

DVP REPORT
BY SUBBULAKSHMI NATARAJAN

Manga & Anime



Table of Contents

<i>Introduction</i>	2
<i>Design Process</i>	3
<i>Implementation</i>	9
<i>Conclusion</i>	14
<i>Bibliography</i>	15
<i>Appendix - Five sheets design</i>	16

Project Title: Manga and Anime Data Visualisation

Introduction

Recently, both children and adults have started to watch and read more anime and manga series. It's said that the characters and storylines of manga and anime differ significantly from those of Marvel, DC, and other novels. Many young people read manga and then watch the corresponding anime to visualise the story they've read. The popularity of anime saw a significant increase in 2019 [1,2], and after the COVID-19 pandemic, it became even more popular. Additionally, this project will explore the distinction between the genre preference of the users watching anime and manga, popularity as well as score. This is why I've chosen to analyse the popular anime and manga adaptations that are currently captivating young teens and adults.

This project aims to create interactive visualisations to understand the distributions and comparisons of various features in manga and anime data. The primary focus is on analysing the distribution of anime and manga types, comparing scores and popularity rankings, examining demographics, and exploring correlations within the datasets.

The main questions addressed are as follows:

1. What is the distribution of various types of anime and manga?
2. How do the scores and popularity rankings compare between anime and manga?
3. What are the demographics for anime and manga?

These visualisations aim to provide clear and interactive insights derived from the cleaned datasets, enabling users to explore and analyse trends effectively.

Design Process

Since we are talking about manga and anime insights we will be making some insights about it. So when the users use my dashboard he should be able to choose the best out of the dashboard when they are using. So, for that I have divided the design section into five parts

- Distribution of Anime and Manga
- Score and Popularity Comparison
- Source to Type of Anime and Manga
- Demographic Analysis

Each part has its own purpose and has been designed meticulously after using five design sheet which is in appendix section of the report the following are the choice of the graphs which i have written in the following section:

❖ Distribution Plot - Anime and Manga

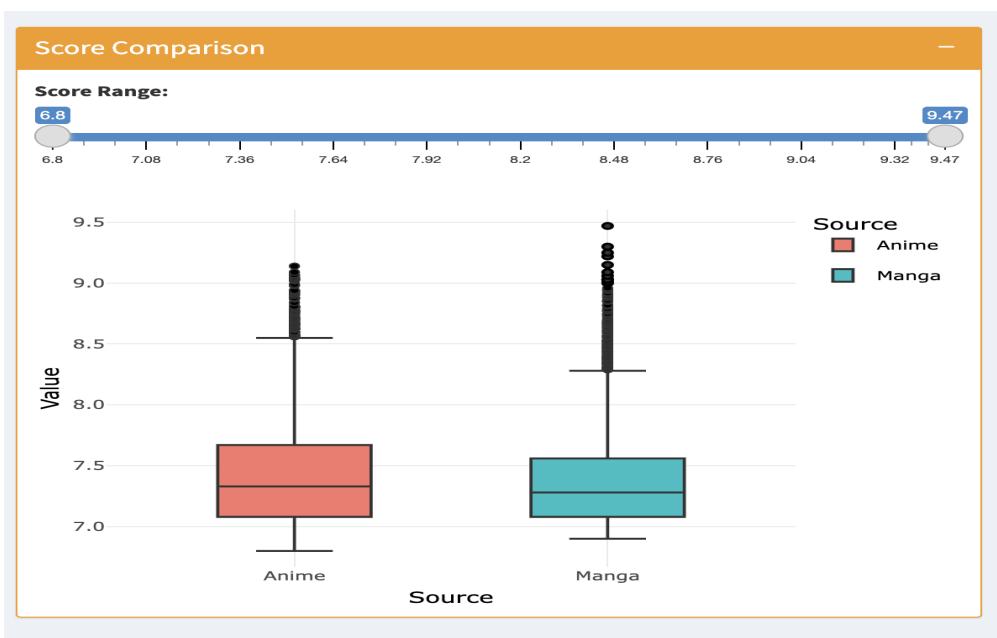
The bar chart named "Distribution of Anime Types" shows how various forms of anime are spread based on their count. The x-axis displays several types of anime, such as movies, music, ONA (Original Net Animation), OVA (Original Video Animation), specials, and TV series, while the y-axis shows how many of each category exist. The blue bars show these numbers, with the height proportional to the number of anime of that category. TV series have the largest count, making them the most popular anime type. Movies and OVAs have high counts, suggesting their popularity. Special anime have a reasonable count, whereas ONAs are less common but still more prevalent than music-related anime, which has the lowest count.

The bar chart named "Distribution of Manga Types" shows how various types of manga are distributed based on their count. The x-axis includes several sorts of manga, such as doujinshi, light novels, manga, manhua, manhwa, novels, and one-shots, while the y-axis displays how many of each type exist. The green bars show these counts, with the height proportional to the number of manga of that category. Manga has the highest count, making it the most popular manga genre by a wide margin. Light novels and manhwa have moderate counts, but doujinshi, manhua, novels, and one-shots have low counts. This chart efficiently demonstrates the overwhelming popularity of traditional manga in comparison to other manga formats.



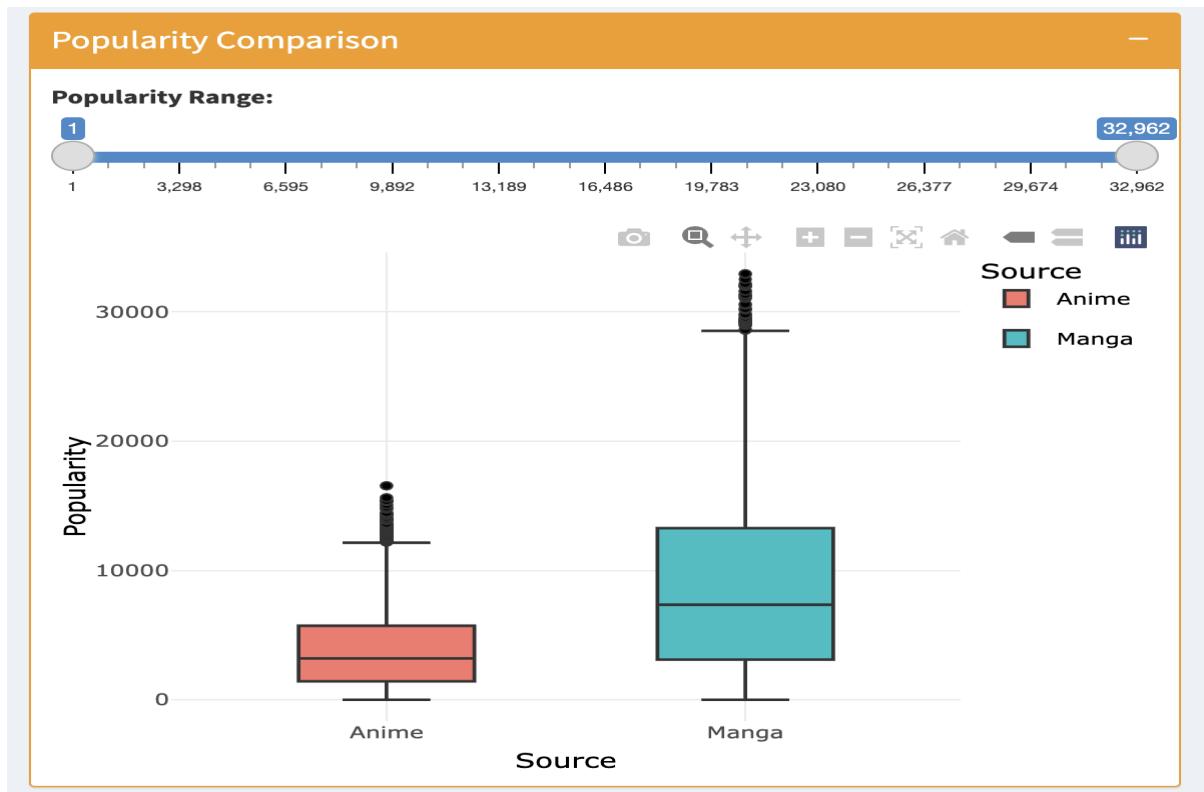
❖ Score Comparison

The graphic shows a "Score Comparison" box plot within a web-based dashboard for analysing and visualising anime and manga data. This graphic contrasts the distribution of scores for anime and manga, which are represented by red and cyan boxes, respectively. The graphic shows numerous essential data for each medium, including the minimum, first quartile, median, third quartile, and maximum. Notably, the anime scores appear to be more diverse, with a cluster of outliers stretching to the higher end. In contrast, manga scores are more compactly distributed but still vary. This box plot is interactive, with a slider above the chart that allows viewers to select the score range they want to see, tailoring the data displayed based on their personal interests. This functionality improves user engagement by allowing for a dynamic investigation of how ratings vary across anime and manga, revealing trends and distributions within the data.



