

```
1 • create database sqlPrjt1;
2 • use sqlPrjt1;
3
```

- Create database sqlprjt1;

This command creates new database in MySQL Server with name sqlprjt1.

- Use sqlPrjt1;

Using this command a database(sqlPrjt1) is used and set to active database. Which stores table, views, procedure and other database objects.

```
4 • create table olympics(
5     ID int,
6     Name varchar(500),
7     Sex varchar(10),
8     Age int,
9     Height int,
10    Weight int,
11    Team varchar(350),
12    NOC varchar(100),
13    Games varchar(500),
14    Year int,
15    Season varchar(400),
16    City varchar(300),
17    Sport varchar(500),
18    Event varchar(400),
19    Medal varchar(200));
```

This command creates table in database.
Which can be seen with command "Show tables".

For creating columns in a table with different Data Types, we have specified datatype with its name.

ID int: ID will be name of column with 'int' data type, which will store integer values for each row.

Name varchar (500): Name column with varchar (500) will store variable character string with up to 500 characters.

Sex varchar (10): Sex column with 10 characters will store gender of athlete 'M' or 'F'.

Age int: Integral column to store age of athletes.

Height int: An integer column to store height of athletes.

Weight int: An integer column to store weight of athletes.

Team varchar (350): Team column with character up to 350 to store Team name (i.e. Country Name)

NOC varchar (100): NOC (National Olympic Committees) with up to 100 characters, to store three letter country code.

Games varchar (500): Games column with characters up to 500 to store name of games played in Olympics.

Year int: Integer column to store Year when game took place.

Season varchar (400): Season column up to 400 characters to stores season of games (e.g. summer, winter)

City varchar (300): City column up to 300 characters to store City name.

Sport varchar (500): Sport column up to 500 character to store Names of the sports.

Event varchar (400): Event column up to 400 characters to store name of different event in Olympics of different sports.

Medal varchar (200): Medal column to store medals won by athletes (Gold, Silver, Bronze) or NoMedal for if no medal is won.

```
1 • show variables like "secure_file_priv";
2 • select * from olympics;
3
4 • load data infile "C:/ProgramData/MySQL/MySQL Server 9.0/Uploads/Athletes_Transformed.csv"
5   into table olympics
6   fields terminated by ','
7   enclosed by '"'
8   lines terminated by '\r\n'
9   ignore 1 rows;
```

show variables like "secure_file_priv";

This command will return the directory path where MySQL is allowed to read and write files. It ensures only files this directory can be accessed for security.

Load Data: Use the LOAD DATA INFILE statement to load the data into your table.

"C:/ProgramData/MySQL/MySQL Server 9.0/Uploads/Athletes_Transformed.csv" is the path of directory with file containing data (Athletes_Transformed.csv).

- **FIELDS TERMINATED BY ',':** Specifies that fields in the file are separated by commas.
- **ENCLOSED BY '"':** Specifies that fields are enclosed by double quotes.
- **LINES TERMINATED BY '\r\n':** Specifies that lines are terminated by carriage return and newline characters.
- **Ignore 1 rows:** to ignore first row which (often) includes header row in csv.

After this step our data will be loaded in our table which we created earlier with column specified with data present in our file. We also have Table Data Import Wizard which is handy to use for small datasets but for large data like this we have to use this method.

```
31 • select * from olympics
32   limit 1000;
```

We can check table with 'select * from Olympics' which will show Olympics table.

I have used **limit** to show only first 1000 rows.

```

34  -- SQL PROJECT
35  |
36  -- 1. Show how many medal counts present for entire data
37  • select medal,count(*) as medalcount
38  from olympics
39  where medal <> 'NoMedal'
40  group by medal;

```

This query is designed to count the number of medals (excluding 'NoMedal') for each type of medal in the olympics table.

1. **Selects the Medal Type:** select medal
2. **Counts the Number of Each Medal Type:** count (*) as medalcount
3. **From the Olympics Table:** from olympics
4. **Excludes Entries with 'NoMedal':** where medal <> 'NoMedal'
5. **Groups the Results by Medal Type:** group by medal

Result Grid			Filter Rows
	medal	medalcount	
▶	Gold	13372	
	Bronze	13295	
	Silver	13116	

```

42  -- 2. Show count of unique sports present in Olympics
43  • select Sport ,count(*) as SportsCount
44  from olympics
45  group by Sport;

```

In this SQL query is we count number of occurrences for each sport in the olympics table. Here's a breakdown of the query does:

- **SELECT Sport, COUNT(*) as SportsCount:** This selects the **Sport** column and counts the number of rows for each group, aliasing the count as **SportsCount**.
- **FROM olympics:** This specifies the table to query, which is **olympics**.
- **GROUP BY Sport:** This groups the results by the **Sport** column, allowing you to count the number of rows for each sport.

