Data Warehousing Project

Topic: Olympic Data Warehouse

By -

Anish Babu Gogineni (U07170480)

Hema Edavalapati (U66945264)

Rithika Kandimalla (U89077236)

Executive Summary

The Olympic Games represent a pinnacle of human achievement and excellence, bringing together nations and athletes from around the world in a celebration of sport and competition. In this project, we delve into the vast historical data spanning 120 years of Olympic history to extract insights that illuminate the dynamics of athletic performance, medal success, and participation trends.

Through advanced data analytics techniques, including data exploration, statistical analysis, and predictive modelling, we aim to uncover patterns and factors that influence medal outcomes across various sports and events. Our analysis seeks to provide actionable insights for sports organizations, national Olympic committees, and stakeholders to optimize their strategies, resource allocation, and athlete development programs.

By harnessing the power of data, we endeavour to empower decision-makers with the knowledge and tools needed to unlock the full potential of athletes on the global stage. From identifying talent to forecasting medal outcomes, our project aims to contribute to the ongoing pursuit of excellence in Olympic competition and inspire future generations of athletes**.**

Problem Statement

The project aims to explore and analyse the extensive historical data from the Olympic Games using data analytics techniques. Specifically, we seek to uncover insights into performance trends, identify factors influencing medal success, and develop predictive models for future medal outcomes. By leveraging the comprehensive "120 Years of Olympic History: Athletes and Results" dataset from Kaggle, we address the challenge of optimizing strategies and decision-making processes for sports organizations, national Olympic committees, and stakeholders.

The significance of this challenge lies in its potential to reshape the landscape of athletic training, talent identification, and strategic planning. By gaining a deeper understanding of the determinants of Olympic success, stakeholders can tailor their training programs, identify emerging talent, and optimize resource allocation effectively. Additionally, by developing predictive models, we can forecast future medal outcomes, enabling proactive planning and strategic decision-making.

Literature Review:

The exploration of existing literature, studies, and articles related to Olympic data analysis and sports analytics provides a foundational understanding for our project's objectives in data warehousing for Olympic data. Drawing upon prior research, we aim to leverage established methodologies and insights to enhance the data warehousing process and drive innovation in the realm of Olympic sports analytics.

Previous studies on sports analytics offer valuable methodologies and approaches that can be adapted to the data warehousing context for Olympic data. By examining these methodologies through a data warehousing lens, we can identify strategies for structuring and integrating Olympic data tables effectively. Furthermore, insights gleaned from studies on Olympic data analysis provide valuable perspectives on factors influencing athletic performance and medal outcomes, which can inform the design of data models and schema to support comprehensive analysis of Olympic performance trends and factors.

The literature review serves as a cornerstone for our project, offering a comprehensive overview of existing research and methodologies relevant to data warehousing for Olympic data. By synthesizing insights from prior literature, we aim to inform our data modeling, ETL processes, and analytical approaches to optimize the data warehousing process for Olympic sports analytics.

Data Collection and Preparation:

Data Set: Olympic Data 🏅 ⛹ 🏟 🏋 🚴 (kaggle.com)

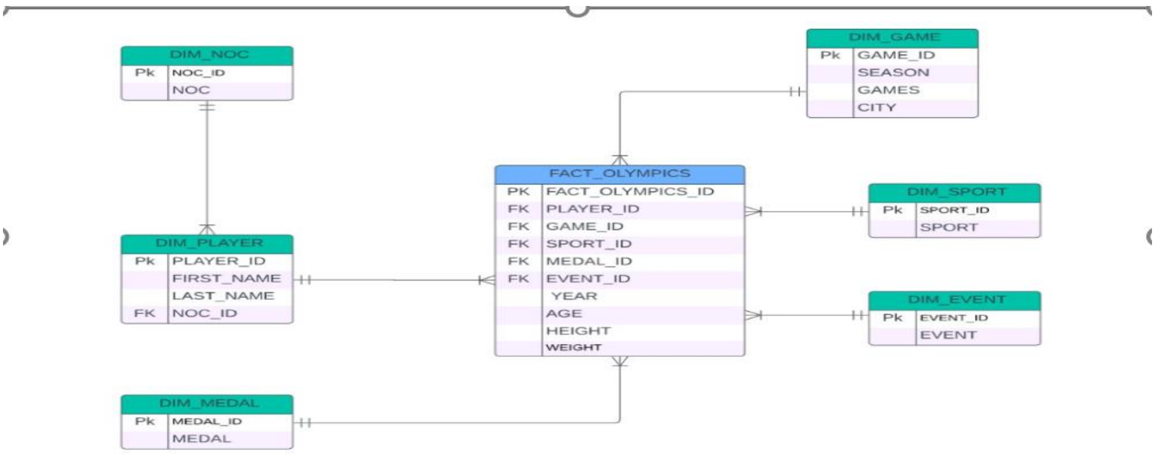
The dataset was gathered from Kaggle, we have performed data cleaning, preprocessing, and transformations to prepare it for analysis. The raw data is in the form of csv files. we have loaded the data in DB performed ETL (Extraction, Transformation and Loading) process,tables were created, and several queries were run to gain insight. The following are the data sets that we have extracted:

Noc_region : This dataset contains the region details such as NOC code and region.

Dataset_olympics : This dataset contains the name of the participants,age,Season, year the game conducted, Medal, sport etc.

Database Design:

Part A: ERD



Part B: Dimensional Modelling

Fact Table:

- **Fact_Olympic_Players**: This is the fact table where data related to Olympic player participation is stored. It contains quantitative data such as Player_ID, Game_ID, Sport_ID, Medal_ID, Event_ID, NOC_ID, Age, Height, Weight, and Year.

Dimension Tables:

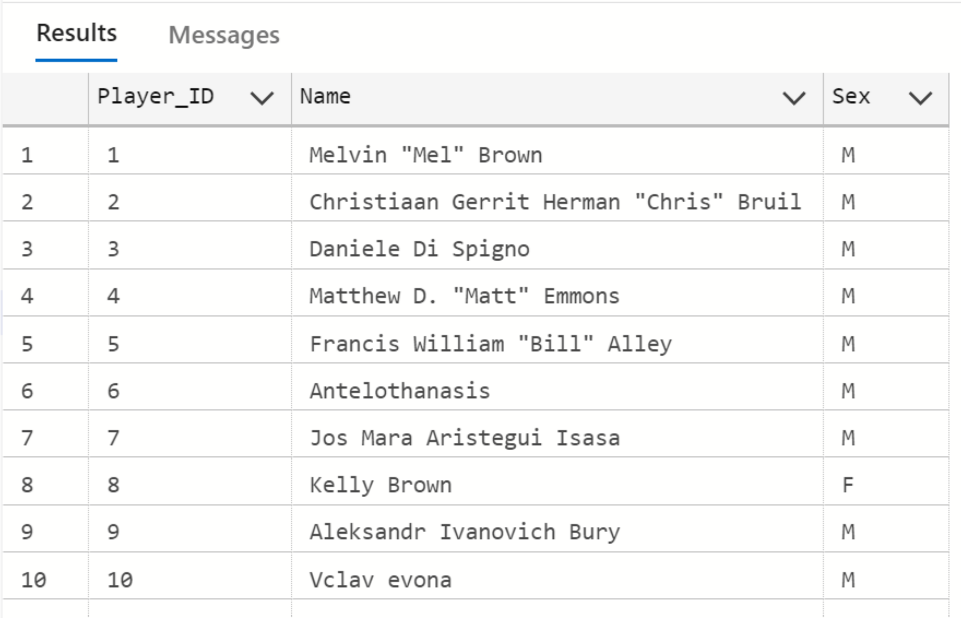
- **DimNOC**: Dimension table containing information about National Olympic Committees (NOCs), including attributes like NOC code and region.
- **DimEvent**: Dimension table storing details about Olympic events, such as the event name.

