

Data Visualisation

ELEMENT 1: STREAMING PLATFORMS VISUALISATIONS

Athira Reghunath - W9561613

TEESSIDE UNIVERSITY | 2023/11

Table of Contents

Introduction	1
Quality Criteria	2
Portfolio	7
“Graph 1”	8
“Graph 2”	10
“Graph 3”	12
“Graph 4”	14
“Graph 5”	16
“Graph 6”	18
“Graph 7”	20
Summary	23
Reference	24

Introduction

It visualisation is the process of placing information into a visual framework, such as a map or graph, to make data easier for the human brain to understand and draw conclusions from. The main goal of data visualisation is to make it easier to see patterns, trends, and outliers in large data sets. Statistics graphics, information visualisation, and information graphics are all phrases that are sometimes used interchangeably. Data visualisation uses visual data to quickly and effectively provide information to all audiences. We need data visualisation since the human brain is not equipped to handle such a massive volume of unstructured, raw data and transform it into something meaningful and understandable. In order to convey data results and identify patterns and trends, graphs and charts are useful.

We are aware that the entertainment industry makes extensive use of online streaming services. With video streaming platforms, users can watch videos online without downloading them. TV shows, motion pictures, YouTube videos, and live streaming content are all examples of streaming media. Services like Netflix and Hulu have been quite successful because they can stream media to customers. This article looks at several publicly available graphs about streaming giants including Netflix, Disney+ Hotstar, Hulu, and Amazon Prime, among others. We are analysing several graphs and grading them using the quality standards that every graph in a data visualisation should meet.

1) Avoid Data Distortion:

Data distortions are one of the main mistakes made in data visualisation. If someone isn't confident with designing, this is an easy mistake to make. When visual components with diverse shapes are resized incorrectly, distortions happen. A distorted image might lead to false relationships between the data values and a decrease in the focal point's information density. Data distortion in visualisation occurs when the data is enhanced with more decorative characteristics (Aesthetics). These decorative features fall short of establishing and illuminating the relationship between data values.

Reasoning

It is agreed among Kirk (2012) and Cairo (2014) on the importance of visualization ethics. Cairo emphasizes that Information alone isn't as crucial as a visual's effectiveness. A false or unclear graphic is unethical since it leads to audience confusion. the problems with deceptive infographics lead to incorrect value interpretation. Visualization has been discovered to include a crucial aesthetic component. But occasionally, these decorative elements are deceptive. the significance of presenting viewers with data rather than depending solely on a beautiful visualisation. So, it is important to avoid data distortion.

2) Avoid Incorrect Colour Use:

The incorrect use of colour in data visualisation is among the most frequent mistakes. The vast colour pallet can cause designers to use either too many or too few colours. Using any colour should be done so with intention. In data visualisation, colour should serve the data story and aid in amplifying or clarifying information for audiences rather than just being attractive.

Reasoning

It is agreed that we should be making certain that the distinctions between the meanings associated with the colours are readable. Utilize colour consistently, be aware of the tone it conveys, be aware of how emotion is evoked by colour, be aware of cultural colour connotations, and design with colour-blind people in mind. Leverage colour selectively. This should be intentional because it can be tempting to let the tool do it for you. These all colour aspects are very important for data visualisation.

3) Avoid Improper use of 3D Graphics:

Occlusion is a significant issue with 3D visualisation. When some marks make it difficult to see other marks, this is known as occlusion. As a result, always cause some areas of a 3D representation to conceal other areas, which lowers efficiency. Users

might need to perform a non-trivial viewpoint shift in order to access concealed things. This might not even find all the things in the visualisation due to inter-object occlusion. Use 3D as necessary to avoid occlusion. Instead, use different colours or sizes to highlight different data points.

Reasoning

Danielle Szafir (2019) and Kirk (2012) both provided examples of data blocking brought on by 3D visualisations. In 3D bar charts and pie charts, Kirk (2012) displayed the occlusion in great detail. The use of 3D results in unclear or cluttered charts. Additionally, Szafir (2019) emphasises the value of just using 3D to visualise molecular surfaces or architectural surfaces.

4) High Data Ink Ratio:

The ink used to depict data on a graph is called "Data Ink". Edward Tufte, a well-known author in the field of data visualisation, developed the idea of the data-ink ratio. Clearly, the data-ink ratio urges chart producers to consider if every piece in the chart is relevant to the message of the chart. The benefits of this approach are; Clear message, saving time and Saving space.

Reasoning

According to Tufte, effective graphical representations increase data-ink and minimise non-data-ink.

He proposed the data-ink ratio, which is determined by subtracting 1 from the percentage of the graph that can be erased without losing data.

He suggests the following 5 guidelines for data ink:

1. Above all else show data.
2. Maximize the data-ink ratio.
3. Erase non-data-ink.
4. Erase redundant data-ink.
5. Revise and edit

Tufte evaluates these ideas in a variety of cases before coming up with a variety of innovative designs that significantly increase the intelligibility of the graphs.

5) Law of Similarity:

According to the concept of similarity, we group things together when they seem to be similar to one another. We also have a tendency to assume that they serve the same purpose. The human brain will classify an object or a specific piece of data or information into the same group or cluster if it is represented using the same shape, colour, or size. The law of similarity is coming under Gestalt principles.

