OLYMPIC

SQL DATA ANALYSIS PROJECT

Created By Rushikesh karpe

PROJECT OVERVIEW

❖ I worked on an **Olympic Medalists Data Analysis** project.

❖ The goal was to find **top countries**, **sports**, and **athletes** from the dataset.

❖ I used **SQL** and **Excel** for data cleaning and analysis.

❖ I found that the **United States** won the most medals, and **Athletics** was the top sport.

❖ This project helped me learn more about **data handling** and **visual skills**.

PROJECT GOAL

❖To analyze Olympic medal data from different years.

❖To find which **countries**, **sports**, and **athletes** won the most medals.

❖To understand trends in **medal wins** over time.

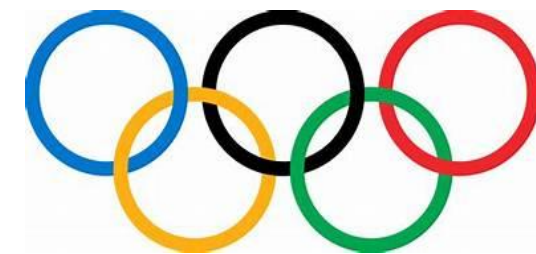❖To improve my skills in **data analysis** using Python and Excel.

TOOL AND TECHNOLOGY

DATASET DESCRIPTION

❖**Source**: Kaggle – All Olympic Medalists Dataset

❖**Total Rows**: 23196 (example – update with actual count if needed)

❖**Columns**: 10

❖**Format**: CSV (Comma Separated Values)

QUERIES OF SQL 1

```
--won the medal win by the india----

select *
from olympic
where   country='India';
```

season text	year integer	medal text	country_code text	country text	athletes text
1	Summer	1900	Silver	IND	India	Norman PRITCHARD
2	Summer	1900	Silver	IND	India	Norman PRITCHARD
3	Summer	1928	Gold	IND	India	Broome Eric PINNIGER, Feroze Uddin KHAN, Maurice GATELEY, Rex A. NORRIS, Leslie HAMMOND, George MARTHINS, Michael ROCQUE, Jaipal SINGH, William John GOODSIR-CULLEN, Richard ALLEN, Dhyan CHAND, Ali S
4	Summer	1932	Gold	IND	India	Carlyle TAPSELL, Sayed Mohammed JAFFAR, Dhyan CHAND, Leslie HAMMOND, Masud MINHAS, Broome Eric PINNIGER, Arthur HIND, Roop SINGH, Lal Shah BOKHARI, Gurmit SINGH KULLAR, Richard CARR
5	Summer	1936	Gold	IND	India	Ahsan Mohomed KHAN, Peter Paul FERNANDES, Gurcharan SINGH GAREWAL, Ahmed Sher KHAN, Sayed Mohomed HUSSAIN, Lionel EMMETT, Mirza Nasir-ud Din MASOOD, Dhyan CHAND, Carlyle TAPSELL, Earnest CULLE
6	Summer	1948	Gold	IND	India	Randhir SINGH GENTLE, Leslie CLAUDIUS, Amir KUMAR, Ranganandhan FRANCIS, Kishan LAL, Keshava DATT, Trilochan SINGH, Patrick JANSEN, Grahanandan SINGH, Latifur REHMAN, Walter D'SOUZA, Kunwar Digvijai SIN
7	Summer	1952	Gold	IND	India	Grahanandan SINGH, Keshava DATT, Govind PERUMAL, Ranganandhan FRANCIS, Randhir SINGH GENTLE, Raghbir LAL, Meldric St. Clair DALUZ, Udham SINGH, Dharam SINGH, Balbir SINGH DOSANJH, Leslie CLAUDIUS, C
8	Summer	1952	Bronze	IND	India	Khashaba Digvijai JHADAV
9	Summer	1956	Gold	IND	India	Shankar LAXMAN, Balkishan SINGH GREWAL, Bakshish SINGH, Udham SINGH, Raghbir LAL, Ranganandhan FRANCIS, Charles STEPHEN, Govind PERUMAL, Raghbir SINGH BHOLA, Amit SINGH BAKSHI, Gurdev SINGH, Hai
10	Summer	1960	Silver	IND	India	Leslie CLAUDIUS, Udham SINGH, John PETER, Raghbir SINGH BHOLA, Joseph ANTIC, Shankar LAXMAN, Prithipal SINGH, Joginder SINGH, Jaswant SINGH, Charanjit SINGH, Jaman LAL SHARMA, Mohinder LAL, Govind SA
11	Summer	1964	Gold	IND	India	Mohinder LAL, Ali SAYED, Prithipal SINGH, Bandu PATIL, Joginder SINGH, Gurbux SINGH, Dharam SINGH, Shankar LAXMAN, Harbinder SINGH, Jagjit SINGH, Haripal KAUSHIK, John PETER, Charanjit SINGH, Udham SINGH,

5

QUEIRES OF SQL 2

--which country have won the most medal in Olympic Histroy--

```sql
Select  country, COUNT(*) AS total_medals
From olympic
Group BY country
Order BY total_medals DESC;
```