- **DimGame**: Dimension table containing information about Olympic games, including attributes like Games, Season, and City.
- **DimMedal**: Dimension table holding information about different types of Olympic medals.
- **DimPlayer**: Dimension table storing attributes related to Olympic players, such as their names and genders.
- **DimSport**: Dimension table containing details about different Olympic sports.

Part C: Create Table Statements

- Creating the Player dimension table

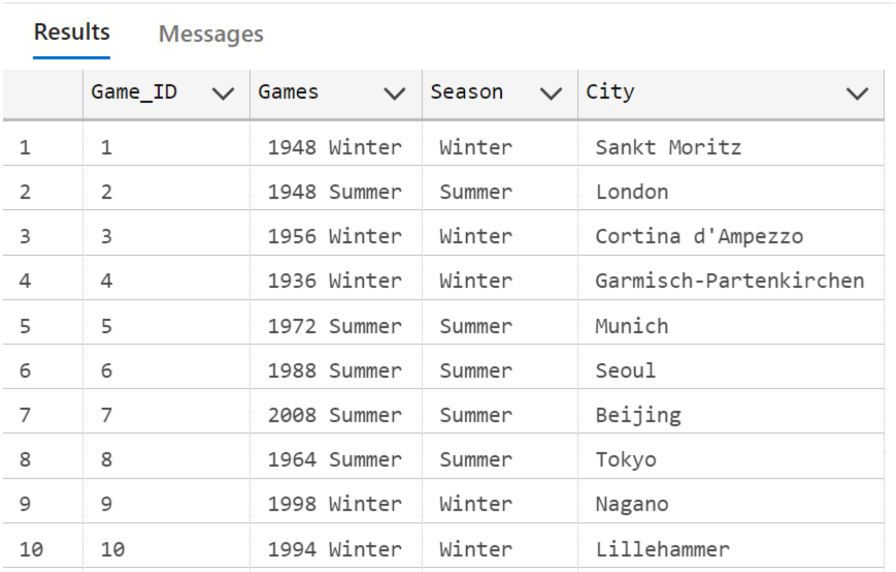
```
CREATE TABLE [Olympic_Players].DimPlayer(
    Player_ID INT PRIMARY KEY IDENTITY(1,1),
    Name VARCHAR(100),
    Sex VARCHAR(50)
);
```



	Player_ID	Name	Sex
1	1	Melvin "Mel" Brown	M
2	2	Christiaan Gerrit Herman "Chris" Bruil	M
3	3	Daniele Di Spigno	M
4	4	Matthew D. "Matt" Emmons	M
5	5	Francis William "Bill" Alley	M
6	6	Antelothanasis	M
7	7	Jos Mara Aristegui Isasa	M
8	8	Kelly Brown	F
9	9	Aleksandr Ivanovich Bury	M
10	10	Vclav evona	M

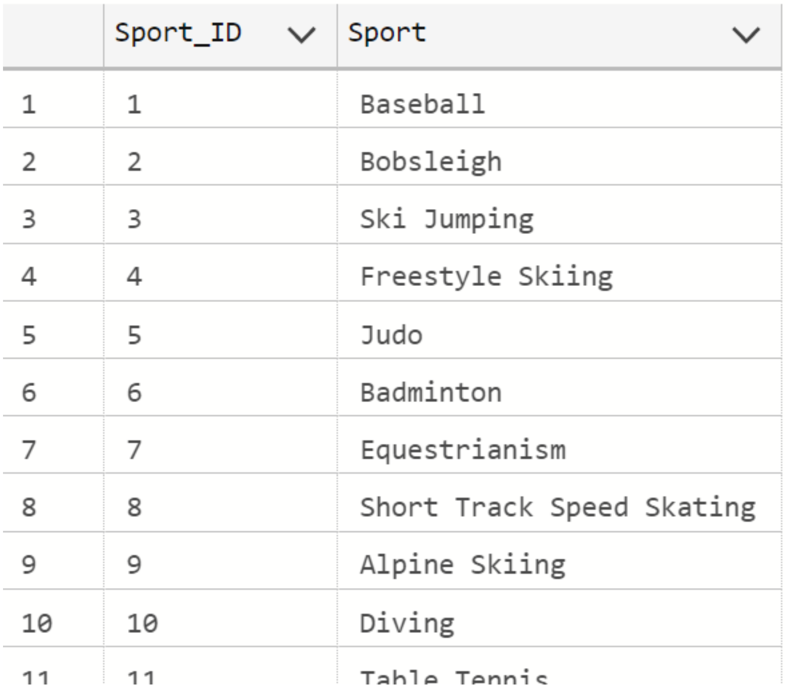
-- Creating the Games dimension table

```
CREATE TABLE [Olympic_Players].DimGame(
    Game_ID INT PRIMARY KEY IDENTITY(1,1),
    Games VARCHAR(100),
    Season VARCHAR(50),
    City VARCHAR(50)
);
```



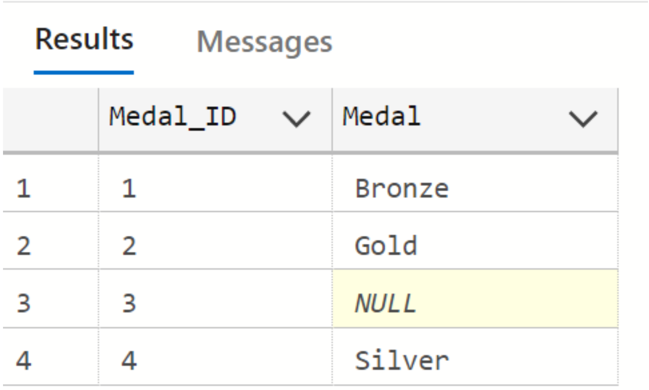
	Game_ID	Games	Season	City
1	1	1948 Winter	Winter	Sankt Moritz
2	2	1948 Summer	Summer	London
3	3	1956 Winter	Winter	Cortina d'Ampezzo
4	4	1936 Winter	Winter	Garmisch-Partenkirchen
5	5	1972 Summer	Summer	Munich
6	6	1988 Summer	Summer	Seoul
7	7	2008 Summer	Summer	Beijing
8	8	1964 Summer	Summer	Tokyo
9	9	1998 Winter	Winter	Nagano
10	10	1994 Winter	Winter	Lillehammer

```
-- Creating the Sport dimension table
 CREATE TABLE [Olympic_Players].DimSport(
    Sport_ID INT PRIMARY KEY IDENTITY(1,1),
    Sport VARCHAR(50)
 );
```



