DATA VISUALIZATION

TERM PROJECT FINAL

Kim, Suchul - 300362625

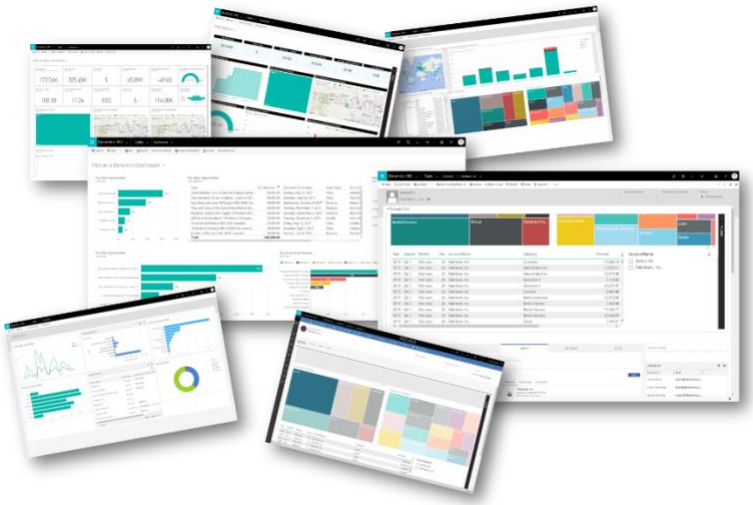
Haridas, Sudarsan - 300353099

Douglas College

CSIS 3860 - 070, Fall 2023: DATA VISUALIZATION

Professor: Rupa Manabala

Dec 06, 2023



This Photo by Unknown Author is licensed under CC BY-SA

PART A

The details of the data set – number of fields, rows, and data size

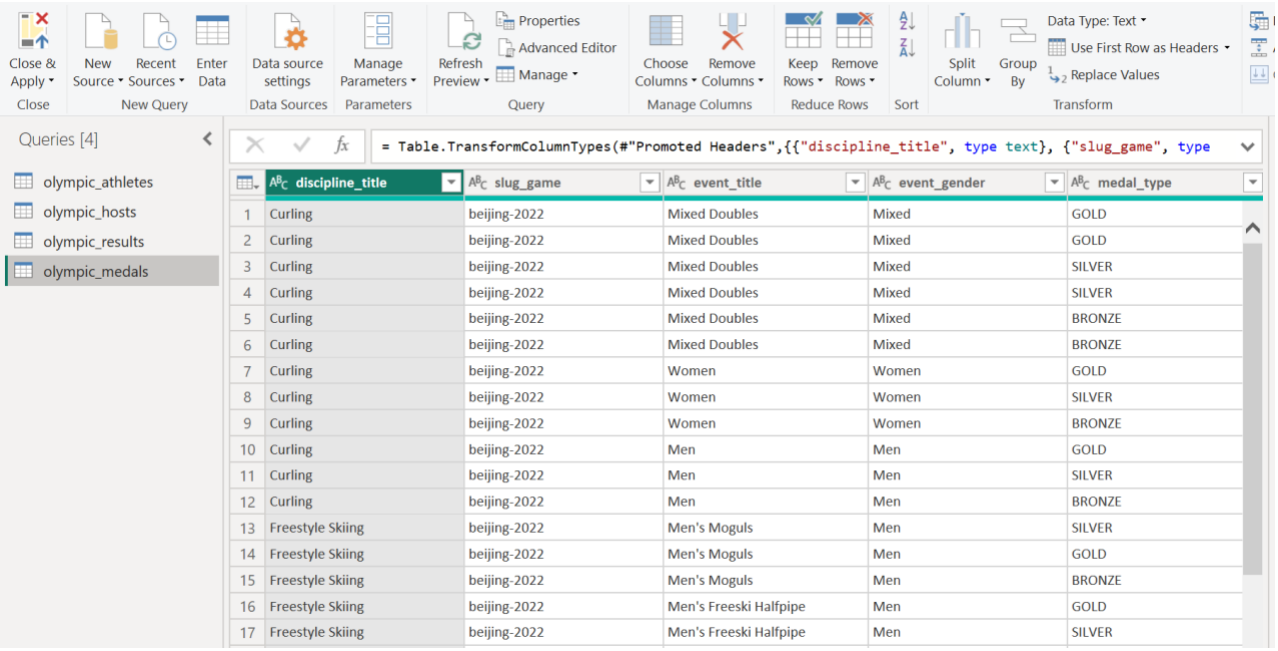
The Olympic dataset contains four tables: olympic_athletes, olympic_hosts, olympic_medals and olympic_results. olympic_athletes has 7 fields and 75904 rows, ,olympic_hosts has 7 fields and 53 rows, olympic_medals has 8 fields and 21697 rows, olympic_results has 9 fields and 162804 rows, Data set size is 45,990KB.