Reasoning

It is agreed because according to Tufte (2001), Few (2012), and Kirk (2019), items are perceived as belonging to the same group if they have similar shapes or colours. This is supported by the concept of similarity. Using the similarity principle, it may be simpler to determine which categories the exhibited data falls into. Qualities like colour, shape, orientation, and form can influence similarity, which aids in classifying objects appropriately.

6) Missing or Misleading Axes:

Always include the X-Axis (horizontal) and the Y-Axis(vertical) of a graph to prevent producing a poor data display. Without an axis, it is simple to interpret data incorrectly, leading to erroneous conclusions about the message.

Additionally, since audiences will be operating under the assumption that both axes start at zero, both should be set to that value. The distance between data points should match the numbers they represent.

Reasoning

It is agreed because the graph should have both the x and y axis for the exact analysis. Visualizations of data convey stories. The narrative that a figure conveys

can be significantly impacted by very subtle decisions, such as the range of the axes in a bar chart or line graph. When seeing data visualisations, you should consider if the graph was created to tell a story that truly reflects the underlying facts or whether the designer was attempting to convince you of a particular viewpoint.

7) Avoid Too Much Data:

When visualisations contain an excessive amount of data, the information becomes overwhelming. It is not the right way to present a graph. It should be attractive and informative. But the excessive data will never give an attractive and easy-understanding graph. So, this should be avoided.

Reasoning

It is agreed because when we are analysing a graph, if it's congested with data it will be totally disturbing and we don't like it also. If the graph only shows the relevant data it's more powerful and catchable.

Portfolio

Provide a list of the graphs chosen related to the streaming giants such as Netflix, Disney+ Hotstar, Hulu and Amazon Prime. These graphs are taken from various resources. There are 8 visualisations related to streaming platforms. Above I already set 8 quality criteria to evaluate each of these graphs.

Graph 1: Number of countries the world's biggest streaming services are available in(as of Jan 2022)

Graph 2: Disney+ Hotstar ARPU has been calculated by Moneycontrol based on the data shared by Disney in their earnings report

Graph 3: Monthly subscription price and latest price increases of selected video streaming services in the U.S(in U.S dollars)

Graph 4: Total Netflix viewership from January 2017 to October 2018

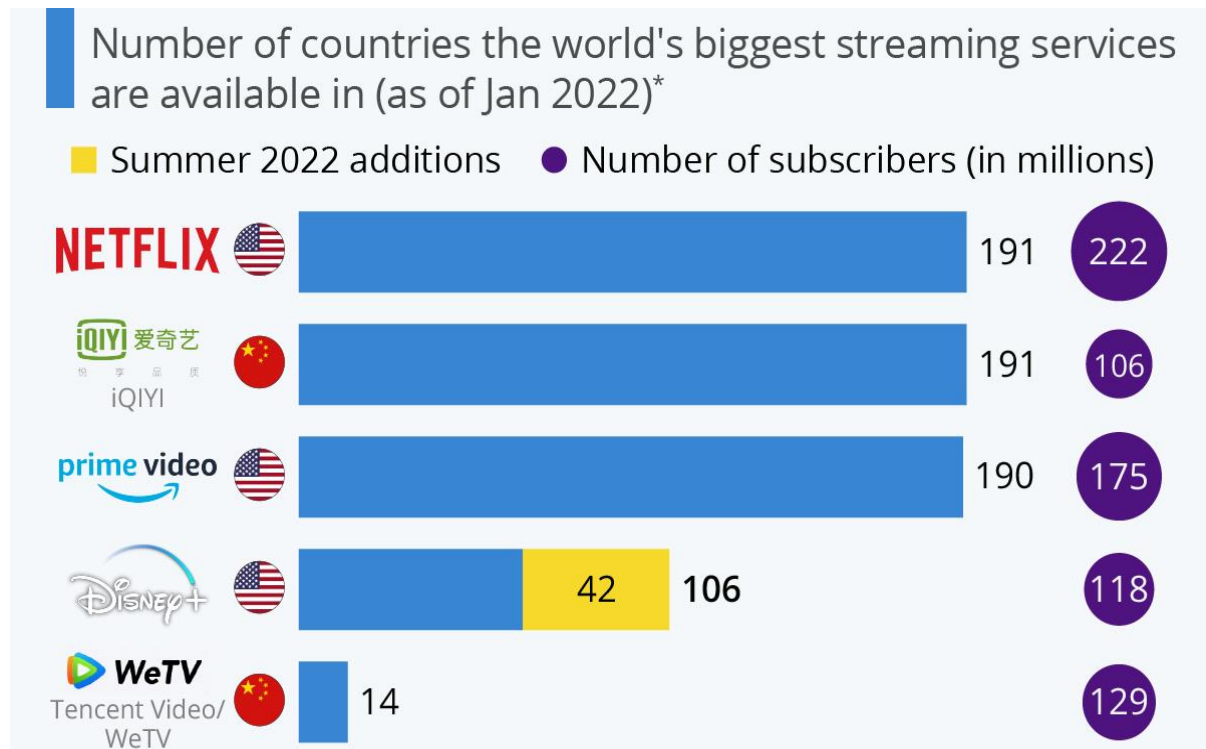
Graph 5: Netflix cumulative quarterly paid subscriber additions worldwide, by year

Graph 6: Rise in Revenues – Disney+ Hotstar

Graph 7: Hulu subscribers from December 2012 to April 2018

Graph 1: Number of countries the world's biggest streaming services are available in (as of Jan 2022)

The horizontal bar graph is showing the number of countries with the world's biggest streaming services available (as of January 2022). Here Number of countries showing in blue colour Summer 2022 additions data is shown in yellow colour and the number of subscribers in million is shown in purple colour. The country flag and the streaming platforms logos are included in the y-axis. This graph is taken from Statista, this site is a continually updated, an online resource for current statistical data on a broad range of topics. The audience is data analysts, data scientists, people working in film or television industries like producers etc.