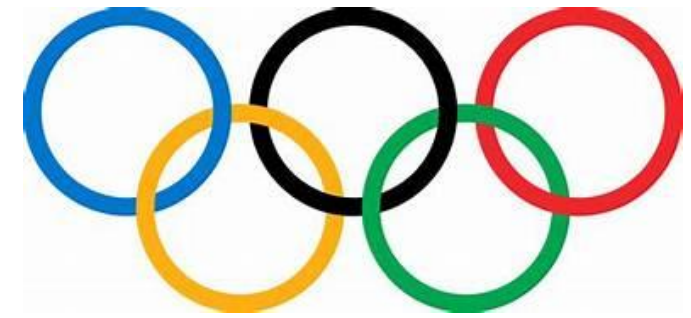|	country text	total_medals bigint
1	United States	2975
2	USSR	1204
3	Great Britain	954
4	France	928
5	Germany	918
6	Italy	759
7	China	715
8	Sweden	681
9	Australia	573
10	Norway	568
11	Japan	560
12	Canada	554

QUERIES OF SQL 3

---Medals won by India in each year---

```sql
Select year, COUNT(*) AS total
from olympic
where country='India'
group by year
order by year;
```

	year integer	total bigint
1	1900	2
2	1928	1
3	1932	1
4	1936	1
5	1948	1
6	1952	2
7	1956	1
8	1960	1
9	1964	1
10	1968	1
11	1972	1
12	1980	1

13	1996	1
14	2000	1
15	2004	1
16	2008	3
17	2012	6
18	2016	2
19	2024	6

QUERIES OF SQL 4

```sql
----Number of male vs female events----
select event_gender, COUNT(*)
From olympic
Group by event_gender;
```

	event_gender text	count bigint
1	Men's	13105
2	Women's	5934
3	Mixed	1208

8

QUERIES OF SQL 5

```sql
--whether Indian male or female athletes won more medals, ----
Select event_gender, COUNT(*) AS total_medals
From olympic
Where country = 'India'
Group BY event_gender
Order BY total_medals ASC;
```

	event_gender text	total_medals bigint
1	Mixed	1
2	Women's	6
3	Men's	27

QUERIES OF SQL 6

---Which Olympic Games have the most medals in history---

```sql
SELECT games,COUNT(*) AS total_medals
FROM olympic
GROUP BY games
ORDER BY total_medals DESC
LIMIT 1;
```

	games text	total_medals bigint
1	2024 Paris	1044

QUERIES OF SQL 7

```sql
--Which Olympic Games India  have the most medals in history---
SELECT games,COUNT(*) AS total_medals
FROM olympic
where country='India'
GROUP BY games
ORDER BY total_medals DESC
LIMIT 1;
```

games text	total_medals bigint
1	2012 London	6

QUERIES FOR SQL 8

```
---Top sports India won medals in ---

select  sport, count(*) AS medals
From olympic
WHERE country = 'India'
Group BY sport
order BY medals desc;
```

	sport text	medals bigint
1	Hockey	12
2	Shooting	7
3	Wrestling Freestyle	6
4	Athletics	3
5	Boxing	2
6	Badminton	2
7	Tennis	1
8	Weightlifting	1

QUERIES FOR SQL 9

```sql
----Indian athletes who won multiple medals-----
SELECT athletes, COUNT(*) AS medals
FROM olympic
WHERE country = 'India'
GROUP BY athletes
HAVING COUNT(*) > 1
ORDER BY medals DESC;
```

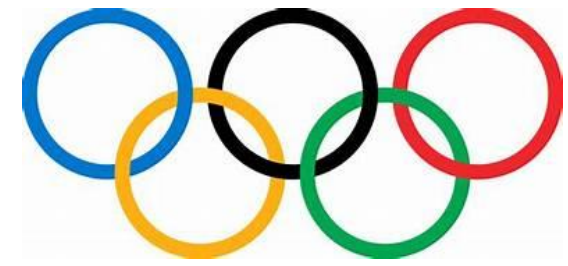|	athletes text	medals bigint
1	Norman PRITCHARD	2
2	Sushil KUMAR	2

QUERIES FOR SQL 10

```
----India's medals in 2024 Olympics----

SELECT COUNT(*) AS medals_2024
FROM olympic
WHERE country = 'India' AND year = 2024;
```

	medals_2024 bigint 🔒
1	6

CHALLENGES AND FUTURE SCOPE

Challenges

❖Some athlete names were missing or wrong in the dataset.

❖Duplicate records caused confusion in counts.

❖No direct column for number of events per athlete.

❖Data only covers winners, not participants.

Future Scope

❖ Add visual dashboards using Power BI or Tableau.

❖Analyze gender-wise medal trends.

❖Compare performance of countries over decades.

❖Use machine learning to predict medal wins.

CONCLUSION

❖ The United States has won the highest number of Olympic medals.

❖ Athletics, Swimming, and Gymnastics are the top medal-winning sports.

❖ Some athletes have won multiple medals, showing strong consistency.

❖ The dataset gives useful insights about country-wise and sport-wise performance.

❖ With more cleaning and visuals, the data can be used for deeper analysis and predictions.

17