	Sport_ID	Sport
1	1	Baseball
2	2	Bobsleigh
3	3	Ski Jumping
4	4	Freestyle Skiing
5	5	Judo
6	6	Badminton
7	7	Equestrianism
8	8	Short Track Speed Skating
9	9	Alpine Skiing
10	10	Diving
11	11	Table Tennis

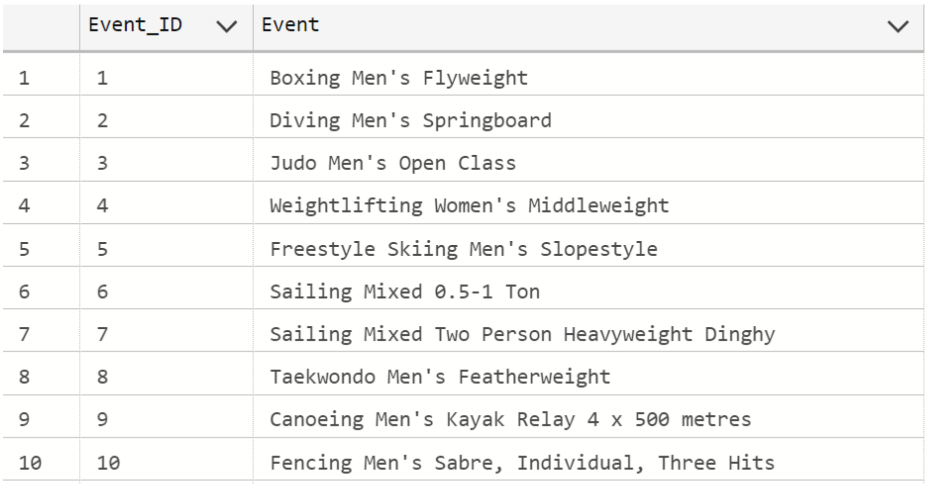
```
-- Creating the Medal dimension table
 CREATE TABLE [Olympic_Players].DimMedal(
    Medal_ID INT PRIMARY KEY IDENTITY(1,1),
    Medal VARCHAR(50)
 );
```



	Medal_ID	Medal
1	1	Bronze
2	2	Gold
3	3	*NULL*
4	4	Silver

-- Creating the Event dimension table

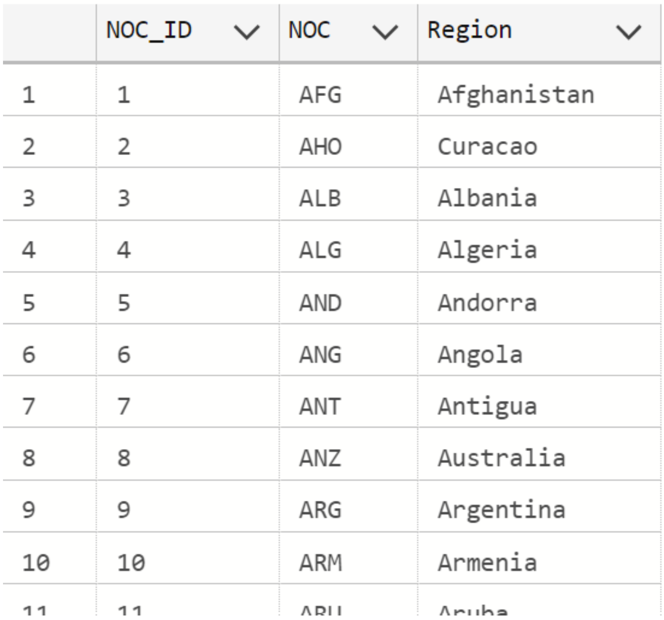
```
CREATE TABLE [Olympic_Players].DimEvent(
    Event_ID INT PRIMARY KEY IDENTITY(1,1),
    Event VARCHAR(100)
);
```



	Event_ID	Event
1	1	Boxing Men's Flyweight
2	2	Diving Men's Springboard
3	3	Judo Men's Open Class
4	4	Weightlifting Women's Middleweight
5	5	Freestyle Skiing Men's Slopestyle
6	6	Sailing Mixed 0.5-1 Ton
7	7	Sailing Mixed Two Person Heavyweight Dinghy
8	8	Taekwondo Men's Featherweight
9	9	Canoeing Men's Kayak Relay 4 x 500 metres
10	10	Fencing Men's Sabre, Individual, Three Hits

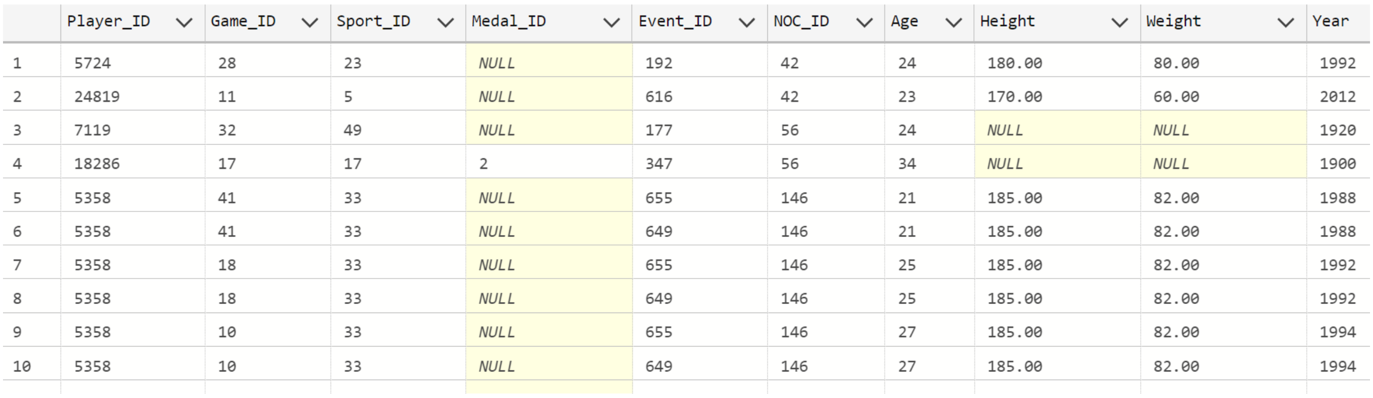
-- Creating the NOC dimension table

```
CREATE TABLE [Olympic_Players].DimNOC(
    NOC_ID INT PRIMARY KEY IDENTITY(1,1),
    NOC VARCHAR(10),
    Region VARCHAR(50)
);
```