Result Grid			Filter Rows
	Sport	SportsCount	
▶	Basketball	4536	
	Judo	3801	
	Football	6745	
	Tug-Of-War	170	
	Speed Skating	5613	


Result 28 ×

```

47 -- 3. Show how many different medals won by team India
48 • select medal,count(*) as medalcount, Team
49 from olympics
50 where medal <> 'NoMedal' and Team='India'
51 group by medal;

```

- **SELECT medal, COUNT (*) as medalcount, Team:** This selects the **medal** and **Team** columns, and counts the number of rows for each group, aliasing the count as **medalcount**.
- **FROM olympics:** This specifies the table to query, which is **olympics**.
- **WHERE medal <> 'NoMedal' AND Team = 'India':** This filters the results to only include rows where the **medal** is not 'NoMedal' and the **Team** is 'India'.
- **GROUP BY medal:** This groups the results by the **medal** column, allowing you to count the number of medals for each type of medal.

Result Grid			Filter Rows:
	medal	medalcount	Team
▶	Gold	138	India
	Silver	19	India
	Bronze	40	India

```

53 -- 4. Show event wise medals won by India show from highest to lowest medals won in order
54 • select Event,medal,count(*) as medalcount
55 from olympics
56 where medal <> 'NoMedal' and Team='India'
57 group by event,medal
58 order by medalcount desc;

```

- **SELECT Event, medal, COUNT (*) as medalcount:** This selects the **Event** and **medal** columns, and counts the number of rows for each group, aliasing the count as **medalcount**.
- **FROM olympics:** This specifies the table to query, which is **olympics**.
- **WHERE medal <> 'NoMedal' AND Team = 'India':** This filters the results to only include rows where the **medal** is not 'NoMedal' and the **Team** is 'India'.
- **GROUP BY Event, medal:** This groups the results by both the **Event** and **medal** columns, allowing you to count the number of medals for each event and medal type.
- **ORDER BY medalcount DESC:** This sorts the results in descending order based on the **medalcount** column, so the events with the most medals will be shown first.

Result Grid	Filter Rows:	Export:	W
Event	medal	medalcount	
Hockey Mens Hockey	Gold	130	
Hockey Mens Hockey	Bronze	30	
Hockey Mens Hockey	Silver	13	
Alpinism Mixed Alpinism	Gold	7	
Shooting Mens Air Rifle 10 metres	Gold	1	

```

60  -- 5. Show event wise medals won by India in order of year
61  •  select Event,Year,medal
62      from olympics
63      where medal <> 'NoMedal' and Team='India'
64      group by event,medal,Year
65      order by Year;

```

Select specific columns: The query selects three columns from the **olympics** table: **Event**, **Year**, and **medal**. This means that only these three columns will be included in the result set.

- **Filter out unwanted rows:** The query uses the **WHERE** clause to filter out rows that don't meet certain conditions. Specifically, it excludes rows where the **medal** column is equal to **'NoMedal'** and includes only rows where the **Team** column is equal to **'India'**.

- **Sort the result set:** The query uses the **ORDER BY** clause to sort the result set in ascending order by the **Year** column. This means that the earliest year will come first, and the latest year will come last.

- **Retrieve the result set:** The query retrieves the result set from the **olympics** table, which includes only the rows that meet the conditions specified in the **WHERE** clause and are sorted in the order specified by the **ORDER BY** clause.

Result Grid	Filter Rows:	Export:
Event	Year	medal
Athletics Mens 200 metres	1900	Silver
Athletics Mens 200 metres Hurdles	1900	Silver
Alpinism Mixed Alpinism	1924	Gold
Hockey Mens Hockey	1928	Gold
Hockey Mens Hockey	1932	Gold

```

66  -- 6. show country who won maximum medals.
67  •  select Team,medal,count(*) as medalcount
68      from olympics
69      where medal <> 'NoMedal'
70      group by Team,medal
71      order by medalcount desc;

```

- **Selects the Team and Medal Type:** SELECT Team, medal
- **Counts the Number of Each Medal Type for Each Team:** COUNT (*) AS medalcount
- **From the Olympics Table:** FROM olympics
- **Excludes Entries with 'NoMedal':** WHERE medal <> 'NoMedal'
- **Groups the Results by Team and Medal Type:** GROUP BY Team, medal
- **Orders the Results by Medal Count in Descending Order:** ORDER BY medalcount DESC