Analysis questions

- Which countries have been the most successful in terms of total medals won across all Olympic Games?
- How has the total number of medals awarded evolved over the years?
- What trends can be observed in gender diversity over the history of the Olympic Games?
- Which country has hosted the most Olympics?
- Do host countries tend to win more gold medals?

Data Connections

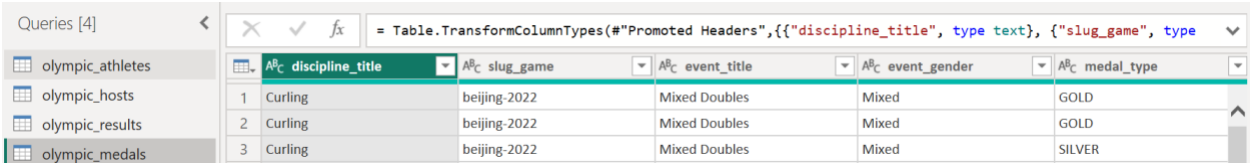
Import data from 4 Olympics csv files into power bi



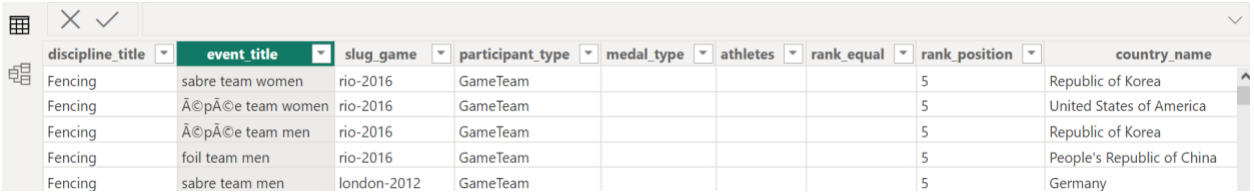
	discipline_title	slug_game	event_title	event_gender	medal_type
1	Curling	beijing-2022	Mixed Doubles	Mixed	GOLD
2	Curling	beijing-2022	Mixed Doubles	Mixed	GOLD
3	Curling	beijing-2022	Mixed Doubles	Mixed	SILVER
4	Curling	beijing-2022	Mixed Doubles	Mixed	SILVER
5	Curling	beijing-2022	Mixed Doubles	Mixed	BRONZE
6	Curling	beijing-2022	Mixed Doubles	Mixed	BRONZE
7	Curling	beijing-2022	Women	Women	GOLD
8	Curling	beijing-2022	Women	Women	SILVER
9	Curling	beijing-2022	Women	Women	BRONZE
10	Curling	beijing-2022	Men	Men	GOLD
11	Curling	beijing-2022	Men	Men	SILVER
12	Curling	beijing-2022	Men	Men	BRONZE
13	Freestyle Skiing	beijing-2022	Men's Moguls	Men	SILVER
14	Freestyle Skiing	beijing-2022	Men's Moguls	Men	GOLD
15	Freestyle Skiing	beijing-2022	Men's Moguls	Men	BRONZE
16	Freestyle Skiing	beijing-2022	Men's Freeski Halfpipe	Men	GOLD
17	Freestyle Skiing	beijing-2022	Men's Freeski Halfpipe	Men	SILVER

Data Transform Strategies

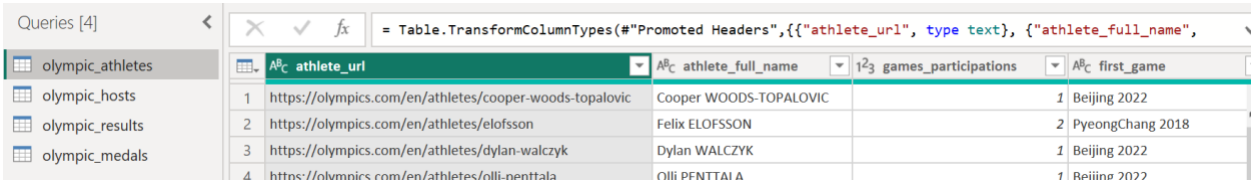
- **Use first row as a Header** : Applies to the first row of Olympic medals.