Graph 1: Number of countries the world's biggest streaming services are available in (as of Jan 2022)

Source: <https://www.statista.com/chart/26719/number-of-countries-the-worlds-biggest-streaming-services-are-available-in/>

Evaluation

1) Avoid Data Distortion:

- Decorative features in the graph do not make any distortion
- No deviations from the focal point
- Summer 2022 additions are clearly shown

2) Avoid Incorrect Colour Use:

- Feels like so many colours are used because of the flags and logos
- Otherwise the colour selection is not making any issue

3) Avoid Improper use of 3D Graphics:

- 3D is not used in this graph
- No kind of occlusion seen
- Nothing is overlapped and all bars are fully visible

4) High Data Ink Ratio:

- Background of the graph is light in colour, so the audience can notice the values clearly
- Images in the graphs create a low data-ink ratio

5) Law of Similarity:

- This graph follows the law of similarity. Here the number of countries is seen in blue colour, summer 2022 additions are displayed in yellow colour and the number of subscribers is in purple colour.
- Here different aspects are displayed in different colours as a group

6) Missing or Misleading Axes:

- No axis is showing in the graph
- In the y-axis more than one data showing

7) Avoid Too Much Data:

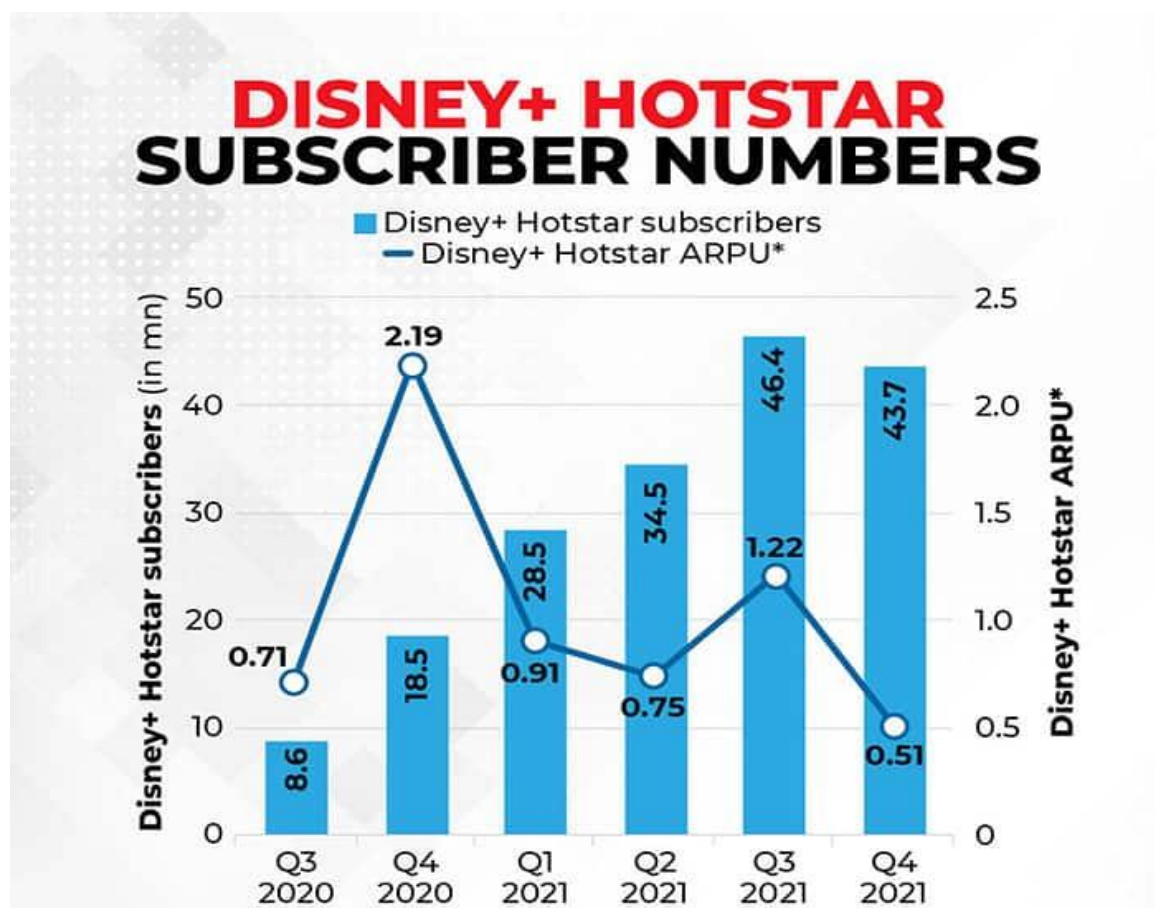
- Feels like too much data are added to a single graph
- The graph looks more complex because of so much data

Recommendations

- The major problem with this graph is the absence of the x and y axis
- The other issue is it includes too much data
- The legends are not seen uniformly, one legend is taller than the other two. It's not even noticed as a legend
- It will be more understandable if use a vertical column chart

Graph 2: Disney+ Hotstar ARPU has been calculated by Moneycontrol based on the data shared by Disney in their earnings report

This bar line graph is showing Disney+ Hotstar ARPU has been calculated by Moneycontrol based on the data shared by Disney in their earnings report. Here the x axis shows the year and the quarter y-axis shows the number of subscribers and Disney+ Hotstar ARPU. Here the blue bars represent the subscribers and the navy-blue line denotes the ARPU. The graph is taken from Moneycontrol and it has so many graphs like this. The audience is the business owners related to Disney plus Hotstar or stakeholders etc.



