DATA VISUALISATION CIA 2

TABLEAU ANALYSIS OF THE 2021 TOKYO OLYMPICS

SUBMITTED BY

1920181 ROHIT DEVANABOINA
1920226 SARVESH AGARWAL
1920449 SONAKSHI MASSEY
1920557 FEBIN RICHARD
1920642 PRACHI GIRIA

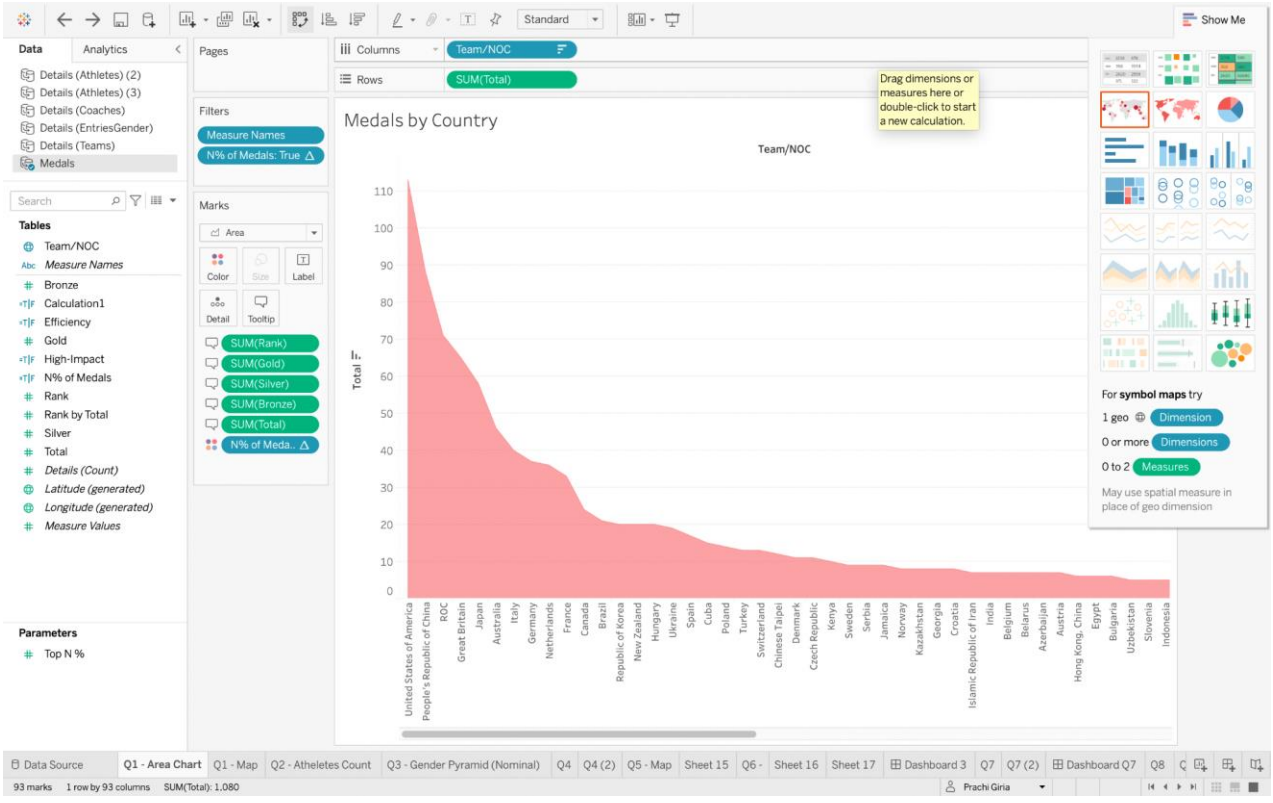
INTRODUCTION TO THE DATASET

This contains the details of over 11,000 athletes, with 47 disciplines, along with 743 Teams taking part in the 2021(2020) Tokyo Olympics. This dataset contains the details of the Athletes, Coaches, Teams participating as well as the Entries by gender. It contains their names, countries represented, discipline, gender of competitors, name of the coaches.

Link to Dataset: [2021 Olympics in Tokyo]

QUESTIONS & ANSWERS

1. How did each country fare in terms of Medals won?



Variables Used:

Sum Total of medals, Rank, Rank Total, Bronze, Silver, Gold, Name of country(NOC)

Calculations done:

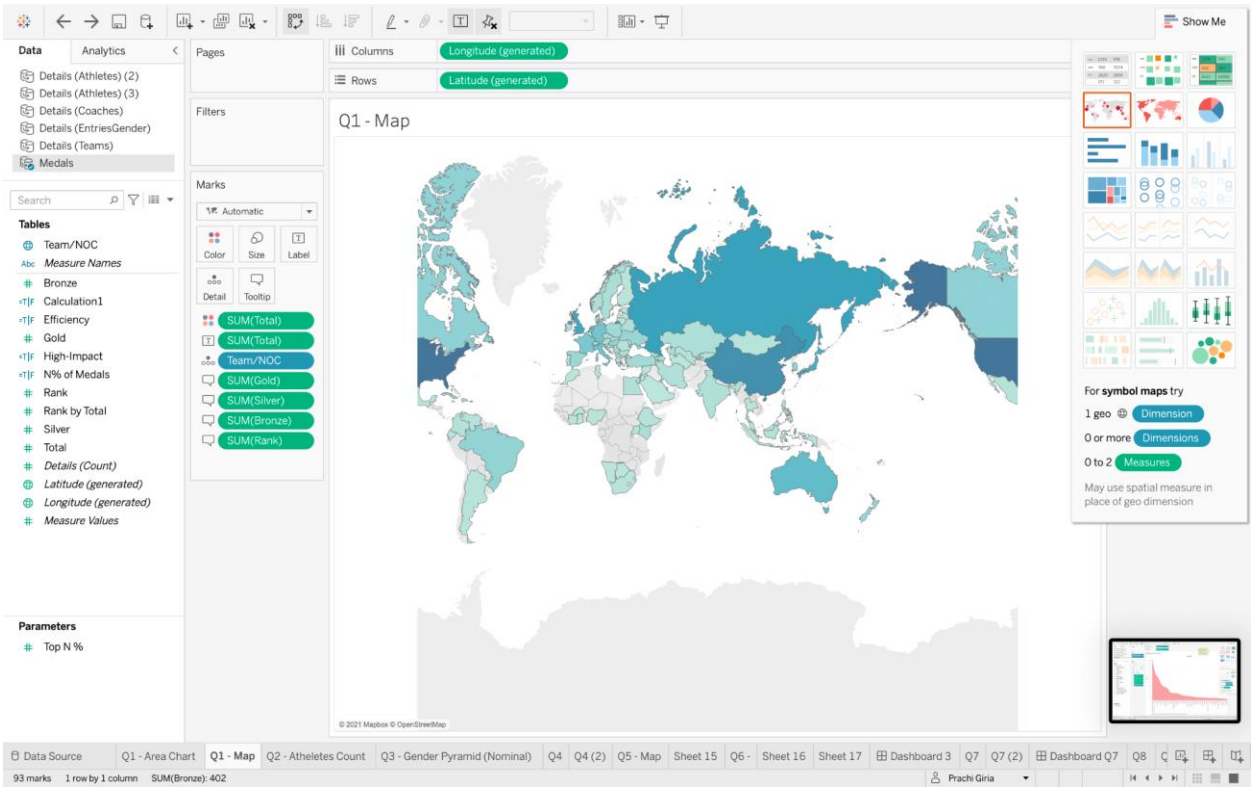
1. Rank Percentile (sum(total))>=1- (Top N%)

