CZ4124 / SC4024 Assignment #2 – Video Presentation Critique #1

Video Number: **V03** **Video Title**: **V03-final**

1. Choice of visuals, visualisation techniques and appropriate use of human visual perception principles

At the timestamp of **[0:52]**, shown in Figure 1, the graph depicting the increased number of athletes in 2024 based on disciplines could benefit from sorting the x-axis by value-based ordering. This approach would improve the visual simplicity of the chart, in line with the **Gestalt principle of simplicity**.

Additionally, the current design makes it difficult to distinguish the actual values of the number of athletes, which could lead to confusion. For example, it's unclear whether there is a gender difference in [Bowling] or whether the low values for [3x3 Basketball] and [Equestrian] indicate actual low counts or reflect missing/unclean data. To address this, adding horizontal grid lines in the background would help viewers more easily estimate the number of athletes, as suggested by **Weber's Law**, which improves the ability to **estimate changes** in values.

Furthermore, adding direct labels showing the value on top of each bar would implement the **principle of perceptual organization**, aiding in visual comparison and improving the clarity of the data for users. This would also enhance the overall interpretability of the graph.

At timestamp **[1:27]**, shown in Figure 2, the left graph indicating the top 30 countries by number of medals and medal type does not effectively utilize the **principles of perceptual organization and the Gestalt principle of proximity**. The bars in the bar chart are positioned too closely together, which may lead to confusion as users could mistakenly group the wrong bars. Although the grouping of medal types within the bars correctly applies the **Gestalt principles of proximity and perceptual organization**, allowing users to recognize the medal types as units, the lack of space between the bars makes it challenging to visually process the data.

To address this issue, we could reduce the scope of the medal analysis from the Top 30 to the Top 10 countries. This would provide enough space to better implement the proximity principle, ensuring that each medal type and country can be easily distinguished. This suggestion follows the principle of relevance (**Goldilocks principle**), offering a balanced amount of data to highlight differences without overwhelming the viewer with too much information.

Additionally, in the right graph of Figure 2, the chart utilizes the **principle of simplicity** well by sorting the data for ease of visualization. However, the color scheme could be improved by following the **principle of compatibility**. The data could be sorted with gold at the top, silver in the middle, and bronze at the bottom. This ordering would help users interpret the data more intuitively, as the color scheme would align with the commonly understood hierarchy of medal values.

At timestamp **[1:47]**, shown in Figure 3, the **Gestalt principle of similarity** could be implemented by replacing the x-axis labels with a color legend to indicate the medal types. This would group the medal types by color, making it easier for the user to visually associate the different categories with their corresponding colors, thereby improving clarity and coherence. This approach also enhances **space efficiency** by reducing unnecessary visual elements.

Additionally, the **principle of compatibility** can be applied by sorting the medals in the following order: gold in the middle, silver on the left side, and bronze on the right. This arrangement aligns with the familiar structure of the Olympic podium, where gold is at the center, silver is second, and bronze is third. This intuitive order helps users quickly interpret the data and relate it to the commonly understood ranking of medals.

Furthermore, the graphs use different y-axis scales, which can make it difficult to compare which country won more medals. The bars may appear similar in length, even though the values differ (e.g., the U.S. has the most medals, located at the bottom right, and Korea has the least, located at the top left). To resolve this, the y-axis scales should be standardized for comparison. This would adhere to the **principle of graphical integrity**, ensuring that visual attributes (bar length) are directly proportional to the data values.

Additionally, following **Weber's law** by using a consistent scale would help users easily perceive the differences in medal counts and estimate changes more effectively (**Estimating Change**).

At timestamp **[2:15]**, shown in Figure 4, the left chart should use a bar chart to display the number of participations. This is because human vision is more adept at estimating magnitude based on visual length (according to **Estimating Magnitude**), and a bar chart would make it easier for users to compare the participation counts across countries.