Graph 2: Disney+ Hotstar ARPU has been calculated by Moneycontrol based on the data shared by Disney in their earnings report

Source: <https://www.moneycontrol.com/news/business/disney-hotstar-sees-first-ever-quarterly-decline-in-its-subscriber-base-7707491.html>

Evaluation

1) Avoid Data Distortion:

- Decorative features in the graph do not make any distortion
- Data for the line graph and the bar graph is a little bit confusing some places it displays side by side

2) Avoid Incorrect Colour Use:

- Colours are properly used
- If the line graph is in a different colour tone it will be noticed separately

3) Avoid Improper use of 3D Graphics:

- 3D is not used in this graph
- No kind of occlusion seen
- Nothing is overlapped and all bars are fully visible

4) High Data Ink Ratio:

- Background of the graph is light in colour, so the audience can notice the values clearly
- Data in the line graph and bar graph is confusing sometimes

5) Law of Similarity:

- Here the Hotstar subscribers are shown in the blue bar graph and the Hotstar ARPU is shown in a navy-blue line graph.
- Data are arranged in a group and following the law of similarity

6) Missing or Misleading Axes:

- X and Y axis are showing correctly
- But in the Y direction different values displayed on both sides

7) Avoid Too Much Data:

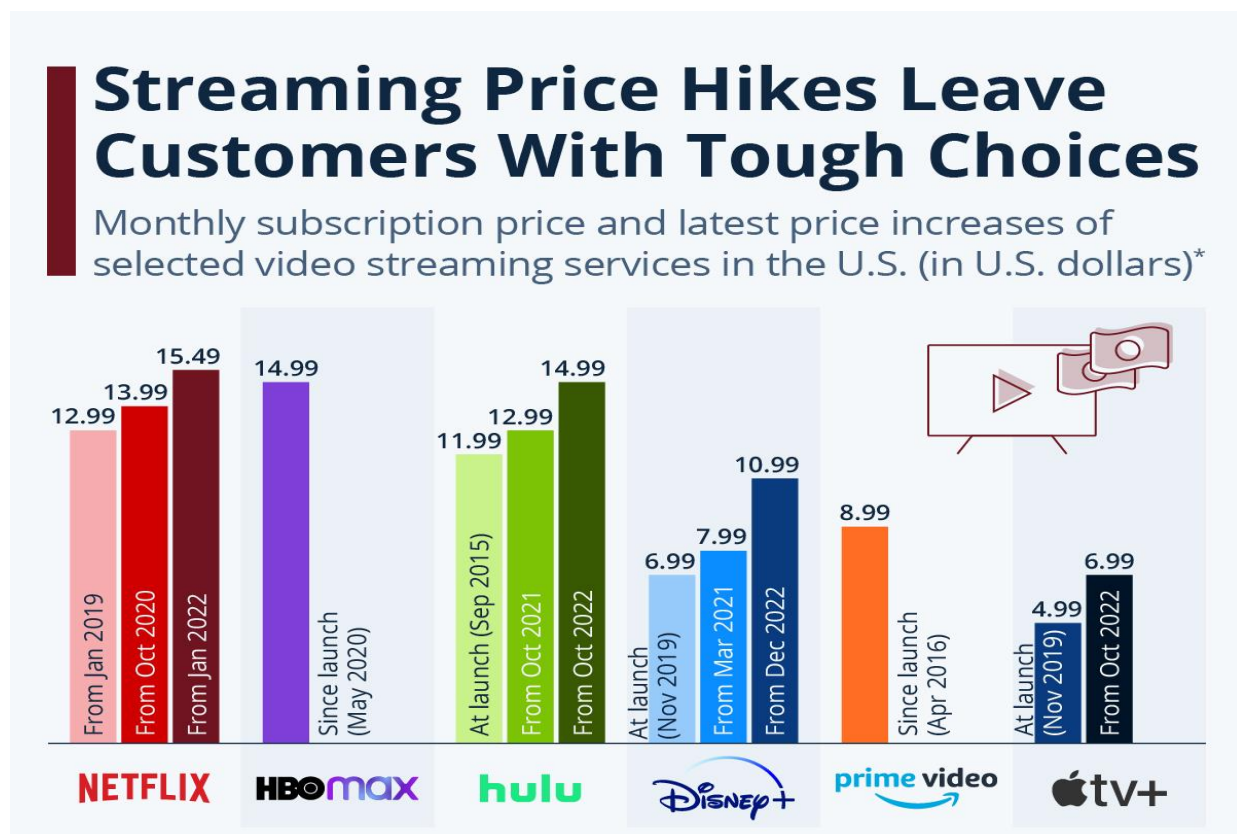
- Data are a little bit more than the usual graphs

Recommendations

- The major problem with this graph is the different values on the y-axis. It makes some confusion
- The other issue is it includes the inclusion of some more data than the usual
- If the colours are in different shades it will be noticeable

Graph 3: Monthly subscription price and latest price increases of selected video streaming services in the U.S

This clustered bar chart is displaying monthly subscription prices and the latest price increases of selected video streaming services in the United States only. Here different streaming services in different years are represented with the different shades of a single colour for each service. This graph is taken from Statista, this site is a continually updated, an online resource for current statistical data on a broad range of topics. The audience is data analysts, data scientists, and people working in film or television industries like producers etc.