	NOC_ID	NOC	Region
1	1	AFG	Afghanistan
2	2	AHO	Curacao
3	3	ALB	Albania
4	4	ALG	Algeria
5	5	AND	Andorra
6	6	ANG	Angola
7	7	ANT	Antigua
8	8	ANZ	Australia
9	9	ARG	Argentina
10	10	ARM	Armenia
11	11	ARU	Aruba

```
-- Creating the Olympics fact table
 CREATE TABLE [Olympic_Players].Fact_Olympic_Players(
    Fact_ID INT PRIMARY KEY IDENTITY(1,1),
    Player_ID INT,
    Game_ID INT,
    Sport_ID INT,
    Medal_ID INT,
    Event_ID INT,
    NOC_ID INT,
    Age INT,
    Height DECIMAL(5, 2),
    Weight DECIMAL(5, 2),
    Year INT,

   CONSTRAINT FK_Player_ID FOREIGN KEY (Player_ID) REFERENCES
[Olympic_Players].DimPlayer(Player_ID),
    CONSTRAINT FK_Game_ID FOREIGN KEY (Game_ID) REFERENCES
[Olympic_Players].DimGame(Game_ID),
    CONSTRAINT FK_Sport_ID FOREIGN KEY (Sport_ID) REFERENCES
[Olympic_Players].DimSport(Sport_ID),
    CONSTRAINT FK_Medal_ID FOREIGN KEY (Medal_ID) REFERENCES
[Olympic_Players].DimMedal(Medal_ID),
    CONSTRAINT FK_Event_ID FOREIGN KEY (Event_ID) REFERENCES
[Olympic_Players].DimEvent(Event_ID),
    CONSTRAINT FK_NOC_ID FOREIGN KEY (NOC_ID) REFERENCES
[Olympic_Players].DimNOC(NOC_ID)
 );
```



	Player_ID	Game_ID	Sport_ID	Medal_ID	Event_ID	NOC_ID	Age	Height	Weight	Year
1	5724	28	23	NULL	192	42	24	180.00	80.00	1992
2	24819	11	5	NULL	616	42	23	170.00	60.00	2012
3	7119	32	49	NULL	177	56	24	NULL	NULL	1920
4	18286	17	17	2	347	56	34	NULL	NULL	1900
5	5358	41	33	NULL	655	146	21	185.00	82.00	1988
6	5358	41	33	NULL	649	146	21	185.00	82.00	1988
7	5358	18	33	NULL	655	146	25	185.00	82.00	1992
8	5358	18	33	NULL	649	146	25	185.00	82.00	1992
9	5358	10	33	NULL	655	146	27	185.00	82.00	1994
10	5358	10	33	NULL	649	146	27	185.00	82.00	1994

Part D: ETL Process

The ETL process ensures that Olympic player data is processed accurately and efficiently, empowering us to extract valuable insights and drive informed decision-making in various areas related to Olympic sports and competitions.

The structured schema enabled us to perform a wide range of analytical tasks, including performance analysis to evaluate player achievements, historical trend analysis to track changes in participation and success over time, and medal distribution analysis to understand patterns and trends in medal distribution across different events, sports, and countries.

Below is the ETL script for populating the data to tables:

We have created two staging tables and populated the data to dimension tables and fact tables accordingly:

```
-- Selecting data from the [Olympic_Players].[Staging_table] and inserting into the [Olympic_Players].[DimEvent] table
INSERT INTO [Olympic_Players].[DimEvent](Event)
(
SELECT DISTINCT Event
FROM [Olympic_Players].[Staging_table]
);



-- Update Event_ID in [Olympic_Players].[Staging_table]
UPDATE [Olympic_Players].[Staging_table]
SET [Olympic_Players].[Staging_table].Event_ID = [Olympic_Players].[DimEvent].Event_ID
FROM [Olympic_Players].[Staging_table]
INNER JOIN [Olympic_Players].[DimEvent]
ON [Olympic_Players].[Staging_table].Event = [Olympic_Players].[DimEvent].Event;



-- Selecting Games, Season and City from [Olympic_Players].[Staging_table] and inserting into [Olympic_Players].[DimGame]
INSERT INTO [Olympic_Players].[DimGame] (Games, Season, City)
(
SELECT DISTINCT Games, Season, City
FROM [Olympic_Players].[Staging_table]
);



-- Update Game_ID in [Olympic_Players].[Staging_table]
UPDATE [Olympic_Players].[Staging_table]
```

```
  SET [Olympic_Players].[Staging_table].Game_ID = [Olympic_Players].[DimGame].Game_ID
  FROM [Olympic_Players].[Staging_table]
  INNER JOIN [Olympic_Players].[DimGame]
  ON [Olympic_Players].[Staging_table].Games = [Olympic_Players].[DimGame].Games
    AND [Olympic_Players].[Staging_table].Season = [Olympic_Players].[DimGame].Season
    AND [Olympic_Players].[Staging_table].City = [Olympic_Players].[DimGame].City;



  -- Selecting Medal from [Olympic_Players].[Staging_table] and inserting into [Olympic_Players].[DimMedal]
  INSERT INTO [Olympic_Players].[DimMedal] (Medal)
  (
  SELECT DISTINCT Medal
  FROM [Olympic_Players].[Staging_table]
  );



  -- Update Medal_ID in [Olympic_Players].[Staging_table]
  UPDATE [Olympic_Players].[Staging_table]
  SET [Olympic_Players].[Staging_table].Medal_ID = [Olympic_Players].[DimMedal].Medal_ID
  FROM [Olympic_Players].[Staging_table]
  INNER JOIN [Olympic_Players].[DimMedal]
  ON [Olympic_Players].[Staging_table].MedaL = [Olympic_Players].[DimMedal].Medal;



  -- Selecting Name, Sex from [Olympic_Players].[Staging_table] and inserting into [Olympic_Players].[DimPlayer]
  INSERT INTO [Olympic_Players].[DimPlayer] (Name, Sex)
  (
  SELECT DISTINCT Name, Sex
  FROM [Olympic_Players].[Staging_table]
  );



  -- Update Player_ID in [Olympic_Players].[Staging_table]
  UPDATE [Olympic_Players].[Staging_table]
  SET [Olympic_Players].[Staging_table].Player_ID = [Olympic_Players].[DimPlayer].Player_ID
  FROM [Olympic_Players].[Staging_table]
  INNER JOIN [Olympic_Players].[DimPlayer]
  ON [Olympic_Players].[Staging_table].Name = [Olympic_Players].[DimPlayer].Name
    AND [Olympic_Players].[Staging_table].Sex = [Olympic_Players].[DimPlayer].Sex;



  -- Selecting Sport from [Olympic_Players].[Staging_table] and inserting into [Olympic_Players].[DimSport]
  INSERT INTO [Olympic_Players].[DimSport] (Sport)
  (
  SELECT DISTINCT Sport
  FROM [Olympic_Players].[Staging_table]
  );



  -- Update Sport_ID in [Olympic_Players].[Staging_table]
  UPDATE [Olympic_Players].[Staging_table]
  SET [Olympic_Players].[Staging_table].Sport_ID = [Olympic_Players].[DimSport].Sport_ID
  FROM [Olympic_Players].[Staging_table]
  INNER JOIN [Olympic_Players].[DimSport]
  ON [Olympic_Players].[Staging_table].Sport = [Olympic_Players].[DimSport].Sport;

-- [Olympic_Players].[DimNOC] has been populated by another ETL script
  -- Update NOC_ID in [Olympic_Players].[Staging_table]
  UPDATE [Olympic_Players].[Staging_table]
  SET [Olympic_Players].[Staging_table].NOC_ID = [Olympic_Players].[DimNOC].NOC_ID
  FROM [Olympic_Players].[Staging_table]
```