❖ Popularity Comparison Box Plot

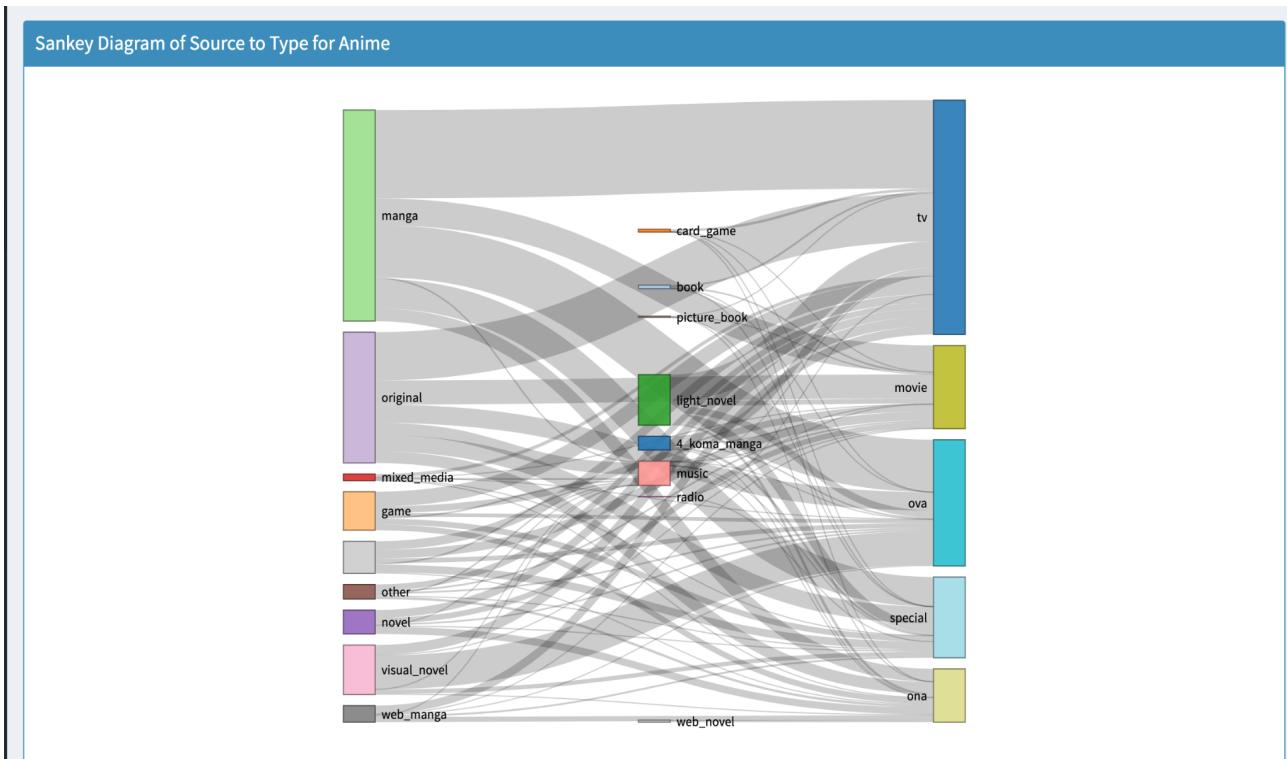
The graphic shows a "Popularity Comparison" box plot from a web-based dashboard used to analyse anime and manga data. This graph shows a visual comparison of popularity ratings for anime and manga, represented by red and cyan boxes, respectively. Lower numbers imply greater popularity. The anime popularity distribution, illustrated in red, has a tighter range and a lower median, indicating that anime titles are more consistently popular than manga. The figure has outliers above the top whisker, showing that certain anime titles are much less popular. In comparison, the manga box (shown in cyan) has a wider range and a higher median, indicating greater diversity in popularity. This visualisation is supplemented with a slider above the chart, which allows users to alter the popularity range they want to analyse, allowing for a dynamic investigation of how different titles compare to one another in terms of popularity.



❖ Sankey Diagrams

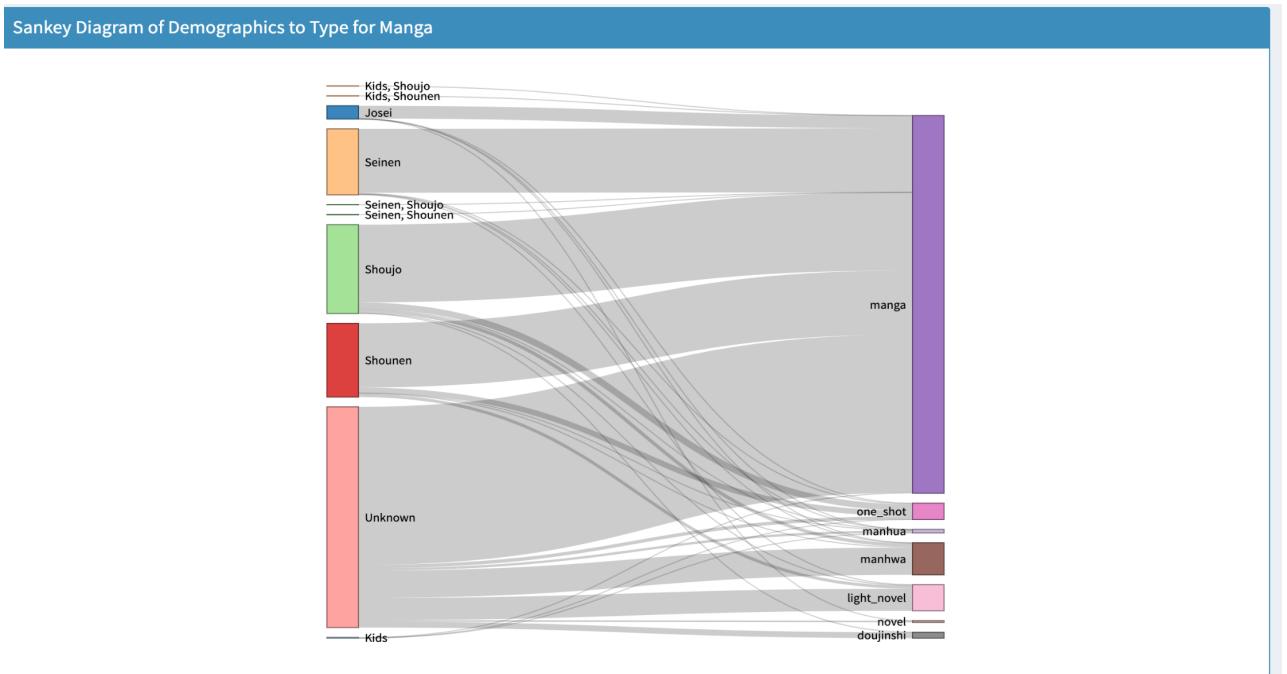
Anime Sankey Diagram

The image depicts a "Sankey Diagram of Source to Type for Anime," which is used to visualise the relationships between various sources of anime content and their adaptations into various anime formats such as TV shows, movies, OVAs, and specials. Each source and anime genre is colour-coded differently—manga is green and TV series is blue, for example—making it easy to distinguish across categories. The thickness of the lines connecting sources to types represents the volume of adaptations, with larger lines signifying more adaptations from a single source to a type. The graphic depicts a wide range of sources, including card games, books, and web novels, demonstrating the vast amount of original content that inspires anime adaptations.



Manga Sankey Diagram

The thickness of the lines in the Sankey figure represents the volume of manga generated for each demographic group assigned to each manga type category. For example, a thicker line connecting 'Shounen' to manga would imply a higher volume of Shounen manga than other categories. This visualisation tool efficiently depicts the target demographic distributions within the manga industry, as well as how distinct manga styles cater to different demographics, providing insights into market trends and content strategies.



