

FIT 3179 – DATA VISUALISATION
ASSIGNMENT 1
BY: VIONNIE TAN
ID: 30092809
TUTOR: KADEK SATRIADI
03-P2

Monash University: Assessment Cover Sheet

Student name	Tan	Student's I.D. number	Vionnie 30092809
School/Campus	Clayton		
Unit name	FIT3179 Data visualisation S2 2020		
Lecturer's name	Prof Bernard Jenny	Tutor's name	Kadek Satriadi
Assignment name	Data Visualisation I Report	Group Assignment: No Note, each student must attach a coversheet	
Lab/Tute Class:	02-02	Lab/Tute Time:	Wed 14:00
Due date: 13-09-2020	Submit Date: 13/09/2020	Word Count:	696
		Extension granted	<input type="checkbox"/>

If an extension of work is granted, specify date and provide the signature of the lecturer/tutor. Alternatively, attach an email printout or handwritten and signed notice from your lecturer/tutor verifying an extension has been granted.

Extension granted until (date):/...../..... Signature of lecturer/tutor:

Late submissions policy	Days late	Penalty applied
Penalties apply to late submissions and may vary between faculties. Please refer to your faculty's late assessment policy for details.		

Patient/client confidentiality: Where a patient/client case study is undertaken a signed [Consent Form](#) must be obtained.

Intentional plagiarism or collusion amounts to cheating under Part 7 of the Monash University (Council) Regulations

Plagiarism: Plagiarism means to take and use another person's ideas and or manner of expressing them and to pass these off as one's own by failing to give appropriate acknowledgement. This includes material from any source, staff, students or the Internet - published and unpublished works.

Collusion: Collusion means unauthorised collaboration on assessable written, oral or practical work with another person. Where there are reasonable grounds for believing that intentional plagiarism or collusion has occurred, this will be reported to the Associate Dean (Education) or nominee, who may disallow the work concerned by prohibiting assessment or refer the matter to the Faculty Discipline Panel for a hearing.

Student Statement:

- I have read the university's Student Academic Integrity [Policy](#) and [Procedures](#)
- I understand the consequences of engaging in plagiarism and collusion as described in Part 7 of the Monash University (Council) [Regulations](#) (academic misconduct).
- I have taken proper care to safeguard this work and made all reasonable efforts to ensure it could not be copied.
- No part of this assignment has been previously submitted as part of another unit/course.
- I acknowledge and agree that the assessor of this assignment may, for the purposes of assessment, reproduce the assignment and:
 - i. provide it to another member of faculty and any external marker; and/or
 - ii. submit to a text matching/originality checking software; and/or
 - iii. submit it to a text matching/originality checking software which may then retain a copy of the assignment on its database for the purpose of future plagiarism checking.
- I certify that I have not plagiarised the work of others or participated in unauthorised collaboration or otherwise breached the academic integrity requirements in the Student Academic Integrity [Policy](#).

Date: 13/09/2020 Signature:  Vionnie *

Privacy Statement:

For information about how the University deals with your personal information go to <http://privacy.monash.edu.au/guidelines/collection-personal-information.html#enrol>

[Table of Contents](#)

INTRODUCTION	4
WHAT.....	4
WHY & HOW	4
BIBLIOGRAPHY	10
APPENDIX: 5 DESIGN SHEET METHODOLOGY.....	11

INTRODUCTION

This report aims to present the visualization based on the streaming service, Netflix. It focuses on the growth of Movies and TV Shows available in Netflix and answers questions such as “Is Netflix focusing more on Movies rather than TV Shows?”. This visualization is primarily aimed at individuals who enjoy watching Netflix and the insights obtained from this report could be useful in determining whether or not producing Movies would be more profitable than TV Shows. The report also demonstrates the use of Munzner’s What/Why/How framework in brief.