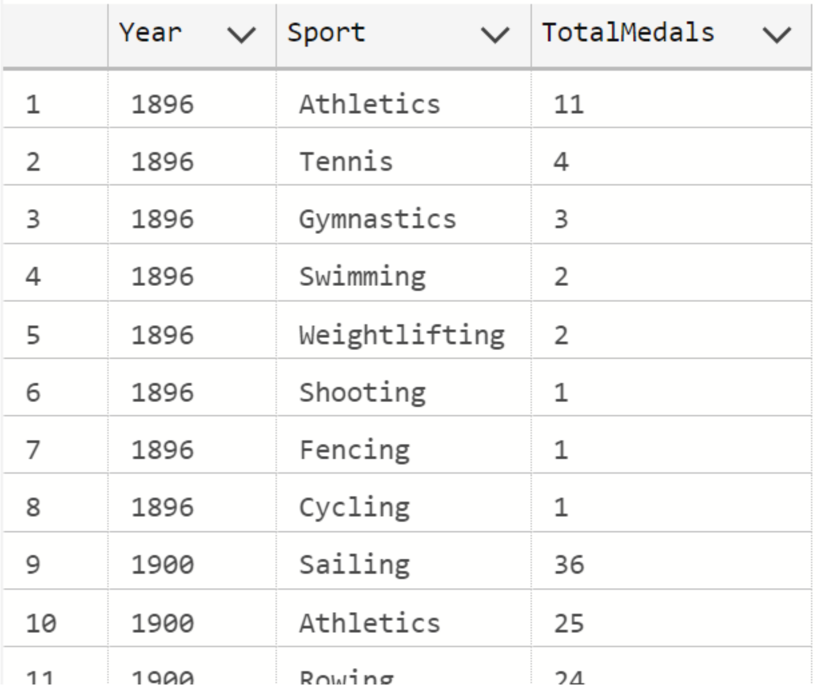
```
INNER JOIN [Olympic_Players].[DimNOC]
ON [Olympic_Players].[Staging_table].NOC = [Olympic_Players].[DimNOC].NOC



-- Load FACT Table with all of the appropriate data from the Staging table
INSERT INTO [Olympic_Players].[Fact_Olympic_Players]
(
    Player_ID,
    Game_ID,
    Sport_ID,
    Medal_ID,
    Event_ID,
    NOC_ID,
    Age,
    Height,
    Weight,
    Year
)
(
    SELECT
        Player_ID,
        Game_ID,
        Sport_ID,
        Medal_ID,
        Event_ID,
        NOC_ID,
        Age,
        Height,
        Weight,
        Year
    FROM [Olympic_Players].[Staging_table]
);



-- Selecting NOC, Region from [Olympic_Players].[NOC_Staging] and inserting into [Olympic_Players].[DimNOC]
INSERT INTO [Olympic_Players].[DimNOC] (NOC, Region)
(
SELECT DISTINCT NOC, Region FROM [Olympic_Players].[NOC_Staging]
);
```

Exploratory Data Analysis:

1.Performance Trends Over Time

```
SELECT
FOP.Year,
DS.Sport,
COUNT(*) AS TotalMedals
FROM
[Olympic_Players].Fact_Olympic_Players AS FOP
JOIN
[Olympic_Players].DimSport AS DS ON FOP.Sport_ID = DS.Sport_ID
JOIN
[Olympic_Players].DimMedal AS Medal ON FOP.Medal_ID = Medal.Medal_ID
WHERE
Medal.Medal IS NOT NULL
GROUP BY
FOP.Year, DS.Sport
ORDER BY
FOP.Year, TotalMedals DESC;
```

Output:



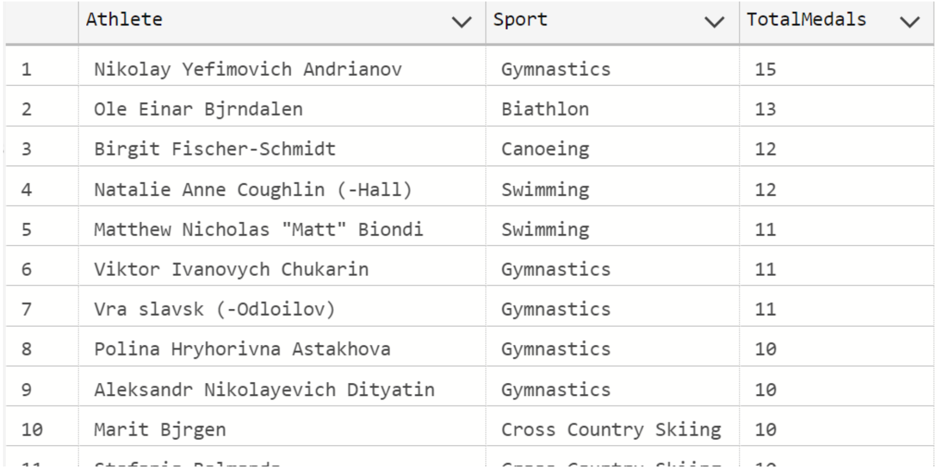
	Year	Sport	TotalMedals
1	1896	Athletics	11
2	1896	Tennis	4
3	1896	Gymnastics	3
4	1896	Swimming	2
5	1896	Weightlifting	2
6	1896	Shooting	1
7	1896	Fencing	1
8	1896	Cycling	1
9	1900	Sailing	36
10	1900	Athletics	25
11	1900	Rowing	24

The above query analyzes performance trends over time by counting the total number of medals won in each sport for each year. It provides insights into how medal counts vary across sports and years, aiding in understanding historical performance trends.

2.Identifying Emerging Talent:

```
SELECT
DP.Name AS Athlete,
DS.Sport,
COUNT(*) AS TotalMedals
FROM
[Olympic_Players].Fact_Olympic_Players AS FOP
JOIN
[Olympic_Players].DimPlayer AS DP ON FOP.Player_ID = DP.Player_ID
JOIN
[Olympic_Players].DimSport AS DS ON FOP.Sport_ID = DS.Sport_ID
JOIN
[Olympic_Players].DimMedal AS Medal ON FOP.Medal_ID = Medal.Medal_ID
WHERE
Medal.Medal IS NOT NULL
GROUP BY
DP.Name, DS.Sport
ORDER BY
TotalMedals DESC;
```

Output:



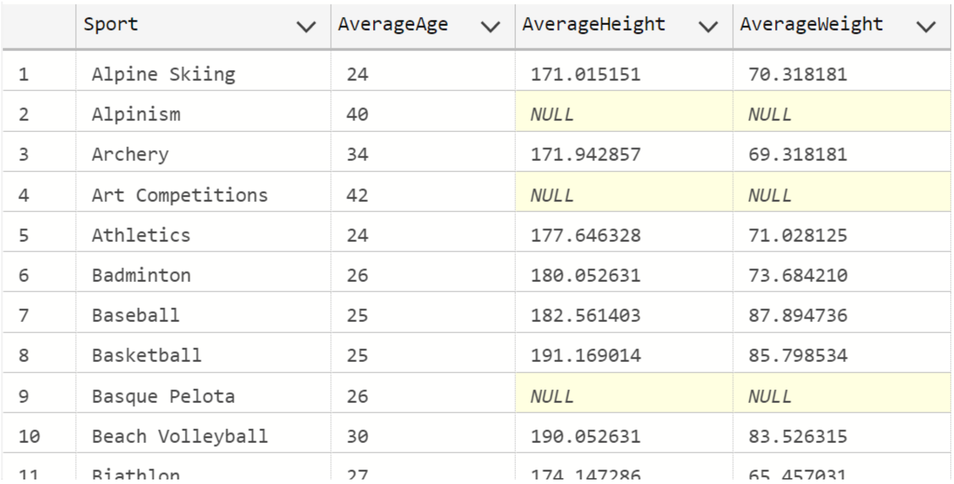
	Athlete	Sport	TotalMedals
1	Nikolay Yefimovich Andrianov	Gymnastics	15
2	Ole Einar Bjrndalen	Biathlon	13
3	Birgit Fischer-Schmidt	Canoeing	12
4	Natalie Anne Coughlin (-Hall)	Swimming	12
5	Matthew Nicholas "Matt" Biondi	Swimming	11
6	Viktor Ivanovych Chukarin	Gymnastics	11
7	Vra slavsk (-Odloilov)	Gymnastics	11
8	Polina Hryhorivna Astakhova	Gymnastics	10
9	Aleksandr Nikolayevich Dityatin	Gymnastics	10
10	Marit Bjrgen	Cross Country Skiing	10
11	Stefania Belmondo	Cross Country Skiing	10

The above query identifies emerging talent by counting the total number of medals won by each athlete in each sport. It helps stakeholders pinpoint athletes who have recently achieved notable success, potentially indicating rising stars in Olympic sports.

3.Optimizing Resource Allocation:

```
SELECT
DS.Sport,
AVG(FOP.Age) AS AverageAge,
AVG(FOP.Height) AS AverageHeight,
AVG(FOP.Weight) AS AverageWeight
FROM
[Olympic_Players].Fact_Olympic_Players AS FOP
JOIN
[Olympic_Players].DimSport AS DS ON FOP.Sport_ID = DS.Sport_ID
WHERE
FOP.Medal_ID IS NOT NULL
GROUP BY
DS.Sport;
```

Output:



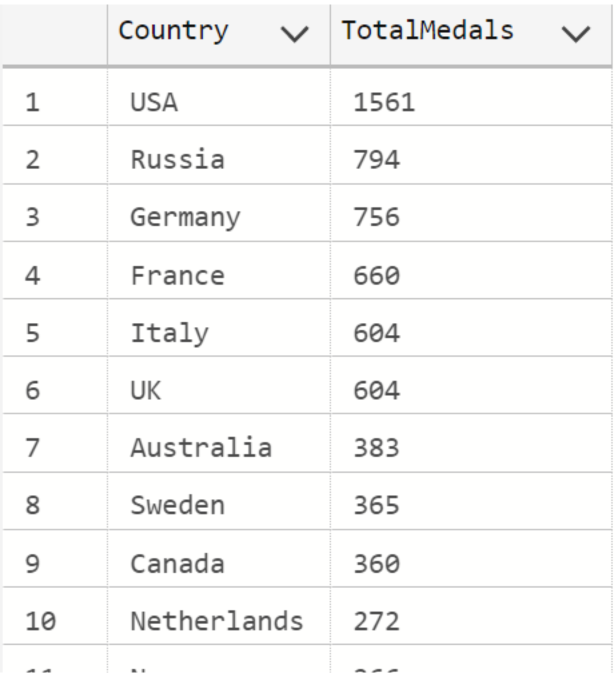
	Sport	AverageAge	AverageHeight	AverageWeight
1	Alpine Skiing	24	171.015151	70.318181
2	Alpinism	40	NULL	NULL
3	Archery	34	171.942857	69.318181
4	Art Competitions	42	NULL	NULL
5	Athletics	24	177.646328	71.028125
6	Badminton	26	180.052631	73.684210
7	Baseball	25	182.561403	87.894736
8	Basketball	25	191.169014	85.798534
9	Basque Pelota	26	NULL	NULL
10	Beach Volleyball	30	190.052631	83.526315
11	Biathlon	27	174.147286	65.457031

The above query optimizes resource allocation by calculating the average age, height, and weight of athletes who have won medals in each sport. It assists in understanding the physical attributes of successful athletes, aiding in tailored resource allocation and training program design.

4. Medal Distribution by Country

```
SELECT
    NOC.Region AS Country,
    COUNT(*) AS TotalMedals
 FROM
    [Olympic_Players].Fact_Olympic_Players AS FOP
 JOIN
    [Olympic_Players].DimNOC AS NOC ON FOP.NOC_ID = NOC.NOC_ID
 JOIN
    [Olympic_Players].DimMedal AS Medal ON FOP.Medal_ID = Medal.Medal_ID
 WHERE
    Medal.Medal IS NOT NULL
 GROUP BY
    NOC.Region
 ORDER BY
    TotalMedals DESC;
```

Output:



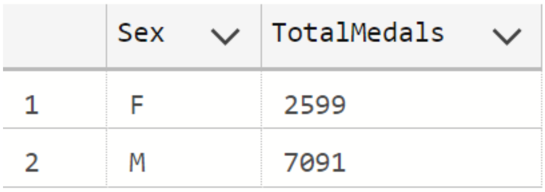
	Country	TotalMedals
1	USA	1561
2	Russia	794
3	Germany	756
4	France	660
5	Italy	604
6	UK	604
7	Australia	383
8	Sweden	365
9	Canada	360
10	Netherlands	272

This query counts the total number of medals won by each country in the Olympic Games. It provides insights into the distribution of medals among nations, revealing which countries have achieved the highest medal counts.

5.Analyzing Medal Success by Gender:

```
SELECT
DP.Sex,
COUNT(*) AS TotalMedals
FROM
[Olympic_Players].Fact_Olympic_Players AS FOP
JOIN
[Olympic_Players].DimPlayer AS DP ON FOP.Player_ID = DP.Player_ID
JOIN
[Olympic_Players].DimMedal AS Medal ON FOP.Medal_ID = Medal.Medal_ID
WHERE
Medal.Medal IS NOT NULL
GROUP BY
DP.Sex;
```

Output:

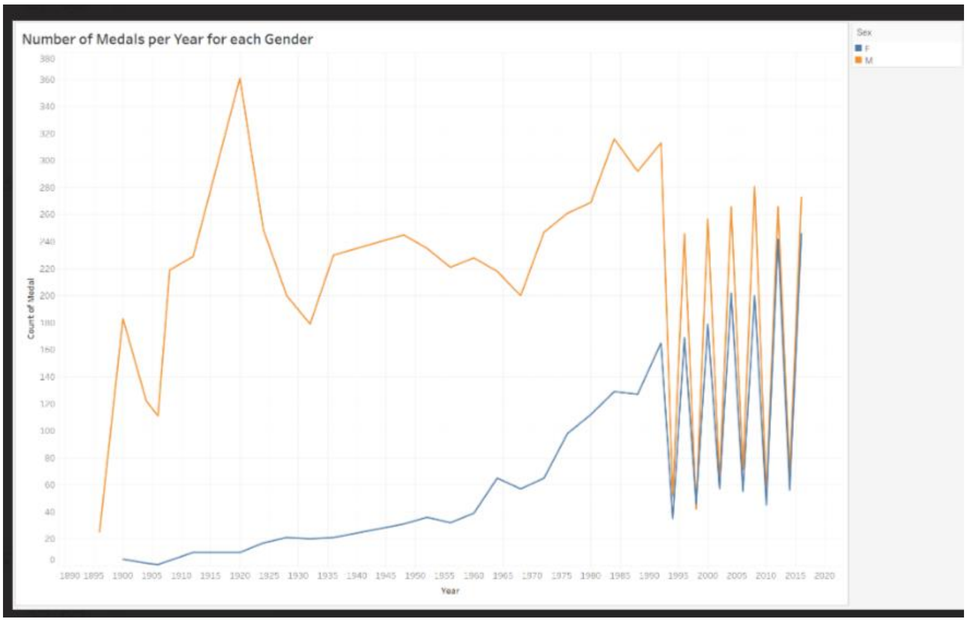


	Sex	TotalMedals
1	F	2599
2	M	7091

The above query analyses medal success by gender by counting the total number of medals won by male and female athletes. It provides insights into gender disparities in Olympic medal counts, aiding in understanding gender representation and performance in Olympic sports.

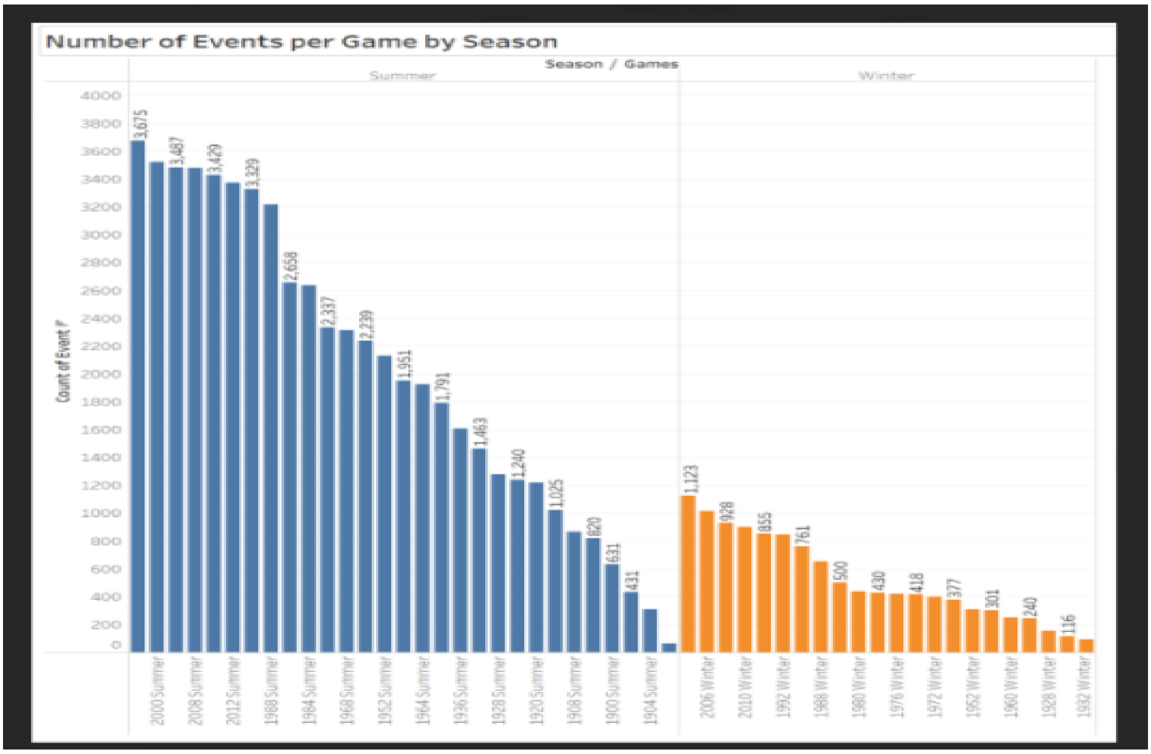
Reporting, Modelling, and Storytelling

1. Medal Count Over Years by Gender:



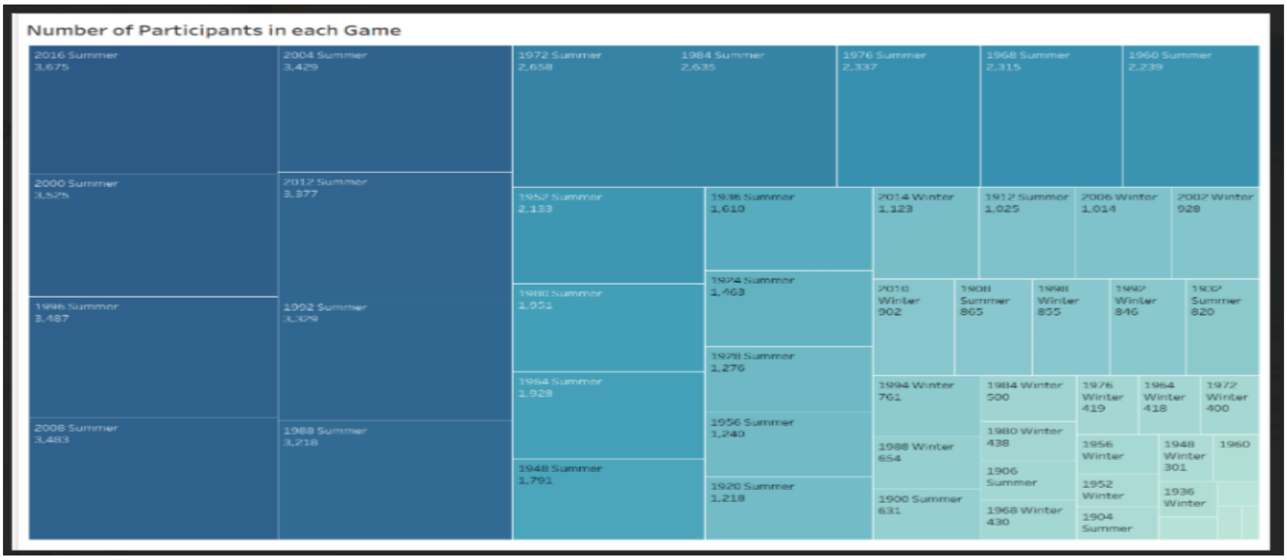
This above line graph illustrates the distribution of Olympic medals over the years, segmented by gender. The graph displays the total count of medals awarded to male and female athletes across various Olympic Games editions. By examining the trends in medal counts for both genders over time, stakeholders can gain insights into gender-specific performance patterns in the Olympics.