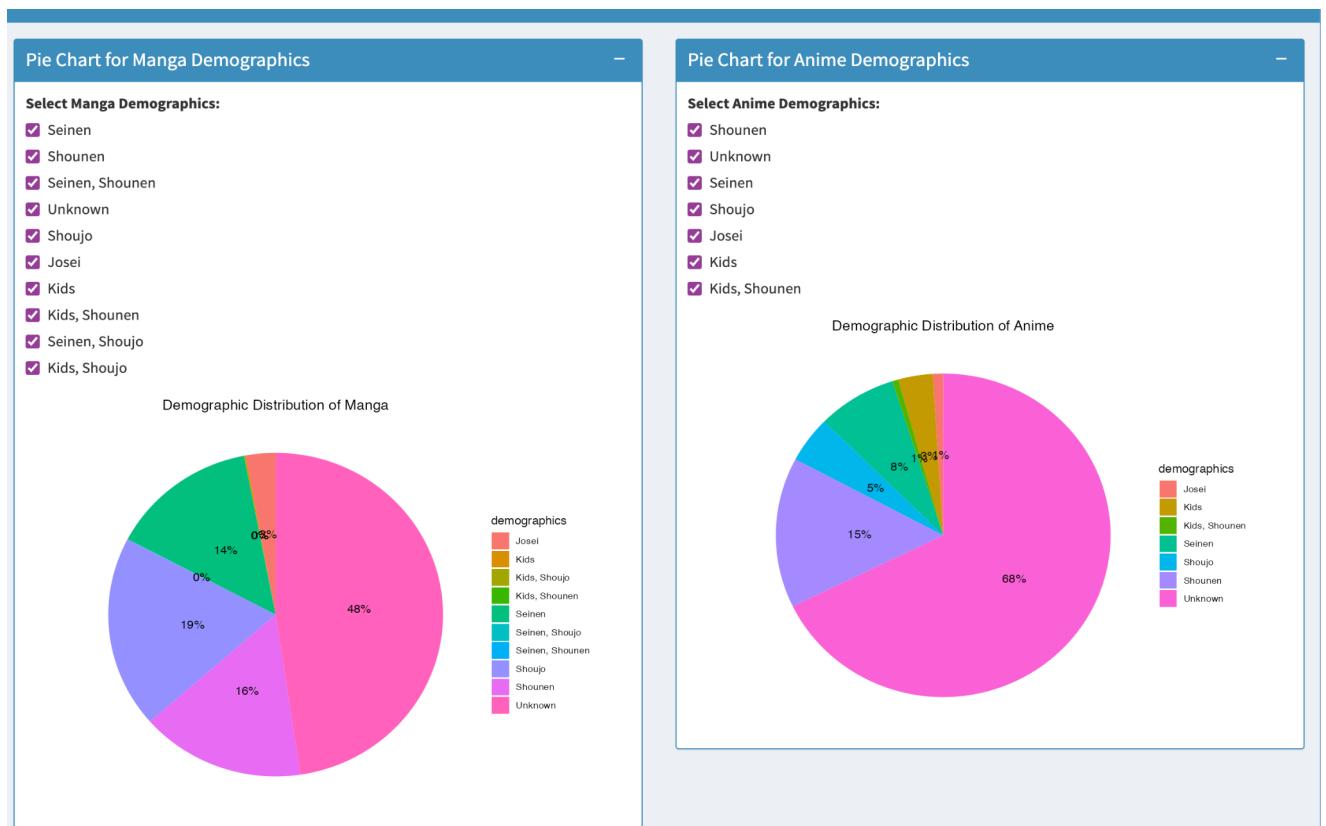
- Demographic Analysis

The illustration shows two pie charts within a Shiny dashboard, each showing the demographic distribution of manga and anime.

On the left side, there is a pie chart captioned "Pie Chart for Manga Demographics." The chart is accompanied with a checkbox group where users can decide which demographics to display. The demographics chosen are Seinen, Shounen, Seinen/Shounen, Unknown, Shoujo, Josei, Kids, Kids/Shounen, Seinen/Shoujo, and Kids/Shoujo. The data shows that "Unknown" is the largest demographic category, accounting for 48% of the Manga distribution. Other large categories include Seinen (19%), Shounen (16%), and Shoujo (14%), with the remaining demographics accounting for lower percentages.

On the right side, there is a pie chart labelled "Pie Chart for Anime Demographics," along with a similar checkbox group for selecting demographics. The demographics chosen are Shounen, Unknown, Seinen, Shoujo, Josei, Kids, and Kids/Shounen. The figure shows that "Unknown" is the most common demographic, accounting for 68% of the Anime distribution. Other important sectors include Shounen (15%) and Seinen (8%), with minor contributions from other demographics.

Both charts use various colours to represent each demographic, and the percentages within the pie slices are clearly labelled.



B. Justification of Final design Using Data Visualization Theory

- Colour Theory

Used contrasting colours (blue for anime and green for manga) to distinguish between genres. The colour selections make the visualisations visually appealing and easy to discern.

- Interaction Design Radio buttons and conditional panels

allow users to choose from various plot types and demographic groupings. This interactivity promotes user engagement and offers a personalised experience based on user preferences.

- Consistency

Ensured consistent design and interactivity across all visualisations, resulting in a coherent and unified user experience. This includes the consistent use of colour schemes, typography, and layout structures.

Implementation

5.1. Technical Implementation

This section provides a high-level overview of the implementation, including the libraries used and the rationale for any variations between the final and implemented designs.

Libraries

- Shiny: Used to construct an interactive online application.
- ggplot2: Used to create the visualisations.
- readxl: Used to read Excel files.
- dplyr and tidyr: Used to manipulate data.
- plotly: Used to make interactive plots.
- reshape2: Simplifies the process of reshaping your data between wide format (with repeated measures spread across wider, sometimes time-based columns) and long format (data is generally "tall", with variables in a single column and repeated instances in numerous rows).
- crosstalk: Improves the interactivity of web-based R visualisations by allowing connected brushing and filtering across several views of the data, which is particularly useful for Shiny web apps.

Challenges

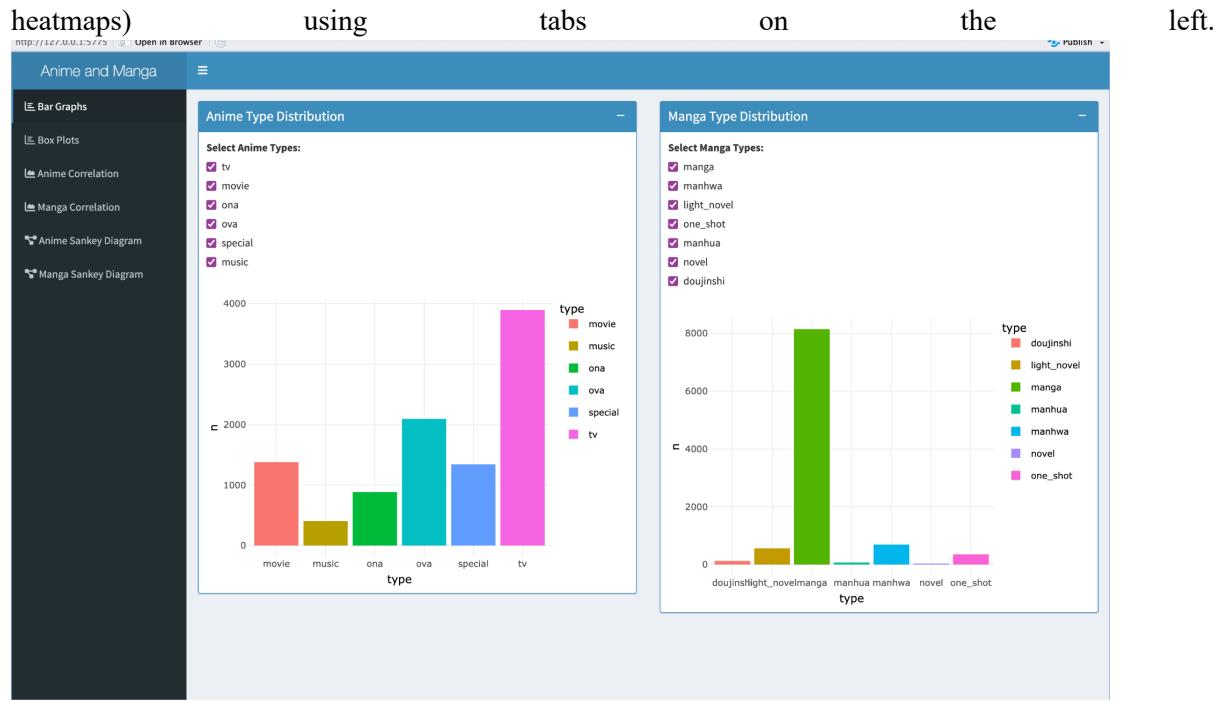
The project required substantial data wrangling to clean and modify the datasets. The integration of many data sources and ensuring smooth performance in Shiny required meticulous planning and execution. To achieve the desired results, advanced Shiny interactivity and high-quality visualisations with ggplot2 were necessary. As well as merging three datasets together to get the relevant graphs as well as to get tabular data such as Sankey data was difficult for this topic. Initially I wanted to use D3 for making a dashboard but then when I saw the complexity of the dataset that was used for this topic I thought I will use the R shiny where I inserted some HTML commands as well as CSS commands for featuring and giving good backgrounds to my dashboard.

5.2. Interactive Narrative Visualisation Implementation

This section describes the final implemented submission, including screenshots and explanations of how the narrative portrays the data insights to the audience.

Overview of the Interface

The application features a sidebar for plot type selection and a main panel for displaying the visualisations. Users can choose between different plot types (anime type distribution, manga type distribution, score comparison, popularity comparison, demographic analysis, and correlation



5.3. Using the Implementation

This section explains how to execute, view, and use the interactive story visualisation. It highlights aspects of the visualisation that a reader may otherwise overlook.

Run the application

To launch the Shiny application, load the required libraries and use the `shinyApp(ui = ui, server = server)` command in R. Ensure that the datasets ('anime_cleaned_data_score' and 'pop.xlsx', 'manga_cleaned_score_pop.xlsx') are in the right directory. I combined all three sources that were cited in the DEP report and made it as one file to make the file run as well as good visualisation which is relevant to the topic.