For the right chart, the country names should be placed above the bubbles rather than inside them. This placement would make it easier for users to read the country names, aligning with the **Gestalt principle of proximity** (grouping related elements together) and the **principle of simplicity** (removing unnecessary visual complexity). By placing the labels outside the bubbles, the chart becomes more user-friendly and legible.

At timestamp **[3:08]**, shown in Figure 5, the heatmap uses the number of medals won as the color code, which leads to an imbalance in the color distribution. This imbalance occurs because, according to the Olympic dataset [1], [Athletics] and [Swimming] have more events compared to other disciplines, as shown in Figure 7. To resolve this issue, the heatmap should use a normalized value, such as the number of medals per event, to provide a more balanced representation of performance across disciplines. This adjustment would account for the varying number of events in each discipline, resulting in a more meaningful comparison of performance.

At timestamp **[4:00]**, shown in Figure 6, the line chart on the left lacks both X and Y axis labels, which could lead to confusion for the viewer. Additionally, the choice of using a line chart to show overall age group participation is not appropriate, as the lines overlap, making it difficult to distinguish between the different age groups and overwhelm the user (not effective according to **principle of relevance/goldilocks principle**). A stacked bar chart would be more suitable for displaying these differences, as it can clearly represent the proportions and changes across age groups. Furthermore, when color-coding for the medals, the **principle of compatibility** suggests using gold for gold medals, silver for silver, and bronze for bronze, as this aligns with users' expectations and improves clarity.

From timestamp **[4:00] to [4:12]**, the presenter concludes that the younger generation excels in gymnastics and swimming, while the older generation excels in equestrian and shooting. However, the box plot chart presented does not highlight complete information, as disciplines such as equestrian and shooting are missing visual elements. This could lead to misleading interpretations and raises concerns about **ethical visualization** and **graphical integrity**, as all relevant data should be accurately represented to avoid bias or omission.

2. Overall Comments

The **strongest features** of the video presentation are its engaging introduction and the focus on relevant factors, such as age and gender, which significantly impact Olympic medal performance. These topics provide valuable insights into the dataset and allow the audience to draw connections between demographic factors and Olympic success.

The **weakest feature** of the presentation is the apparent lack of a comprehensive understanding of the dataset. For instance, in Figure 5, the color coding is based on the total number of medals rather than medals per event. This leads to an imbalanced color distribution, particularly because disciplines like Athletics and Swimming, which have more events (as shown in Figure 7), are overrepresented. This skews the analysis and diminishes its impact.

Additionally, there are instances where the presentation does not adequately clarify missing information. A notable example is in Figure 6, where the conclusions drawn in the video conflict with the data shown in the visualization. This discrepancy can lead to confusion and misinterpretation.

References

[1] https://www.kaggle.com/datasets/piterfm/paris-2024-olympic-summer-games/data

Figures

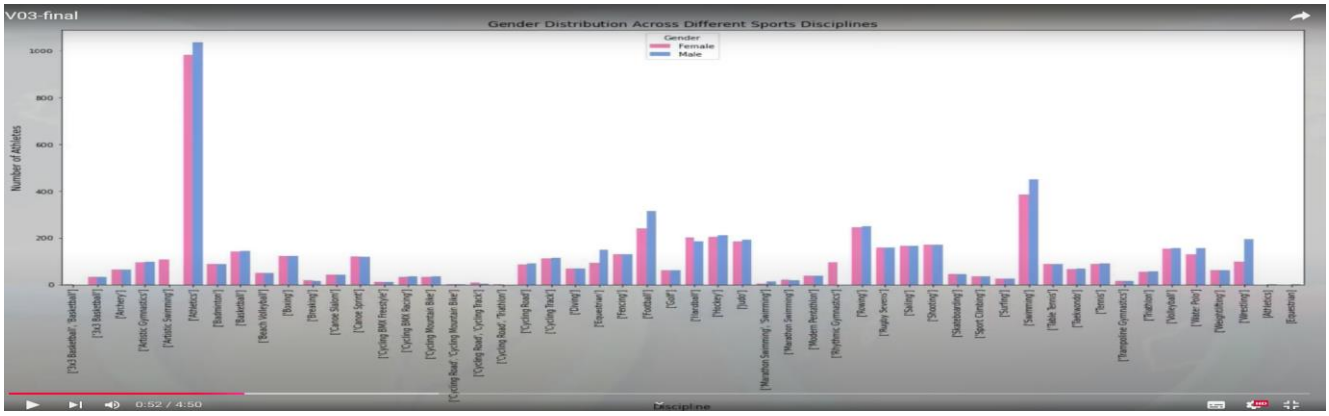

Figure 1. timestamp [0:52]