2. Olympic Events Distribution Over Seasons:



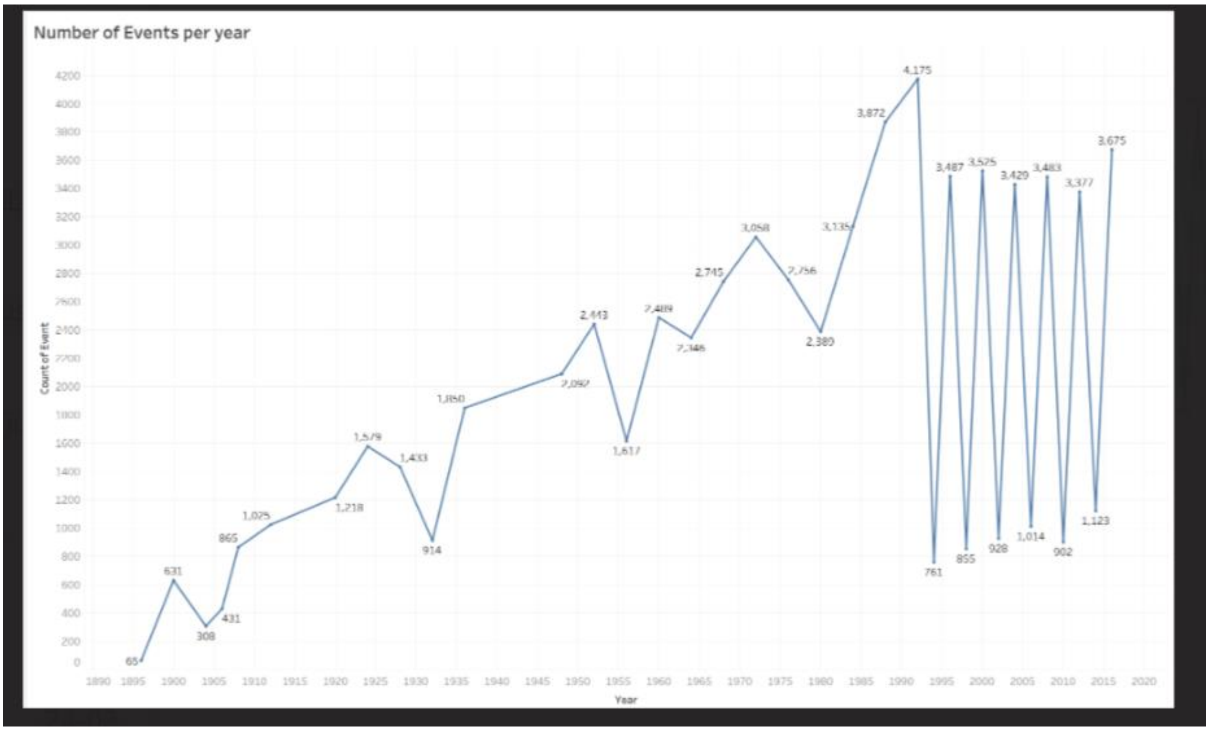
This visualization displays the distribution of events across the two seasons of the Olympic Games. The histogram reveals patterns in event distribution, indicating fluctuations in event frequency across seasons and games. This insight is crucial for optimizing strategies related to Olympic participation, event scheduling, and resource allocation.

3. Participant Count Across Olympic Games



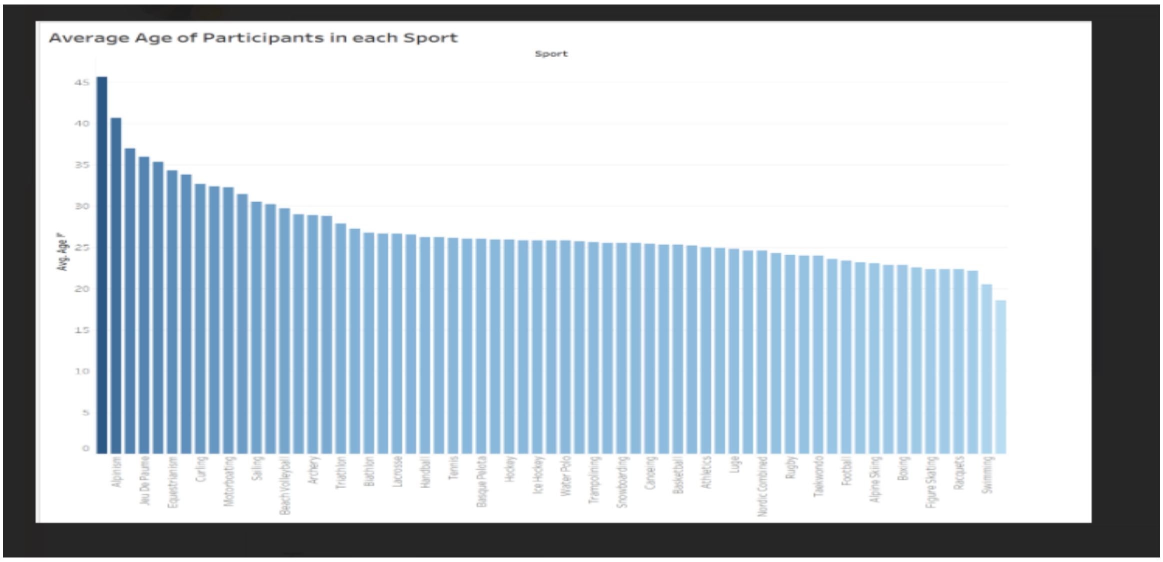
The visualization above presents the count of participants in each Olympic game. It provides insights into the popularity and scale of different Olympic events. This allows stakeholders to identify games with high participation rates and those that may require more attention or promotion to increase participation.

4. Annual Distribution of Olympic Events :



This visualization showcases the count of events held each year in the Olympics. It provides an overview of the yearly distribution of Olympic activities. This allows to track the growth or decline of Olympic events over time and identify years with significant changes in event frequency. Analyzing the number of events per year helps understand the evolution of the Olympic Games and anticipate future trends in event scheduling.

5. Average Age of Participants by Olympic Sport



The visualization illustrates the average age of participants in each Olympic sport. It offers insights into the demographic characteristics of athletes in different sports. This information can be valuable for understanding the age profile of Olympic athletes and identifying trends or patterns in age distribution across sports.

Conclusion

The project aimed to analyze historical Olympic data using data analytics techniques to extract insights beneficial for sports organizations and stakeholders. Leveraging the "120 Years of Olympic History: Athletes and Results" dataset from Kaggle, it explored performance trends, identified emerging talent, optimized resource allocation, and analyzed medal distribution by country and gender.

Through various SQL queries and visualizations, we have analyzed performance trends over time, identified rising stars, optimized resource allocation, examined medal distribution by country and gender, and presented insights through compelling visualizations such as line graphs, histograms, and bar charts. Overall, we aimed to demonstrate the power of data analytics in enhancing athletic performance and strategic decision-making in Olympic sports.

References

1. Data Set: Olympic Data 🏅 ⛷️🥅🏋️🚴 (kaggle.com)
2. The Data Warehouse Toolkit, 3rd Edition - Kimball Group