Extracted Dataset from Kaggle

Imported csv A screenshot of a computer

Description automatically generated

Transformed data by removing ID, renaming NOC to Country for readability, deleted rows where theres NA for height weight and age as these will be used for analysis   
Preservation of Data Quality:Deleting rows with missing values ensures that the remaining data is complete and of high quality. This can be crucial if the analysis heavily relies on the accuracy and completeness of the data.  
, using Replace Values, I replaced NA with participated (as it makes more sense), renamed city with Host City for clarity

Renamed M in gender to Male, F in gender to female

Removed Team to prevent confusion as there are cases of east Germany and west Germany

Before:

A screenshot of a computer

Description automatically generated

imported another excel file to join the dataset so that the full name of the country is shown, using merge queries using left join as we want to see only the names of the country in the athlete table

Renaed column 1 as Country Code and Renamed Column 2 as country name.

A screenshot of a computer

Description automatically generated  
Removed the column 3 as its not related.

Renamed USA to United States of America, UK to United Kingdom and remove errors to maintain data integrity.

After:  
  
A screenshot of a computer

Description automatically generated

Added Title Olympic Games Visualisation

**Medal distribution by country**

Used the map visualisation feature and use specific columns like country and medal to show distribution of medals.  
  
Need to reorder the medal bubbles in Gold Silver and Bronze

A map of the world

Description automatically generated

Distribution of Medals by Gender  
  
A pie chart with numbers and a percentage

Description automatically generated

Added in the Sex and Count of medals columns into this particular visualisation  
  
A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

Under details details, placed the values inside for aesthetic purposes

**Medal distribution by country**

Using the metrics table

Drag the country name to rows, medals to columns and medals to values and change it to count so that it will show the number.  
In the filtering section, filter country name by TOP N so that u can select the top 10 countries with the most number of medals and change the by value to count of medal.

Then in the medal filtering, select bronze silver and gold only as we are only concerned about medals won.

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

**Top 10 athletes of all time**

A screenshot of a computer

Description automatically generatedA chart with text and numbers

Description automatically generated

Added Athlete Name to X axis and Count of medal to Y axis.   
Turned On Data Labels to show the number of medals won by a particular athlete.  
Changed the colour of the medals to bronze silver and gold respectively for coherency  
Filtered with top N where N is 10 for the best 10 performing athletes of all time

**Top 15 medals distribution by sport**

**A chart of different sports

Description automatically generated with medium confidence**

Used conditional formation for the colours where dark Red is the highest number of medals and light red is the lowest number of colours

Turned on data labels to show the number of medals in each of the treemap box