



CAPSTONE PROJECT : SPORTS DATA ANALYSIS (OLYMPIC DATASET)

MECE FRAMEWORK

ATHLETE ANALYSIS

DEMOGRAPHIC INSIGHTS	CAREER ANALYSIS
Analyse athletes Gender, Age, Height, Weight and region.	Analysis related to Total no of season's played, Game and event that athlete is playing and Total no of medals won.

SEASON ANALYSIS

GAMES OCCURRENCE PATTERN	GAMES AND CITY	NOC related to games year
Analyse games year and its season.	Identifying which city had max no of games season.	Find out which National Olympic committee organised games in which year.

EVENT ANALYSIS

EVENT AND SPORT	ATHLETE RELATION WRT EVENT	TRENDS RELATED TO EVENTS
Analysis of events that fall into particular sport category.	Identifying which athlete plays under which event.	Evaluating which event have been added recently and which one is oldest.

SPORTS ANALYSIS

SPORTS WRT SUMMER/WINTER	SPORTS INSIGHTS ACC TO GAMES	LEAST PLAYED SPORTS
Identifying which sports were played in summer and winters respectively	Finding the total no of sports played in different Olympic games	Identifying the sports that have been played least among all the games.

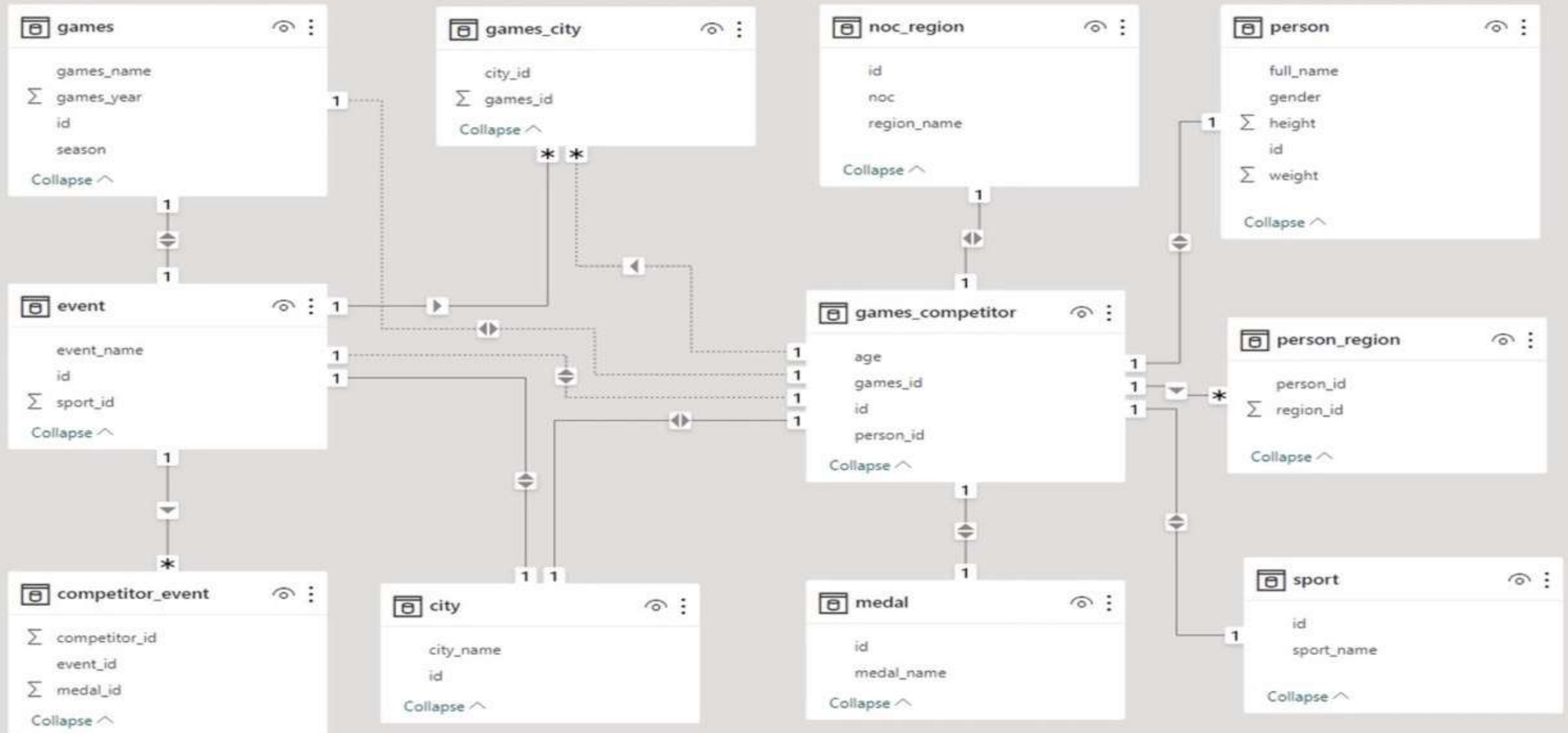
REGION ANALYSIS

REGION AND SPORTS/EVENT	REGION ANALYSIS
Identify which region participated in max and min events	Total no of regions participated in different games and which region has highest participation rate.

MEDAL ANALYSIS

ATHLETES AND REGION WRT MEDALS	GENDER AND MEDAL	AGE WRT MEDALS
Identifying relation between medals won by region and diff events related to it.	Finding relation between gender and no of medals.	Medal distribution according to age i.e oldest and youngest.

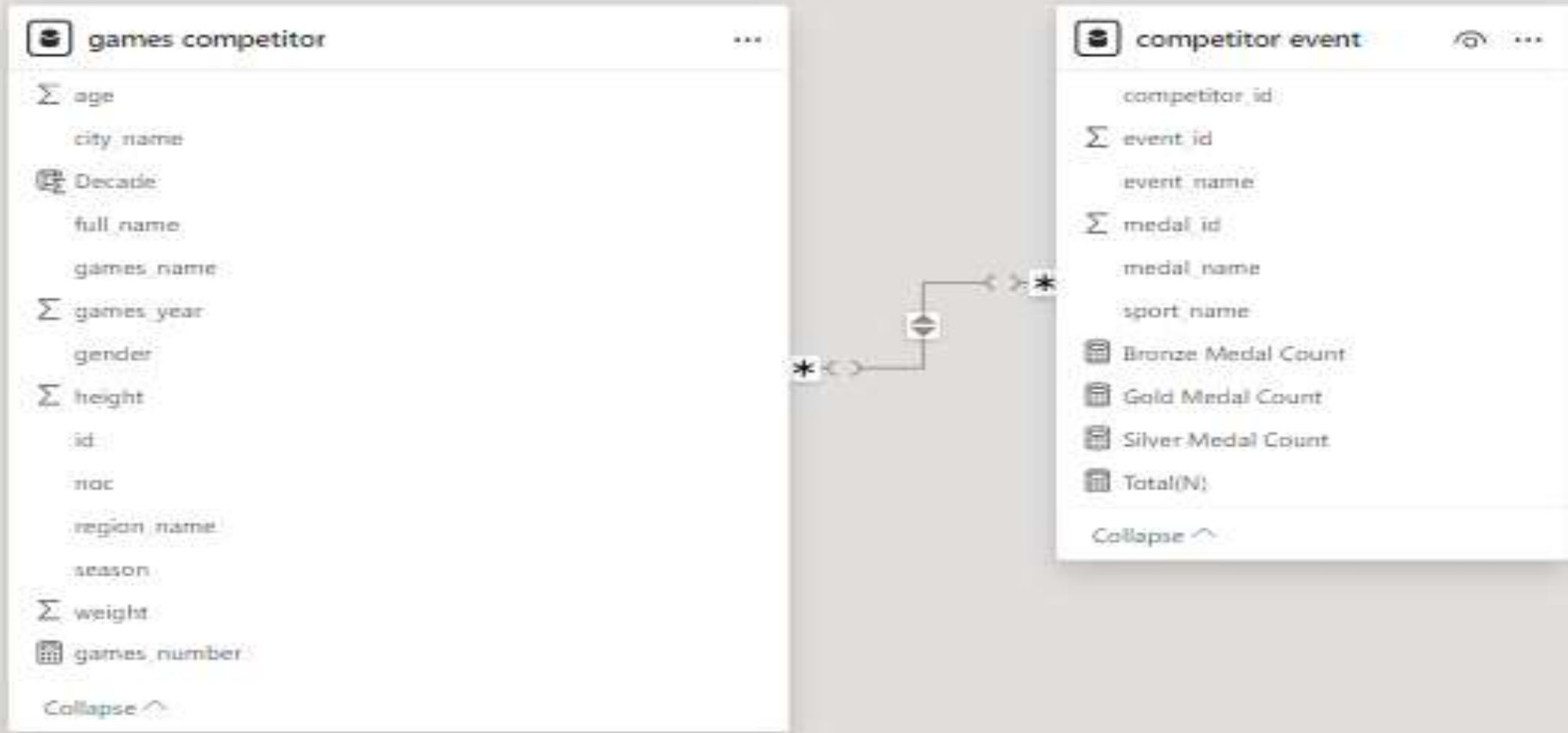
ER DIAGRAM



POWER BI QUESTIONS

The background is a gradient of teal and blue, with a subtle pattern of small white dots. Faint, light blue circular diagrams are visible in the corners. In the top right, there is a large circular diagram with concentric circles and radial lines, resembling a gauge or a complex chart. In the bottom left, there is a smaller circular diagram with a dashed line and an arrow. In the bottom right, there is another circular diagram with concentric circles and a dashed line.

