OLYMPIC SPORTS DATASET ANALYSIS

Outcome:

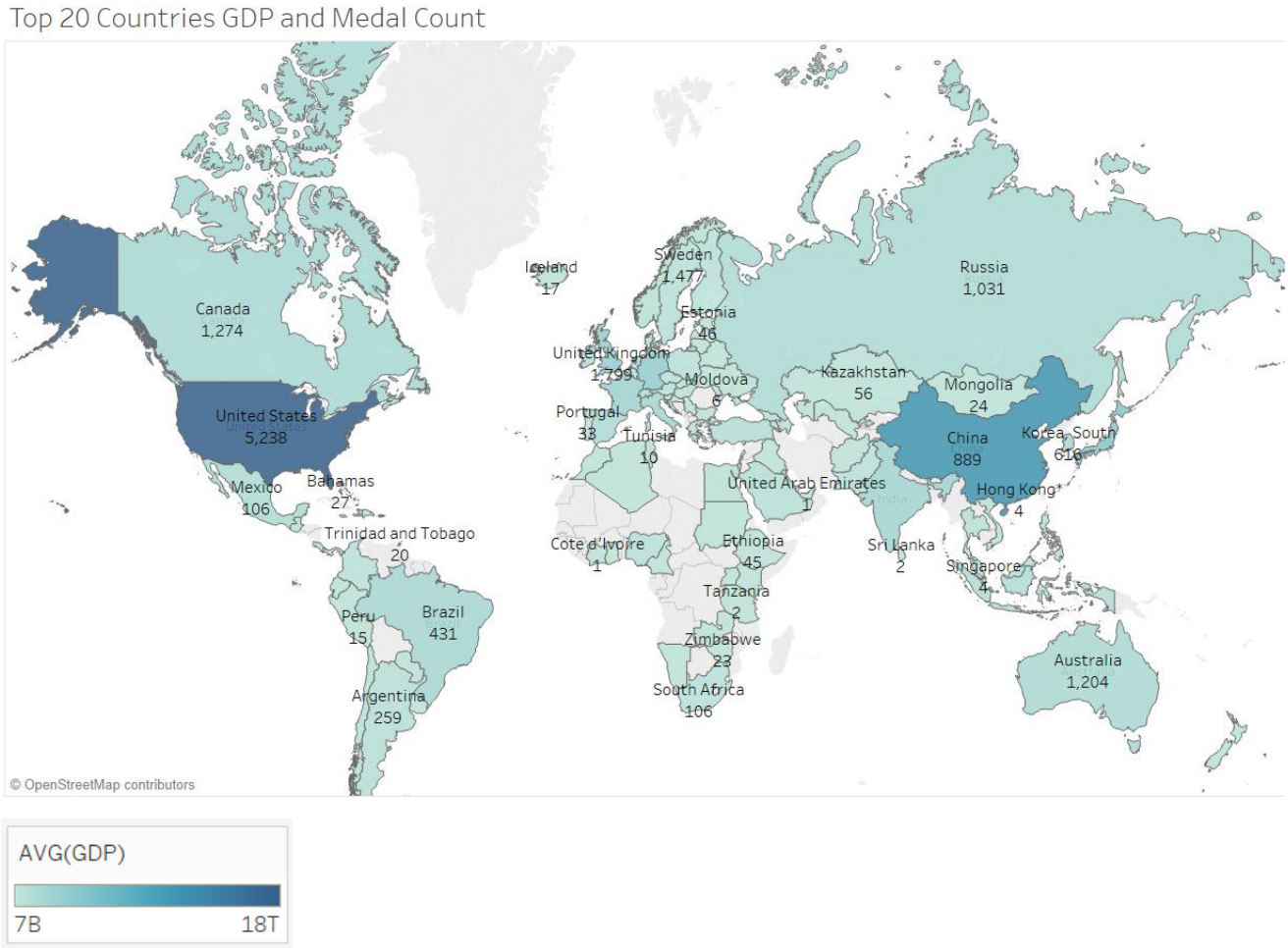
An initial manual analysis of the dataset reveals 3 outcomes from our perspective:

1. Higher GDP of the country leads to bagging more medals.
2. Minimal to zero participation from women during the beginning years.
3. Higher participation in the Summer Olympics than the Winter Olympics.

