**SQL Project for Tokyo 2020 Olympics Analysis**

**by Nivedita Aneja**

**Source of the Data:** <https://www.kaggle.com/datasets/piterfm/tokyo-2020-olympics/data?select=medals.csv>

**Data Sets Used:**

**Athletes:** includes some personal information of all athletes

**Medals:** includes general information on all athletes who won a medal

**Purpose of the Project:** The goal of the project is to analyse athlete participation and medal distribution at the Tokyo 2020 Olympics using SQL. Insights derived include total athlete participation, country-wise performance, and athlete-level achievements.

**Tables created:** Athletes and Medals

**Problem Statements and Analytical Questions**

1. **Total Participation:**  
   How many athletes participated in the Tokyo 2020 Olympics?
2. **Countries Participated:**How many countries took part in the Games?
3. **Country-wise Athlete Representation:**How many athletes represented each country?
4. **Gender-wise Participation:**What is the distribution of athletes by gender across all countries?
5. **Female Participation in Sports:**Which sports saw participation from at least one female athlete?
6. **Youngest Medal Winner:**Who was the youngest athlete to win a medal?
7. **Country Medal Performance:**Which countries won the most medals?
8. **Medal Type Distribution:**How many Gold, Silver, and Bronze medals were awarded across all events?

**Example Queries and Insights**

**1. Total Participation**

**SELECT COUNT(DISTINCT name) AS athletes\_count**

**FROM athletes;**

***Insight:*** There were 11,629 athletes who participated in the Olympics

**2. Countries Participated**

**SELECT Count(distinct country) as country\_count**

**from athletes;**

***Insight:*** 206 countries participated.

**3. Country-wise Athlete Representation**

**SELECT country, COUNT(DISTINCT name) AS total\_athletes**

**FROM athletes**

**GROUP BY country**

**ORDER BY total\_athletes DESC;**

***Insight:*** Shows the number of athletes who participated from each country.

**4.** **Gender-wise Participation**

**SELECT gender, count(gender) as gender\_cat\_count**

**from athletes**

**group by gender;**

***Insight:*** Shows the distribution of athlete participation on the basis of gender.

5. **Female Participation in Sports**

**SELECT DISTINCT discipline**

**from athletes**

**where gender = 'Female';**

***Insight:*** Shows the sports for which women have participated.

**6. Youngest Medal Winner**

**SELECT name, birth\_date**

**FROM athletes**

**ORDER BY birth\_date DESC**

**LIMIT 1;**

***Insight:*** Presents the youngest Medal winner.

**7. Country Medal Performance**

**SELECT country, count(\*) as total\_medal**

**FROM medals**

**GROUP BY country**

**ORDER BY total\_medal desc;**

***Insight:*** Highlights the performance of each country in terms of medals earned.

1. **Medal Type Distribution**

**SELECT medal\_type, medal\_code, COUNT(\*) AS medal\_count**

**FROM medals**

**GROUP BY medal\_type, medal\_code**

**ORDER BY medal\_count DESC;**

***Insight:*** Shows the total number of all types of medals won by all the countries.

**Key Takeaways**

* **Extensive Participation:**Over 11,000 athletes participated across multiple disciplines, representing diverse countries and sporting events.
* **Country-wise Insights:**Countries with larger delegations generally won more medals, highlighting a link between participation volume and success.
* **Gender Representation:**Both male and female athletes were well represented, with participation spanning nearly all disciplines.
* **Youngest Medal Winner:**The youngest medallist demonstrated that elite performance is not restricted by age, underscoring the rise of young talent in global sports.
* **Country Medal Performance:**A few countries dominated the medal table — notably the U.S., R.O.C, and China — reflecting strong athletic programs and resources.

**Conclusion:***The analysis of the Tokyo 2020 Olympics datasets highlights global diversity in participation and performance. With over 11,000 athletes representing more than 200 countries, the Games demonstrated both inclusivity and competitive excellence. The near gender balance and widespread female participation across sports reflected continued progress toward equality in athletics.*

*The data-driven exploration provided clear insights into how participation scale, talent distribution, and sporting investment shape Olympic success.*