2. Added the exceptions of ROC as Russia and Chinese Taipei as Taiwan

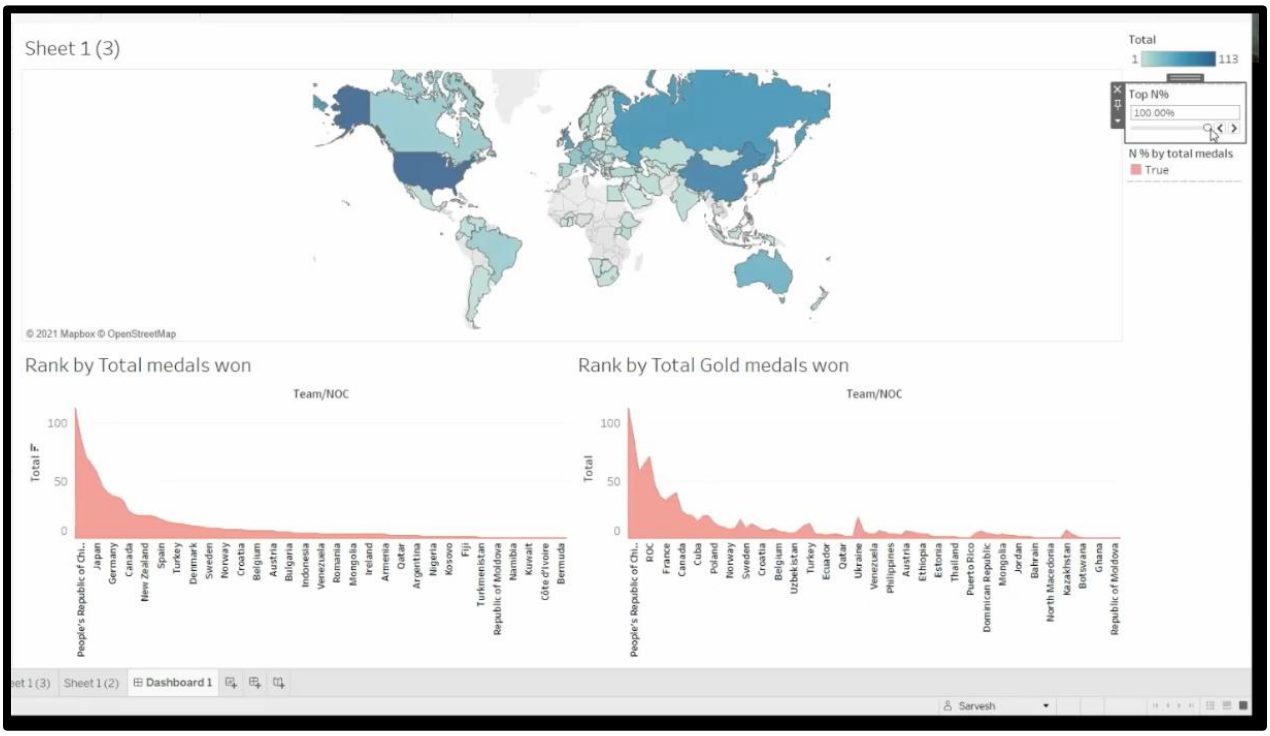
Key Observations:

1. Most of the medals are won by the top 15%-20% of countries
2. Ranking is based on total gold medals won and not only total medals won and hence the curve had some sharp edges which represented the distortion.
3. The Worldwide view shows how many countries participated and how many won how many medals. It is observed that USA, JAPAN, and CHINA remain with highest number of medals

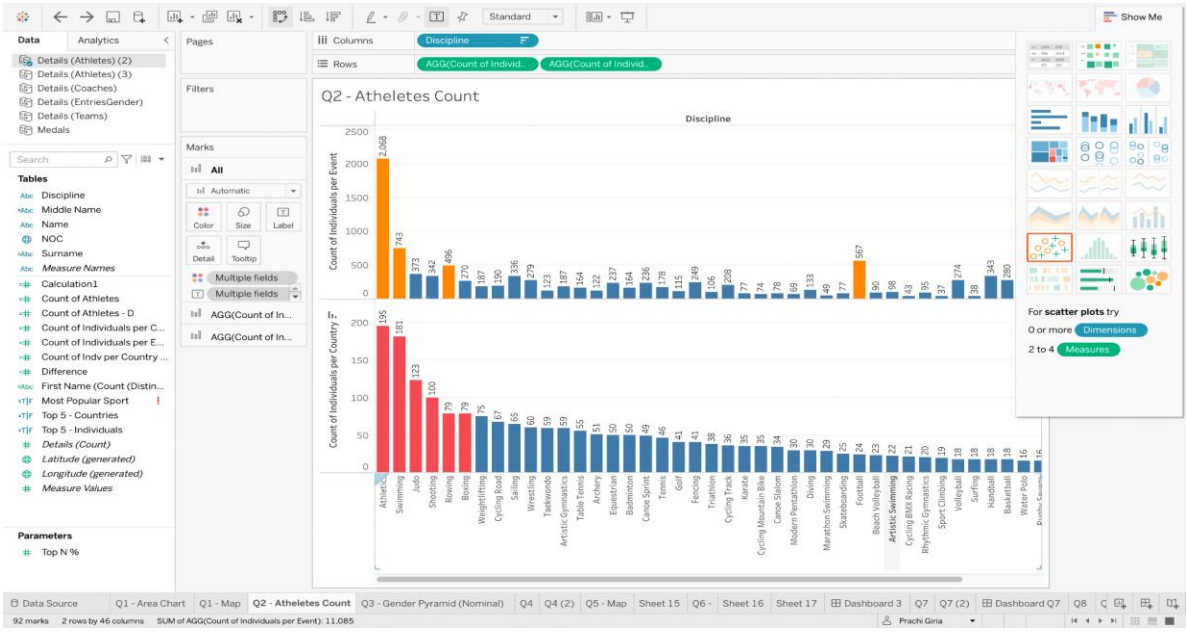
Another way of answering this question is by representing the countries on maps:



Dashboard for Question 1



2. Which events were the most popular? Highlight the top 5 most popular events by Country & Individual participation?



Variables Used:

Discipline, Athletes Count, NOC count

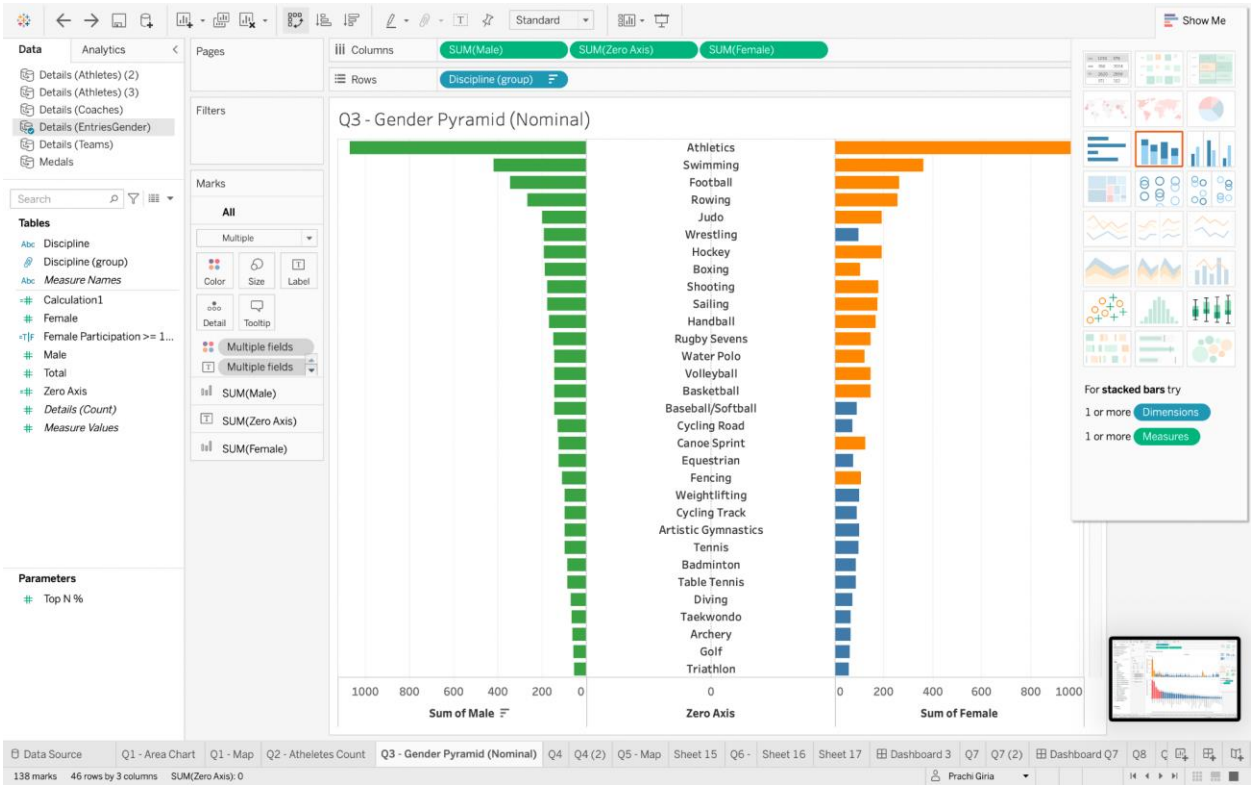
Calculations done:

1. Rank (Athletes count)<6
2. Rank (NOC count)<6

Key Observations:

1. The number of participants in team games is higher even though fewer countries have participated in it.
2. The highest participation is in Athletics and Swimming, both in terms of Country and Individual Participation

3. What was the Gender Distribution like for all events? | Highlight events with more than 100 female participants.



Variables Used:

Discipline, Count of Male, Count of Female

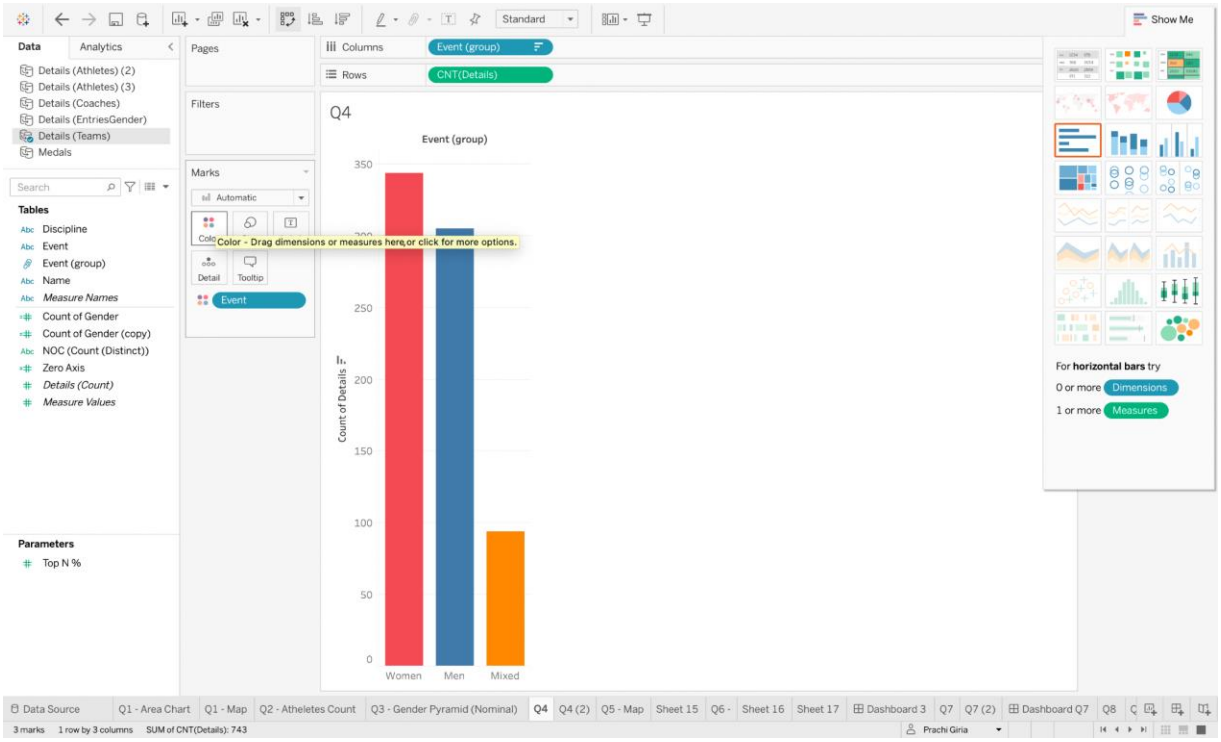
Calculations done:

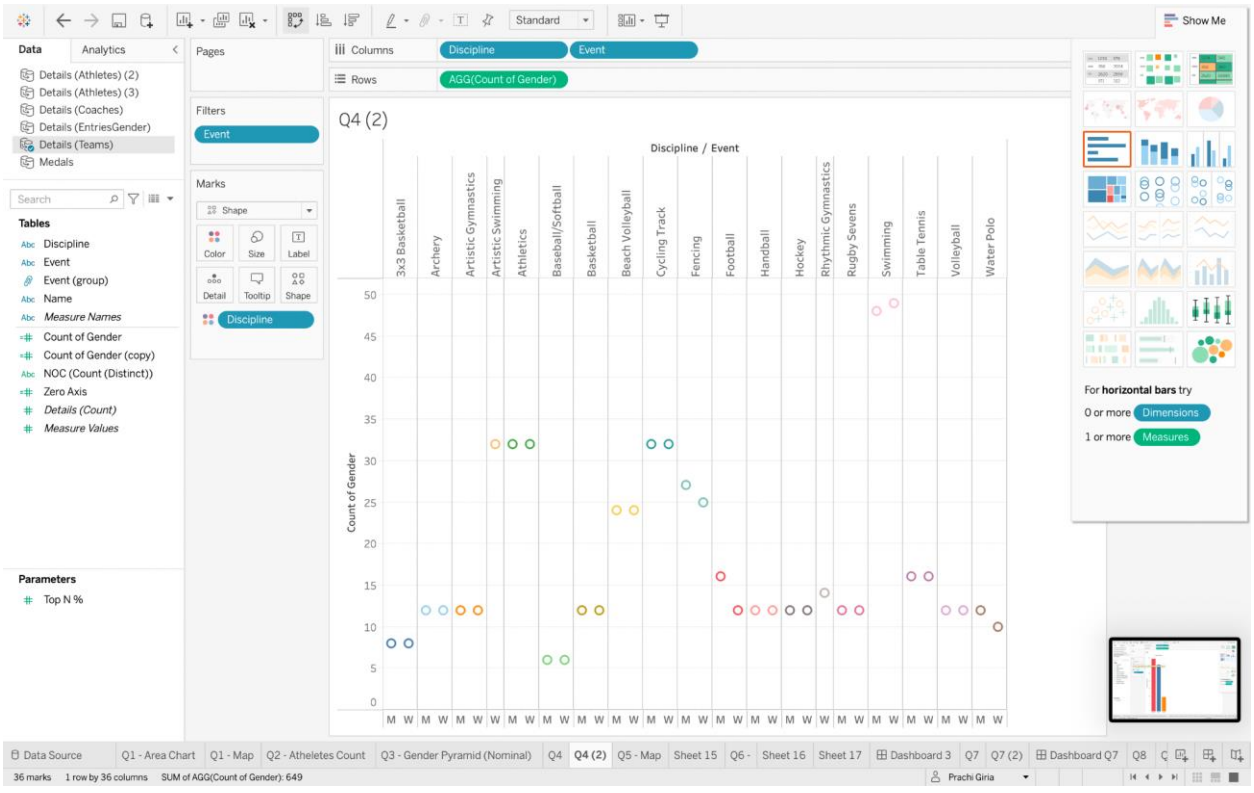
1. Zero Axis: 0
2. Female Participation > 100: sum([Female]) > 100

Key Observations:

1. Male-Female gender gaps are negligible in a majority of cases, with a few exceptions
2. Wrestling and other combative sports are still highly Male-dominated
3. Performance art type sports like Artistic Swimming and Rhythmic Gymnastics are restricted to Female participants

4.What was the Gender Distribution like for Team events? | Highlight Female Oriented events (Female Participation > Male Participation)





Variables Used:

Discipline, Event, Count of Gender (Men/Women)

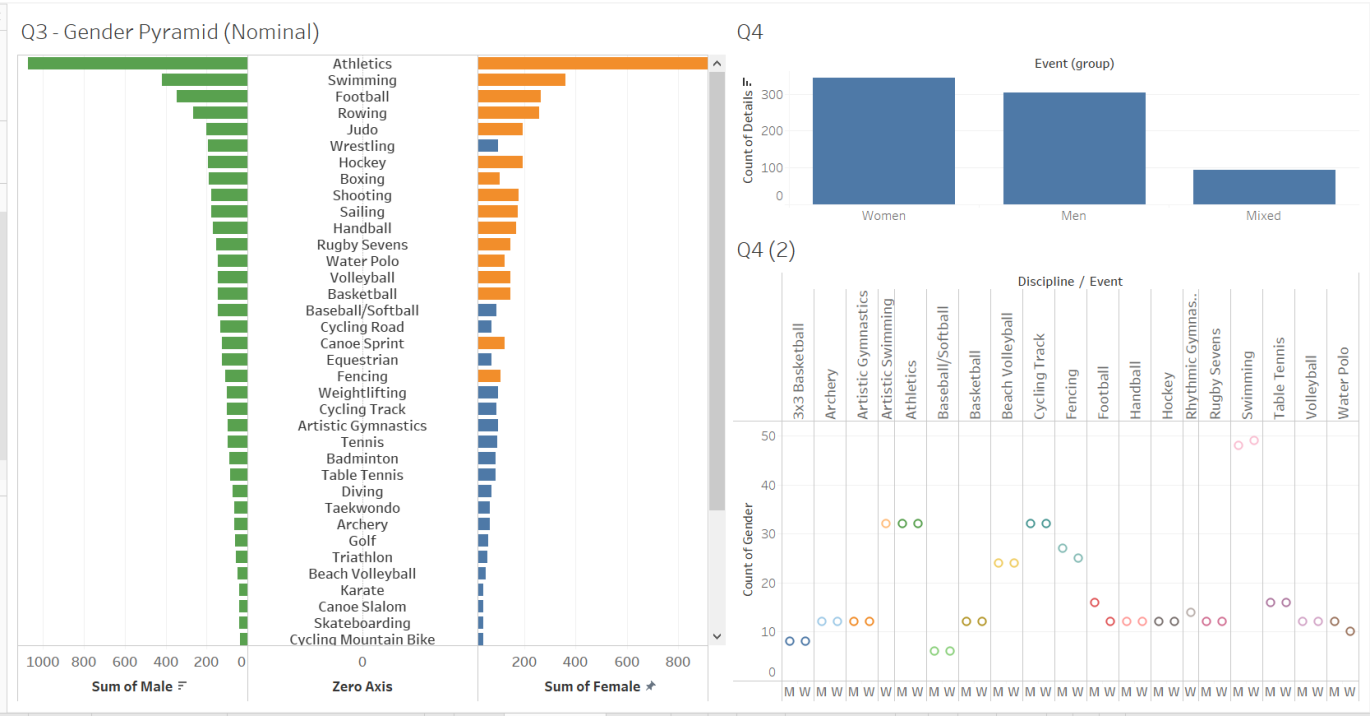
Calculations done:

1. Count of Gender: count ([Event])

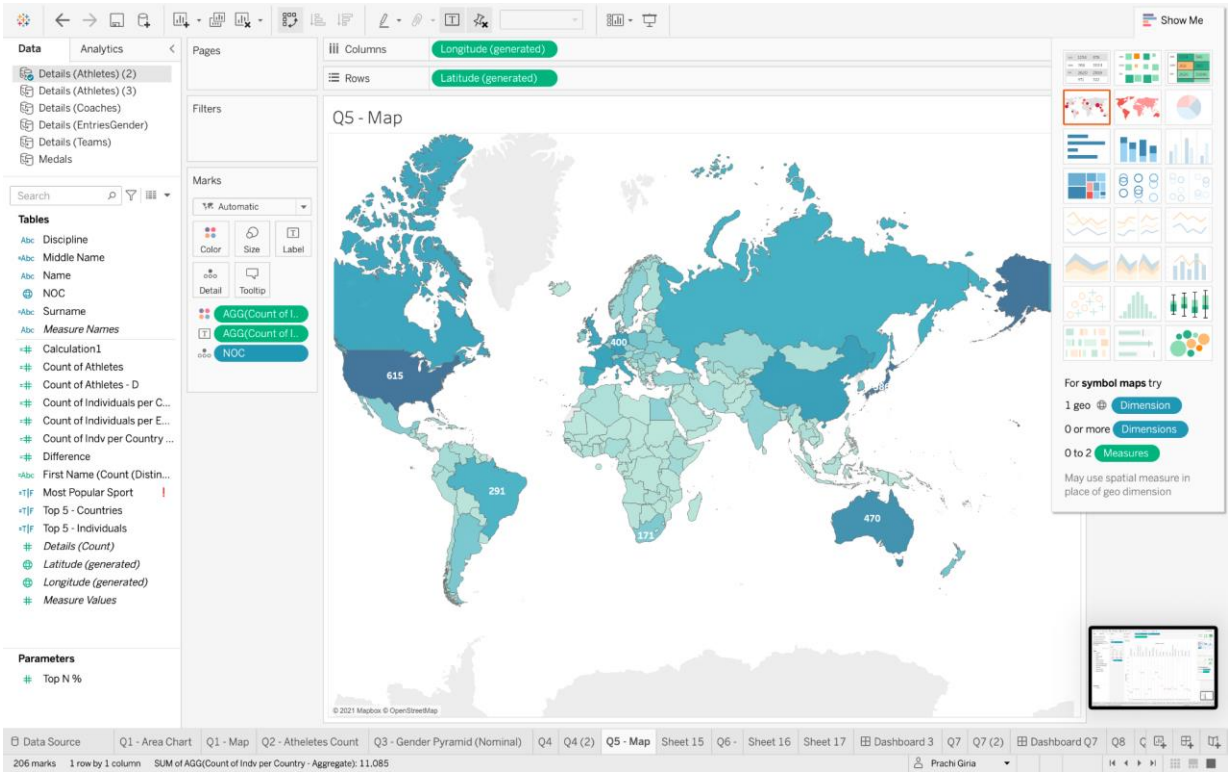
Key Observations:

1. Swimming is the only Mixed Event where FEmale participation is greater than male participation
2. A majority of events are neutral - same level of participation from both genders
3. Fencing, Football and Water Polo are the only team events with Male Participation > Female Participation

Dashboard for Q3 & Q4



5. What was participation like from the various countries? Highlight the country with the most participation in each continent.



Variables used:

Longitude, Latitude, NOC (Name of countries), Count of Individuals per Country

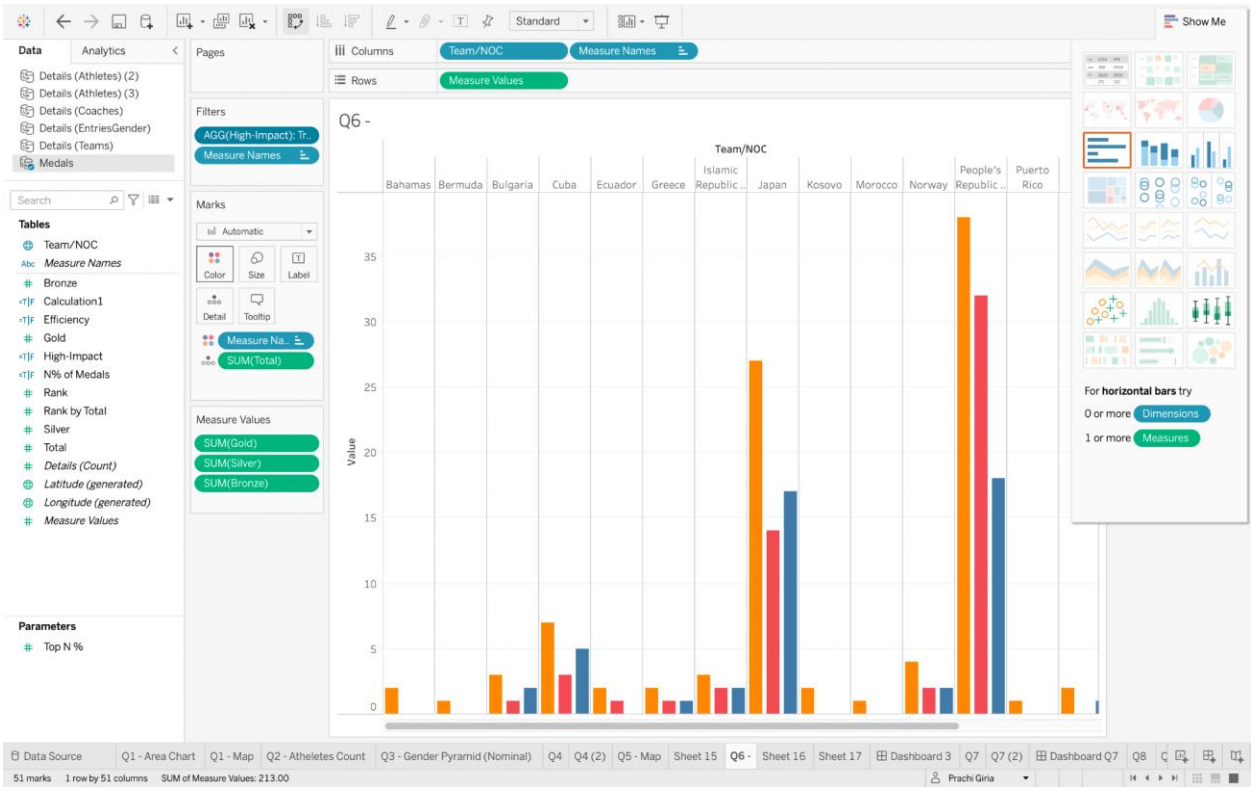
Calculations done:

1. COUNT([NOC])

Key Observations:

1. Countries with more medals tended to have higher participation
2. A large proportion of participants came from the top 20% of countries
3. Japan had an above-average participation level as it was the host nation

6. What countries had a "high-impact" performance? (Golds > Silver/Bronze)



Variables used:

Team/NOC (Name of Countries), Measure Names, Measure Values, Total(Sum)