The Dataset for this visualization can be obtained from:

<https://www.kaggle.com/shivamb/netflix-shows>

The URL for this visualization can be obtained from this link:

[https://public.tableau.com/profile/vionnie.tan#!/vizhome/30092809_Assignment1V1_1/Netflix Visualisation?publish=yes](https://public.tableau.com/profile/vionnie.tan#!/vizhome/30092809_Assignment1V1_1/Netflix%20Visualisation?publish=yes)

WHAT

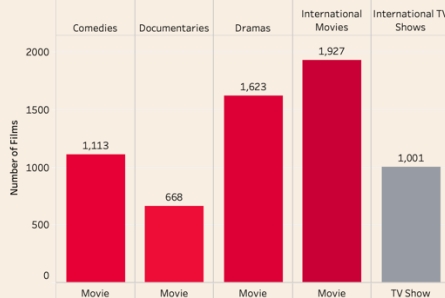
The author of this Netflix Dataset is Shivam Bansal and was obtained from Kaggle. It contained all the necessary fields such as Type, Country, Release Year, Rating and Genre. These fields would be the core of the visualization as they help in generating the graphs. However, one problem that arose was that some columns had several values separated by a comma in each row, which meant that some data cleansing had to be done via Python. The creation process of the visualization was straightforward and used the methods that were taught in the tutorials and took around 7 days to complete.

WHY & HOW



Movies or TV Shows?

Founded in 1997, Netflix is a renowned streaming service known for its variety of shows. In recent years, Netflix has expanded its catalogue to over 13,900 titles with most of these titles coming from their original content such as Movies and TV Shows (Cook, 2020). This visualisation explored the growth of these content over the past decade



Top 5 Genres of Movies and TV Shows in Netflix

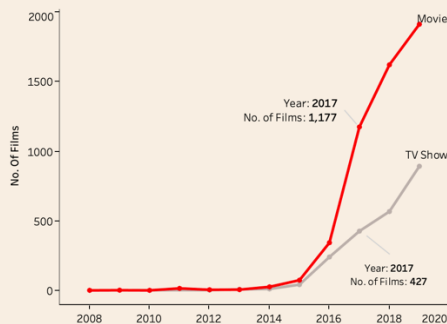
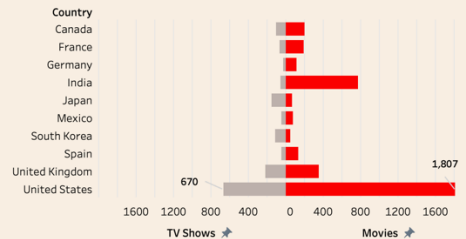
Netflix hosts a variety of genre in its catalogues. Its most prominent genre is **International Movies**, holding up a count of **1926 (25.6%)** of most genres. **Dramas** and **Comedies** are also some of the renowned genres, each having a count of **1113** and **668** respectively.

For TV Shows, its most popular genre is **International TV Shows** with a count of **1001 (13.3%)** proportion of total genres.

Top 10 Countries where Movies and TV Shows were produced

As of **2019**, Netflix has produced much of its TV Shows and Movies in the United States with both sectors reaching counts of **670** and **1800** respectively.

Other countries such as India (**775**) and United Kingdom (**335**) also contribute to much of Netflix's movie production. Majority of TV Shows were also produced in Japan (**151**) and United Kingdom (**220**).



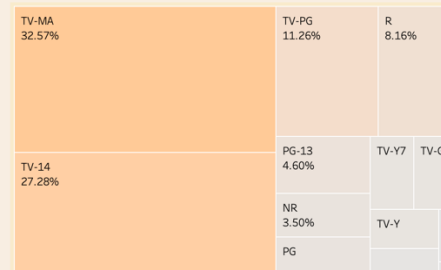
Release of Movies and TV Shows over the decade

Netflix's catalogue breakthrough came at **2017** when it released **3x** the number of movies compared to the previous year. However, its TV Show count that year was only a mere **427**, nearly **3x** off its movie count.

In **2018**, the TV Show release also jumped up by almost **double** compared to its previous year. However, the figure recorded is still overshadowed by the magnitude of movie releases Netflix made that year.

Ratings of Movies and TV Shows in Netflix

Netflix's catalogue of films has various of ratings designed to cater towards its subscribers. It's most common rating is **TV-MA** accumulating **32%** of Netflix's total catalogue. **TV-14** and **TV-PG** are also common to see as it holds a steady **27%** and **11%** total share respectively.