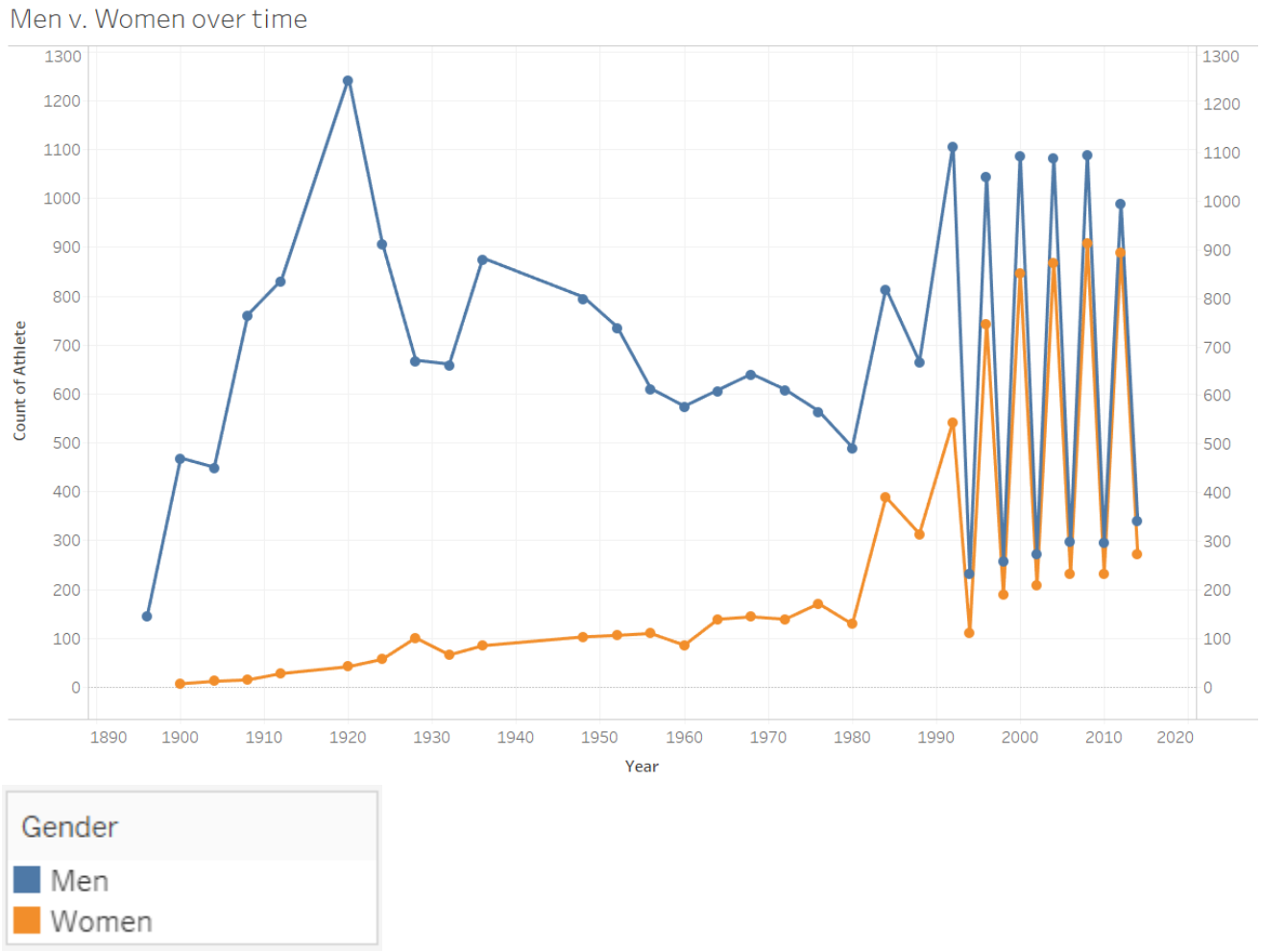
Evidence:

We used Tableau Public to visualize the data.

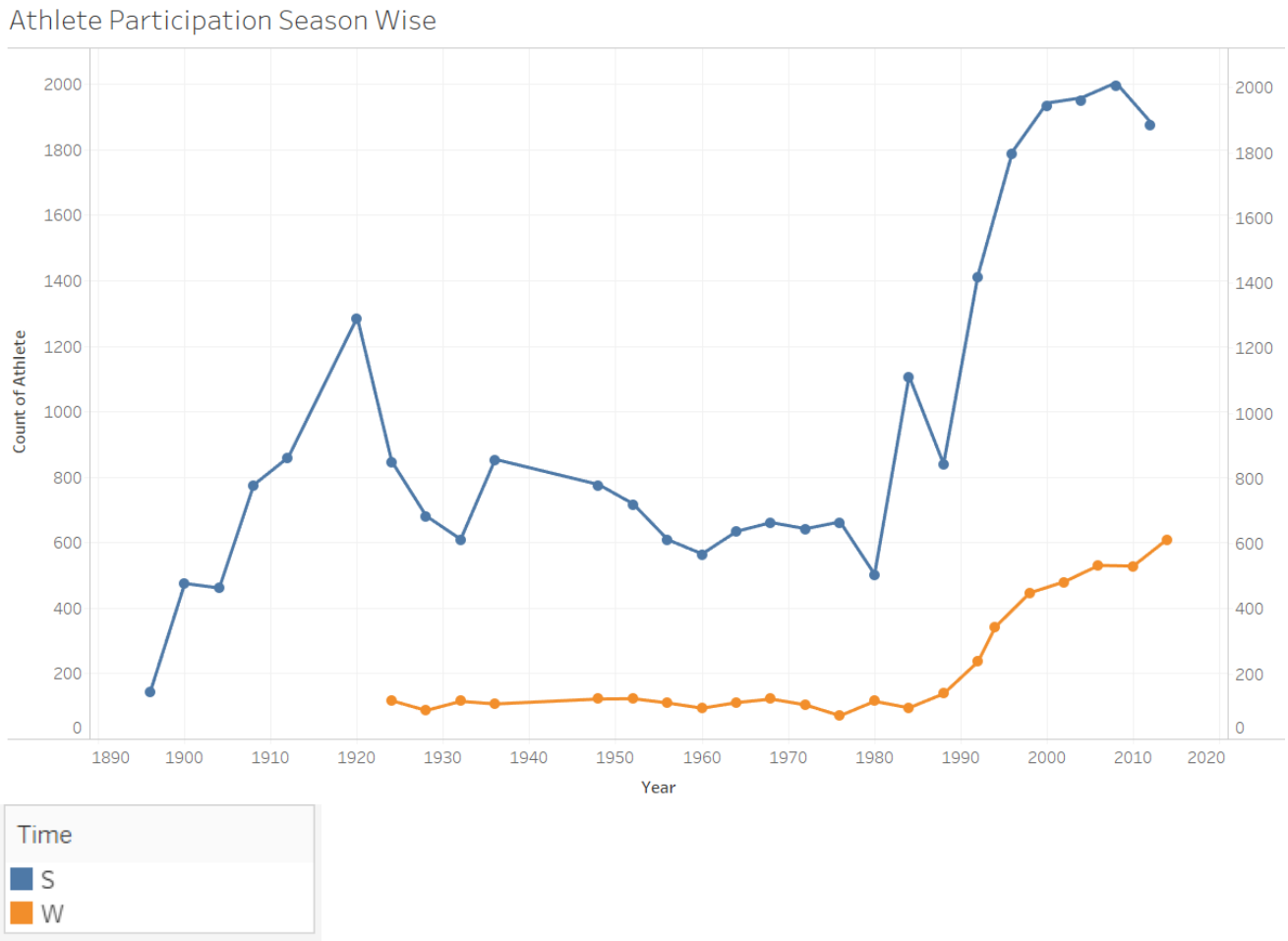
1. The first outcome we wanted to verify was the GDP and total medal count.



