1 Introduction

In this assignment, I accomplished a novel visualization on the dataset of Tokyo Olympics 2020. Tableau was used as the main tool in this assignment and the dataset was obtained from the Kaggle community. The main purpose of this assignment is to create a dashboard to display the data collected during the Olympics more intuitive and readable for users and also create a platform for exploratory data analysis.

2 Description

- **Tools:**

  Tableau was chosen to be the main tool to complete the assignment since it contains powerful functionality to create novel visualization with dashboards. I initially tried to complete the assignment on Google Colab and finally found it extremely hard to merge a visualization product on it. Since the dataset was obtained from Kaggle community and preprocessed, I did not need to have any steps for preprocessing the dataset

- **The dataset:**

  This dataset was obtained from the Kaggle community and was preprocessed. It contains 5 csv files. Only 3 of them were used for implementation of this assignment(Athletes, Entries Gender, and Medals) All the values of the attributes are types of integers. In the file of the Medals, it contains the counts of the medals in 3 different types(gold, silver and bronze) of all the teams participated in. And in the file of Entries Gender, it contains the counts of female and male players in each discipline. In the file of Athletes, it contains the information of every athlete who joined the event.

  All the data types are quantitative.

- **Tasks:**
  a. The visualization firstly displayed the top 20 countries with most medals in the Olympics game in Tokyo 2020 by a ranked bar chart. It looked up and compared the values of the medals of the top 20 countries in the ranking.
  b. Gender distribution of every discipline was implemented by a pie chart to compare the distribution of male and female players in each discipline. It discovered a part-whole relationship.

c. The distribution of the most participated disciplines was visualized by a pie chart for comparison of different disciplines. It displayed a part-whole relationship of the disciplines.
d. A geographic map was established to display the total medals of each country on the locations of those on the map with different levels of color for comparison. It mapped the colors to the map for comparison of the values of the medals.
e. The visualization of the number of players participated in each country was implemented for ranking. It compared the values.

- **Encoding channels and idioms:**
  - For the ranking of the medals of top 20 countries, it is a colored bar chart and it expresses the values of total medals in the aligned length and color was used to encode the numerical values of the total medals of each country for better comparison of the values.
  - For the gender distribution of each discipline, a pie chart is used. And Area marks (wedges) with angle channel; radial layout. It is intuitive for comparing the gender distribution of each discipline.
  - In the visualization of most participating disciplines, a colored pie chart is used. The area of each discipline marks angle channel; radial layout to display the part-whole relationship. The color is used to encoding the categorical attributes of the types of disciplines to distinguish different disciplines.
  - For total medals won by each country, a colored geographic map is implemented. In the location of each country, the total number of medals won in Tokyo 2020 is displayed and the color shades are used to express the quantitative attribute of the medals for comparison of values.
  - For the ranking of participation levels of the countries, a ranked and colored bar chart is used, it expresses the values of total participated athletes in the aligned length and color shades were used to encoding the numerical values. It helped to compare the number of total participated athletes of each country.

- **Discussion**

  Since the visualization is implemented in Tableau Desktop, a dashboard is established to display the data. The most complex part is to create the interactive parts of the dashboard. I decided to create the floating chart when putting the pointer on each bar of the total medals bar chart to display a detailed distribution of medals types of each country. For the gender distribution of each discipline, there is a selection button for the user to select which discipline to display on the pie chart. For the map, the detailed distribution of medals of each country can also be displayed by putting the

pointer on the location of the country. To check the participation levels of countries, a filter was used.