Graph 3: Monthly subscription price and latest price increases of selected video streaming services in the U.S(in U.S dollars)

Source: <https://www.statista.com/chart/27983/prices-of-video-streaming-subscriptions-in-the-us/>

Evaluation

1) Avoid Data Distortion:

- Decorative features in the graph do not make any distortion
- All values are clearly visible and do not making any confusion

2) Avoid Incorrect Colour Use:

- Here year wise data for each streaming platform are seen in the different shades of a single colour. That's good but the differentiation between the two services is confusing

3) Avoid Improper use of 3D Graphics:

- 3D is not used in this graph
- No kind of occlusion seen
- Nothing is overlapped and all bars are fully visible

4) High Data Ink Ratio:

- Background of the graph is light in colour, so the audience can notice the values clearly
- Data in the bars feels a little congested

5) Law of Similarity:

- Here each streaming platform's yearly data is considered as a group and given different shades of a single colour for those bars
- Some bars are showing the data outside and most of the other is displayed inside

6) Missing or Misleading Axes:

- X axis is denoted by the logo and the y-axis is absent
- No y-axis and label for the y-axis

7) Avoid Too Much Data:

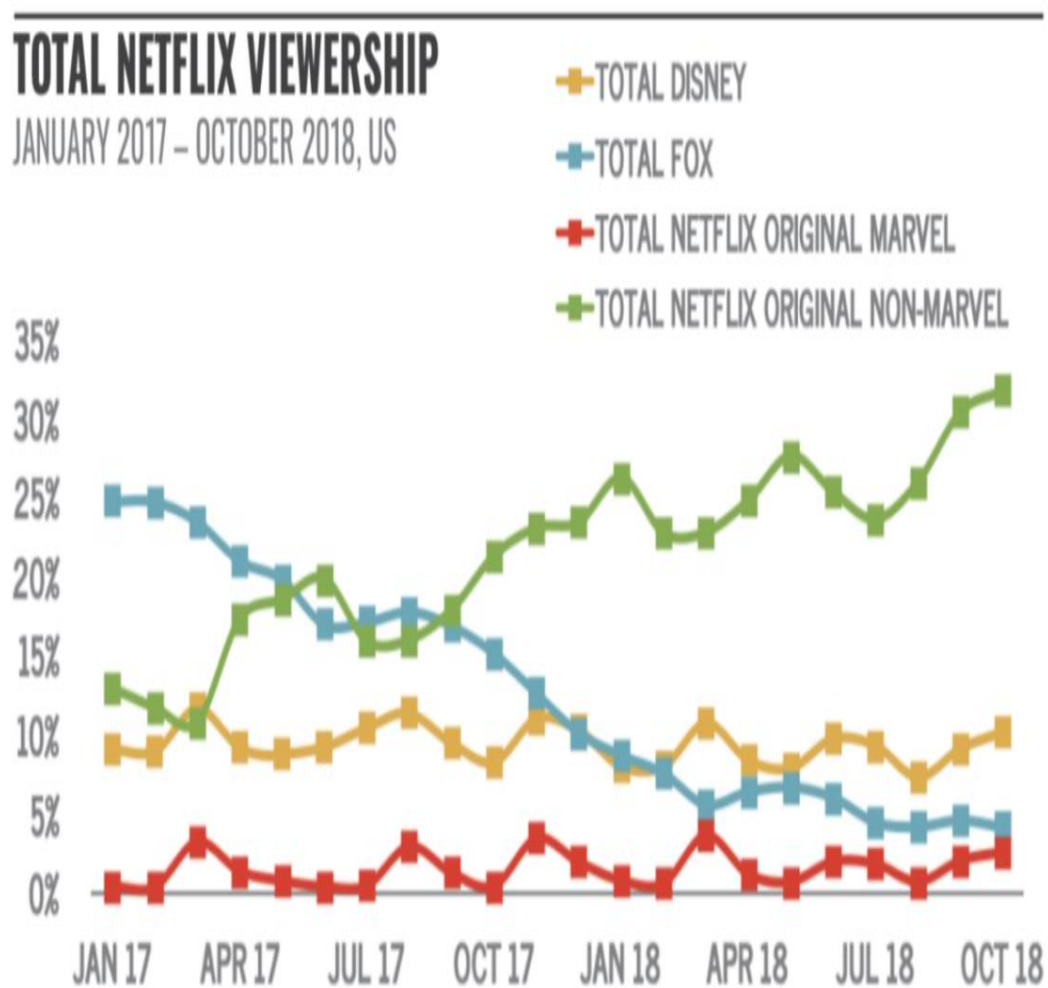
- Data is enough but the data in the bars seems to be congested

Recommendations

- The major problem with this graph is the absence y-axis
- The other issue is the separation is not feeling between the two streaming service bars
- It can be visualised using a stacked bar chart with legends
- Data is clearly understood and every bar has its own data on the top

Graph 4: Total Netflix viewership from January 2017 to October 2018

This dotted line chart is displaying the total Netflix viewership from January 2017-October 2018 in the United States. Here each line with different colours are showing the viewership percentage of Disney, Fox, Netflix original marvel and Netflix original non-marvel. This graph is taken from bgr.com, it's a journalistic site. The audiences are data analysts, data scientists, and people working in film or television industries like producers etc.