Viewing the visualisation

The dashboard shown in the screenshot is intended for an interactive visualisation of anime and manga data. It has an easy-to-use interface, with a navigation sidebar and a main panel that displays the visualisations. Users can choose from a variety of visualisation options on the sidebar, such as "Anime Type Distribution" and "Manga Type Distribution." Each selectable visualisation, such as the "Anime Type Distribution" depicted, displays data in the form of colour-coded bar charts that graphically

illustrate the counts of various types of anime or manga, allowing for simple comparison and analysis. This design allows for dynamic and fascinating data exploration, improving user awareness of the distribution and popularity of various anime and manga categories.



Interactive Features

1. Selection Filter

The screenshot shows the Shiny dashboard interface, with an emphasis on the "Anime Type Distribution" and "Manga Type Distribution" sections under the "Bar Graphs" category. Each plot includes a selection filter to increase user interaction: for anime, options include TV, movie, ONA, OVA, special, and music; for manga, possibilities include manga, manhwa, light novels, one-shots, manhua, novels, and doujinshi. These filters enable users to customise the displayed data by selecting or deselecting kinds, with the bar charts dynamically updating to reflect their preferences. The charts are colour-coded to identify between the types, with the height of each bar denoting the count, allowing for easy and quick visual comparison.



- Score Filter

The screenshot shows two interactive box plots from your Shiny dashboard: one for "Score Comparison" and another for "Popularity Comparison" of anime and manga. These visualisations provide a better grasp of how scores and popularity metrics are distributed throughout these two media forms. Each figure has a slider that allows viewers to filter the data range dynamically. The "Score Comparison" slider modifies the score range, narrowing the data shown in the box plot to scores within a specific range. This feature improves user involvement by allowing for extensive analysis of score distributions and comparisons across multiple levels. The plots use colour coding—red for anime and cyan for manga—to visually distinguish between the two mediums, allowing for clear and rapid comparisons.



- Hover and click

Hover over the plots to reveal tooltips containing extra information.



Conclusion

This project successfully created interactive and intelligent visualisations of manga and anime data, providing a comprehensive perspective of numerous aspects of the datasets, including type distributions, scores, popularity rankings, demographics, and correlations. The design and implementation followed data visualisation principles, resulting in clear narratives and engaging user experiences. Using Shiny, the project added dynamic interactivity, allowing users to explore and analyse data in a personalised and relevant way.

The focused design approach, which includes the Five Design Sheet Methodology, was critical to producing the final polished product. This process emphasised the value of feedback and continual development, ensuring that the visualisations were not only visually appealing but also effective in expressing the intended insights. The integration of different data sources, as well as the rigorous cleaning and transformation of datasets, demonstrated the technical complexity required to obtain the desired outcomes.

One of the key strengths of this project is the seamless user interface that balances functionality and simplicity. The use of contrasting colours, consistent design elements, and intuitive interactive features like radio buttons, sliders, and tooltips significantly enhance the user experience. These elements ensure that users can easily navigate through different visualisations, filter data, and access detailed information, thus making the exploration process both enjoyable and informative.

Looking back, there are various opportunities for development and future growth. One area for improvement may be the addition of more complex interactive features, such as animations or linked visualisations, which would increase user engagement and give a more dynamic exploring experience. For example, using animations to display patterns over time or linking multiple visualisations to highlight relevant data points could provide more in-depth insights and a more engaging user experience.

Furthermore, increasing the dataset to incorporate more recent data or other relevant qualities may broaden the reach of the visualisations. This could entail combining new genres, demographic information, or user ratings from other platforms to create a more complete picture of the manga and anime scene. By doing so, the visualisations may reveal new trends and patterns, making the study more complete and up to date.

To summarise, this study demonstrated the effectiveness of interactive visualisations in making complex data more accessible and intelligible. Shiny and ggplot2, paired with a thorough design approach, produced a tool that is both functional and interesting. While there is always room for improvement, the current implementation serves as a solid platform for further investigation and analysis of manga and anime data. The project's findings not only contribute to a greater understanding of the manga and anime scene, but also demonstrate the possibilities of interactive visualisations in data analysis and storytelling. Future research in this field can build on these foundations, investigating new dimensions and broadening the reach to provide even more detailed insights.

Bibliography

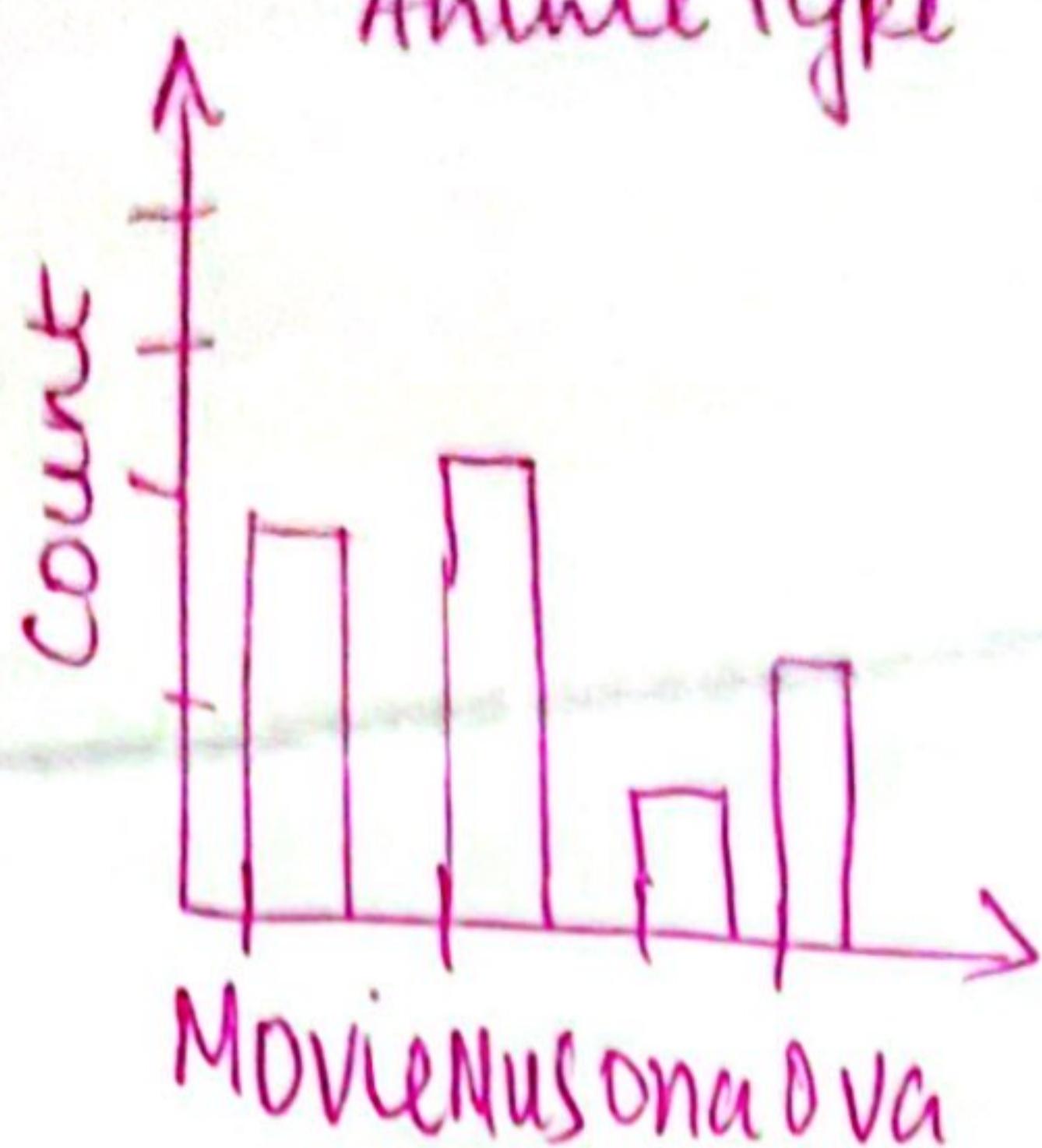
- Chang, W. (n.d.). Shinydashboard. RStudio. Retrieved from <https://rstudio.github.io/shinydashboard/>
- Cheng, J. (n.d.). *Crosstalk: Inter-Widget Interactivity for HTML Widgets*. RStudio. Retrieved from <https://rstudio.github.io/crosstalk/>
- HTML Color Codes. (n.d.). Retrieved from <https://htmlcolorcodes.com/>
- W3Schools. (n.d.). *CSS Tutorial*. Retrieved from <https://www.w3schools.com/css/>
- Japan Craft. (n.d.). The Difference Between Manga and Anime. Retrieved from <https://japancraft.co.uk/blog/manga-anime-difference/>
- Uchi Japan. (n.d.). Japanese Culture: The World of Manga and Anime. Retrieved from <https://uchijapan.com/news/japanese-culture-the-world-of-manga-and-anime>
- Allaire, J., Cheng, J., Xie, Y., McPherson, J., Chang, W., Allen, J., ... & Borges, B. (2022). networkD3: D3 JavaScript Network Graphs from R. R package version 0.4. Retrieved from <https://cran.r-project.org/web/packages/networkD3/networkD3.pdf>
- Wickham, H. (2022). reshape2: Flexibly Reshape Data: A Reboot of the Reshape Package. R package version 1.4.4. Retrieved from <https://cran.r-project.org/web/packages/reshape2/index.html>
- Plotly. (n.d.). Plotly R Library. Plotly Technologies Inc. Retrieved from <https://plotly.com/r/>
- Holtz, Y. (n.d.). *Sankey diagram – The R Graph Gallery*. The R Graph Gallery. Retrieved from <https://r-graph-gallery.com/sankey-diagram.html>
- RStudio. (n.d.). *SliderInput*. shiny 0.14.1. Posit. Retrieved from <https://shiny.posit.co/r/reference/shiny/0.14.1/sliderinput>
- Stack Overflow. (2020). *How to create a simple drop-down menu in shiny*. Retrieved from <https://stackoverflow.com/questions/61044496/how-to-create-a-simple-drop-down-menu-in-shiny>
- Perrier, V., & Meyer, F. (n.d.). *Dropdown*. shinyWidgets. Retrieved from <https://dreamrs.github.io/shinyWidgets/reference/dropdown.html>