Dataset Obtained from <https://www.kaggle.com/shivamb/netflix-shows> Visualisation by: Vionnie Tan 30092809

Figure 1: Final Netflix Visualization

Top 5 Genres of Movies and TV Shows in Netflix

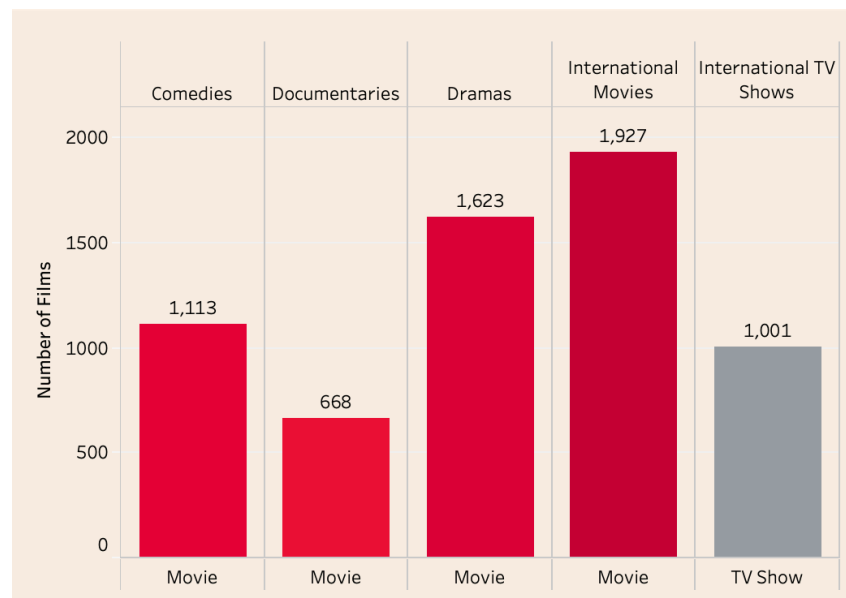


Figure 2: Bar Chart depicting various genres

The visualization idiom used here is a Bar Chart, which depicts the variety of genres available across Netflix Movies and TV Shows. The Bar Chart only emphasizes on the top 5 genres as increasing this number would only lead to broader genres and readers may not be able to distinguish between the vast choices of the genre as there would be a broader focus point. Each bar is labelled with its genre and type and provides insight towards the type of genres movies or tv shows tend to have.

Top 10 Countries where Movies and TV Shows were produced

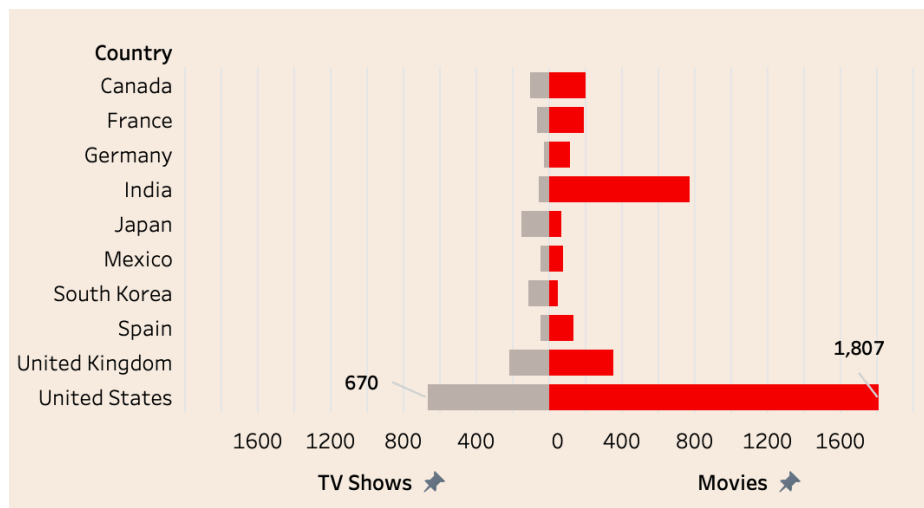


Figure 3: Diverging Bar Chart to differentiate content in various countries

The visualization idiom used here is a Diverging Bar Chart, with the axis representing the types – Movies and TV Shows. This graph indicates the count of movies and tv shows produced in each country. Here, only the top 10 is shown as increasing this number would only result in non-distinguishable graphs. Users will be able to see the growth in the production of movies and tv shows sectors in each country