2. The second outcome we wanted to validate was the performance of men and women over the years, 1896 to 2014.



3. The third outcome we verified was the athlete participation during the Summer and Winter Olympics



Method:

We started by searching a sport or business-related datasets and after riffling through various Stock market and Sports datasets, we eventually decided on the dataset of Olympics. The dataset we are using has been taken from Kaggle and has been provided by the IOC Research and Reference Service and published by The Guardian's Datablog. It describes which countries and athletes have won the most medals at Olympics from 1896 – 2014 cleaved in three subsets as follows:

- Dictionary:
  a) Country
  b) Country Code
  c) Population of that Country
  d) GDP per Capita

- Summer:
  a) Year of Olympics
  b) City where Olympics was held
  c) Sport played
  d) Discipline of the corresponding Sport
  e) Athlete
  f) Country the athlete represented
  g) Gender
  h) Event (single or pairs)
  i) Medal

- Winter:
  a) Year of Olympics
  b) City where Olympics was held
  c) Sport played
  d) Discipline of the corresponding Sport
  e) Athlete
  f) Country of the athlete represented
  g) Gender

h) Event (single or pairs)
i) Medal

We, then proceeded to load all the subsets in RStudio and combined them into one file to effectively create visualizations in Tableau. RStudio was also used to get a count of number of events in the Summer and Winter Olympics. We then proceeded to answer our first question:

1. Is GDP of a country relevant to the performance of their athletes?

The Gross Domestic Product (or GDP) of countries is one of the primary indicators used to gauge the health of country's economy. Higher the GDP, better the economy. To derive an answer for this, we loaded our final combined dataset in Tableau and selected to view the data in maps format as we needed GDP and athlete comparisons by Country. We added the Longitude and Longitude in the sheet and color marked the countries, eventually changing the color charts from darker to lighter corresponding to descending GDP per capita. Post that, we put the dimension of medals in text and adjusted the label by Counts which led us to understand which country had won the greatest number of medals as well.

2. Did Women Participate in the same ratio as men since the Olympics began?

Our data logs information about Olympics from 1896 to 2014, i.e. from the first modern Olympics that was held in Greece. We expect to see minimal to zero performance from women as they were probably not allowed to participate at the time. For this evaluation we plotted the total count of Gender participating in the Olympics over time. We also color marked the Gender for easy reading. This showed us that women, although started participating only 4 years later, participated in very low numbers. But as time went on and female participation was normalized (and encouraged) the numbers weren't far apart.

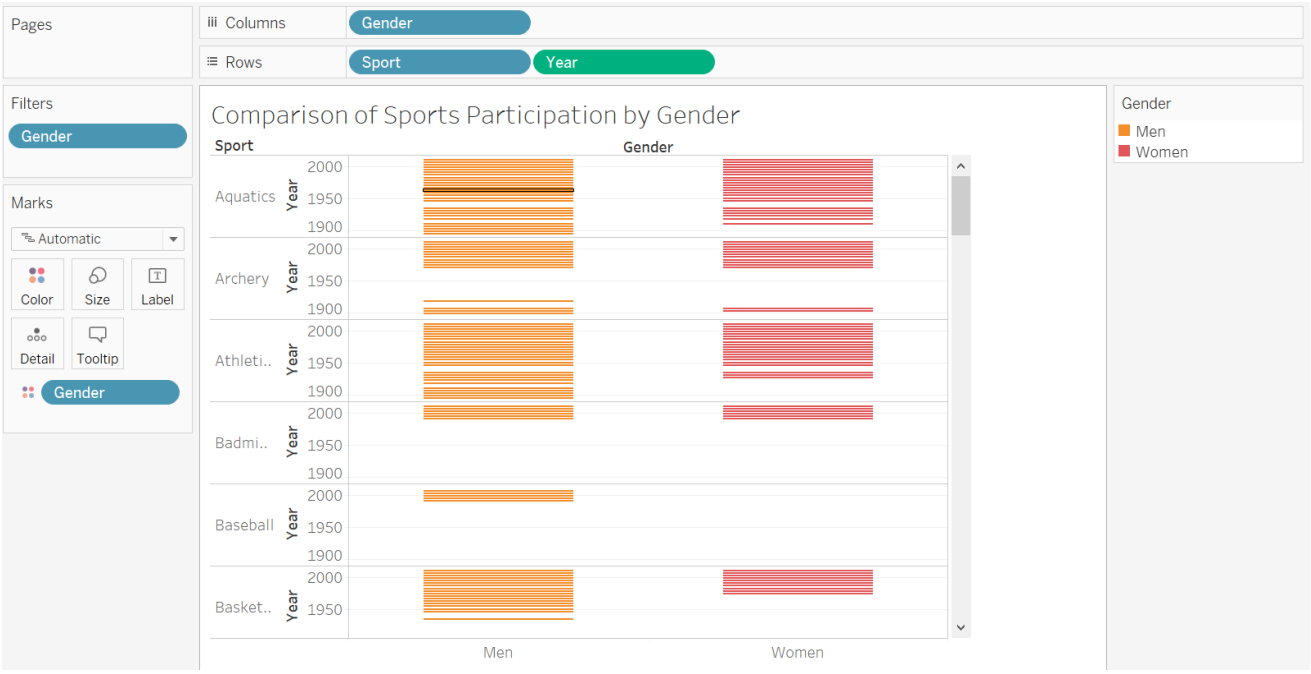
3. Is Summer Olympics more preferred than Winter?