Appendix - Five sheets design

①

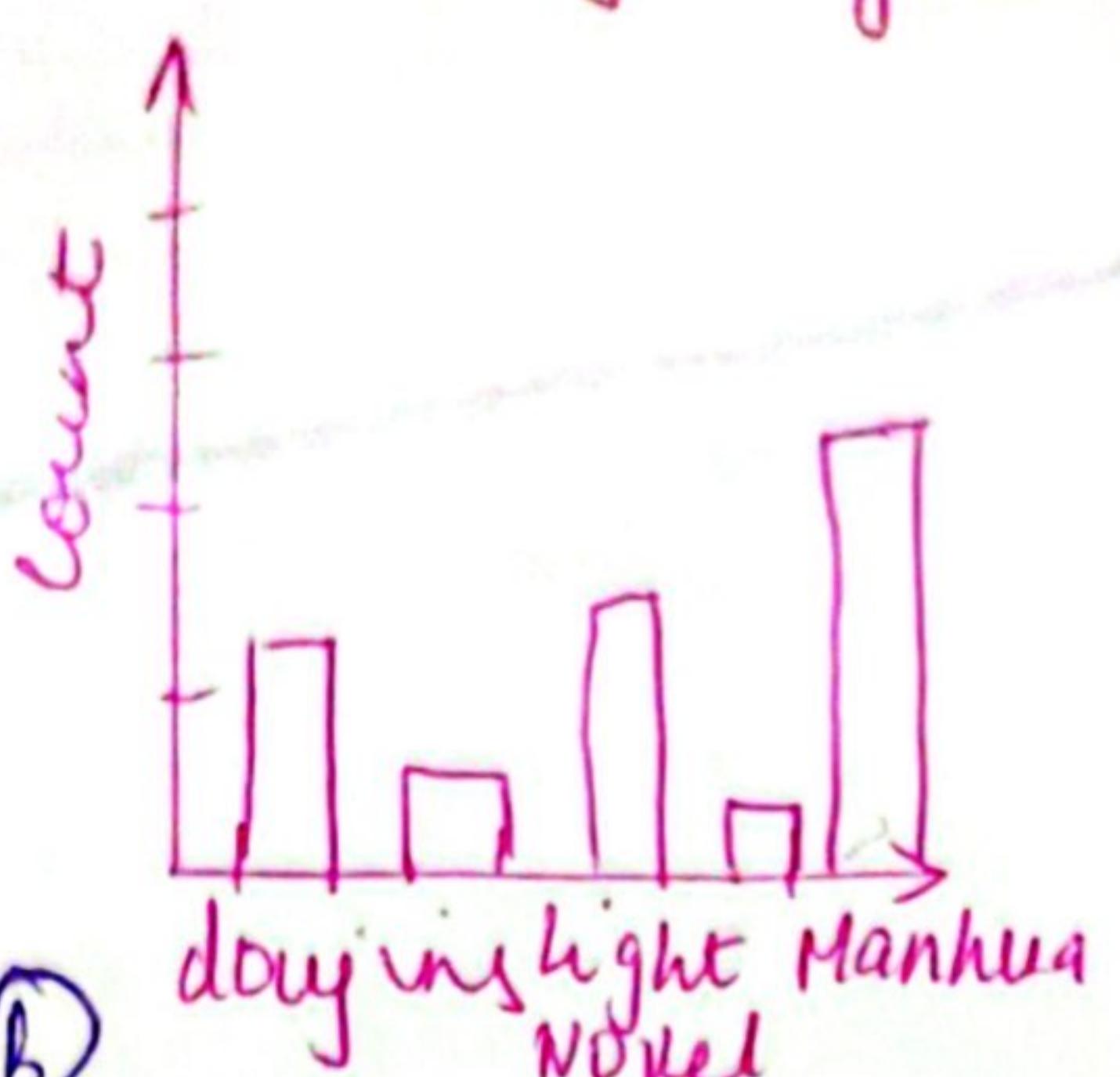
Sheet 1

Anime Type

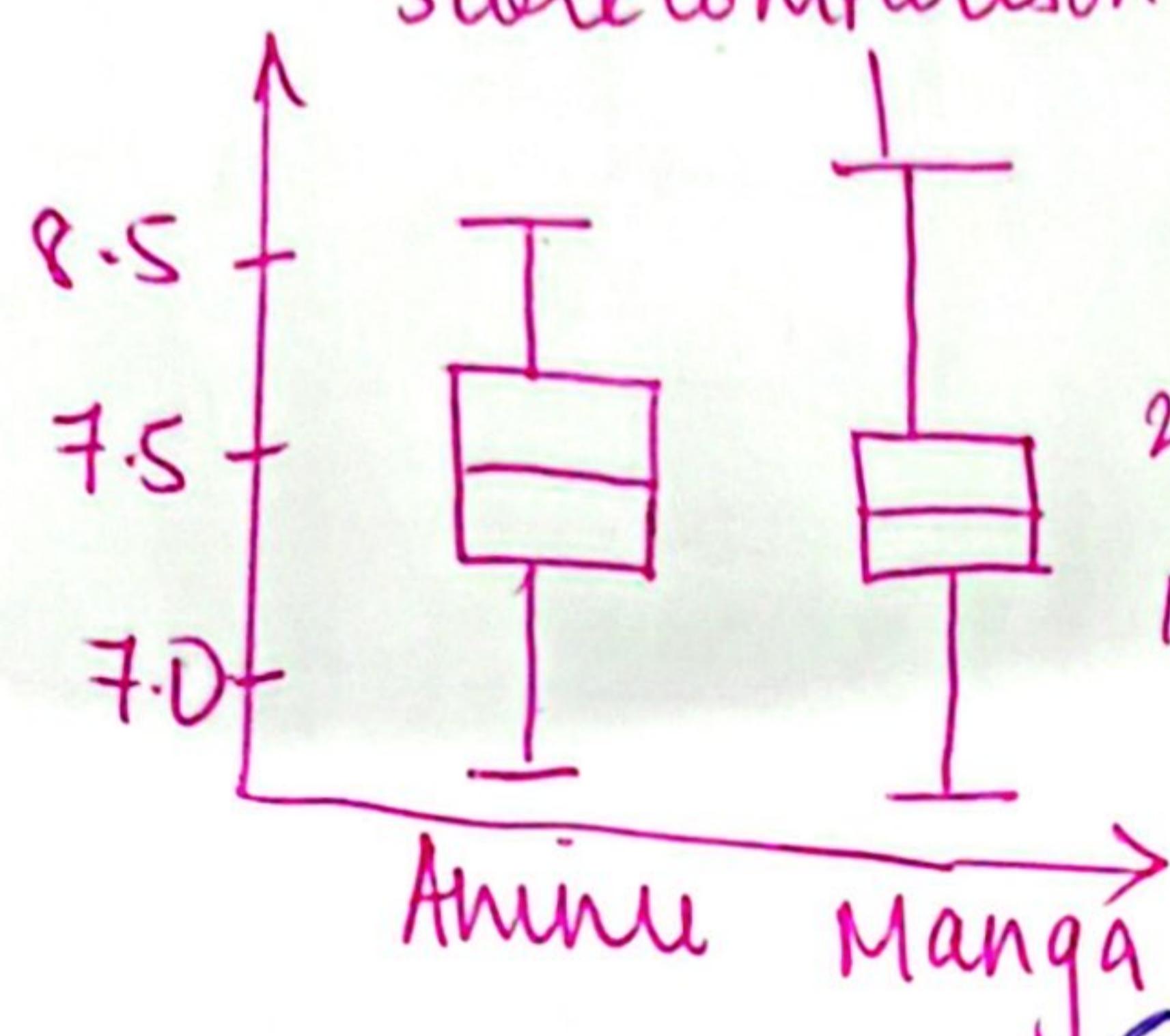


②