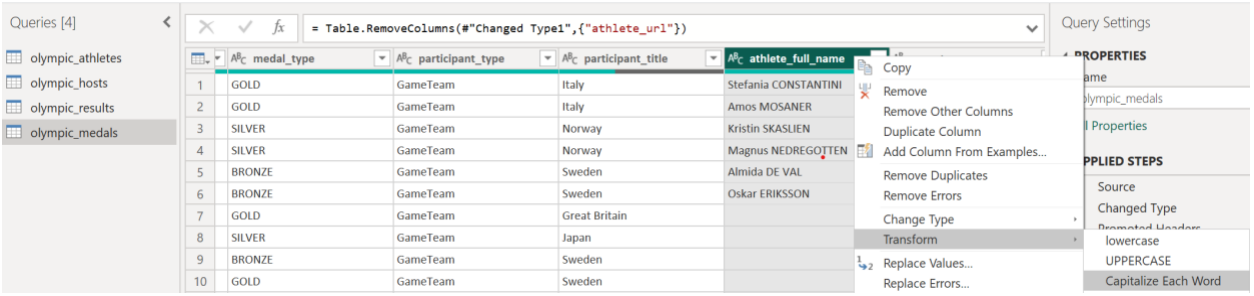
- **Replace Value** : Delete erroneous characters, such as Ã‰pÃ©e.



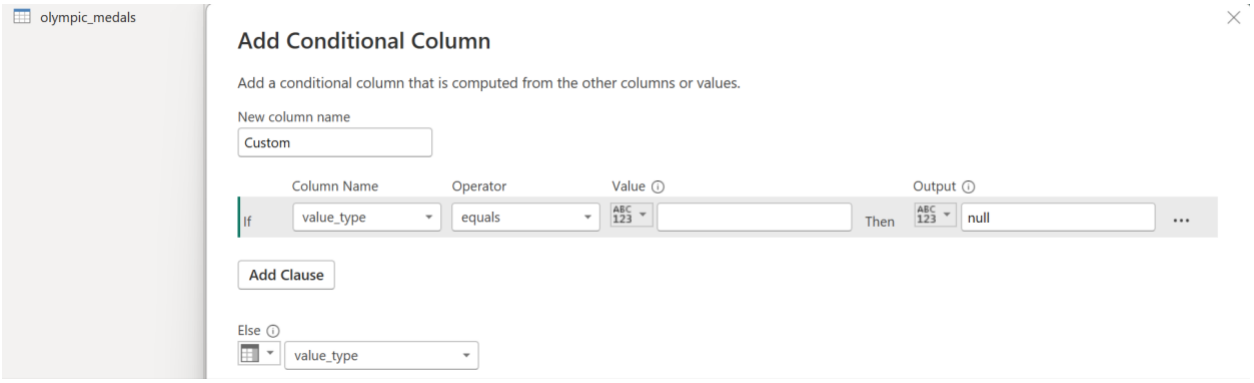
- **Remove Columns** : Delete unnecessary column that will not be used.



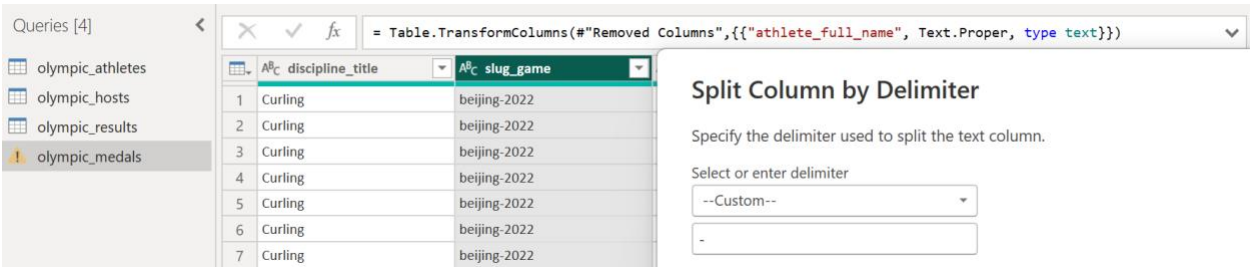
- **Capitalize Each Word** : Change the case of the athlete full name.