MERGED TABLE AND NEW RELATIONSHIP



GAMES ANALYSIS



HOW MANY OLYMPIC GAMES HAVE BEEN HELD IN EACH SEASON (SUMMER VS. WINTER)?

ANALYSIS SUMMARY

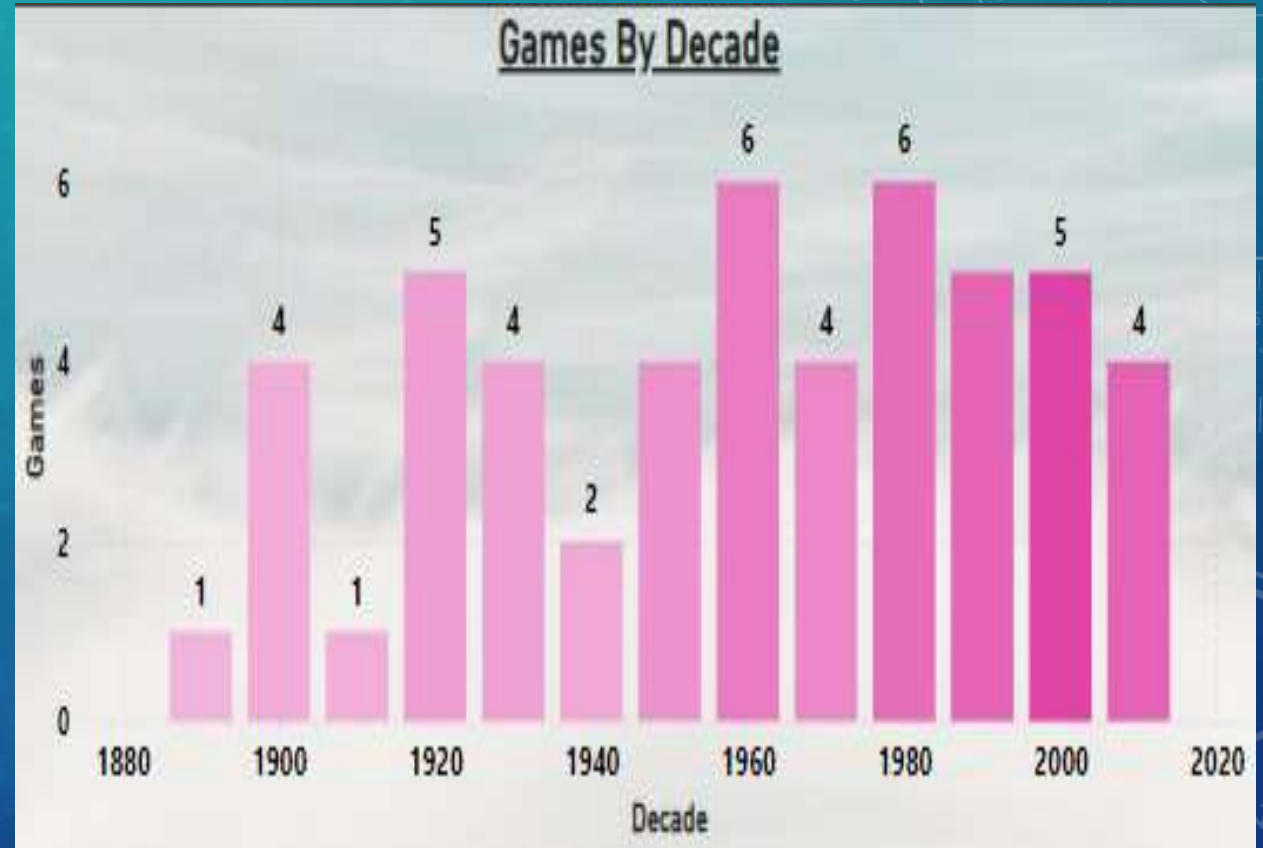
GLEANING FROM THE VISUALIZATION
TOTAL GAMES SEASON OF OLYMPIC
HAVE TAKEN PLACE MORE IN SUMMER
RATHER THAN IN WINTER.