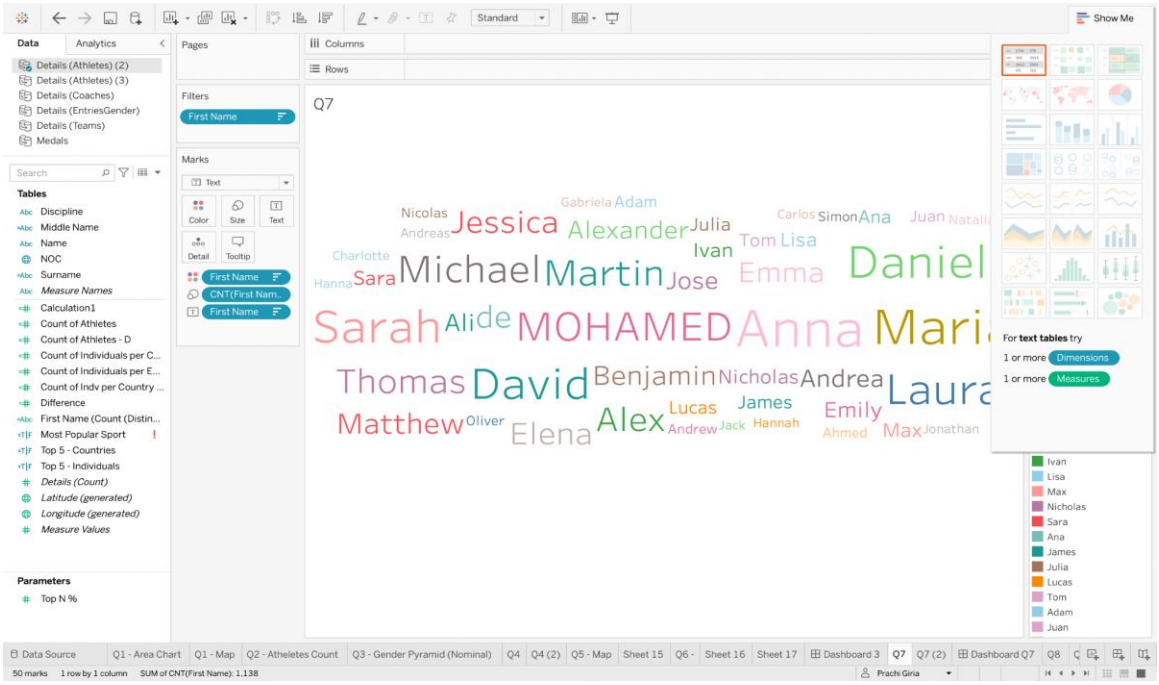
Calculations done:

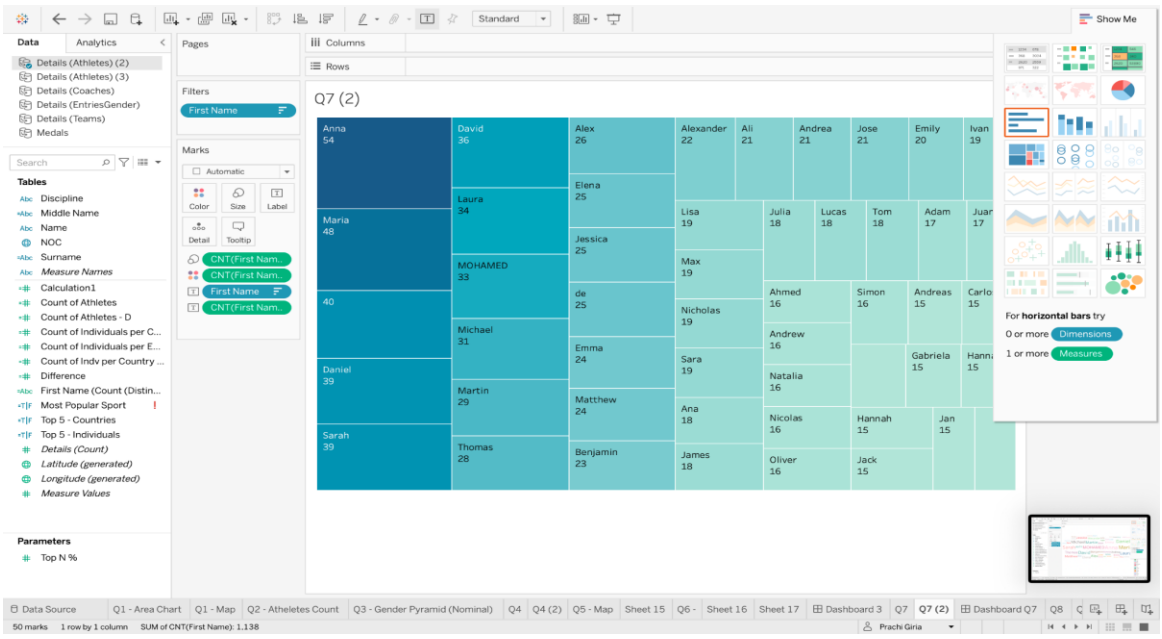
1.SUM([Gold]) > SUM([Silver]) AND sum([Gold]) > sum([Bronze])

Key Observations:

1. The top performers - USA & JAPAN - were high-impact nations despite the fact that they were also leaders in terms of total medals
2. The high impact nations tend to be smaller than a majority of Olympic Nations - this could indicate a strategy of focusing on dominating a few select events by pooling all resources into that event
    a. Example of this would be Jamaica in Running Events

7)What are the most common names at the Olympics?





Variables Used:

Names: First Name

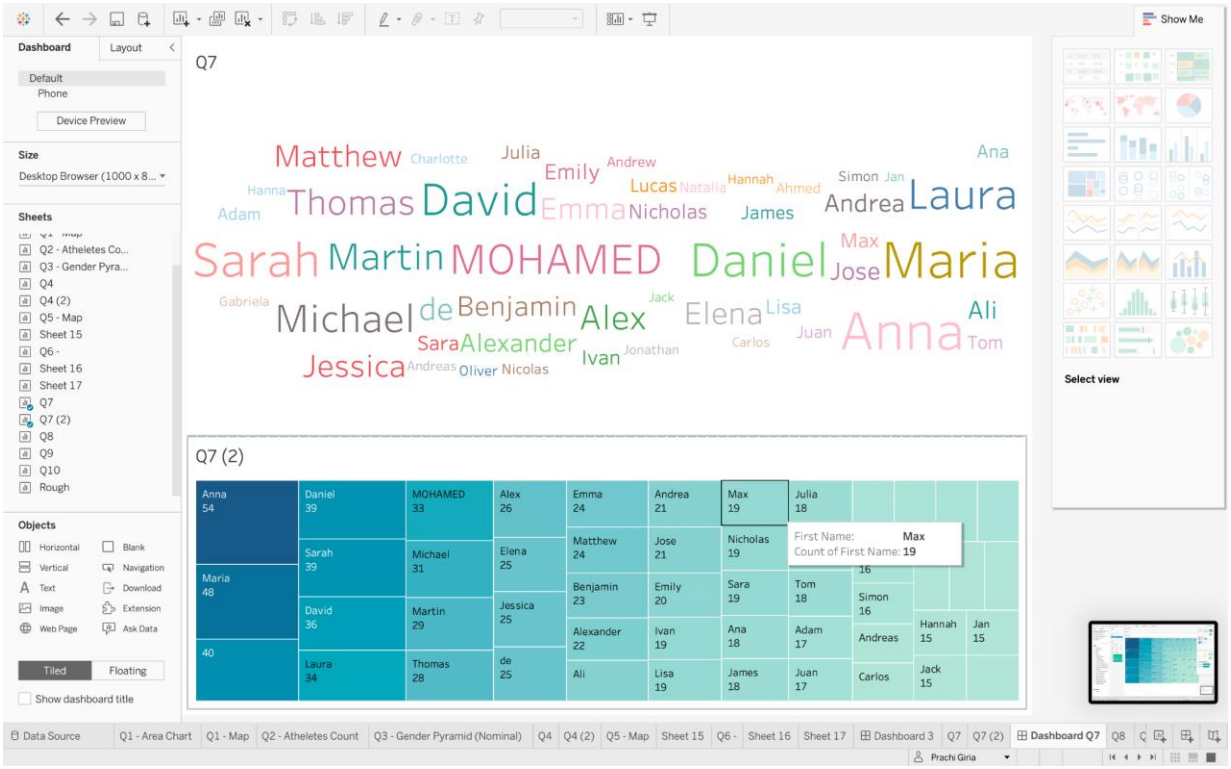
Calculations done:

count([name])

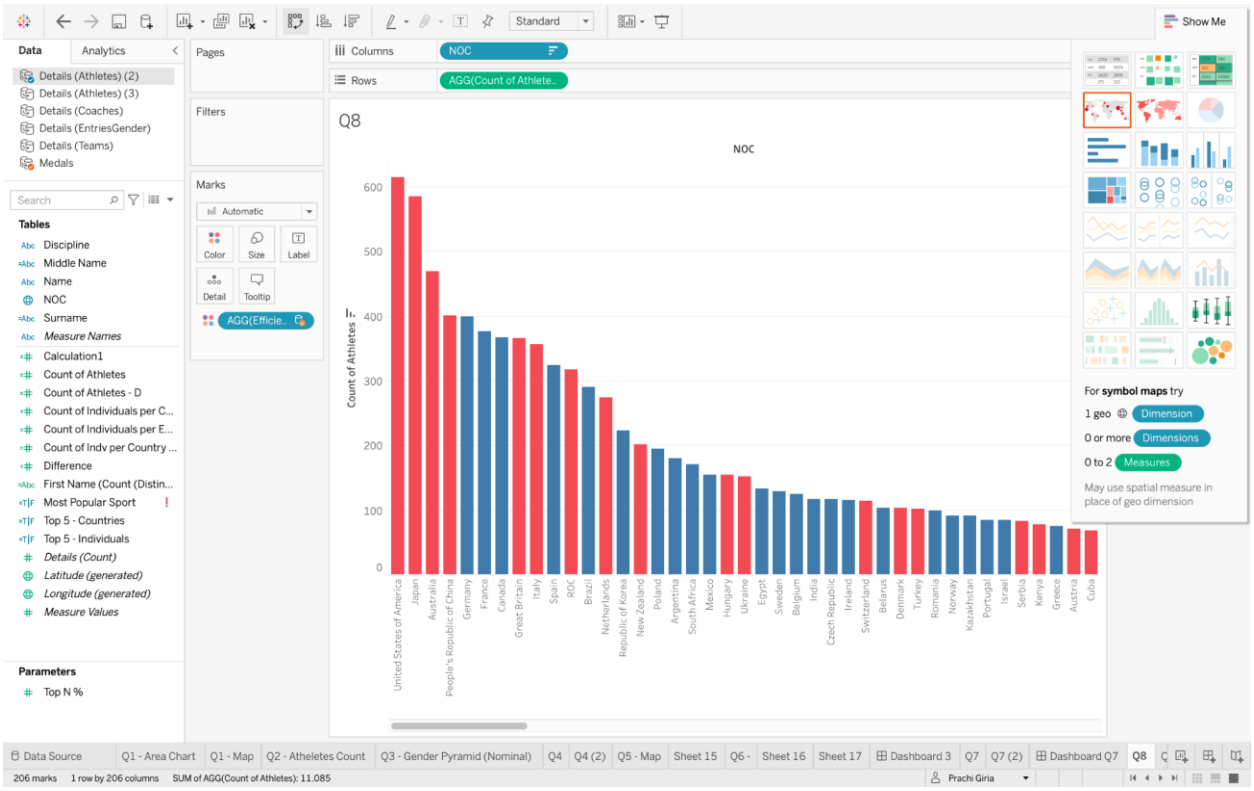
Key Observations:

1. Anna, Maria were the most common names of the athletes with a count of 54 and 48 persons respectively
2. Daniel was the most common Male name
3. The most popular names are all western in origin - alluding to the possibility that such nations send the largest contingents

DASHBOARD



8)Which countries' athletes are the most "efficient"? (Highest Medals/Participation ratio)



Variables Used:

Name of Country, Count of Athletes, Count of Medals

Calculations done:

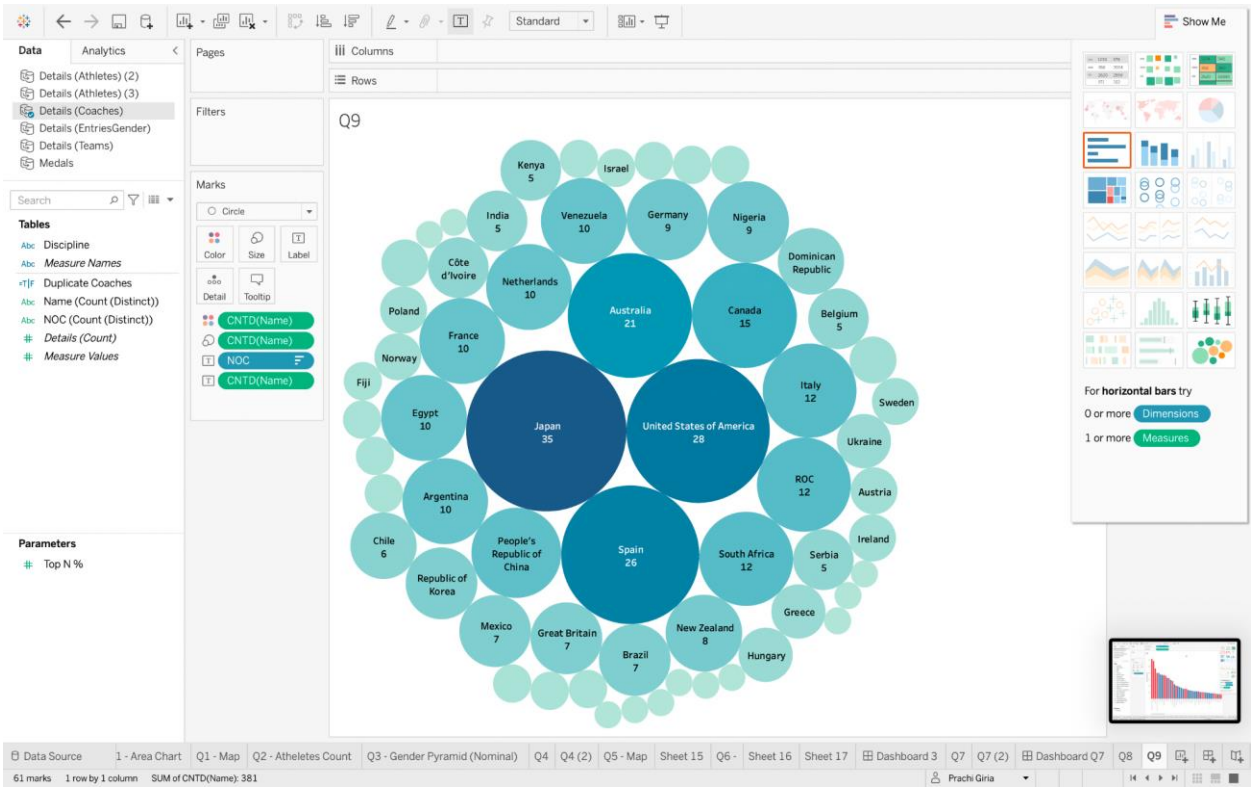
1. Total Participants = around 11,067 and Total Medals = 1080
2. Average Efficiency = 11067/1080 = ~ 0.097
3. sum([Total])/COUNT([ATHLETES (Athletes)].[ATHLETES])>0.097 (above average efficiency nations)