Release of Movies and TV Shows over the decade

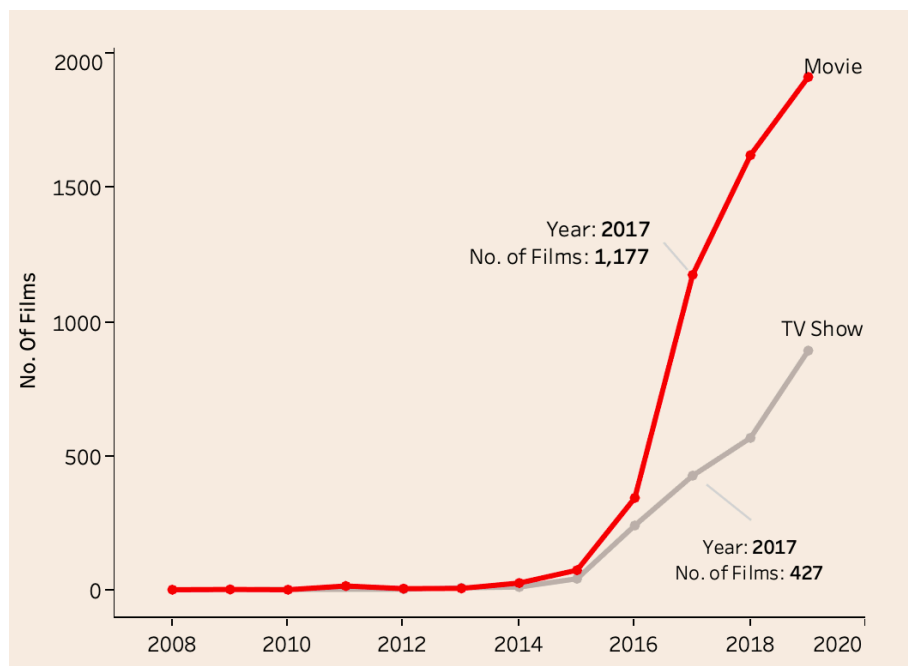


Figure 4: Line Chart depicting various content in Netflix

The visualization idiom used here is a Line Chart, with quantitative attributes representing the count and ordinal attributes representing the year. This graph is the core of the visualization, as it shows the growth of both movies and tv shows across all years. It uses points as its marks which aids in further solidifying the clear difference between movies and tv shows. The line chart is also interactive as it changes depending on which country, we are viewing in Figure 3. Gridlines are removed in this line chart to reduce chartjunk.