What is the distribution of games across different decades?

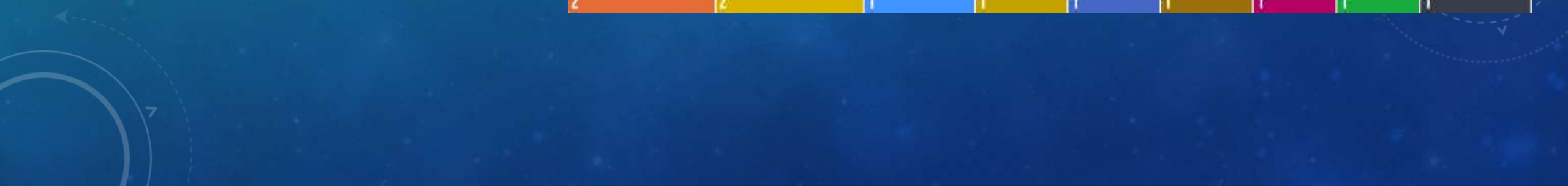
ANALYSIS SUMMARY

GLEANNING FROM THE VISUALIZATION
MAX NO OF GAMES HAVE TAKEN PLACE
IN 60'S AND 80'S AND MIN NO OF
GAMES WERE IN 1890'S AND 1910'S
BUT ONE CAN SAY THERE WAS ALMOST
CONSISTENCY IN GAMES AFTER 1960'S.



hosted the most Olympic Games?

GLEANNING FROM THE VISUAL
ATHINA, LONDON AND
INNSBRUCK HAVE HOSTED
GAMES THRICE WHEREAS
MAX NO OF CITIES HAVE
HOSTED GAMES ONLY ONCE.