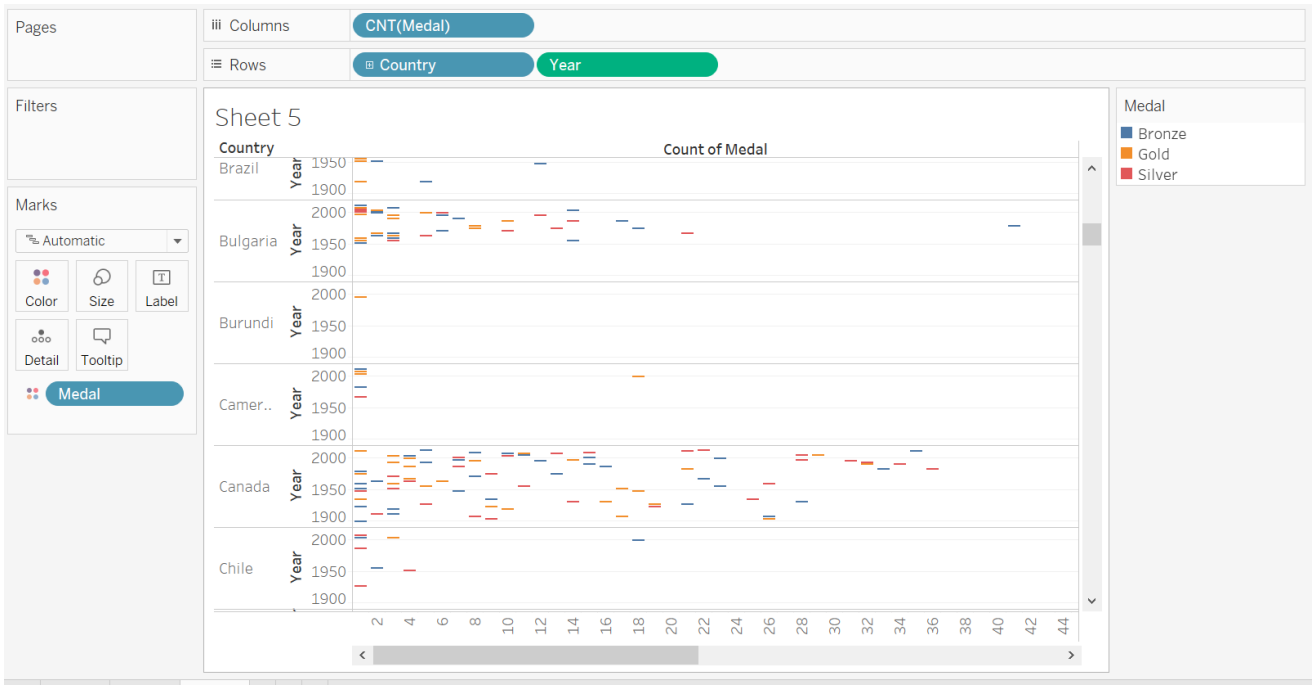
Using the loaded dataset in RStudio, we took counts of events in Summer and Winter Olympics and found that Summer Olympics has 666 unique events while Winter Olympics has only 83 unique events. To get a better understanding of the data, we then went on to create an athlete participation graph in Tableau where we brought in the

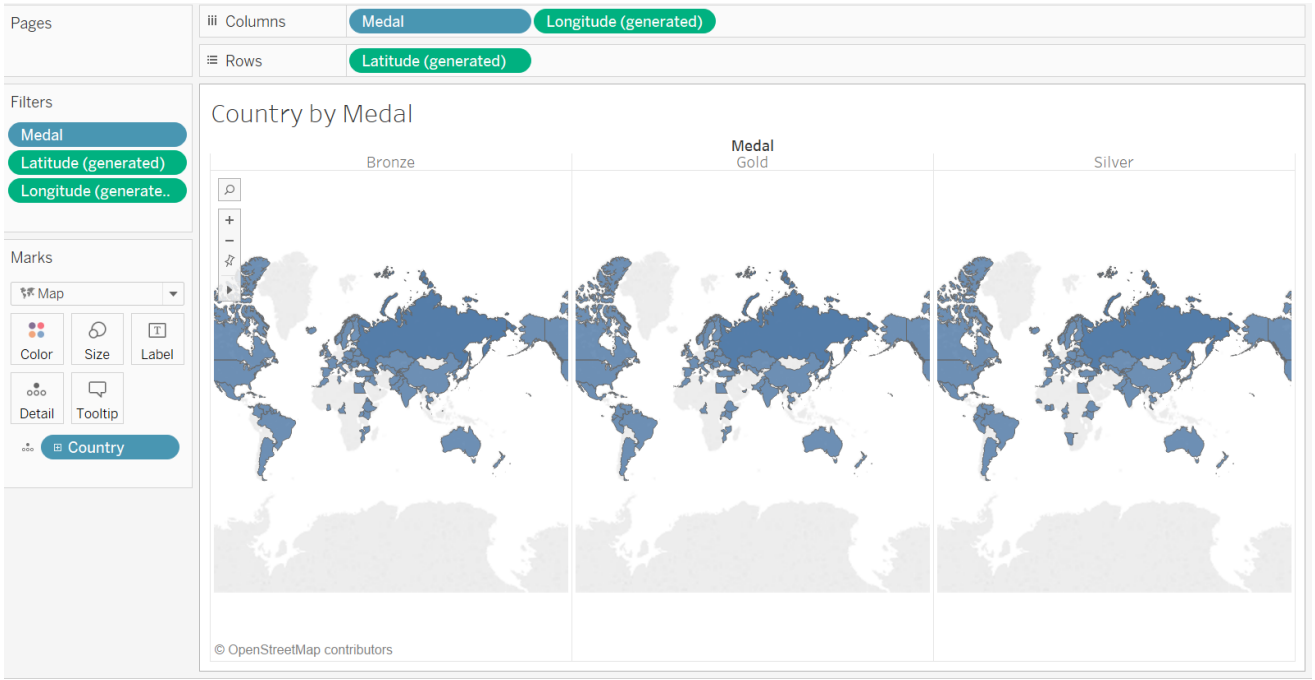
events, years and counts of athletes to eventually conclude that athlete participation is much higher in summer Olympics. Another reason why we might expect more participation and events is because Winter Olympics requires the presence of snow and ice, which is not found in many countries hence, athletes wouldn't be able to practice for it as easily as they can for the Summer Olympics.

