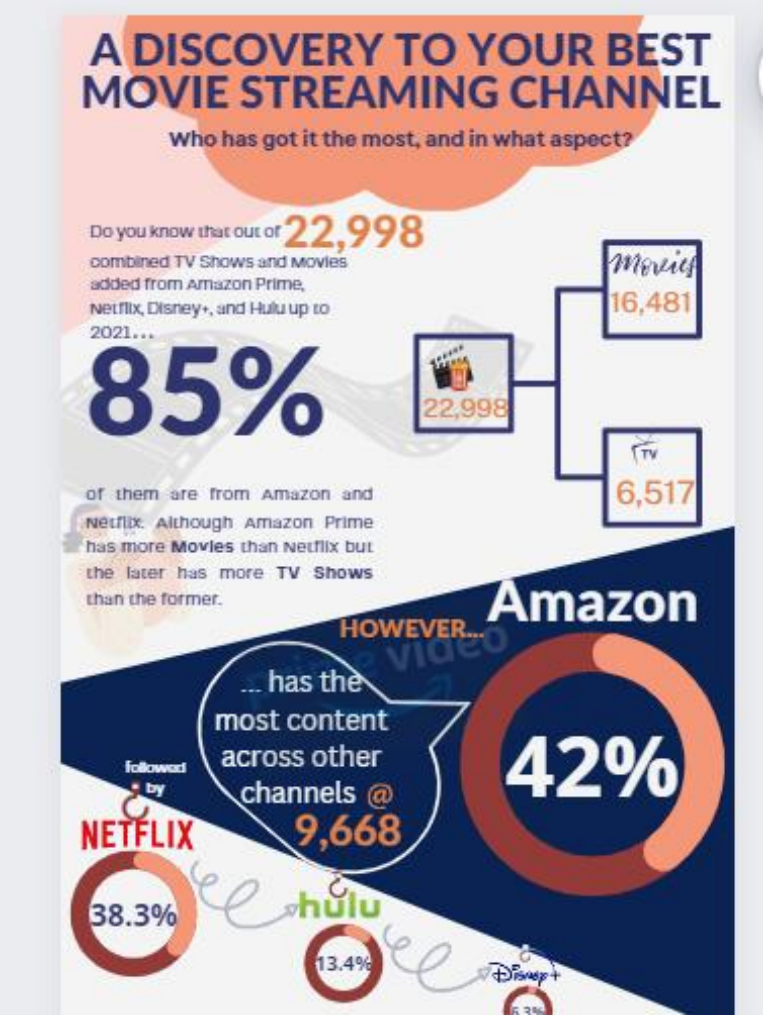


- Changed to recognizable datatypes
- Sorted the rows according to 'date added' column
- Re-categorized ratings columns into four viewing level as a column
- Split columns with more than one value into sub columns
- Removed nulls, errors, unwanted columns, and rows with conflicting values

We worked on the four datasets in csv – combined as streaming giants in xlsx format (Excel):

- Each from Kaggle Dataset Repository
- Datatype: Predominantly Qualitative Data (Categorical)
- Contains 13 columns (one extra column for Channel)
- 22,998 total unique rows
- Used the count of unique shows as the only measure

Dataset	#Unique Rows	#Columns	#Categorical Columns	#Numerical Columns
Amazon	9668	12	11	1
Disney+	1450	12	11	1
Hulu	3073	12	11	1
Netflix	8807	12	11	1



[Click here to see our live infographics]

3.0 Demo: Infographics

In the next slides, we will discuss the individual visualizations in our infographics:

- The charts we visualized will also be justified in the following slides
- Charts will be demonstrated in chronological order of the data story on the infographics.
- Each of the team members will give contributions to the data story – going from general, drilling down into the specific datapoints used in the charts.