Key Observations:

1. Only 3 of the top 6 nations by medal count are above the average efficiency - indicating that these nations may be earning more medals because of the volume of athletes they send to the Olympics

2. Smaller nations are either highly efficient or have 0 efficiency (0 medals won), indicating that they either perform well above their weight or don't perform at all.

9)Which countries sent the largest number of coaches to the Olympics?



Variables Used:

NOC, Names

Calculations done:

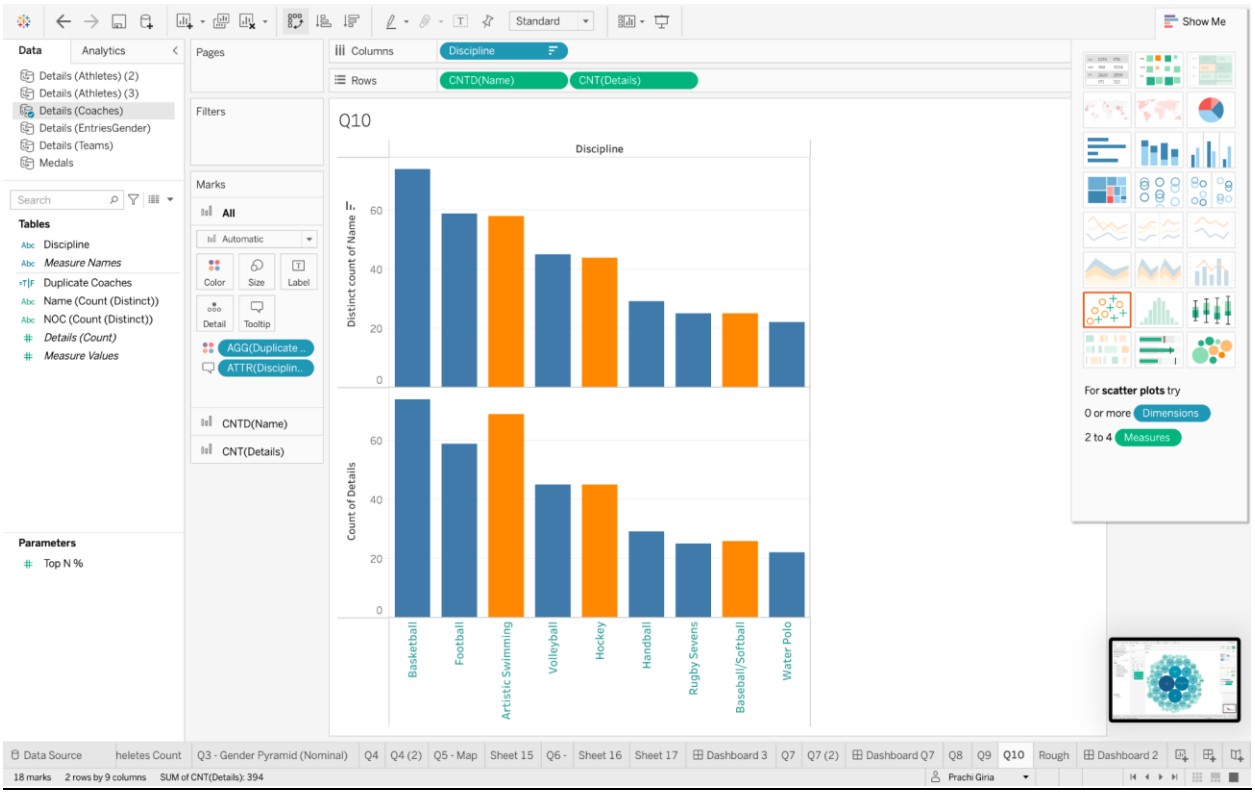
1. name(count(district))

Key Observations:

1. Japan has the largest number of couches(35) , the United States of America (28), Spain (26)

2. Japan has the largest as they had introduced new event into the Olympics (being the Host) which gave them a larger number of events they can effectively participate in

10)Which events have the highest number of coaches?



Variables Used:

Discipline, Name

Calculations done:

1. count(name), countd(name)
2. Countd([Name]) <> count([Name])

Key Observations:

1. Most events did not have any coaches that coaches multiple teams (one coach per team)

2. Artistic Swimming (all female event), Hockey & Baseball all had coaches that coached multiple teams, indicating that contingents believed their coaches were able to handle the additional burden of coaching multiple teams in such cases

Conclusion

The Olympics behave like a microcosm of the world ands its nations

- The developed dominate the scoreboard, the developing continue to strive to give their best at as many events as they can, and the smaller nations surprise the world by triumphing over their larger neighbours
- The Olympics are truly a global phenomenon - with around 195 nations participating in the latest edition
- Gender discrepancies are much less pronounced compared to the real world - perhaps a sign of what's to come in our daily lives
- The old is mixed with the new - Athletics straight from the Greek Olympics is put right next to Skateboarding, the latest addition to the event roster

Tableau and its functionality has proven to be versatile and easy to learn:

- Geographic data was automatically detected, making mapping a breeze
- Required calculations were often automatically carried out, requiring little modification
- The programming component (Functions & Parameters) proved to be simple yet effective
- The variety of ready to use visualization formats helped convey the dataset and its patterns with ease
- The flexibility of the software enabled us to experiment with innovative visualization formats
- The sharing functionality allowed us to share our finding with each other