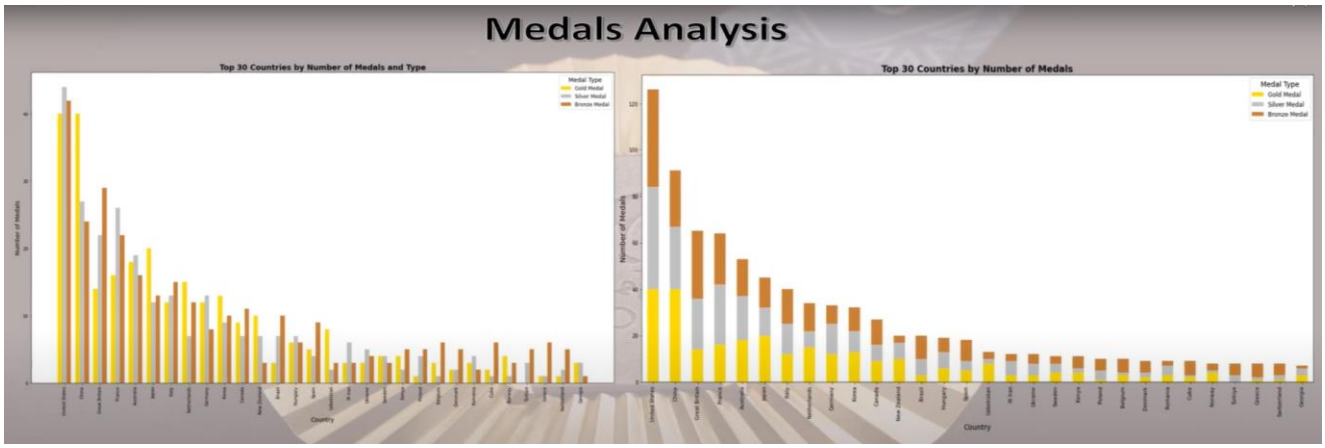

Figure 2. timestamp [1:27]