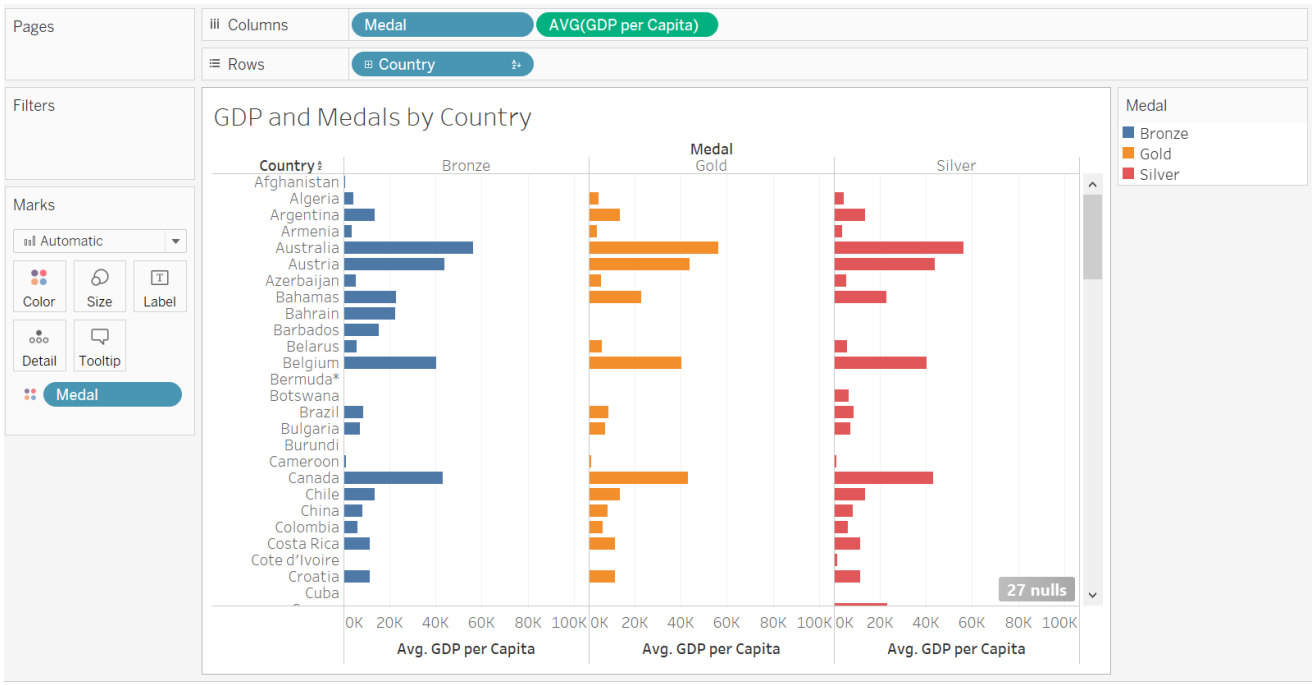
Through this process, we utilized both RStudio and Tableau for evaluations, data visualizations and interpretations. Attached below are remaining visualizations which were crucial to our understanding of the data:

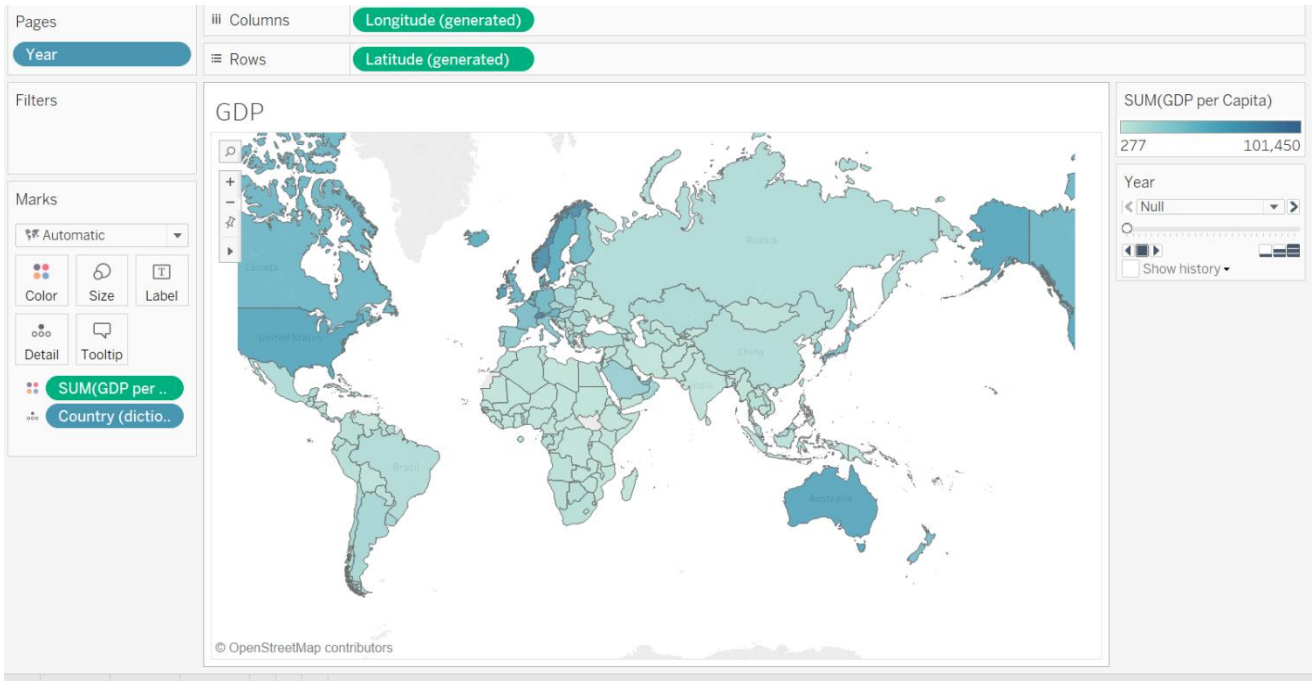
Snapshots:

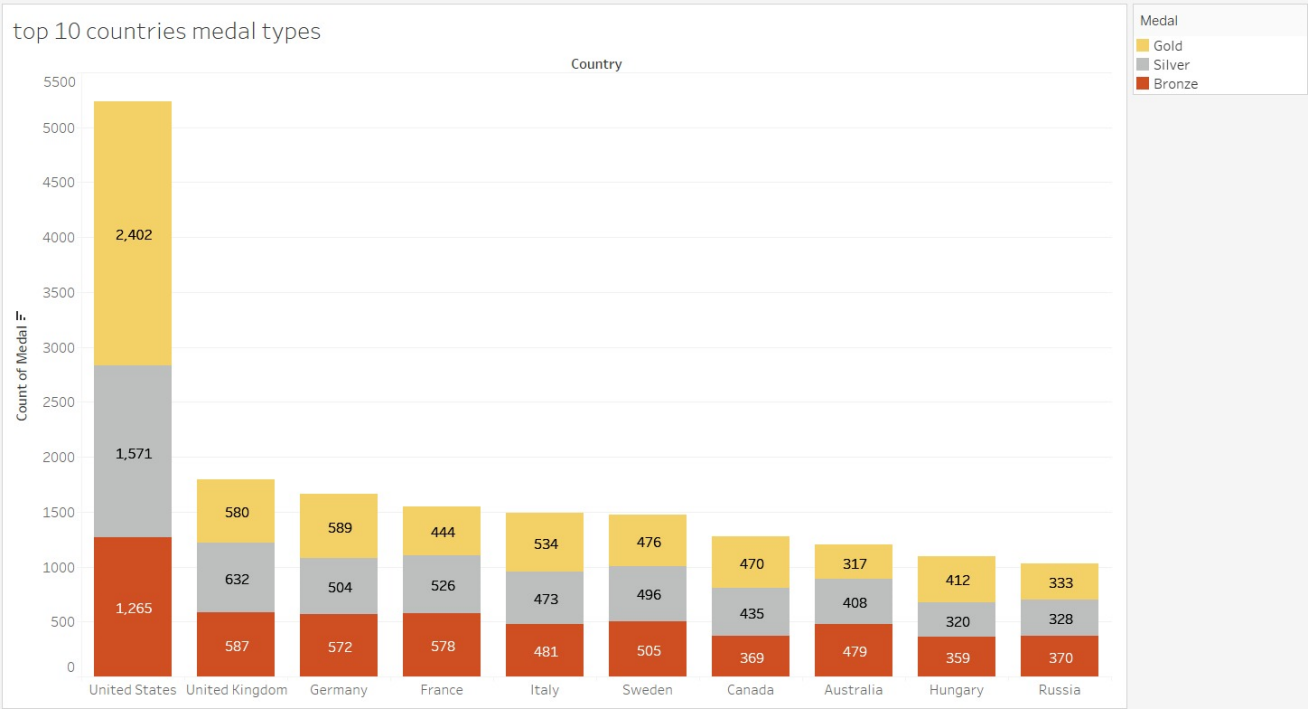


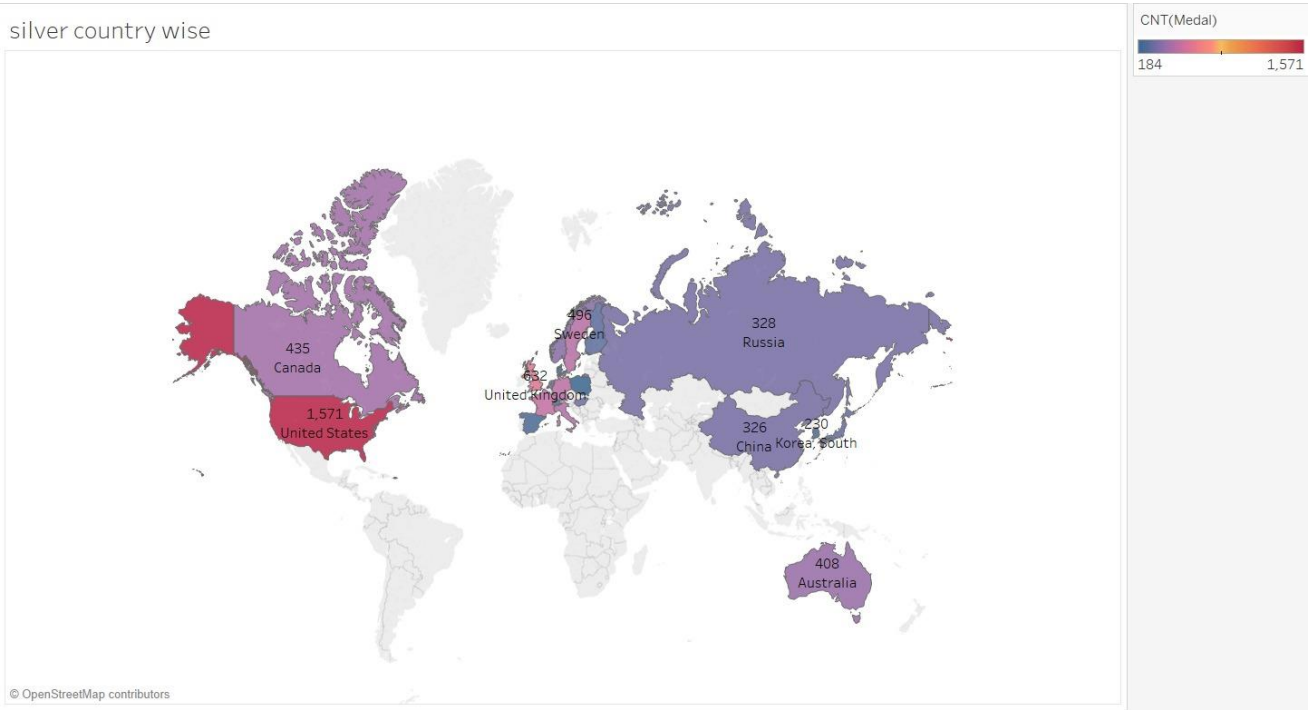

Pages
Columns
CNT(Medal)
Rows
Country
Year
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Medal
Sheet 5
Country
Count of Medal
Brazil
Bulgaria
Burundi
Camer..
Canada
Chile
Year
1900
1950
2000
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44
Medal
Bronze
Gold
Silver