Ratings of Movies and TV Shows in Netflix

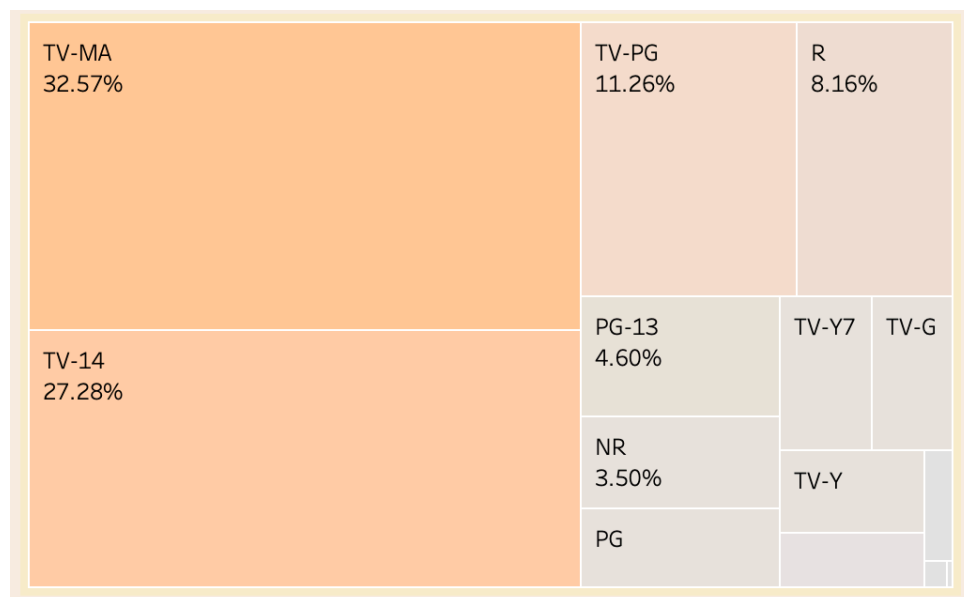


Figure 5: Tree Map depicting various ratings

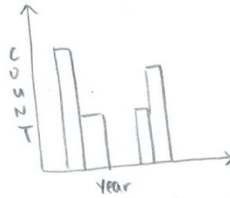
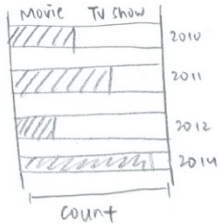
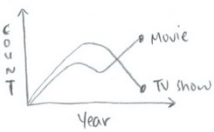
The visualization idiom used here is a Treemap. The Treemap was chosen because ratings didn't have many categories, which would help showcase a clear hierarchical structure (FusionCharts, 2020). Here, the quantitative attribute being focused is the percentage of the ratings – with larger areas depicting higher percentages. The difference in proportions of movies and tv shows could be viewed from the tooltip as it shows how much of the rating is occupied by movies or tv shows. The Treemap is also interactive as its proportion changes depending on which country, we are viewing. Gaining a clear hierarchical structure could inform the users regarding how well movies or tv shows perform with the specified rating.

BIBLIOGRAPHY

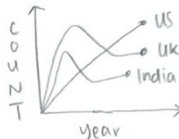
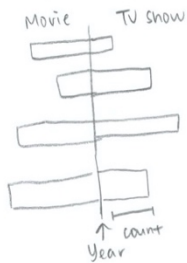
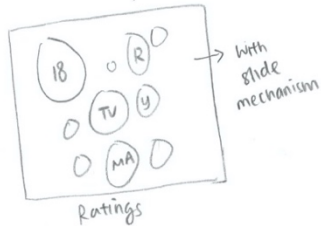
- Agarwal, A. (2018). *Tableau Tutorial 109 - How to Create Divergent Bar Chart in Tableau*. Retrieved from <https://www.youtube.com/watch?v=dxqQlvmH-e4>
- Bansal, S. (2019). *Netflix Movies and TV Shows | Kaggle*. Retrieved from Movies and TV Shows listings on Netflix: <https://www.kaggle.com/shivamb/netflix-shows>
- Cook, S. (2020, 7 29). *50+ Netflix statistics and facts stats that define the company's dominance [2020 version]*. Retrieved from comparitech: <https://www.comparitech.com/blog/vpn-privacy/netflix-statistics-facts-figures/#:~:text=Netflix%20has%20produced%20over%201%2C500,producing%20original%20content%20in%202013>
- Ester, L. (n.d.). *Interactive Visualization of Airbnb versus Rent in Melbourne, Victoria*. Retrieved from https://d3cgwrxphz0fqu.cloudfront.net/fb/cc/fbcc720bf80f302d6ff54da95ddac4c300f68c0c?response-content-disposition=inline%3Bfilename%3D%225DS_Melbourne_Airbnb_vs_Rental_Liliana_Ester.pdf%22&response-content-type=application%2Fpdf&Expires=1599948701&Signature=1599948701&Signature=1599948701&Signature=1599948701
- FusionCharts. (2020). *Treemap Chart - A Complete Guide | FusionCharts*. Retrieved from <https://www.fusioncharts.com/resources/chart-primers/treemap-chart>
- Jenny, B. (2020). *FIT3179 Data Visualisation - Week 03: Five Design Sheet Methodology*. Retrieved from https://d3cgwrxphz0fqu.cloudfront.net/3d/c8/3dc8e40b4025c7fb0504fdcaea62f850cca86aea?response-content-disposition=inline%3Bfilename%3D%22FIT3179_Week03%204%20Five%20Design%20Sheet%20Methodology.pdf%22&response-content-type=application%2Fpdf&Expires=159994
- Netflix. (2020). *Netflix | Brand Assets*. Retrieved from <https://brand.netflix.com/en/assets/>