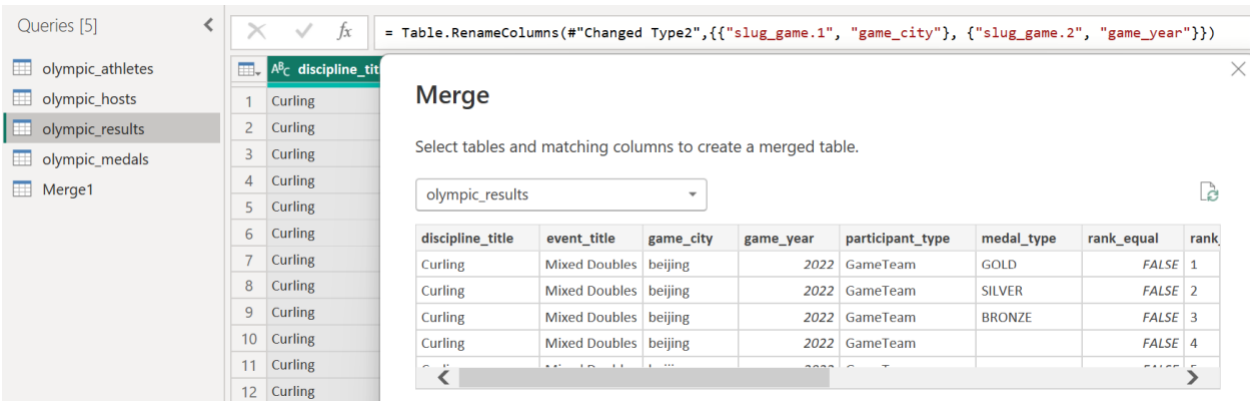
- **Add Conditional Column** : Set the blank field of value type to null.



- **Split Column by Delimiter** : Split the slug game into columns by city and year.

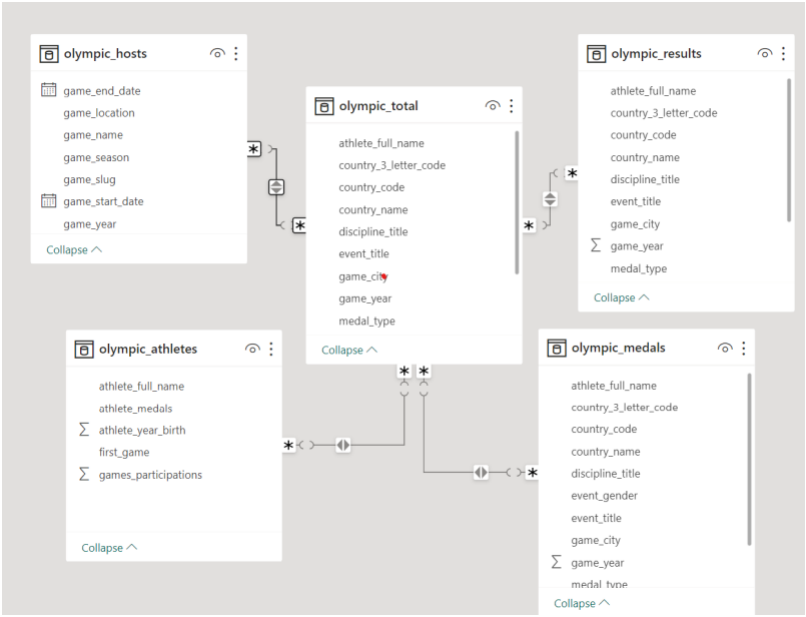


- **Merge Query** : Merge the tables to create an Olympic total table.



Building the Data Model

The tables are connected based on the athlete's full name. The olympic host table functions as a date.



Leveraging DAX

Game year and game month from the olympic_hosts table to create a Year-Month.

```
Month Year = 'olympic_hosts'[game_month] & "-" & 'olympic_hosts'[game_year]
```

game_end_date	game_start_date	game_location	game_name	game_season	game_year	game_month	Month Year
2/20/2022 4:00:00 AM	2/4/2022 7:00:00 AM	China	Beijing 2022	Winter	2022	2	2-2022
8/8/2021 7:00:00 AM	7/23/2021 4:00:00 AM	Japan	Tokyo 2020	Summer	2020	8	8-2020

The average number of medals by sport

```
Total Medal Count = COUNTROWS(
    FILTER(
        'olympic_results',
        'olympic_results'[medal_type] = "Gold" || 'olympic_results'[medal_type] = "Silver" || 'olympic_results'[medal_type] =
        "Bronze"
    )
)
```

The number of rows in the olympic_results table where the medal type is "Gold."

```
Gold Medal Count = COUNTROWS(
    FILTER(
        'olympic_results',
        'olympic_results'[medal_type] = "Gold"
    )
)
```