SPORTS ANALYSIS

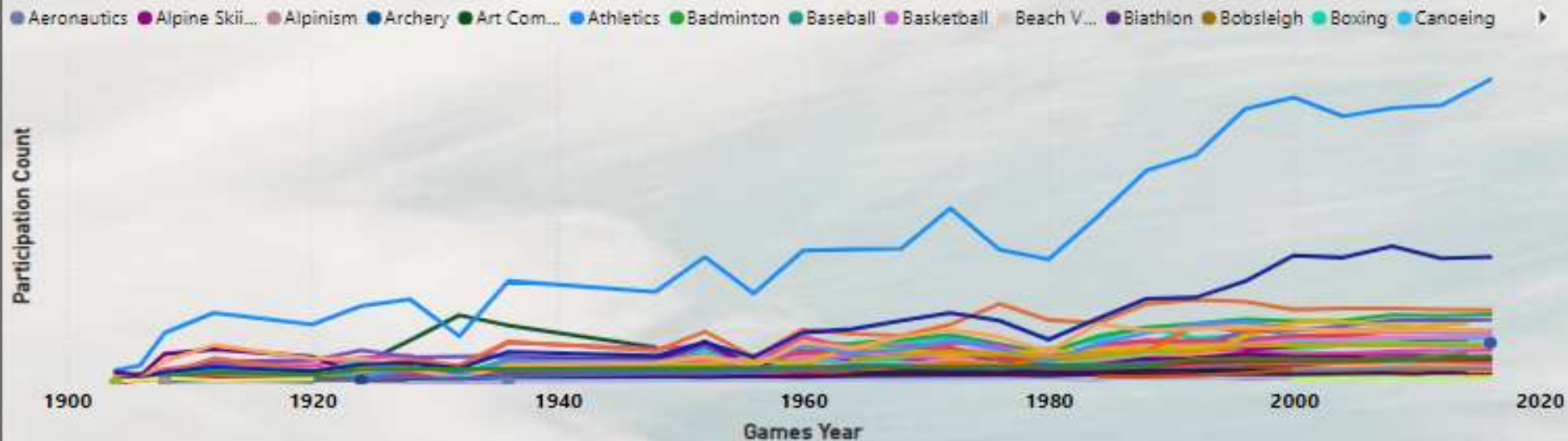
Sports Played By Season



Number Of Events Played In Sports



Participation Count In Different Sports wrt GameYears



Aeronautics

Alpine Skiing

Alpinism

Archery

Art Competitions

Athletics

Badminton

Baseball

Basketball

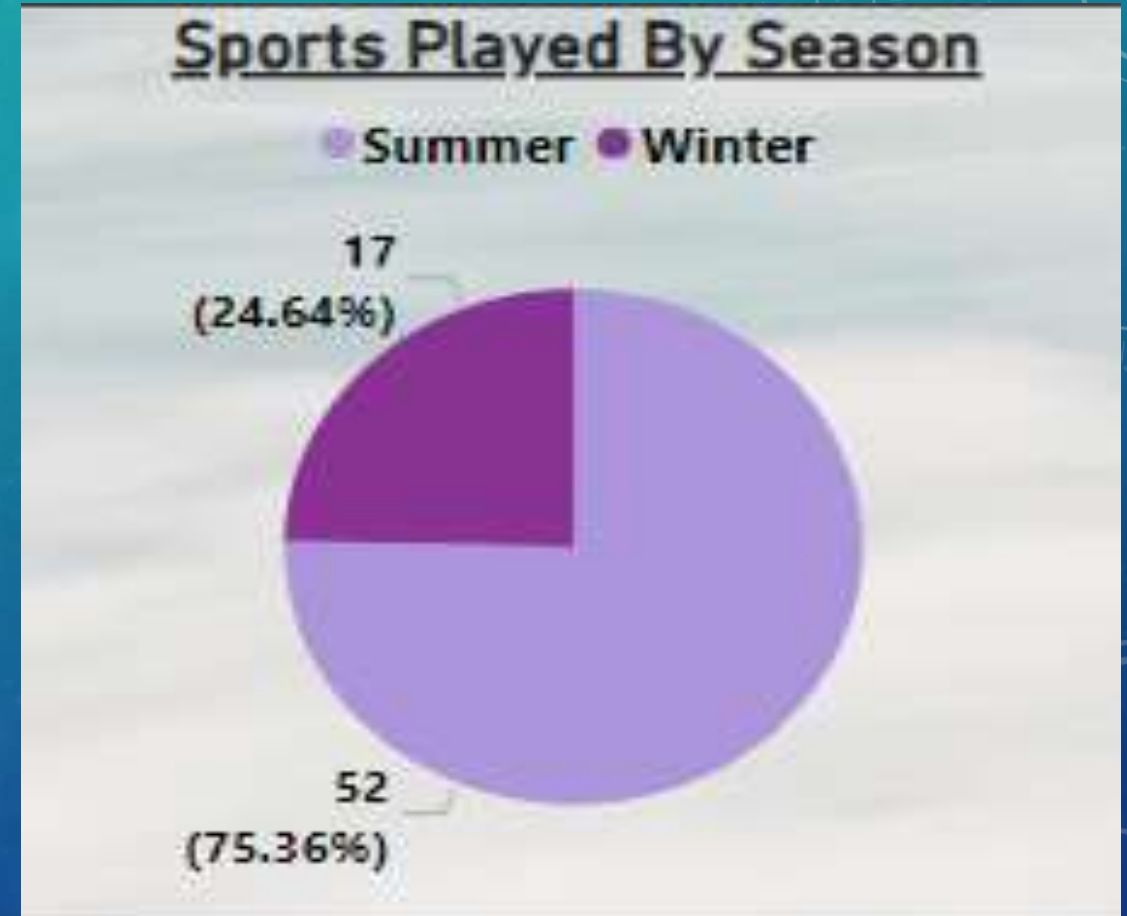
Basque Pelota

Beach Volleyball

What is the distribution of sports between Summer and Winter Olympics?