Result Grid	Filter Rows:	Export:
Team	medal	medalcount
United States	Gold	2474
United States	Silver	1512
United States	Bronze	1233
Soviet Union	Gold	1058
Soviet Union	Silver	716

```

73      -- 7.show top 10 countries who won gold
74 •    select Team,medal,count(*) as medalcount
75      from olympics
76      where medal = 'Gold'
77      group by Team,medal
78      order by medalcount desc
79      limit 10;

```

- Select specific columns: The query selects the Team, medal, and the count of gold medals (medalcount).
- Filter out unwanted rows: The WHERE clause includes only rows where the medal is 'Gold'.
- Group the results: The GROUP BY clause groups the results by Team and medal.
- Sort the result set: The ORDER BY clause sorts the result set in descending order by medalcount.
- Limit the result set: The LIMIT clause limits the result set to the top 10 rows.

	Team	medal	medalcount
►	United States	Gold	2474
	Soviet Union	Gold	1058
	Germany	Gold	500
	Italy	Gold	359
	Great Britain	Gold	519

Result 41 ×

```

81      -- 8. show in which year did United states won most gold medal
82 •    select Team,Year,medal,count(*) as medalcount
83      from olympics
84      where medal = 'Gold' and Team='United States'
85      group by Team,medal,Year
86      order by medalcount desc;
87

```

- **SELECT Team, Year, COUNT (*) AS medalcount:** Selects the team, year, and counts the number of gold medals for each year.
- **FROM olympics:** Specifies the table to query from.
- **WHERE medal = 'Gold' AND Team = 'United States':** Filters the records to include only gold medals won by the United States.
- **GROUP BY Team, Year:** Groups the results by team and year so that COUNT (*) can aggregate the number of medals per year.
- **ORDER BY Year:** Orders the results by year, which is usually more intuitive to see the progression over time.

	Team	Year	medal	medalcount
►	United States	1984	Gold	190
	United States	1996	Gold	157
	United States	2012	Gold	139
	United States	2016	Gold	137
	United States	2000	Gold	128

Result 46 ×


```

89      -- 9. In which sports United States has most medals.
90 •    SELECT Team, Sport, medal, COUNT(*) AS medalcount
91      FROM olympics
92      WHERE medal <> 'NoMedal' AND Team = 'United States'
93      GROUP BY Team, Sport, medal
94      ORDER BY medalcount DESC;

```

- **SELECT Team, Sport, medal, COUNT (*) AS medalcount:** Selects the team, sport, type of medal, and counts the number of medals for each combination.
- **FROM olympics:** Specifies the table to query from.
- **WHERE medal <> 'NoMedal' AND Team = 'United States':** Filters the records to include only those where the medal is not 'NoMedal' and the team is 'United States'.
- **GROUP BY Team, Sport, medal:** Groups the results by team, sport, and medal type so that COUNT (*) aggregates the number of medals for each combination.
- **ORDER BY medalcount DESC:** Orders the results by the count of medals in descending order.

Team	Sport	medal	medalcount
United States	Swimming	Gold	645
United States	Athletics	Gold	537
United States	Athletics	Silver	313
United States	Basketball	Gold	281
United States	Swimming	Silver	250

```

96      -- 10. Find top three players who won most medals along with their sports and country
97 •    select Name,Team,Sport,medal,count(*) as medalcount
98      from olympics
99      where medal <>'NoMedal'
100     group by Name,medal,Sport,Team
101     order by medalcount desc
102     limit 3;

```

- **SELECT Name, Team, Sport, medal, COUNT (*) AS medalcount:** Selects the athlete's name, team, sport, type of medal, and counts the number of medals won.
- **FROM olympics:** Specifies the table to query from.
- **WHERE medal <> 'NoMedal':** Filters out records where the medal is 'NoMedal', focusing only on actual medal wins.
- **GROUP BY Name, medal, Sport, Team:** Groups the results by athlete name, medal type, sport, and team so that COUNT (*) can aggregate the number of medals per grouping.
- **ORDER BY medalcount DESC:** Orders the results by the number of medals won, in descending order.
- **LIMIT 3:** Limits the results to the top 3 rows based on the highest medal counts.