Total medal count up to the current context's maximum year

```
Medal Count by Country = CALCULATE(
    [Total Medal Count],
    ALLEXCEPT('olympic_results', 'olympic_results'[country_name])
)
```

discipline_title	game_city	game_year	participant_type	medal_type	rank_equal	rank_position	country_name	country_code	country
Tennis	rio	2016	GameTeam			9	Colombia	CO	COL
Tennis	rio	2016	GameTeam			9	Poland	PL	POL

Total number of medals by counting the rows where the medal type is either "Gold," "Silver," or "Bronze."

```
Average Medals by Sport = AVERAGEX(
    SUMMARIZE('olympic_results', 'olympic_results'[discipline_title], "Medal Count", [Total Medal Count]),
    [Medal Count]
)
```

discipline_title	game_city	game_year	participant_type	medal_type	rank_equal	rank_position	country_name	country_code	country
Tennis	rio	2016	GameTeam			9	Colombia	CO	COL
Tennis	rio	2016	GameTeam			9	Poland	PL	POL

Total number of medals up over the game year.

```
Medals Over Time = CALCULATE(
    [Total Medal Count],
    FILTER(
        ALL('olympic_results'[game_year]),
        'olympic_results'[game_year] <= MAX('olympic_results'[game_year])
    )
)
```

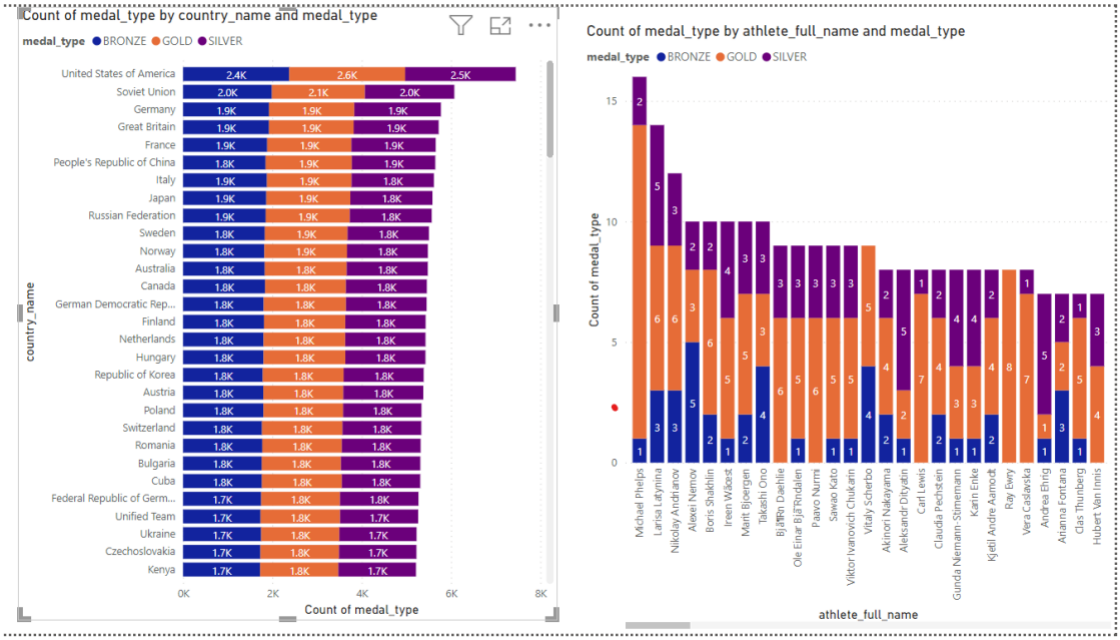
discipline_title	game_city	game_year	participant_type	medal_type	rank_equal	rank_position	country_name	country_code	country
Tennis	rio	2016	GameTeam			9	Colombia	CO	COL

Visualization Data

Bar-Chart

United States of America in medal_type GOLD made up 0.59% of Count of medal.