- **Critical Analysis:**

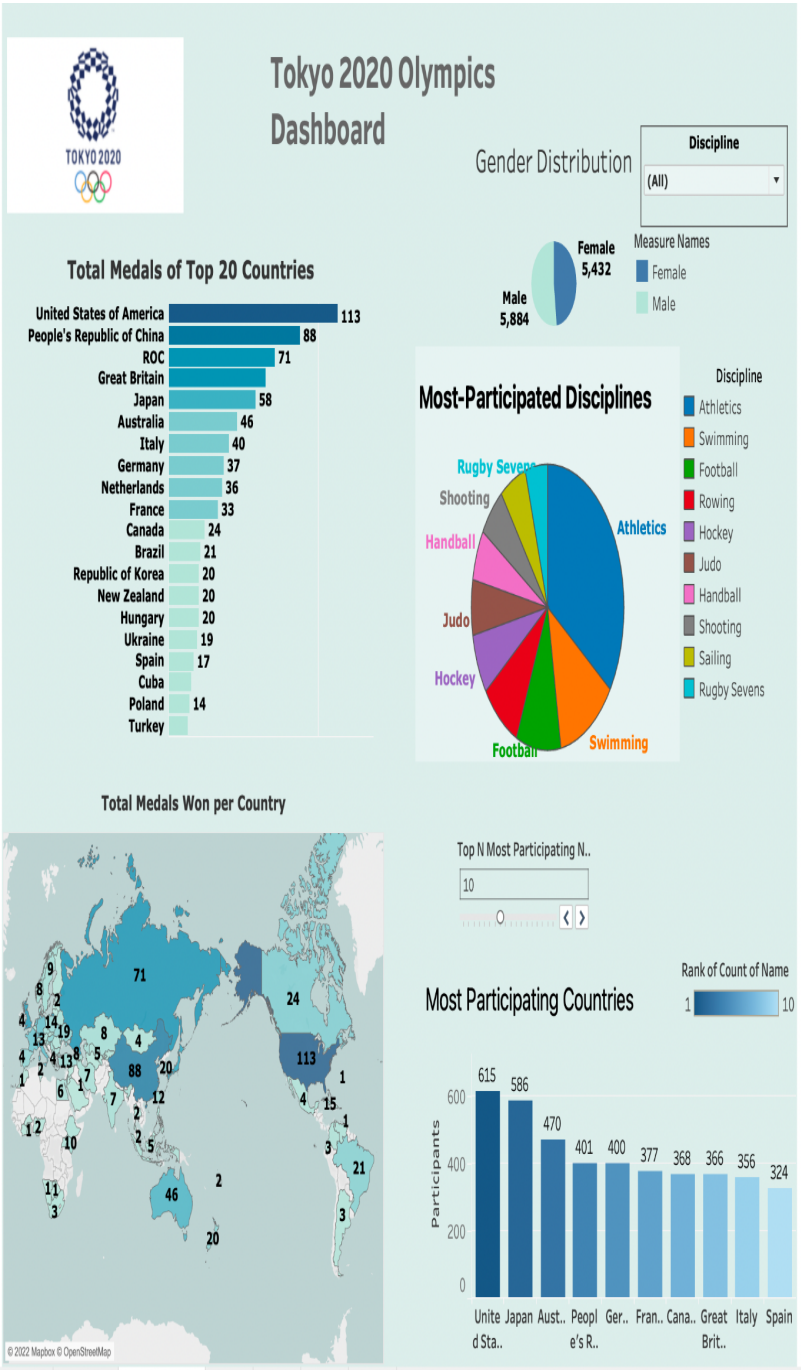
  In my visualization, the variety of encoding channels can be done by more options. And for layout, it can be cleaner and with a more dynamic display of the dataset. For the interaction between users and the product, the user experience can be improved by giving more user-friendly guides to the users.

References

[1] SARKHEL, A. P. (2021). *2021 Olympics in Tokyo*. Kaggle. https://www.kaggle.com/datasets/arjunprasadsarkhel/2021-olympics-in-tokyo

[2] “Tableau,” [Online]. Available:https://www.tableau.com/.

[3]



Sheet 1 | Sheet 2 | Sheet 3 | Sheet 4 | Sheet 2 (2) | Sheet 1 (2) | Dashboard 1 | Sheet 5 | Sheet 6 | Sheet 7 | Sheet 8