Name	Team	Sport	medal	medalcount
Michael Fred Phelps II	United States	Swimming	Gold	23
Raymond Clarence Ray Ewry	United States	Athletics	Gold	10
Frederick Carlton Carl Lewis	United States	Athletics	Gold	9

```

105 -- 11. Find player with most gold medals in cycling along with his country.
106 • select Name,Team,Sport,medal,count(*) as medalcount
107 from olympics
108 where medal <>'NoMedal' and Sport='Cycling'
109 group by Name,medal,Sport,Team
110 order by medalcount desc
111 limit 1;

```

- **SELECT Name, Team, Sport, medal, COUNT (*) AS medalcount:** Retrieves the athlete's name, team, sport, type of medal, and counts the number of medals won.
- **FROM olympics:** Specifies the table to query from.
- **WHERE medal <> 'NoMedal' AND Sport = 'Cycling':** Filters records to include only those where the medal is not 'NoMedal' and the sport is 'Cycling'.
- **GROUP BY Name, medal, Sport, Team:** Groups the results by athlete name, medal type, sport, and team to count the number of medals for each grouping.
- **ORDER BY medalcount DESC:** Orders the results by the count of medals in descending order.
- **LIMIT 1:** Limits the results to the top 1 row, giving the athlete with the highest number of medals in the specified sport.

Result Grid Filter Rows: Export: Wrap Cell Con					
	Name	Team	Sport	medal	medalcount
▶	Christopher Andrew Chris Hoy	Great Britain	Cycling	Gold	6

```

114 -- 12. Find player with most medals (Gold+Silver+Bronze) in Basketball also show his country.
115 • select Name,Team,Sport,count(*) as totalmcount
116 from olympics
117 where medal in ('Gold','Silver','Bronze') and Sport='Basketball' and medal <> 'NoMedal'
118 group by Name,Team,Sport
119 order by totalmcount desc
120 limit 1;

```

- **SELECT Name, Team, Sport, COUNT (*) AS totalmcount:** Selects the athlete's name, team, sport, and counts the number of medals.
- **FROM olympics:** Specifies the table to query from.
- **WHERE medal IN ('Gold', 'Silver', 'Bronze'):** Filters the records to include only those where the medal is either 'Gold', 'Silver', or 'Bronze'.
- **AND Sport = 'Basketball':** Further filters the results to include only records where the sport is 'Basketball'.
- **AND medal <> 'NoMedal':** Excludes records where the medal is 'NoMedal'.
- **GROUP BY Name, Team, Sport:** Groups the results by athlete name, team, and sport to aggregate the number of medals.
- **ORDER BY totalmcount DESC:** Orders the results by the count of medals in descending order.
- **LIMIT 1:** Limits the results to the top 1 row, which will be the athlete with the highest number of medals.

Result Grid Filter Rows: Export:				
	Name	Team	Sport	totalmcount
▶	Teresa Edwards	United States	Basketball	5


```

122 -- 13. Find out the count of different medals of the top basketball player Teresa Edwards
123 • select Name,Team,Sport,medal,count(*) as Medalcount
124 from olympics
125 where name = 'Teresa Edwards' and medal <> 'NoMedal'
126 group by Name,Team,Sport,Medal;

```

- **SELECT Name, Team, Sport, medal, COUNT (*) AS Medalcount:** Selects the name of the athlete, their team, sport, medal type, and counts the number of each type of medal.
- **FROM olympics:** Specifies the table to query from.
- **WHERE Name = 'Teresa Edwards' AND medal <> 'NoMedal':** Filters the records to include only those where the athlete's name is 'Teresa Edwards' and the medal is not 'NoMedal'.
- **GROUP BY Name, Team, Sport, medal:** Groups the results by athlete name, team, sport, and medal type to aggregate the number of medals.

Result Grid Filter Rows: <input type="text"/> Export: Wrap					
	Name	Team	Sport	medal	Medalcount
▶	Teresa Edwards	United States	Basketball	Gold	4
	Teresa Edwards	United States	Basketball	Bronze	1

```

128 -- 14. Find out medals won by male,female each year , Export this data and plot graph in excel
129 • select Year,Sex,count(*) as Medalwon
130 from olympics
131 where medal in ('Gold','Silver','Bronze') and medal <> 'NoMedal'
132 group by Year,Sex
133 order by Year;

```

- **SELECT Year, Sex, COUNT (*) AS Medalwon:** Selects the year, sex, and counts the number of medals.
- **FROM olympics:** Specifies the table to query from.
- **WHERE medal IN ('Gold', 'Silver', 'Bronze'):** Filters records to include only those with medals of 'Gold', 'Silver', or 'Bronze'.
- **AND medal <> 'NoMedal':** Ensures 'NoMedal' entries are excluded.
- **GROUP BY Year, Sex:** Groups the results by year and sex to count the number of medals per combination.
- **ORDER BY Year:** Orders the results by year to see the progression over time.

Result Grid Filter Rows: <input type="text"/>			
	Year	Sex	Medalwon
▶	1896	M	143
	1900	F	13
	1900	M	591
	1904	F	10
	1904	M	476
Result 69 x			