APPENDIX: 5 DESIGN SHEET METHODOLOGY

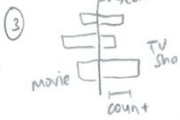
IDEAS



Movie	TV Show	
10%		2010
5%	25%	2011
2%		2012
2%		2013



FILTER



CATEGORIZE

Categorize into movies / TV shows

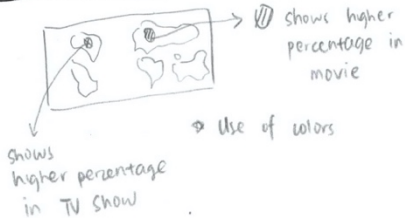
Compare the visualisation

Delve deeper into the countries

See if

Rating
Genre
Cast
Has any effect

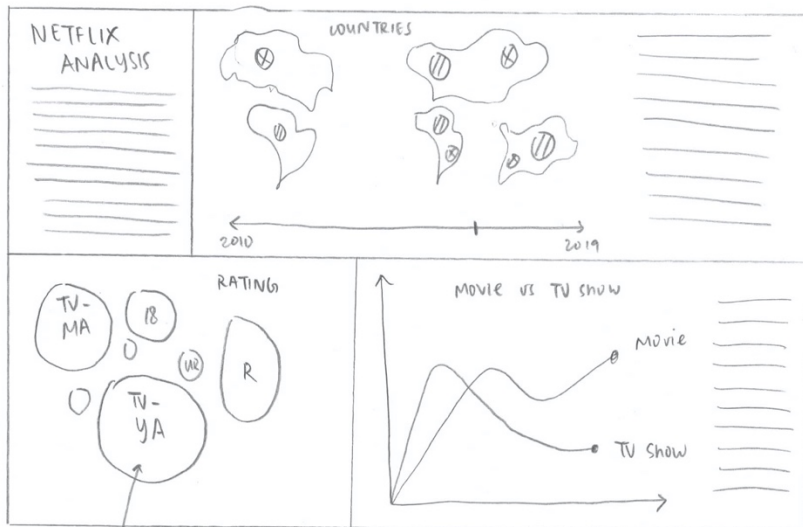
COMBINE & REFINE



QUESTION

- Does it answer whether or not Netflix should focus on movies / TV shows.
- Is the task at hand comparison between Type or between country
- Is the topic the main focus?

LAYOUT



Hover will show more information

Colour scheme: Black + Red
(same with Netflix)

Typography: Netflix font (?),
sans-serif

Title: Dashboard view

Author: Vionnie Tan

Date: 5/9/2020

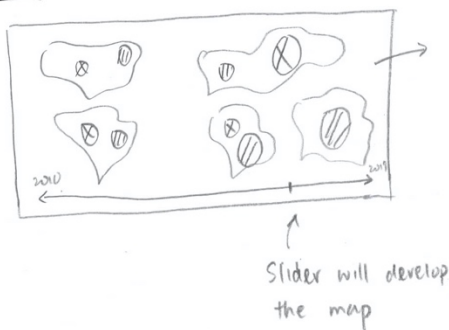
Sheet: 2

Task: Netflix visualization

OPERATIONS

- * Slider on map
Will show at which interval, which country has most TV / movies
- * Hover on each country on the map will show percentage of movie / TV show
- * Hover on the rating bubble will show information
- * Can include slider in rating to show different ratings per filter

Focus



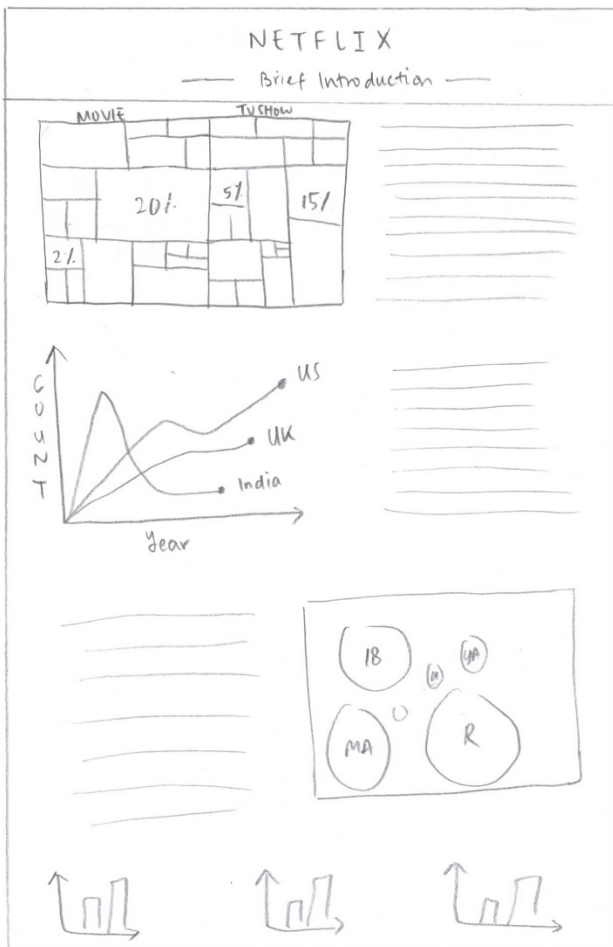
Shows the trend in movies / TV shows in a specific country