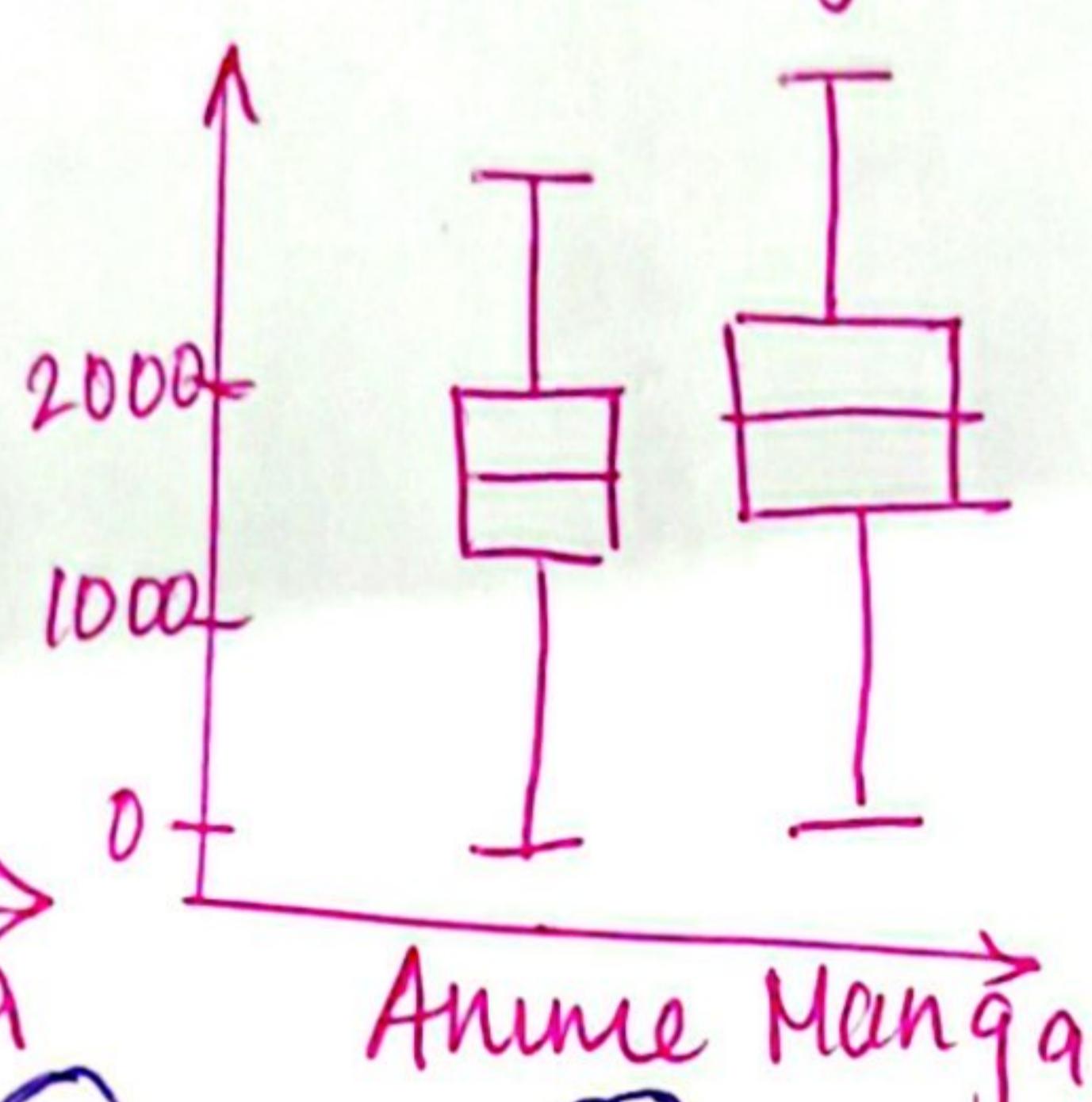
Manga Type



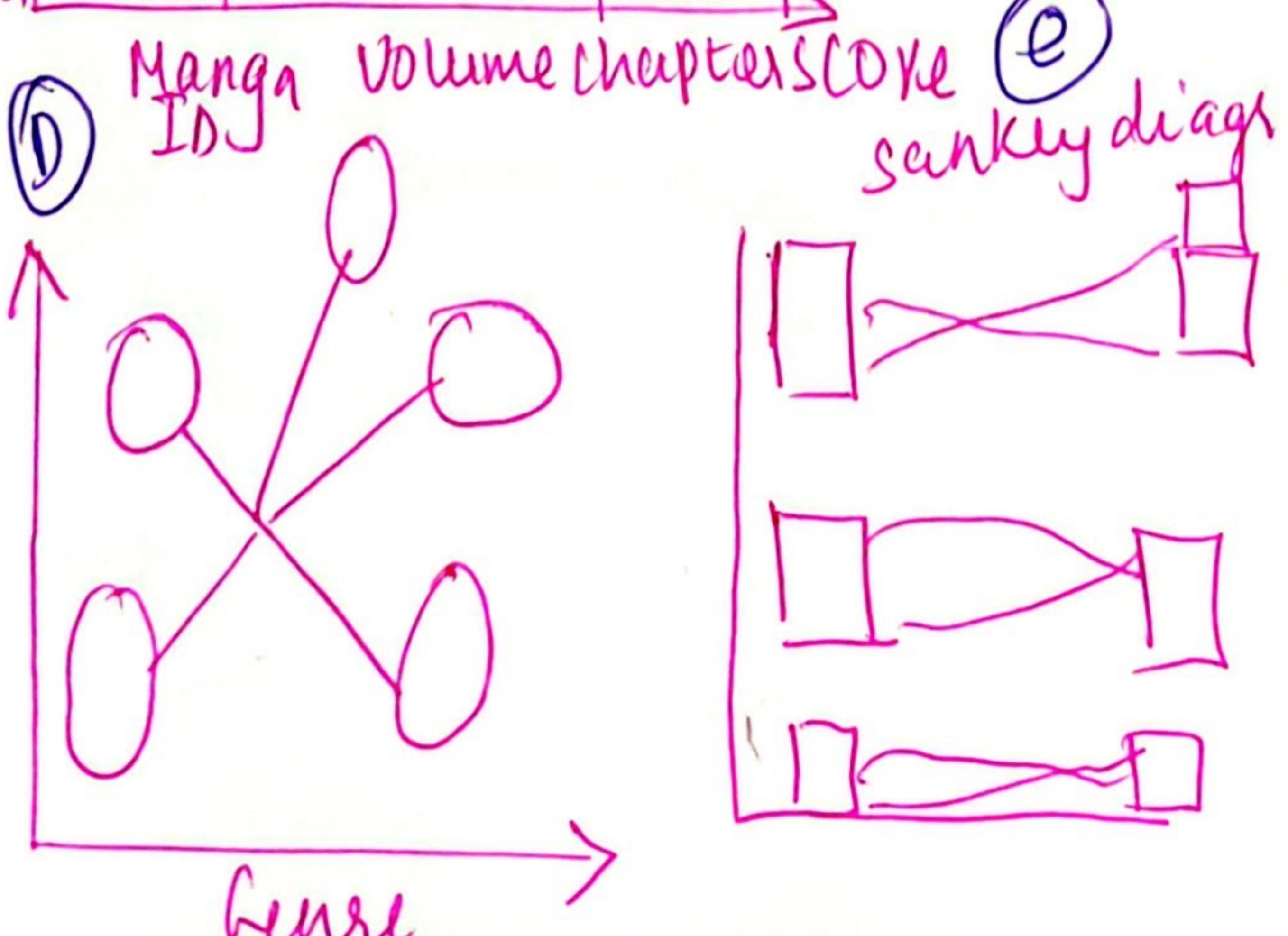
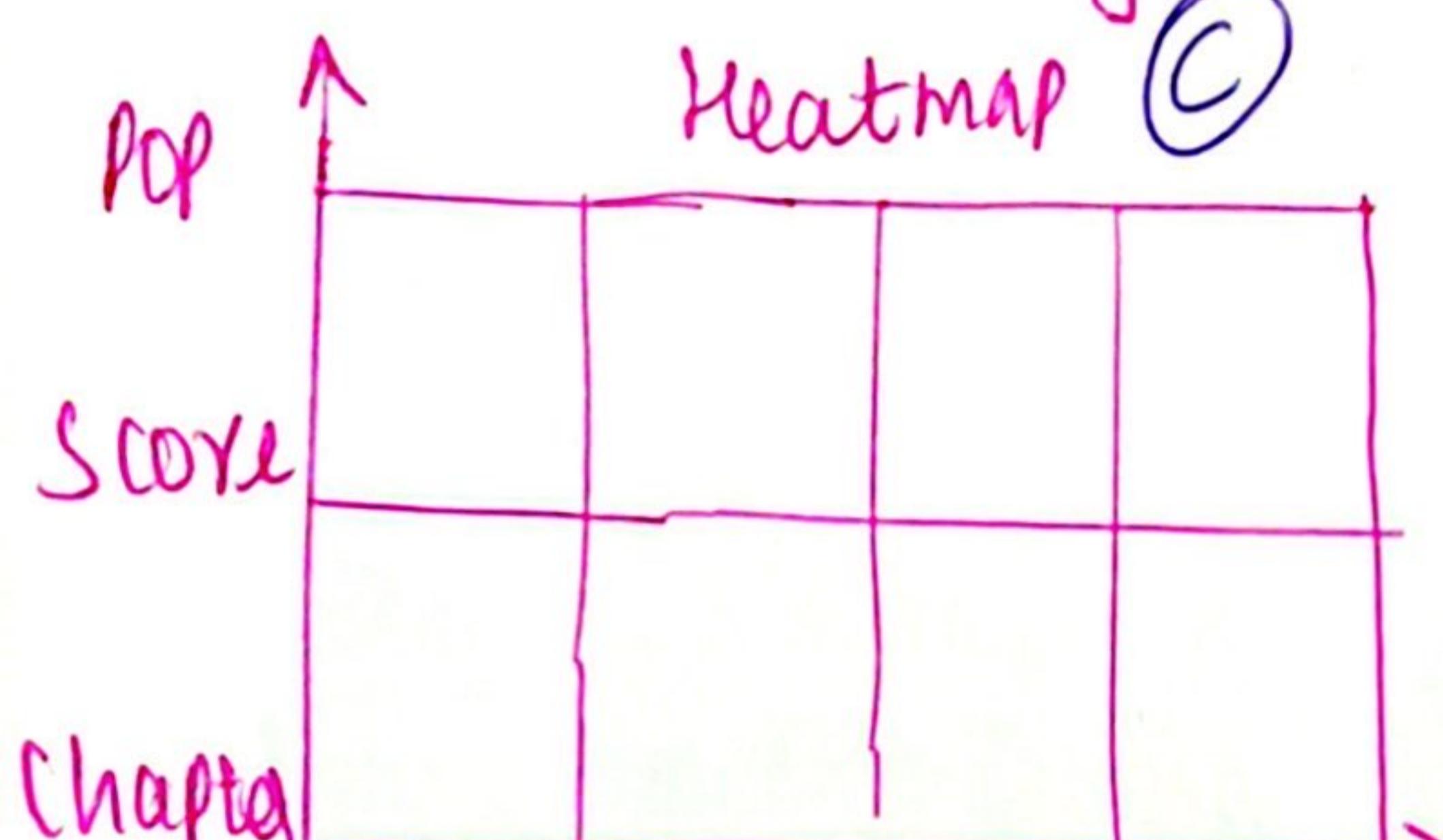
Score Comparison



