**Olympics Data analysis using PostgreSQL**

* **By katari Pavan**

CREATE TABLE OLYMPICS\_HISTORY (

ID INT,

NAME VARCHAR,

SEX VARCHAR,

AGE VARCHAR,

HEIGHT VARCHAR,

WEIGHT VARCHAR,

TEAM VARCHAR,

NOC VARCHAR,

GAMES VARCHAR,

YEAR INT,

SEASON VARCHAR,

CITY VARCHAR,

SPORT VARCHAR,

EVENT VARCHAR,

MEDAL VARCHAR

);

CREATE TABLE OLYMPICS\_HISTORY\_NOC\_REGIONS (NOC VARCHAR, REGION VARCHAR, NOTE VARCHAR);

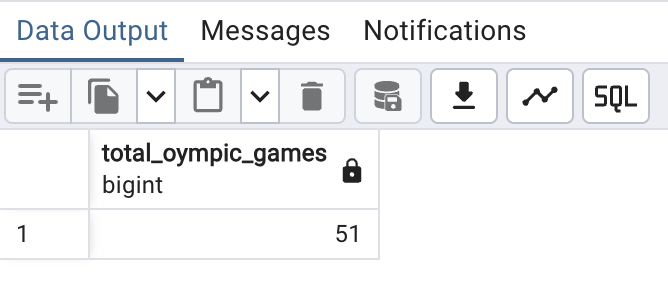
SELECT \* FROM OLYMPICS\_HISTORY;

SELECT \* FROM OLYMPICS\_HISTORY\_NOC\_REGIONS;

**Q1 find out the total number of olympic games have been played**

SELECT COUNT(DISTINCT GAMES) AS TOTAL\_OYMPIC\_GAMES

FROM OLYMPICS\_HISTORY;



**Q2 list down all olympic games held so far**

SELECT DISTINCT YEAR, SEASON, CITY

FROM OLYMPICS\_HISTORY ORDER BY YEAR;