SUMMARY ANALYSIS

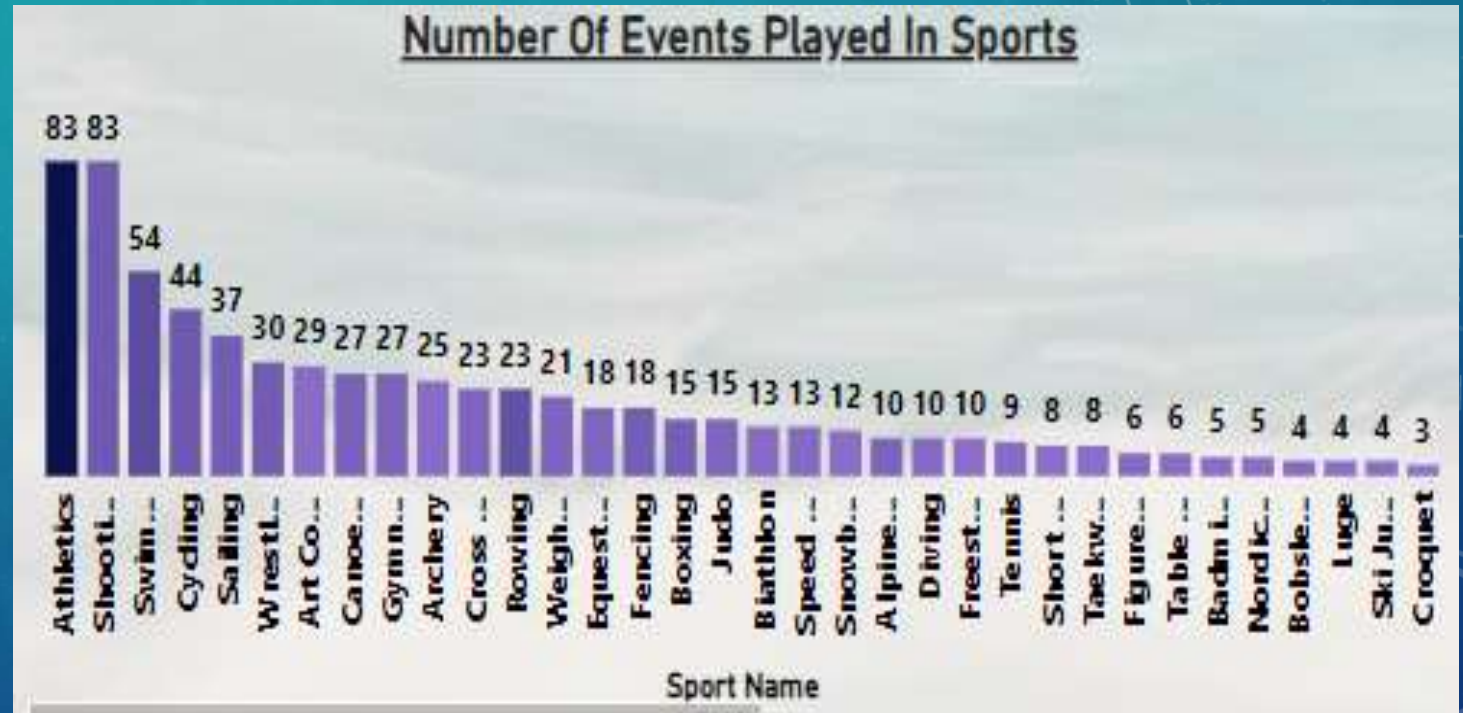
GLEANNING FROM VISUAL TOTAL THERE ARE 2 SEASON FOR OLYMPIC GAMES AND IN SUMMER SEASON MAX NO OF SPORTS ARE PLAYED I.E 52 AND IN WINTER ONLY 17 SPORTS ARE PLAYED.



Which sports have the highest number of events in the Olympics?