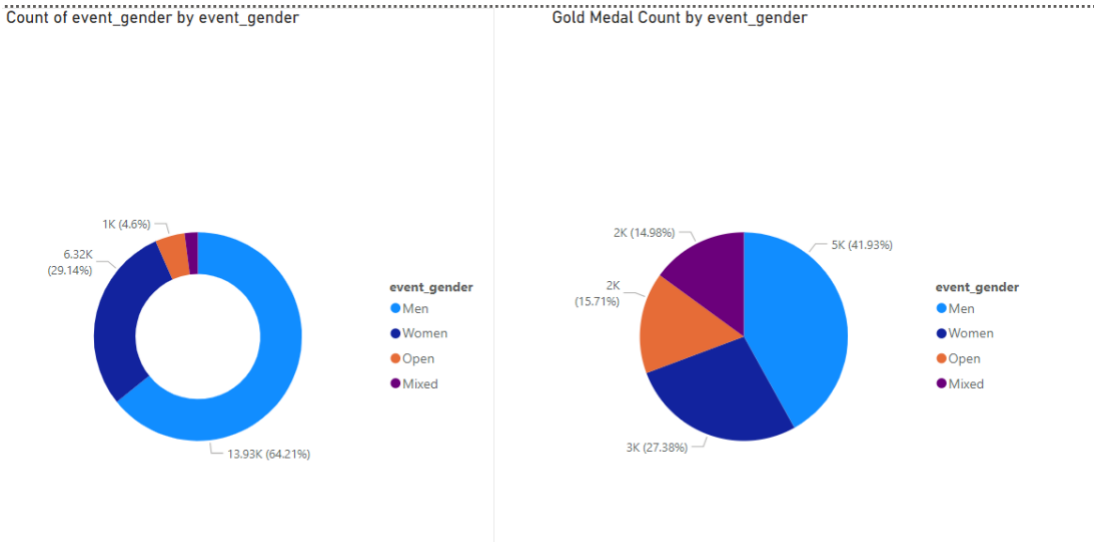
Michael Phelps in medal_type GOLD made up 0.09% of Count of medal.



Pie-Chart

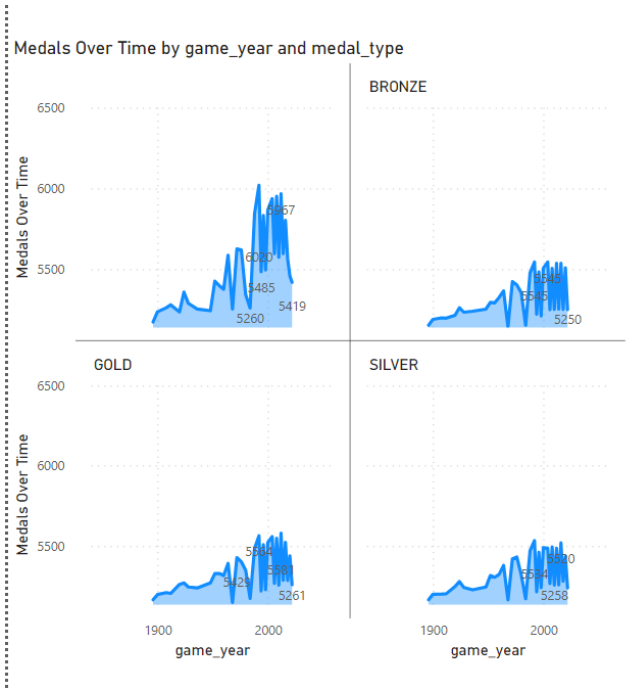
Men had the highest Count of event_gender at 13,932(64%), followed by Women(29%)

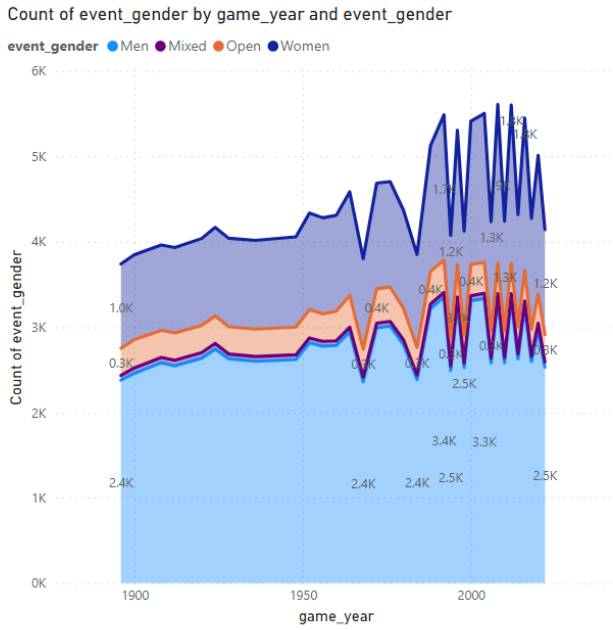
Men had the highest Gold Medal Count at 4856(42%), followed by Women(27%)



Time-Series

Depending on the year, the number of medals for gold, silver, and bronze is all on the rise, but has been stagnant since 2000. The number of men and women is increasing by year, and has become saturated since 2000.