Popularity Comp



Heatmap



N. Subbulakshmi
34069178
DATA Exploration
and Visualization
Date : 31/06/2024

3 Questions

- Various Types of anime and manga
- Score and popularity
- Demographic

Best viz

- (A) Best for type
- (B) Best for demographic
- (C) Best for score and pop
- (D) Types of manga and anime

Categorize

1-D Dimension

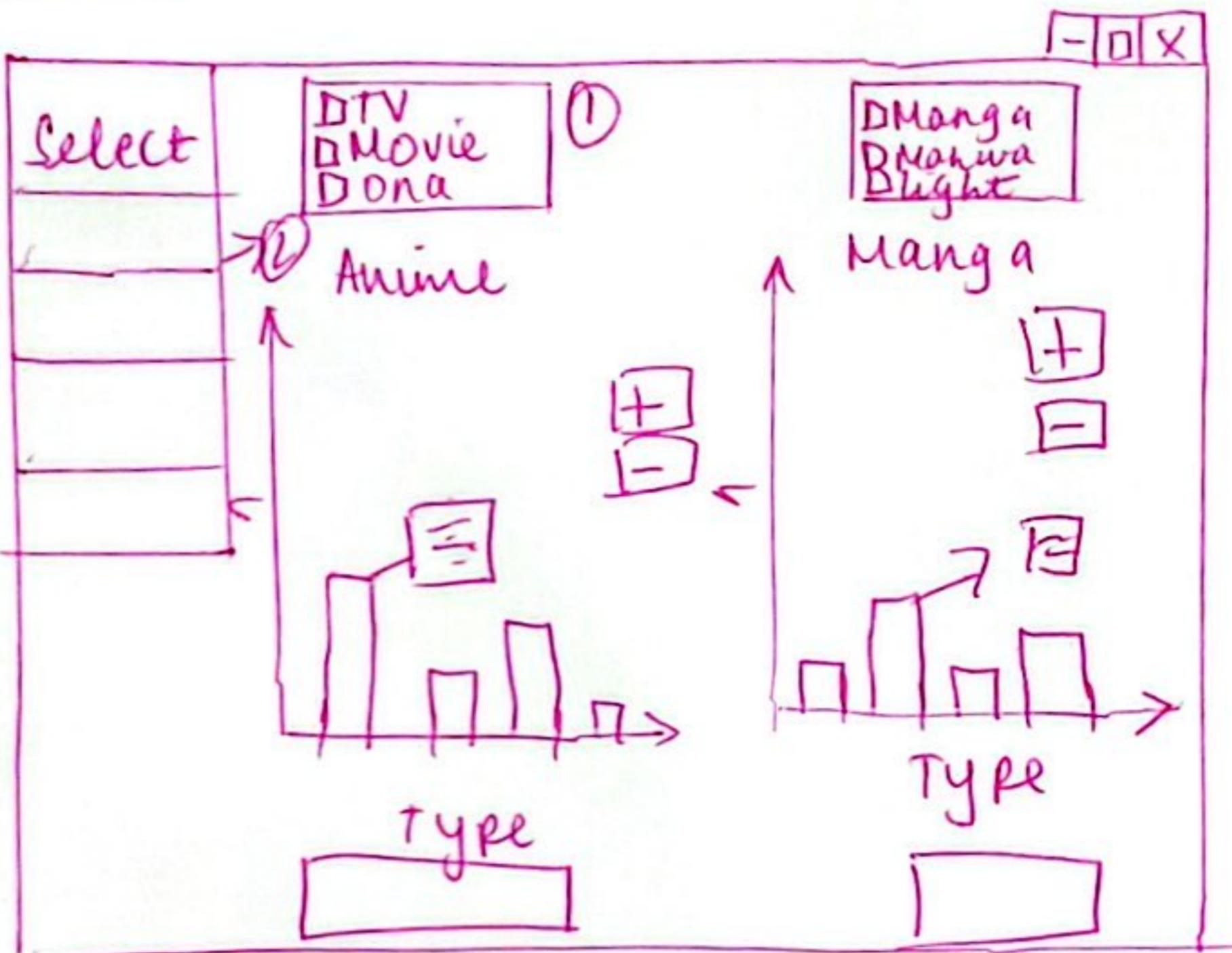
- Bar chart
- Pie chart

2-D Dimension

Stacked bar chart

SHEET 2

Layout



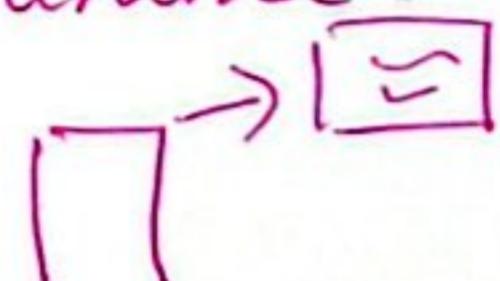
Operation

- ① Check box: For making user to select the type of media they want
- ② Tab: To change the tab for each graph
- ③ Zoom: To zoom the graphs using the plotly function
- ④ TOOL tip: To know the count of type.

Subbulakshmi
34069178
Date: 31/06/2024

zoom/Focus

The bar chart would allow us to give more visualize the type of manga and anime.