SUMMARY ANALYSIS

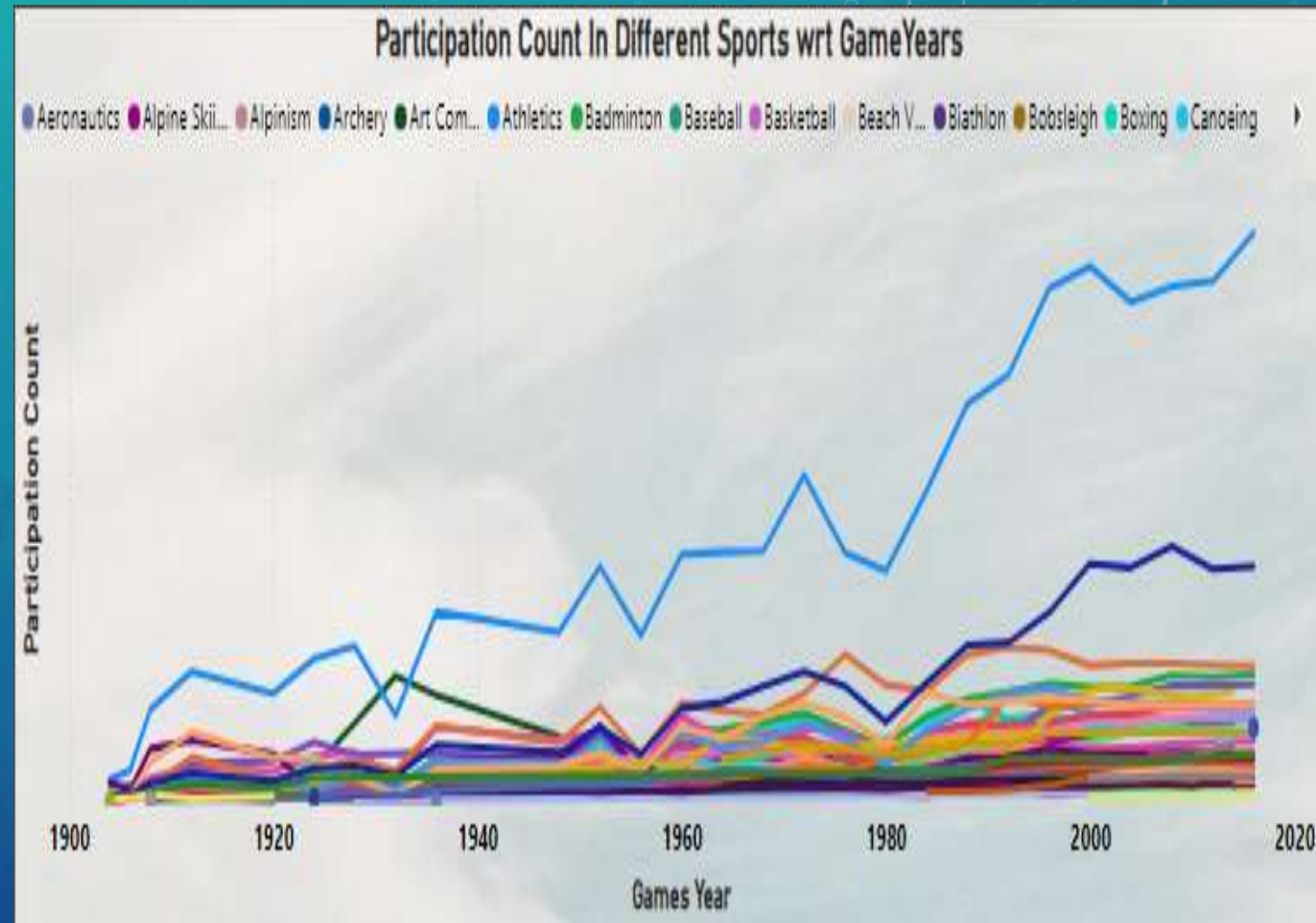
GLEANNING FROM VISUAL EVERY SPORT HAVE DIFFERENT NO OF EVENTS AND ATHLETICS/SHOOTING HAVE MAX NO OF EVENTS I.E 83 AND AFTER THAT SECOND HIGHEST IS SHOOTING.



How has the participation in each sport evolved over time?

SUMMARY ANALYSIS

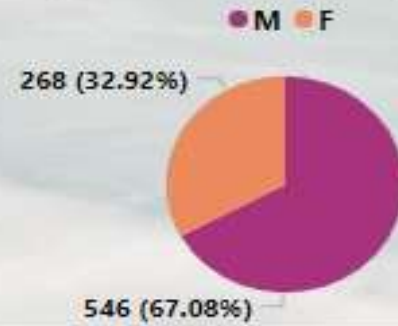
GLEANNING FROM VISUALISATION
PARTICIPATION COUNT IN DIFFERENT
SPORTS HAVE BEEN INCREASING WITH TIME
AND THE GRAPH OF ATHLETICS
PARTICIPATION IS QUITE PROGRESSIVE IN
COMPARISON TO OTHER SPORTS.



EVENT ANALYSIS

Sports Name	Total Events
Aeronautics	1
Alpine Skiing	10
Alpinism	1
Archery	25
Art Competitions	29
Athletics	83
Badminton	5
Baseball	1
Basketball	2
Basque Pelota	1
Beach Volleyball	2
Biathlon	13
Bobsleigh	4
Boxing	15
Canoeing	27
Cricket	1
Croquet	3
Cross Country Skiing	23
Curling	2
Cycling	44
Diving	10
Equestrianism	18
Fencing	18
Figure Skating	6
Football	2
Freestyle Skiing	10
Golf	3
Gymnastics	27
Handball	2
Hockey	2

Proportion Of Events By Gender



Aeronautics
Mixed Aeron...

Alpine Skiing
Men's Giant ...

Alpine Skiing
Women's Co...

Alpine Skiing
Women's Sla...

Alpine Skiing
Men's Combi...

Alpine Skiing
Men's Slalom

Alpine Skiing
Women's Do...

Alpine Skiing
Women's Su...

Alpine Skiing
Men's Down...

Alpine Skiing
Men's Super G

Alpine Skiing
Women's Gia...

Alpinism
Mixed Alpin...

Events By Games Year



How many events are there in each sport?