Graph 4: Total Netflix viewership from January 2017 to October 2018

Source: <https://bgr.com/entertainment/netflix-originals-tv-shows-leaving-friends/>

Evaluation

1) Avoid Data Distortion:

- No data distortion is shown in the graph
- All values are clearly visible and do not making any confusion

2) Avoid Incorrect Colour Use:

- Colours are properly used here

3) Avoid Improper use of 3D Graphics:

- 3D is not used in this graph
- No kind of occlusion seen
- Nothing is overlapped and all bars are fully visible

4) High Data Ink Ratio:

- Background of the graph is light in colour, so the audience can notice the values clearly

5) Law of Similarity:

- Here each streaming service's yearly data as a group and each of them showing in different colours

6) Missing or Misleading Axes:

- No label for both the axis
- X-axis displays the data with three months gap
- Y axis is not in the graph only values are there

7) Avoid Too Much Data:

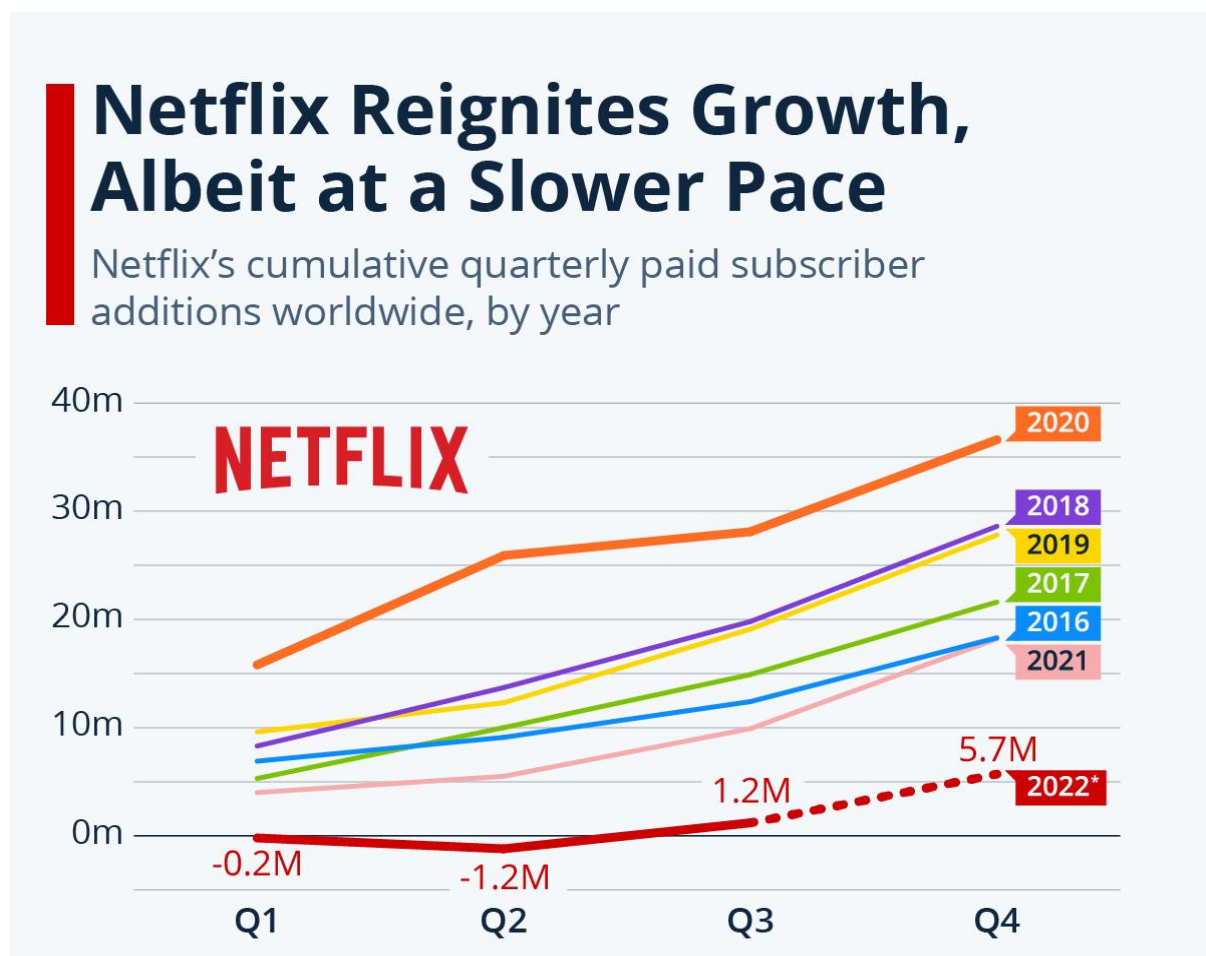
- Data is enough and easily understood but the points in the line have no data values

Recommendations

- The major problem with this graph is no y-axis, no grid lines and no value for each point in the line
- The title is misleading because the graph is not only showing the viewership of Netflix
- Instead of a line graph we can use an area graph or scatter plot here

Graph 5: Netflix cumulative quarterly paid subscriber additions worldwide, by year

This line chart is displaying the Netflix cumulative quarterly paid subscriber additions worldwide, by year. Here each line with different colours are showing the quarterly paid subscriber for each year. This graph is taken from Statista, this site is a continually updated, an online resource for current statistical data on a broad range of topics. The audience is data analysts, data scientists, and people working in film or television industries like producers, Netflix shareholders etc.