**Q3 mention the total no.of nations who participated in each olympics game**

WITH ALL\_COUNTRIES AS (

SELECT GAMES,NR.REGION

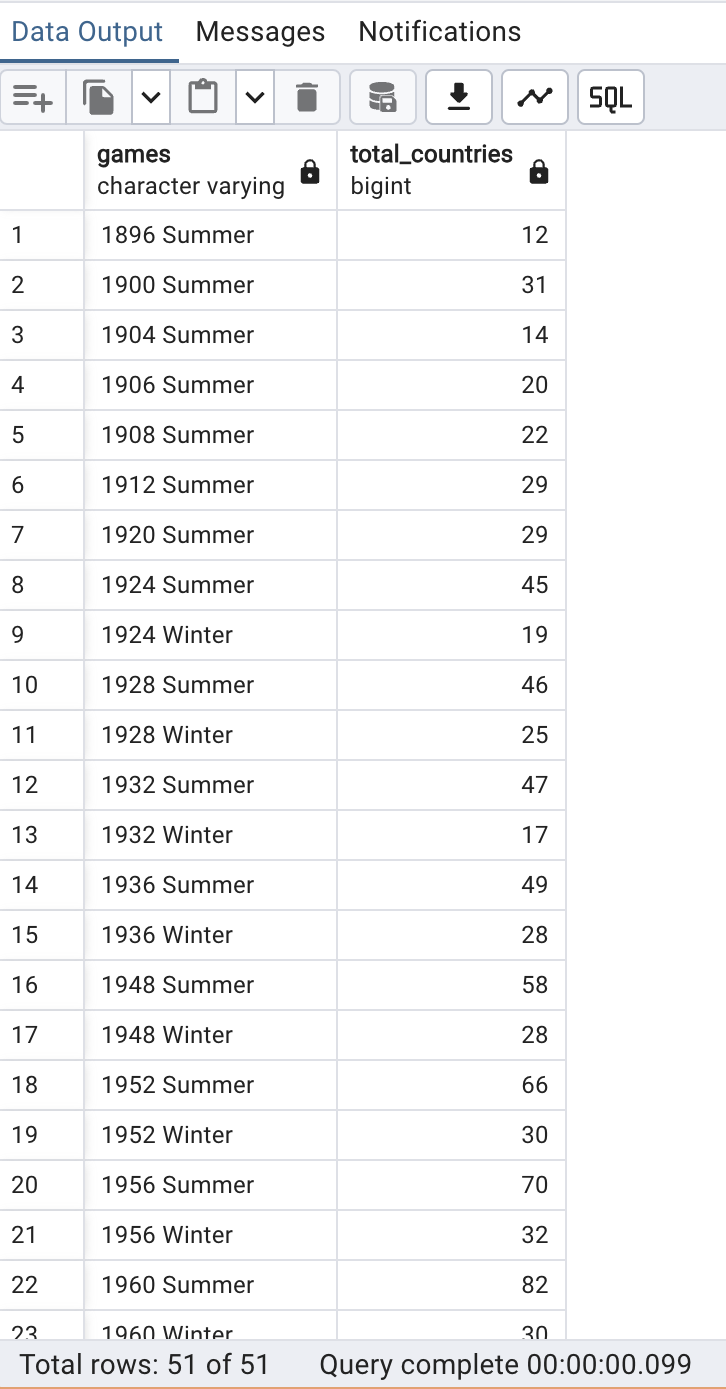
FROM OLYMPICS\_HISTORY OH

JOIN OLYMPICS\_HISTORY\_NOC\_REGIONS NR ON NR.NOC = OH.NOC

GROUP BY GAMES,NR.REGION )

SELECT GAMES, COUNT(1) AS TOTAL\_COUNTRIES

FROM ALL\_COUNTRIES GROUP BY GAMES ORDER BY GAMES;



**Q4 which nation participated in all olympics**

WITH TOT\_GAMES AS (

SELECT COUNT(DISTINCT GAMES) AS TOTAL\_GAMES

FROM OLYMPICS\_HISTORY),

COUNTRIES AS (

SELECT GAMES, NR.REGION AS COUNTRY

FROM OLYMPICS\_HISTORY OH

JOIN OLYMPICS\_HISTORY\_NOC\_REGIONS NR ON NR.NOC = OH.NOC

GROUP BY GAMES, NR.REGION ),

COUNTRIES\_PARTICIPATED AS (

SELECT COUNTRY,

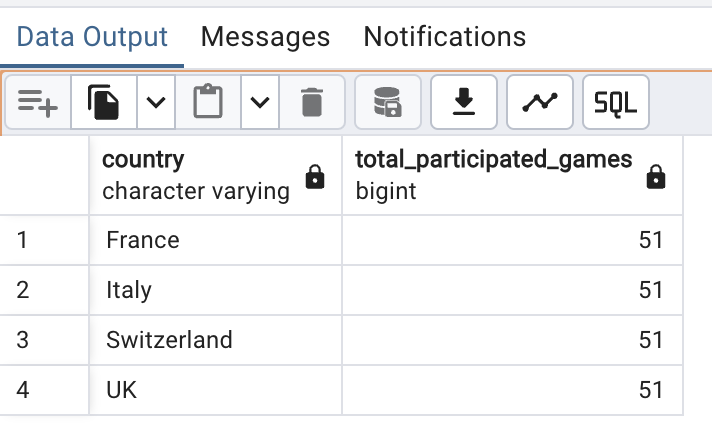
COUNT(1) AS TOTAL\_PARTICIPATED\_GAMES

FROM COUNTRIES GROUP BY COUNTRY)

SELECT CP.\* FROM COUNTRIES\_PARTICIPATED CP

JOIN TOT\_GAMES TG ON TG.TOTAL\_GAMES = CP.TOTAL\_PARTICIPATED\_GAMES

ORDER BY 1;



-- Q5 identify the sport which was played in all summer olympics

WITH T1 AS (

SELECT COUNT(DISTINCT GAMES) AS TOTAL\_SUMMER\_GAMES

FROM OLYMPICS\_HISTORY

WHERE SEASON = 'Summer'),

T2 AS ( SELECT DISTINCT SPORT, GAMES

FROM OLYMPICS\_HISTORY

WHERE SEASON = 'Summer'