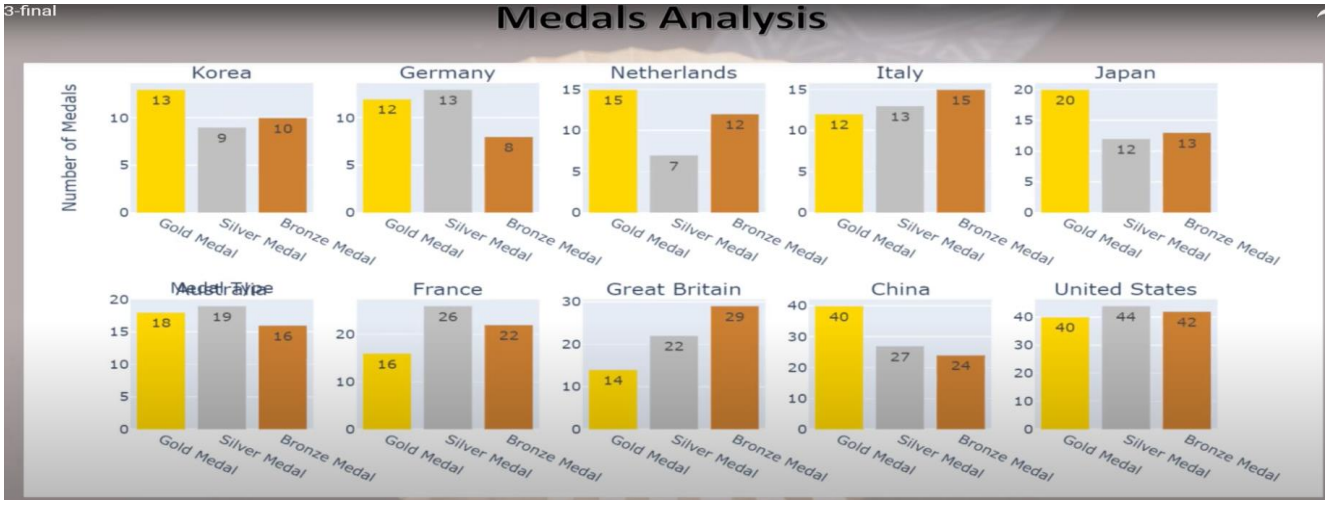

Figure 3. timestamp [1:47]

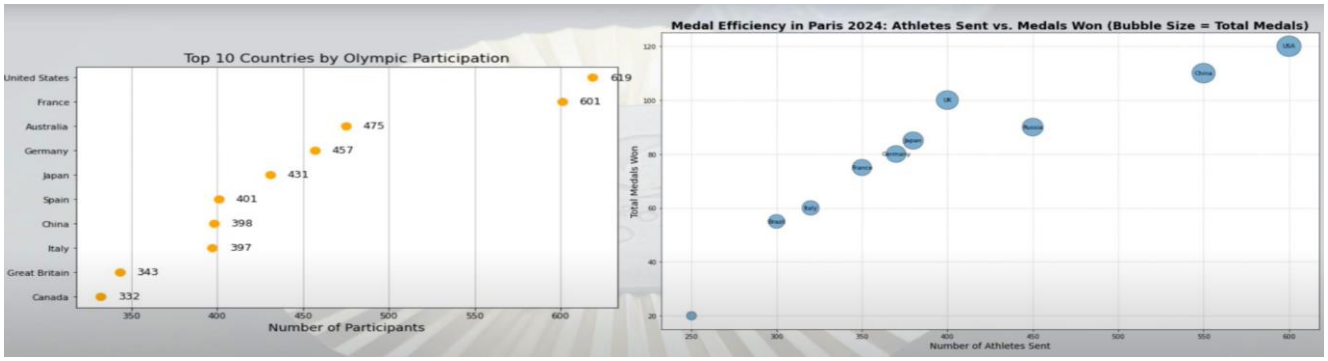

Figure 4. timestamp [2:15]

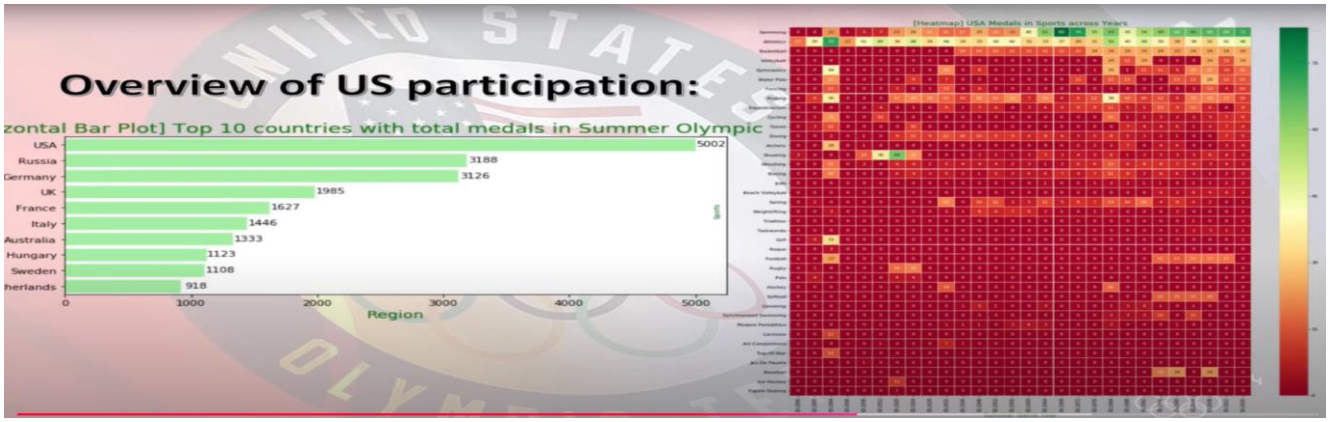

Figure 5. timestamp [3:08]