The Type of Anime and Manga count

Discussion

Pros

- Easy to distinguish
- User friendly
- Very readable and simple

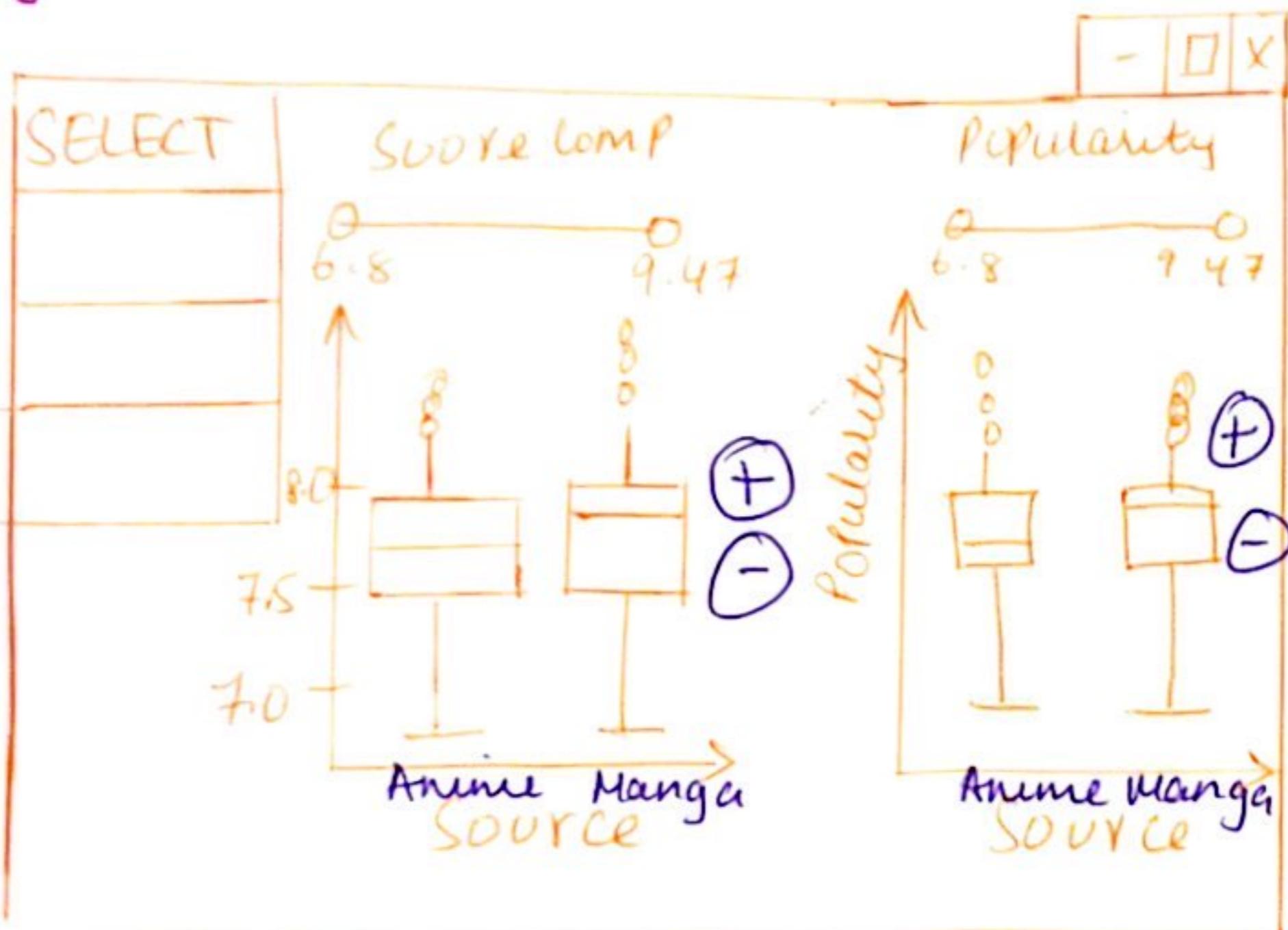
Cons

- Doesn't give in detail information but we can handle it.

SHEET 3

Subbulakshmi
34069178
Date: 31/06/2024

Layout



Operation

① Sliders input : To allow users to adjust the Score for both Anime and Manga as well as popularity .

② Remove the Label : The user can remove the label and put parameters they want

zoom/Focus
The box plot will show the score and popularity so you can zoom and look too.

Discussion

Pros :
- Very flexible
- Give choice to the users
- more flexibility

Cons :
- Can be inappropriate to choose the data

Sheet 4

Subbulakshmi
34069178
Date: 3/06/2024

Layout

SELECT	Heatmap	BB	- O X																		
	<table border="1"> <tr><td>00</td><td>00</td><td>DD</td></tr> <tr><td>00</td><td>00</td><td>D</td></tr> <tr><td>D</td><td>D</td><td>ID</td></tr> </table> <p>Manga Type Demographic ID</p>	00	00	DD	00	00	D	D	D	ID	<table border="1"> <tr><td>11</td><td>00</td><td>00</td></tr> <tr><td>00</td><td>11</td><td>00</td></tr> <tr><td>00</td><td>00</td><td>11</td></tr> </table> <p>Anime Type Demographic ID</p>	11	00	00	00	11	00	00	00	11	
00	00	DD																			
00	00	D																			
D	D	ID																			
11	00	00																			
00	11	00																			
00	00	11																			

Operation

No operations which is not considered good visualization so we are not going to ahead with this visualization type.

But I have put a button box choosing the relevant

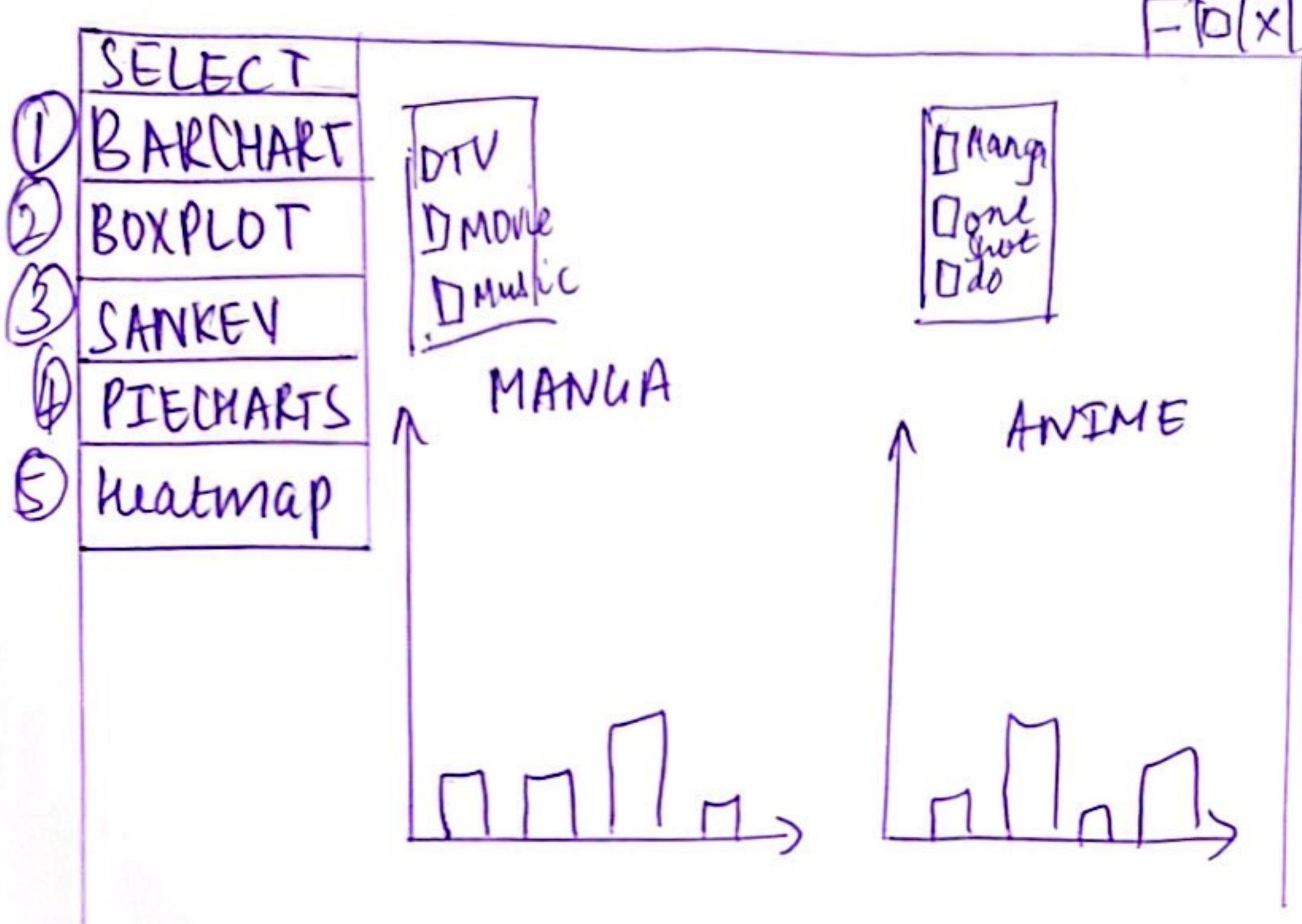
Zoom/Focus
Heatmap is good choice but not really good for visualization

Discussion Pros

- Simple
- Cons
- Users will get

Sheet 5

Subbulakshmi
34069178
Date : 3/06/2024



zoom/Focus
①, ②, ③, ④
and ⑤
Click the
graph you
want to
choose

Tool required
① D3, R Shiny
② Lots of
Data, Tables
from Kaggle
③ Excel
④ Coding

operation
① Slider
input
② Check
box
③ Drop box

Discussion

The initial screen shows the types of manga and anime (Just rough drawing)

Time Required

1 year .(MAX)