ORDER BY GAMES),

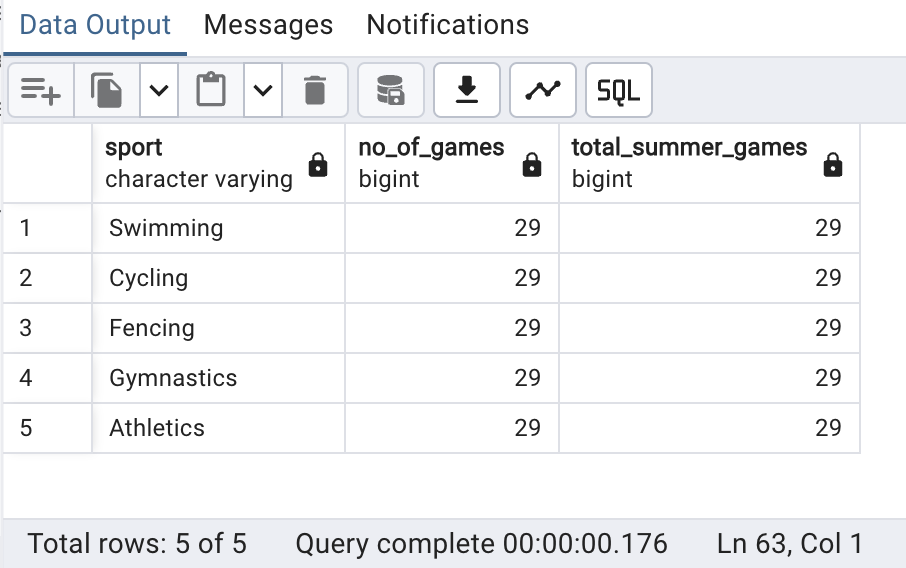
T3 AS ( SELECT SPORT, COUNT(GAMES) AS NO\_OF\_GAMES

FROM T2 GROUP BY SPORT)

SELECT \* FROM

T3

JOIN T1 ON T1.TOTAL\_SUMMER\_GAMES = T3.NO\_OF\_GAMES;



**Q6 identify the sport which was played only once in Olympics**

WITH T1 AS (

SELECT DISTINCT GAMES, SPORT

FROM OLYMPICS\_HISTORY),

T2 AS (

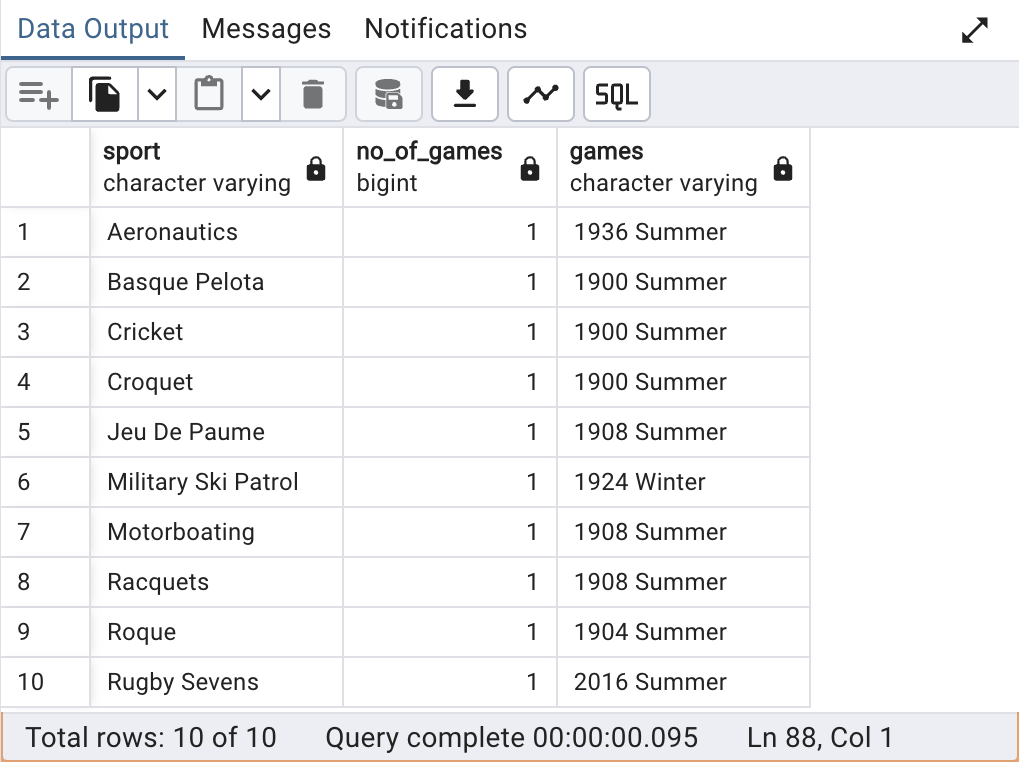
SELECT SPORT, COUNT(1) AS NO\_OF\_GAMES

FROM T1 GROUP BY SPORT )