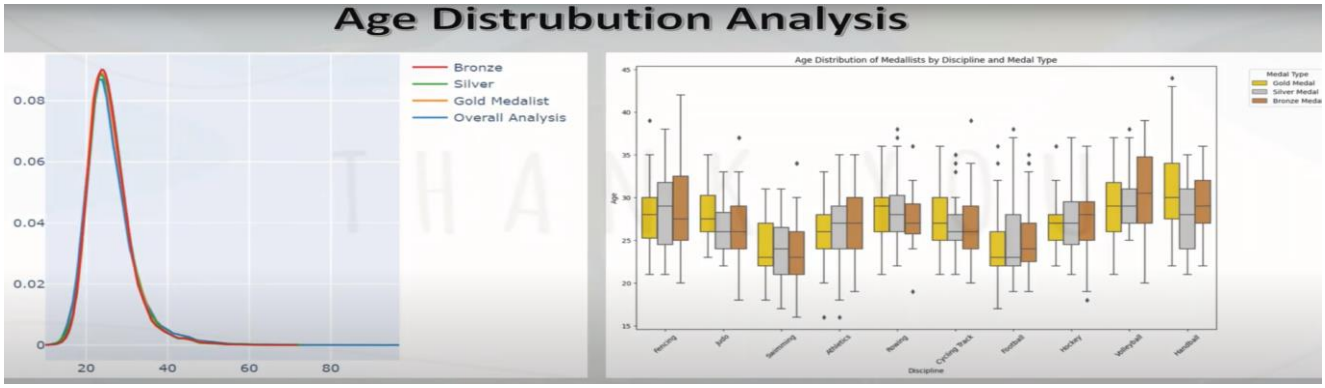

Figure 6. timestamp [4:00]

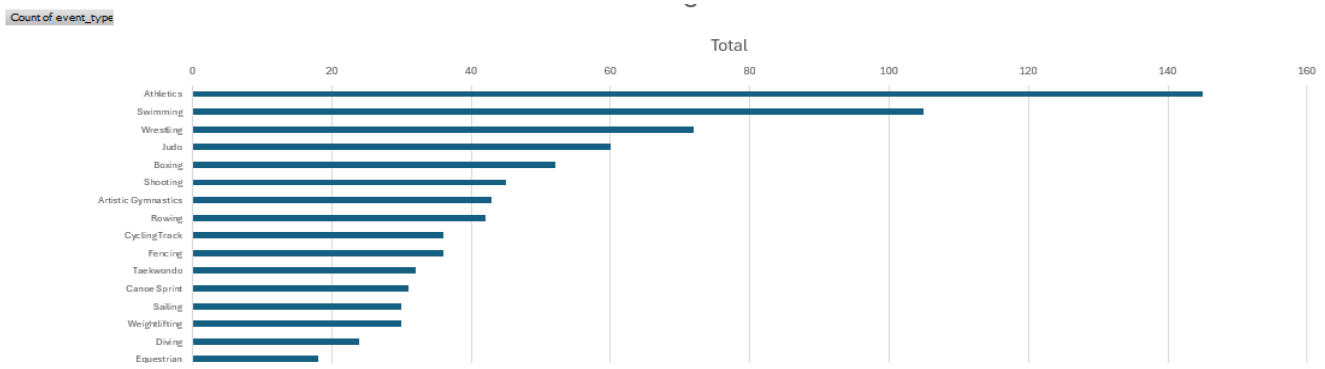

Figure 7. Number of events in each disciplines (construct using Excel using Reference [1] dataset).