4.0 Choice of Visualizations

S/N	Charts	Chart Types	Justification
1	Progress Ring Chart	Categorical – Comparison (Quantitative)	To show the level of content of each Channel against the overall using shaded part or arc in comparison with the unshaded arc
2	Doughnut Chart	Hierarchical - Part-to-Whole (Quantitative)	To demonstrate subscribers in nominal arrangement of different channels (constituent categories) by percentage of the total as seen reflecting in center.
3	Vertical Stacked Bar Chart	Hierarchical - Part-to-Whole (Quantitative)	To show the level in nominal arrangement of each tile within a bar – where each channel is mapped by the tiles, and each year in intervals by the bars
4	Pictogram	Categorical – Comparison (Quantitative)	To show average count of same-year released contents using shaded icons or symbols in comparison with the unshaded icons or symbols.
5	Radial Progress	Categorical – Comparison (Quantitative)	To show percentage of total contents added 1 year after they were released using shaded part or arc in comparison with the unshaded arc
6	Progress Bar Chart	Categorical – Comparison (Quantitative)	To show percentage of total contents added more than 1 year after they were released using shaded part or arc in comparison with the unshaded arc
7	Multi-series Line Chart	Time-Series – Trends (Quantitative)	To show the trends across the given period for the different categories (channels) in multi coloured lines
8	Vertical & Horizontal Clustered Bar	Categorical – Comparison (Quantitative)	To show the comparison in nominal arrangement of each bar within a cluster in the same nominal arrangement
9	Horizontal Bar Chart	Categorical – Comparison (Quantitative)	To show the comparison in nominal arrangement of each bar (channel) – where the lengths of each bars give the difference, and categorical label in full display

5.0 Critical Reflection

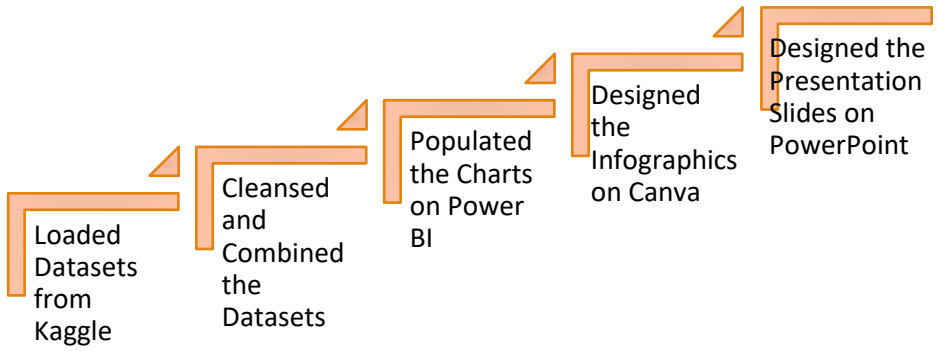
There were 11 categorical columns out of 12 in each of the dataset. This limited us from having enough measures to analyse apart from counts of unique shows.

The Netflix dataset had an issue when it was combined with the rest of the datasets; there were conflicting values – some categorical data entered in date columns

Another issue we faced was being limited by some of the few chart designs that are in Canva.

In the end, we overcame these challenges by: using count of shows as measure; removing rows conflicting the columns; and exploring the Charts in Canva.

6.0 Project Workplan



Saif Ur Rehman Bhutta

Team's Organizer and Convener

- Saif created the Team Group Calls and Trello Tasks
- He preprocessed the data in Excel and gave ideas on some of the charts to use

Canva

- Populated 1 Chart
- He had huge impact on all the designs

Adebiyi Ebenezer

Team's Secretary

- Ade handled all the action plans and documented every tasks
- He was instrumental with external resources and researches

Canva

- Populated 2 Charts
- He equally had huge impact on all the designs

Monica Chintala

Team's Active Contributor

- Monica's suggestions helped with the team's presentation slides

Canva

- Populated 2 Charts
- She was very useful in the colour palates

Chukwujekwu Joseph Ezema

Power BI and Infographics Designer

Chukwujekwu used Power BI to create early static visualisations to give an idea of how to populate the chart on the Infographics.

He equally used it to do the data pre-processing as demonstrated in 4.0

Canva

- Populated 3 Charts
- Designed the aesthetics and flow of the infographics.