SELECT T2.\*, T1.GAMES

FROM T2 JOIN T1 ON T1.SPORT = T2.SPORT

WHERE T2.NO\_OF\_GAMES = 1 ORDER BY T1.SPORT;



**Q7 calculate total number of sports played in each olymics**

WITH T1 AS (

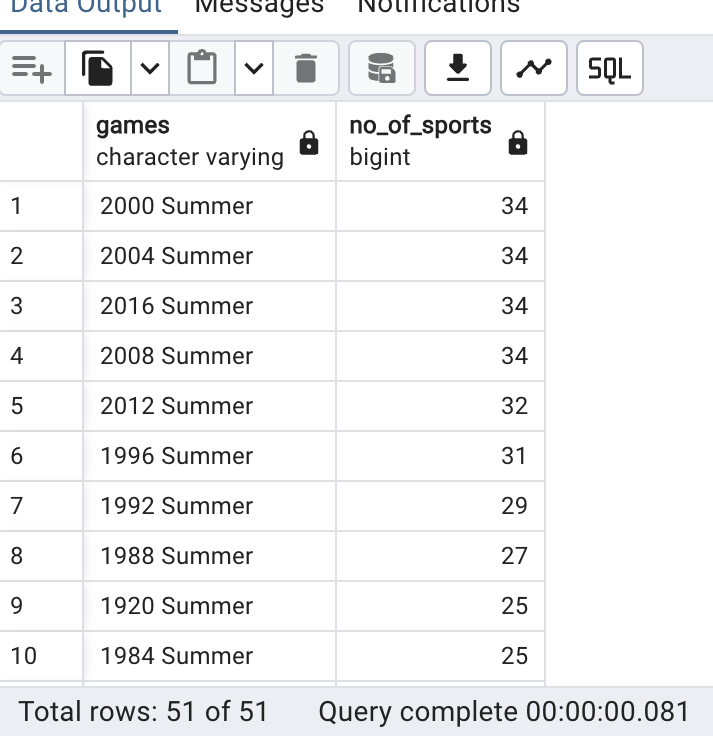
SELECT DISTINCT GAMES,SPORT

FROM OLYMPICS\_HISTORY),

T2 AS (SELECT GAMES,COUNT(1) AS NO\_OF\_SPORTS

FROM T1 GROUP BY GAMES)

SELECT \* FROM T2 ORDER BY NO\_OF\_SPORTS DESC;



**Q8 find out top 5 athletes who had won most gold medals**

WITH T1 AS (

SELECT NAME, TEAM, COUNT(1) AS TOTAL\_GOLD\_MEDALS

FROM OLYMPICS\_HISTORY WHERE MEDAL = 'Gold'