Graph 5: Netflix cumulative quarterly paid subscriber additions worldwide, by year

Source: <https://www.statista.com/chart/21465/global-paid-net-subscriber-additions-by-netflix/>

Evaluation

1) Avoid Data Distortion:

- Focal point of the graph is confusing
- Plot seems to be incomplete
- Netflix logo is placed inside the graph it's not a good practice

2) Avoid Incorrect Colour Use:

- Colours are properly used here

3) Avoid Improper use of 3D Graphics:

- 3D is not used in this graph
- No kind of occlusion seen

4) High Data Ink Ratio:

- Background of the graph is light in colour, so the audience can notice the values clearly
- It will be better if add a legend for each year instead of giving labels inside the graph

5) Law of Similarity:

- Here quarterly data for each year is considered as a group and each year in separate colours

6) Missing or Misleading Axes:

- No label for both the axis
- X-axis displays the data with three months gap
- Y axis is not showing -10m in the graph

7) Avoid Too Much Data:

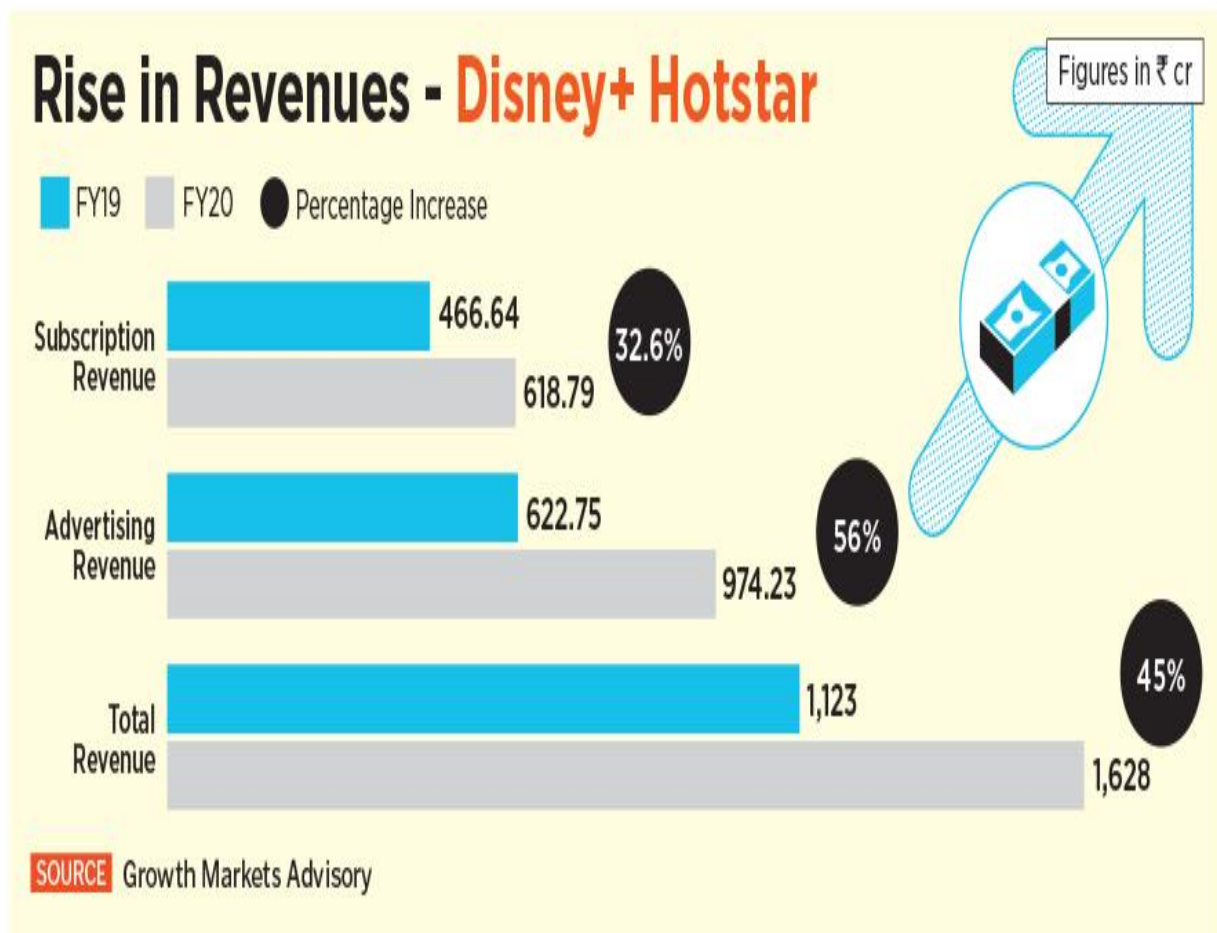
- Data is enough and easily understood but the points in the line have no data showing

Recommendations

- The major problem with this graph is y-axis not showing -10m but the axis is in the graph
- Only 2022 shows each data values
- It will be better to add a year as legends
- Instead of a line graph we can use an area graph or clustered bar chart here

Graph 6: Rise in Revenues – Disney+ Hotstar

This clustered bar chart is displaying the Disney+ Hotstar Rise in Revenue and the percentage of increase. Here year is given as legends. Blue bars are showing data for the financial year 2019 data and white bars are showing the financial year 2020 data. The black rounds are displaying the increasing percentage of Subscription Revenue, Advertising Revenue and Total Revenue. This graph is taken from Forbes India. The audiences are data analysts, data scientists, and people working in film or television industries like producers, Disney+ Hotstar shareholders etc.