Slider will develop the map

Discussion

- * Enough information?
- * Are we presenting or exploring?
- * Comparing TV show / movie using a map might not be a good idea. — No focus
- * Information may be spread out in an unnatural way
- * Simple, gets the job done
- * Can we use something other than

LAYOUT



Small Bar Charts to show cast, actors,

Title: Alternative Dashboard
 Author: Vionnie Tan
 Date: 5/9/2020
 Sheet: 3
 Task: Netflix visualization

OPERATIONS

- ★ No sliding mechanism
- ★ Hover will show additional information
- ★ Line graph will have a certain focus,
- ★ Hover over will show the highest proportions of movies/TV shows per year
- ★ Interactivity between tree map and line chart. When press on the tree map, line chart will change
- ★ Bottom charts are also interactive and can view their progress over time

FOCUS



Hover will show information regarding how well/worse it did compared to the year prior

The treemap (drawn roughly) can be used to show the percentages of movies / TV shows per year

DISCUSSION

- ★ Treemap can show percentages well, but the separation of movie and TV show makes it hard to differentiate between them
- ★ Can show more visualisations. Overall, more clarity
- ★ Too similar to the example given at moodle

Layout

Netflix
movies vs TV shows

Genres
of movies and
TV shows

"The most
common
genres
are
---, ---, ---"

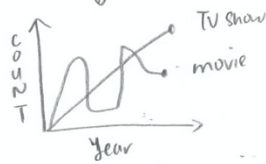


Ratings of different
movies & TV shows



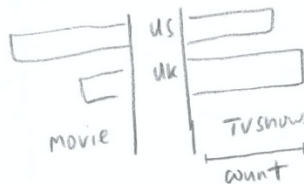
Whether
ratings were
more prominent
in movies / TV shows

Growth of movies
compared to TV
shows from
2008 - 2018



TV show raised
by 3X over past
3 years

The country that
produced the
most movies was
US, and TV shows
was UK



Title: Narrative visualisation idea

Author: Vionnie Tan

Date: 5/9/2020

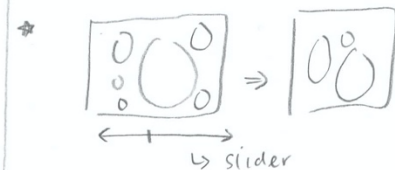
Sheet: 4

Task: Narrative visualisation

OPERATIONS

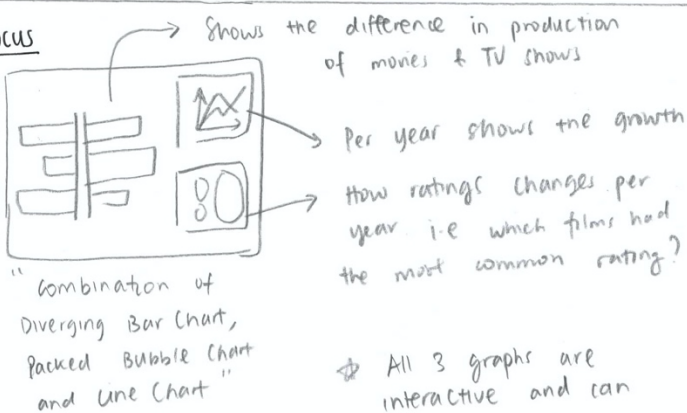
★ Filter / country will change
all proportions in the
graph