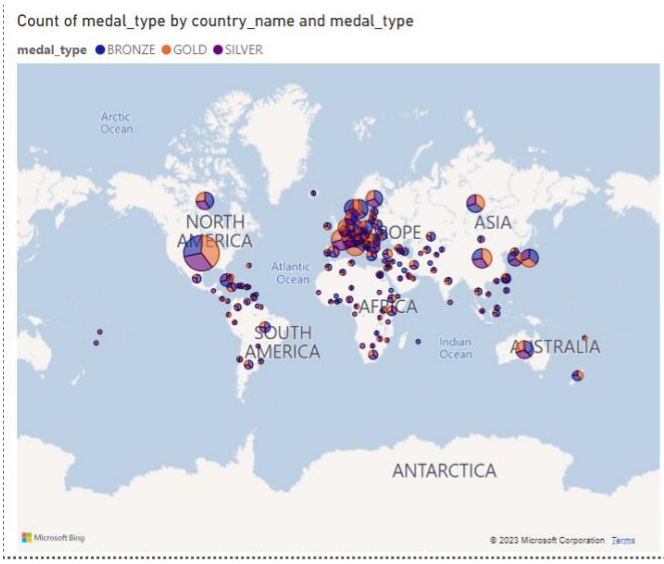




Geographic

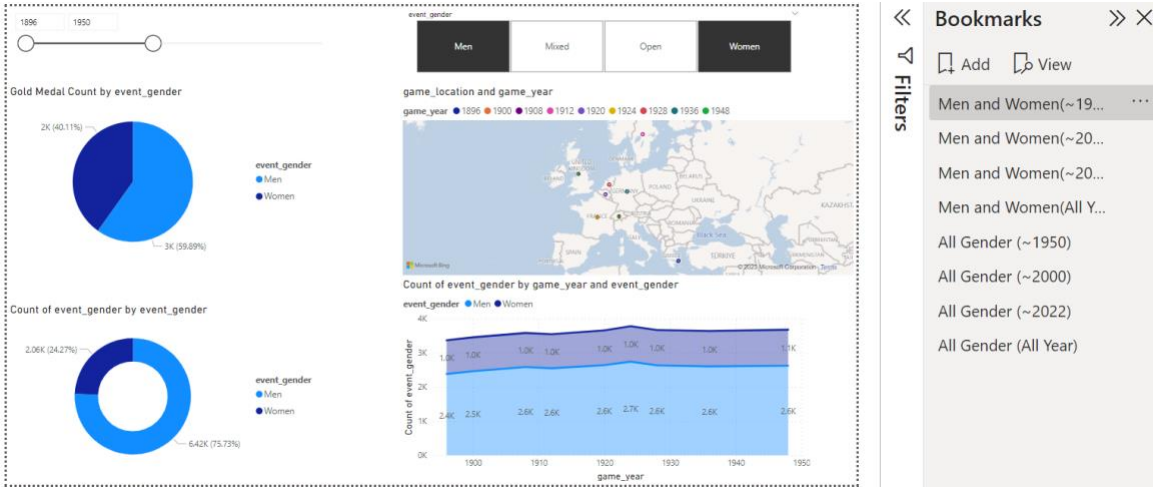
The United States has won the most gold, silver, and bronze medals in Olympic history.

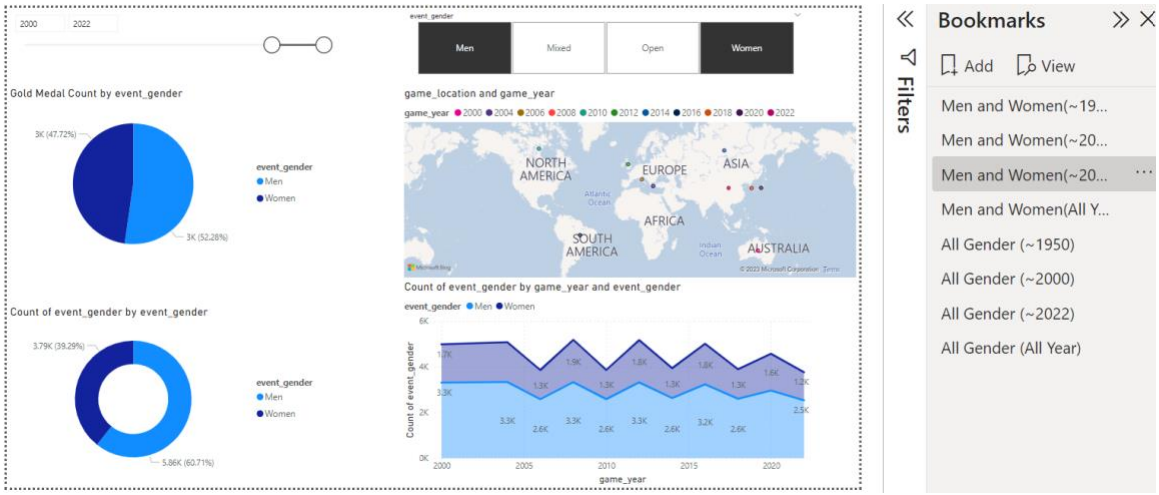
France, Japan, and the United States hosted the most Olympic Games, four each.