Graph 6: Rise in Revenues – Disney+ Hotstar

Source: <https://www.forbesindia.com/article/take-one-big-story-of-the-day/over-600-million-indians-have-a-smartphone-star-india-wants-disney+-hotstar-to-be-on-every-one-of-them/69495/1>

Evaluation

1) Avoid Data Distortion:

- Decorative features in the graph do not make any distortion
- All values are clearly visible and do not making any confusion

2) Avoid Incorrect Colour Use:

- Colours are properly used here
- Light colours are used and it used in a meaningful way

3) Avoid Improper use of 3D Graphics:

- 3D is not used in this graph
- No kind of occlusion seen
- Nothing is overlapped and all bars are fully visible

4) High Data Ink Ratio:

- Background of the graph is light in colour, so the audience can notice the values clearly
- Picture on the side of the graph shows that Figures in crores increasing

5) Law of Similarity:

- Here each yearly data is considered as a group and the percentage of growth is considered as another group
- Both groups are displayed in three different colours so follow Law similarity by using the same colour for similar data

6) Missing or Misleading Axes:

- X axis has no data or label
- Y axis has no label but revenue is the data
- No grid lines are seen

7) Avoid Too Much Data:

- Data is enough and easy to understand

Recommendations

- The major problem with this graph is no y-axis and no grid lines
- The y-axis value has no units but it showing in the side picture. It's not easily noticeable
- Instead clustered bar chart we can use clustered column chart or stacked bar chart and revenue as the legend

Graph 7: Rise in Revenues – Disney+ Hotstar

This is a bar chart displaying data from Hulu subscribers from December 2012 to April 2018. Here each bar is showing the number of subscribers in millions. This graph is taken from the business insider site. It's a very simple graph with only one colour used. The audience is data analysts, data scientists, and people working in film or television industries like producers, Hulu shareholders etc.

Hulu Subscribers

In millions

hulu



Graph 7: Hulu subscribers from December 2012 to April 2018

Source <https://www.businessinsider.com/hulu-announces-new-offerings-to-lure-advertisers-2018-5?r=US&IR=T>

Evaluation

1) Avoid Data Distortion:

- No decorative features are used in this graph
- All values are clearly visible and no distortions

2) Avoid Incorrect Colour Use:

- Colours are properly used here
- The bar has given the same colour as the Hulu logo

3) Avoid Improper use of 3D Graphics:

- 3D is not used in this graph
- No kind of occlusion seen

4) High Data Ink Ratio:

- Background of the graph is light in colour, so the audience can notice the values clearly
- Only one colour is given to the graph as the data demands

5) Law of Similarity:

- Here data is less so the subscriber group is showing in a similar colour. So, the graph follows the law of similarity

6) Missing or Misleading Axes:

- No label for both the axis
- X-axis displays the data without an equal interval
- No Y axis is showing

7) Avoid Too Much Data:

- Data is enough and easily understood

Recommendations

- The major problem with this graph is no y-axis and no grid lines
- The Y-axis data showing the top of the graph has no units. It's showing under the title as in millions but it's not easily noticeable.
- The X-axis values are not shown in equal intervals and the graph seems to be a very plane.

Summary

Data Visualization aims to communicate heavy data in an efficient, easy-to-understand way by using graphic representations. Visualising something allows our brains to internalize and reflect on it. Seeing trends and making sense of data is much easier and more effective than trying to figure out what's going on in a lengthy document that would take a lot of attention to comprehend. As such, a data Visualisation is a form of storytelling aimed at assisting us in making informed decisions. So, the correct visualisation should follow some criteria.

Here in this report, we are taken seven criteria for reviewing seven different graphs. These graphs are taken from different sources and all are related to the streaming platforms. Data visualisation is very much important in the field of steaming platforms because the statistics of the number of subscribers, movies and Tv shows related details etc for each platform are really important for business analysis and making improvements according to it. Here I included different types of visualisation related to streaming platforms. Each one is sharing different information and the visualisations are also different. The graphs made me realize, not the style, colour, or alignment of a graph, but the data it represents that is most important. While aesthetics is good for attracting audiences, they can distort data if used inappropriately so use those according to the criteria.

Data visualisation should establish two aspects of the data: Show data relationships that are too complex to express in words, and make it easier for audience members to comprehend and explore the information presented.

References

Andy Kirk – 2016 - Data visualisation: A handbook for data driven design

Nathan Yau - 2013 - Data points: visualization that means something (Online)

<https://venngage.com/blog/misleading-graphs/> - 5 Ways Writers Use Misleading
Graphs to Manipulate You

<https://www.forbesindia.com/article>

<https://www.statista.com/chart/>

<https://www.businessinsider.com>

<https://bgr.com/entertainment/>

<https://www.moneycontrol.com/>