★ Annotations on line
chart



★

Focus



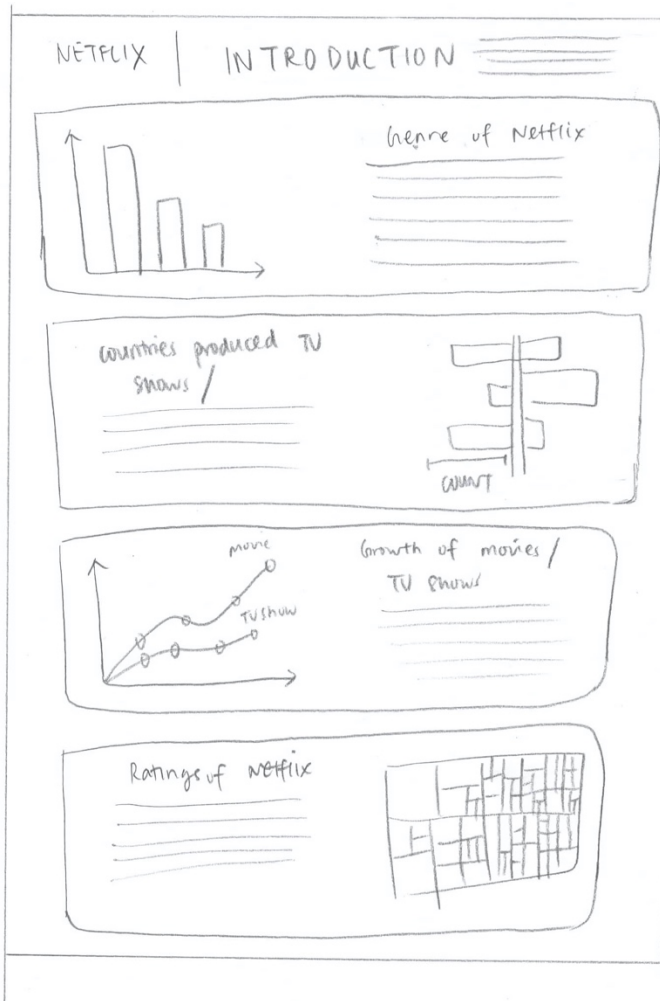
★ All 3 graphs are
interactive and can
change depending on
the country of focus

Discussion

★ Could the difference in
rating and genre have
more connection towards
the topic?

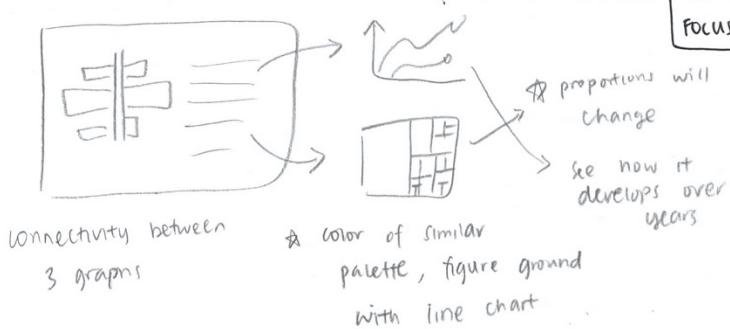
★ Could more interaction
be possible?

★ Are focus being given
to compare TV shows
& movies?



Conclusion

↳ Netflix should invest more in movies / TV shows as it is projected to generate a count of ~~more~~ more movies / TV shows



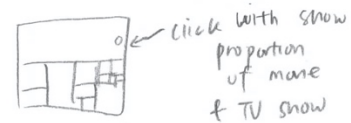
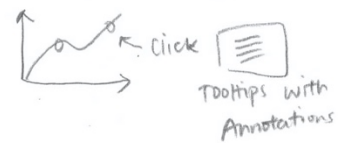
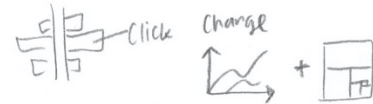
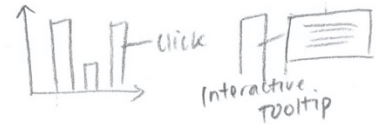
Title: Final Design sheet

Author: Vionnie Tan

Date: 5/9/2020

Sheet: 5

Task: Final Implementation Design



OPERATIONS

DETAIL

- ★ Data cleansing to be done with Python
- ★ Remove nulls and 2020 (incomplete data)
- ★ Construction Time: 7 Days

LAYOUT

FOCUS

- ★ See notes on how to make diverging bar chart
- ★ Dataset download as CSV