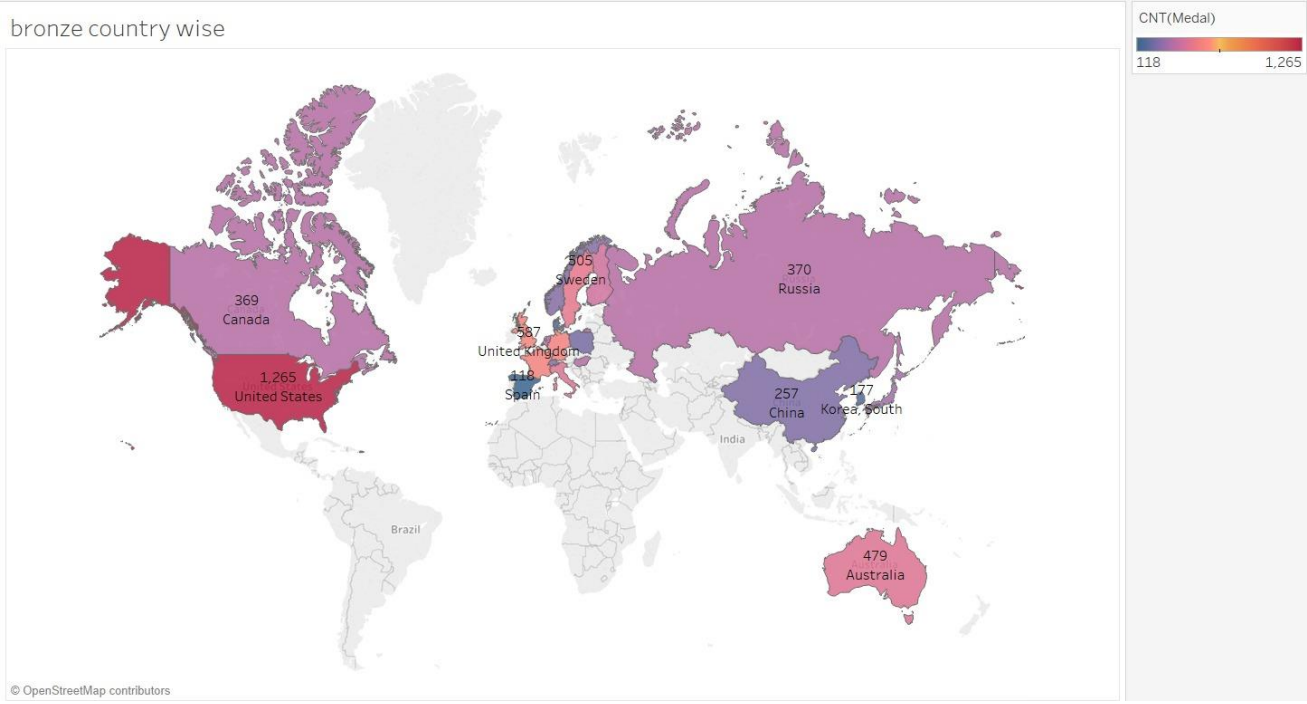

Pages
Columns
Medal
Longitude (generated)
Rows
Latitude (generated)
Filters
Medal
Latitude (generated)
Longitude (generate..
Marks
Map
Color
Size
Label
Detail
Tooltip
Country
Country by Medal
Medal
Bronze
Gold
Silver
© OpenStreetMap contributors