GROUP BY NAME, TEAM ORDER BY TOTAL\_GOLD\_MEDALS DESC),

T2 AS (

SELECT \*, DENSE\_RANK() OVER (

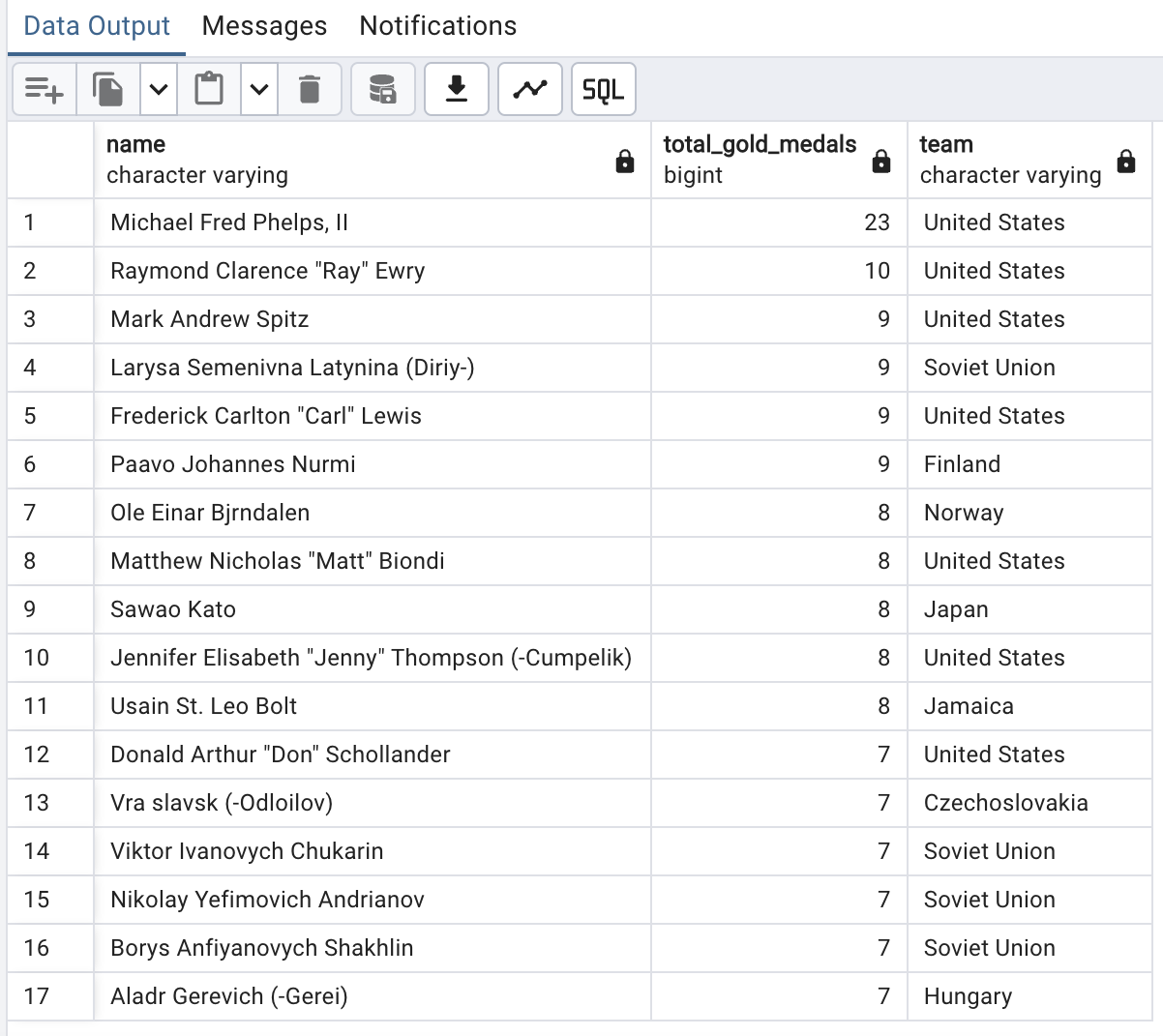
ORDER BY TOTAL\_GOLD\_MEDALS DESC

) AS RNK

FROM T1 )

SELECT NAME, TOTAL\_GOLD\_MEDALS, TEAM

FROM T2 WHERE RNK <= 5;



**Q9 find out top 5 athletes who had won most medals**

WITH T1 AS (

SELECT NAME, TEAM, COUNT(1) AS TOTAL\_MEDALS

FROM OLYMPICS\_HISTORY

WHERE MEDAL IN ('Gold', 'Silver', 'Bronze')