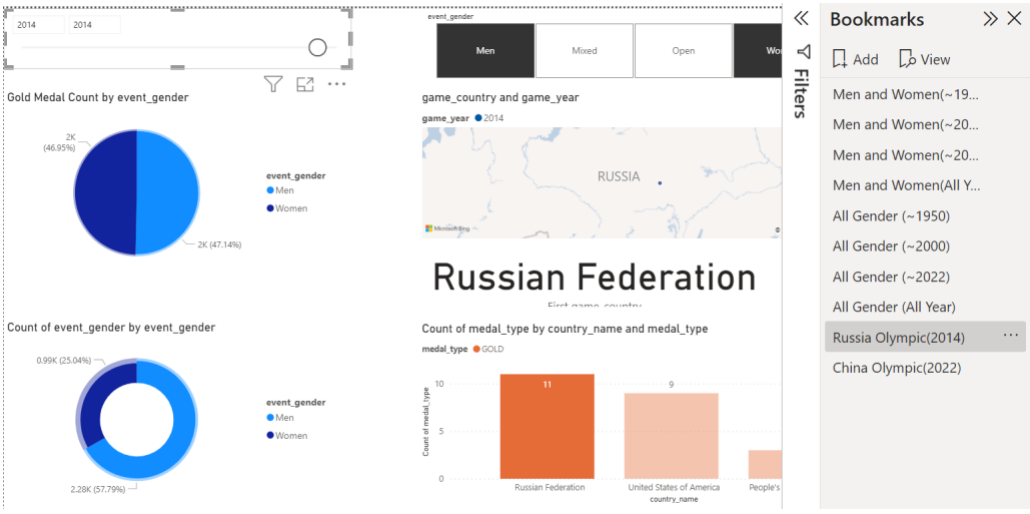


Bookmarks

Using bookmarks, we can easily observe changes in the number of medals and people according to gender and year.







Answer your analysis questions through your reports

Q : Which countries have been the most successful in terms of total medals won across all Olympic Games?

Bar Chart: This chart shows the percentage of gold medals obtained by each country. The United States has won the most gold(2595ea), silver(2475ea), and bronze medals(2376ea) in Olympic history.

Q. How has the total number of medals awarded evolved over the years?

Time-Series Chart: This chart displays the trend of gold, silver, and bronze medals over the years. The number of medals for gold, silver, and bronze is on the rise, but the growth has been stagnant since the year 2000.

Q. What trends can be observed in gender diversity over the history of the Olympic Games?

Pie Chart and Time-Series Chart : This chart compares the number of medals by gender in the Olympics. The number of medals and participants for both men and women is increasing by year, and has become saturated since the year 2000. Men had the highest count of event_gender at 13,932 (64%), followed by women (29%). Men also had the highest Gold Medal Count at 4856 (42%), followed by women (27%).

Q. Which country has hosted the most Olympics?

Geographic Analysis: Through geographic analysis, it is observed that France, Japan, and the United States have each hosted the Olympics four times. This can be confirmed through a map or bar chart displaying the number of Olympics hosted by each country.

Q. Do host countries tend to win more gold medals?

Bookmarks show that hosts Russia and China won the most golds at their 2014 and 2022 Olympics, respectively, hinting at a home advantage.

Power BI Explanation Video

https://youtu.be/zl7aJWBjHn0

References

Microsoft Power BI Quick Start Guide – Preferably 3rd edition- Devin Knight, Mitchell Pearson, Bradley Schacht,, et al – ISBN 9781804613498 – Packt Publishing – eBook available from publisher's website (https://www.packtpub.com/product/microsoft-power-biquick-start-guide-second-edition/9781804613498)

Microsoft Ignite

https://learn.microsoft.com/en-us/power-bi/