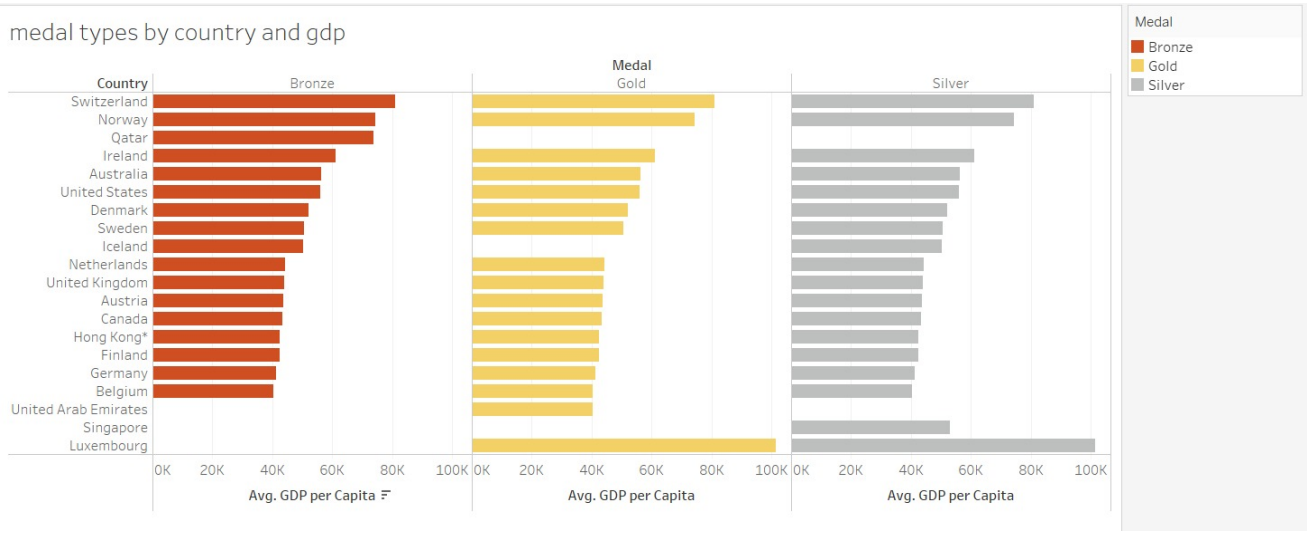

Pages
Columns
Medal
AVG(GDP per Capita)
Rows
Country
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Medal
GDP and Medals by Country
Medal
Bronze
Gold
Silver
Country
Afghanistan
Algeria
Argentina
Armenia
Australia
Austria
Azerbaijan
Bahamas
Bahrain
Barbados
Belarus
Belgium
Bermuda*
Botswana
Brazil
Bulgaria
Burundi
Cameroon
Canada
Chile
China
Colombia
Costa Rica
Cote d'Ivoire
Croatia
Cuba
27 nulls
0K 20K 40K 60K 80K 100K
Avg. GDP per Capita

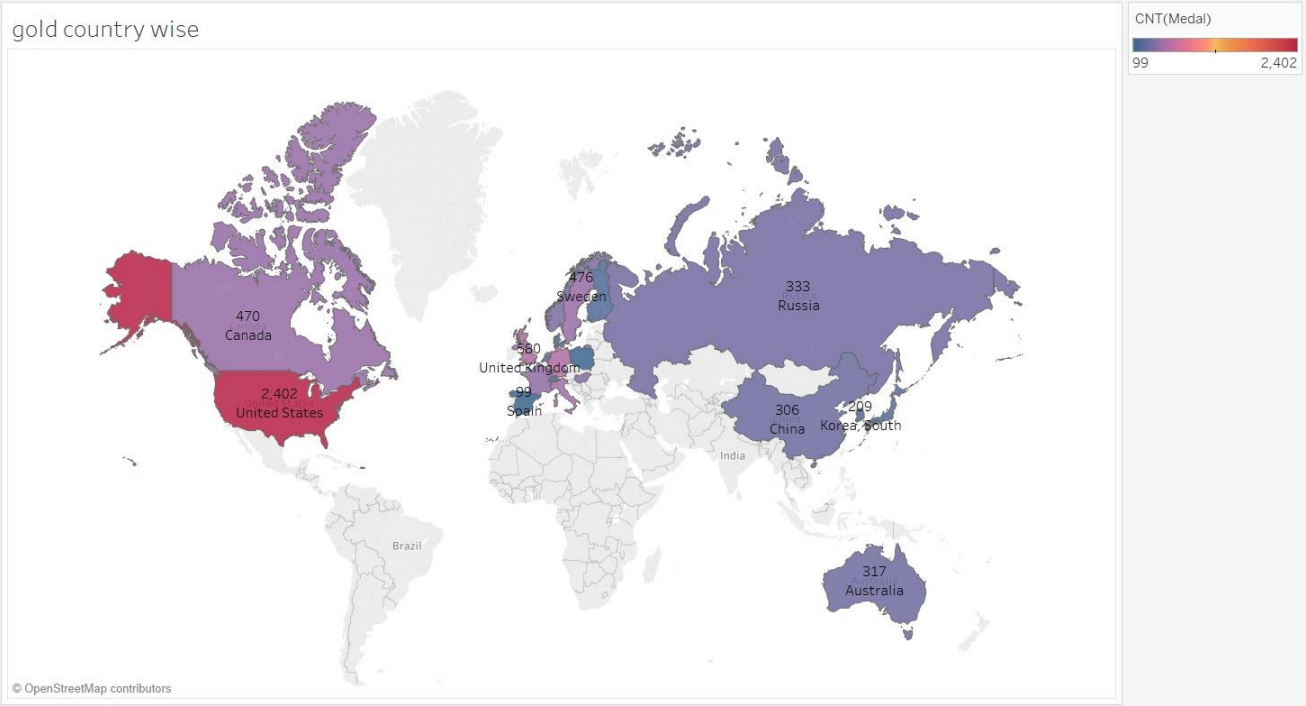

Pages
Year
Columns
Longitude (generated)
Rows
Latitude (generated)
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
SUM(GDP per ..
Country (dictio..
GDP
SUM(GDP per Capita)
277
101,450
Year
Null
Show history
© OpenStreetMap contributors


top 10 countries medal types
Country
Count of Medal
Medal
Gold
Silver
Bronze
United States
United Kingdom
Germany
France
Italy
Sweden
Canada
Australia
Hungary
Russia
2,402
1,571
1,265
580
632
587
589
504
572
444
526
578
534
473
481
476
496
505
470
435
369
317
408
479
412
320
359
333
328
370


silver country wise
CNT(Medal)
184
1,571
435
Canada
1,571
United States
496
Sweden
632
United Kingdom
328
Russia
326
China
230
Korea, South
408
Australia
© OpenStreetMap contributors


bronze country wise
CNT(Medal)
118
1,265
369
Canada
1,265
United States
505
Sweden
370
Russia
587
United Kingdom
118
Spain
257
China
177
Korea, South
India
Brazil
479
Australia
© OpenStreetMap contributors


medal types by country and gdp
Medal
Bronze
Gold
Silver
Country
Switzerland
Norway
Qatar
Ireland
Australia
United States
Denmark
Sweden
Iceland
Netherlands
United Kingdom
Austria
Canada
Hong Kong*
Finland
Germany
Belgium
United Arab Emirates
Singapore
Luxembourg
0K 20K 40K 60K 80K 100K
Avg. GDP per Capita


gold country wise
CNT(Medal)
99
2,402
476
Sweden
333
Russia
470
Canada
280
United Kingdom
2,402
United States
99
Spain
306
China
209
Korea, South
India
Brazil
317
Australia
© OpenStreetMap contributors