GROUP BY NAME, TEAM

ORDER BY TOTAL\_MEDALS DESC ),

T2 AS (

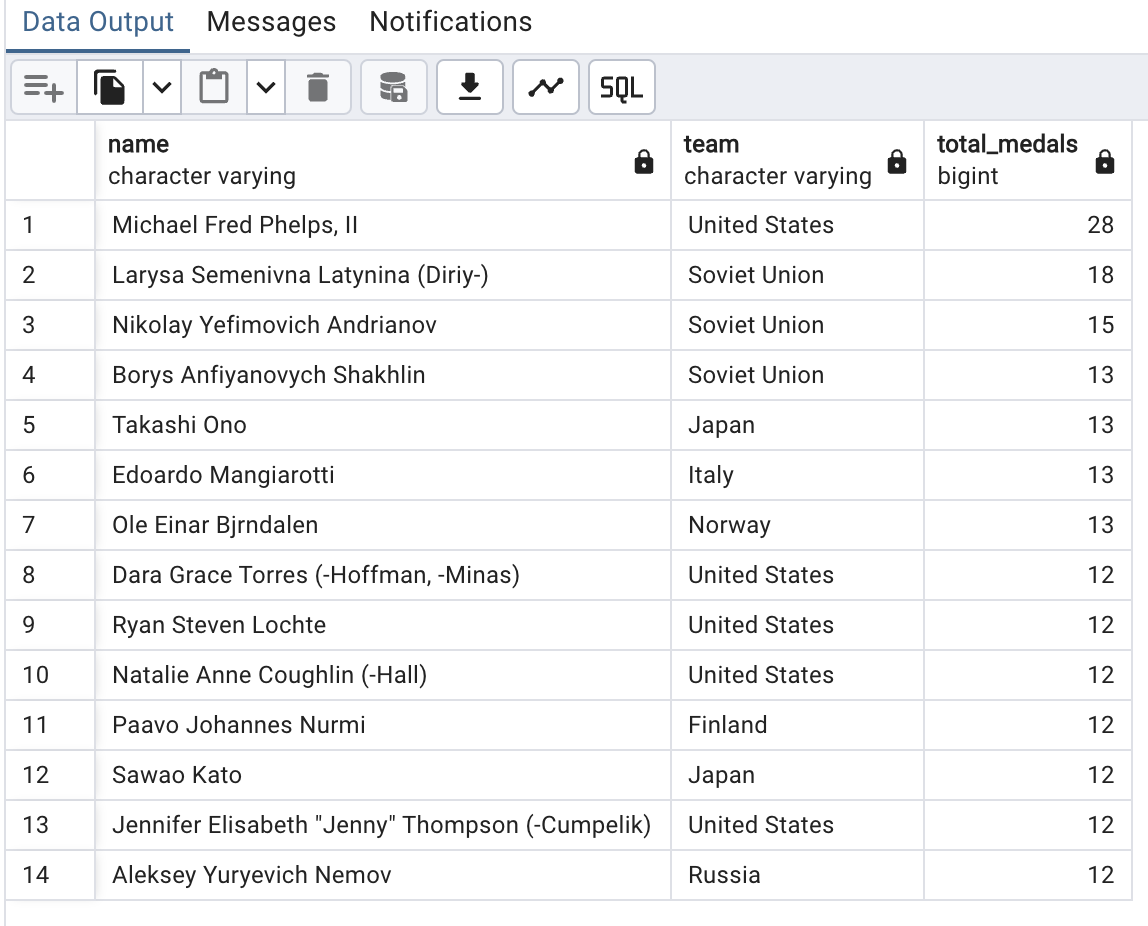
SELECT \*, DENSE\_RANK() OVER (

ORDER BY TOTAL\_MEDALS DESC ) AS RNK

FROM T1 )

SELECT NAME, TEAM, TOTAL\_MEDALS

FROM T2 WHERE RNK <= 5;



**Q10 find top 5 successful countries all olympics**

WITH T1 AS (

SELECT NR.REGION, COUNT(1) AS TOTAL\_MEDALS

FROM OLYMPICS\_HISTORY OH

JOIN OLYMPICS\_HISTORY\_NOC\_REGIONS NR ON NR.NOC = OH.NOC

WHERE MEDAL <> 'NA'

GROUP BY NR.REGION

ORDER BY TOTAL\_MEDALS DESC ),

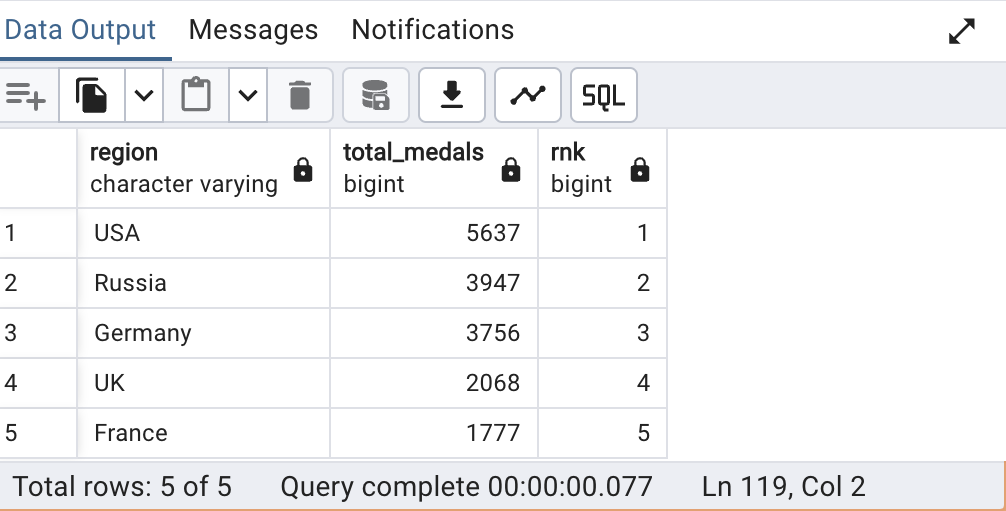
T2 AS (

SELECT \*, DENSE\_RANK() OVER (

ORDER BY TOTAL\_MEDALS DESC ) AS RNK

FROM T1)

SELECT \* FROM T2 WHERE RNK <= 5;



**Q11 find the total gold, silver and bronze medals won by each country**

SELECT

NR.REGION AS COUNTRY,

MEDAL,

COUNT(1) AS TOTAL\_MEDALS

FROM OLYMPICS\_HISTORY OH

JOIN OLYMPICS\_HISTORY\_NOC\_REGIONS NR ON NR.NOC = OH.NOC

WHERE MEDAL != 'NA'

GROUP BY NR.REGION, MEDAL

ORDER BY NR.REGION, MEDAL;

SELECT COUNTRY,

COALESCE(GOLD, 0) AS GOLD,

COALESCE(SILVER, 0) AS SILVER,

COALESCE(BRONZE, 0) AS BRONZE

FROM CROSSTAB (

'select nr.region as country,medal,count(1) as total\_medals

from OLYMPICS\_HISTORY oh

join OLYMPICS\_HISTORY\_NOC\_REGIONS nr on nr.noc = oh.noc

where medal != ''N''

group by nr.region,medal

order by nr.region,medal',

'values (''Bronze''),(''Gold''),(''Silver'')'

) AS RESULT (

COUNTRY VARCHAR,

BRONZE BIGINT,

GOLD BIGINT,

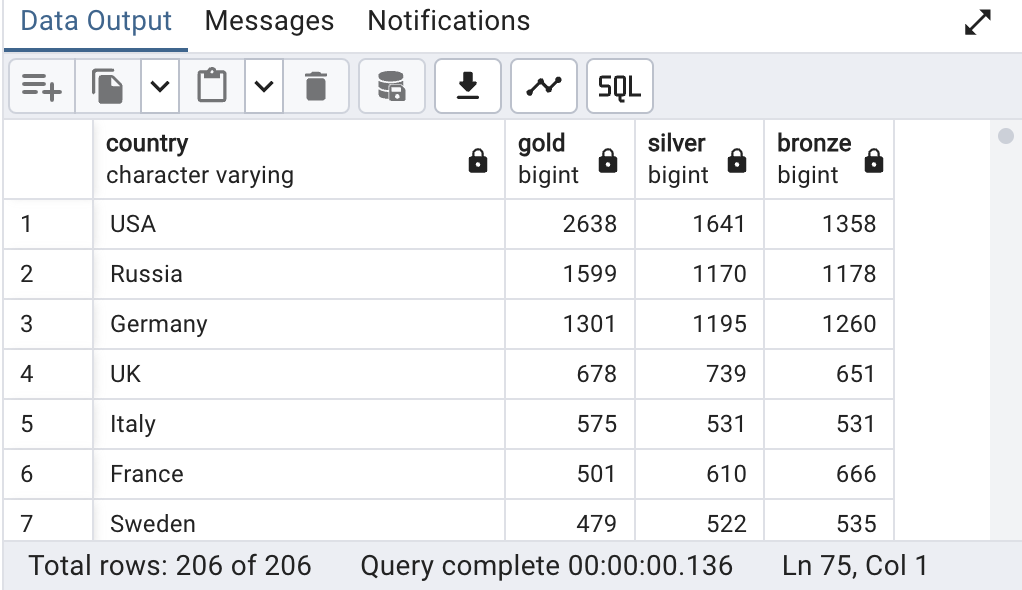
SILVER BIGINT

) ORDER BY

GOLD DESC,

SILVER DESC,

BRONZE DESC;



**Q12 In which event india has won most medals**

WITH T1 AS (

SELECT SPORT, COUNT(1) AS TOTAL\_MEDALS

FROM OLYMPICS\_HISTORY

WHERE MEDAL <> 'NA' AND TEAM = 'India'

GROUP BY SPORT

ORDER BY TOTAL\_MEDALS DESC ),

T2 AS (

SELECT \*, RANK() OVER (

ORDER BY TOTAL\_MEDALS DESC ) AS RNK

FROM T1 )

SELECT SPORT, TOTAL\_MEDALS

FROM T2 WHERE RNK = 1;