SUMMARY ANALYSIS

FOR REPRESENTING NO OF EVENTS IN EACH SPORT TABLE HAS BEEN USED AS IT BECOME QUITE EASY TO ANALYSE THE NUMBERS.ATHLETICS AND SHOOTING SPORT HAVE MAX NO OF EVENTS AND SPORTS SUCH AS ALPHINISM,BASQUE PELOTA,CRICKET,LACROSSE,POLO,ROQUE,RUGBY AND SOFTBALL HAVE SINGLE EVENT ONLY.

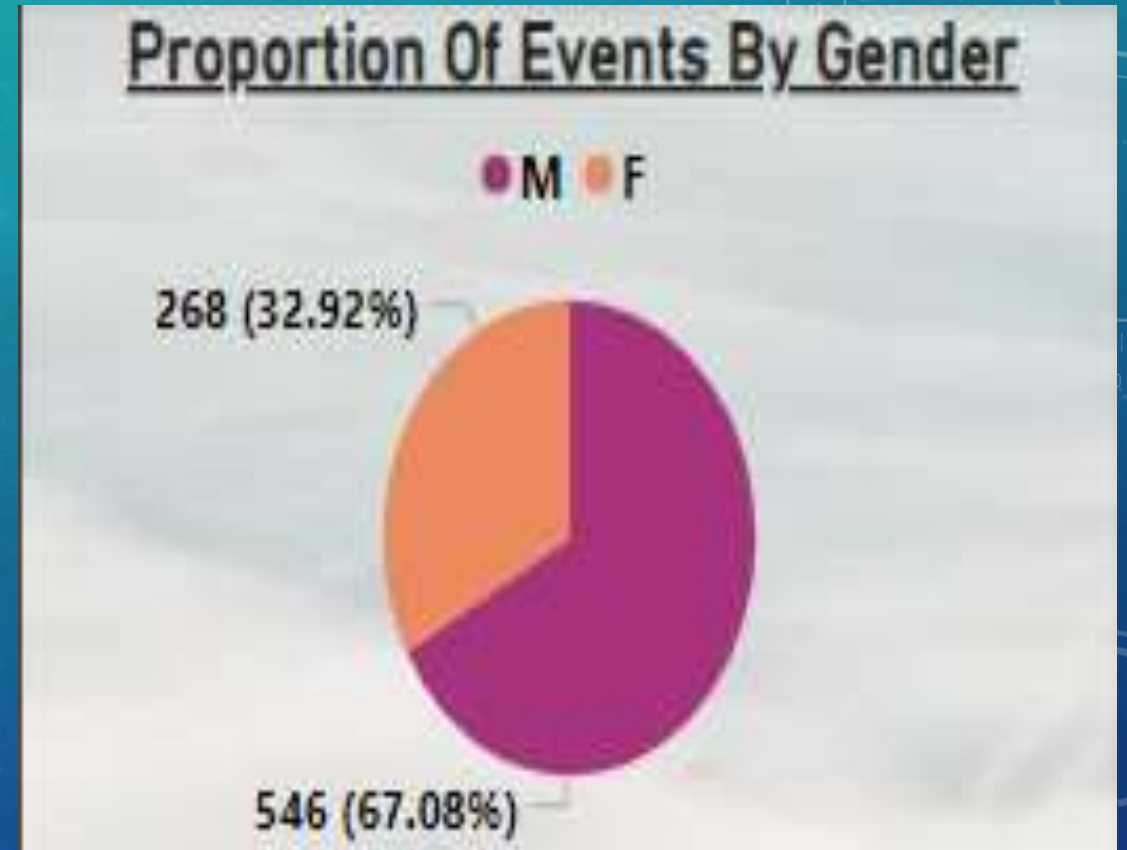
Sports Name	Total Events
Aeronautics	1
Alpine Skiing	10
Alpinism	1
Archery	25
Art Competitions	29
Athletics	83
Badminton	5
Baseball	1
Basketball	2
Basque Pelota	1
Beach Volleyball	2
Biathlon	13
Bobsleigh	4
Boxing	15
Canoeing	27
Cricket	1
Croquet	3
Cross Country Skiing	23
Curling	2
Cycling	44
Diving	10
Equestrianism	18
Fencing	18
Figure Skating	6
Football	2
Freestyle Skiing	10
Golf	3
Gymnastics	27
Handball	2
Hockey	2

Sports Name	Total Events
Ice Hockey	2
Jeu De Paume	1
Judo	15
Lacrosse	1
Luge	4
Military Ski Patrol	1
Modern Pentathlon	3
Motorboating	3
Nordic Combined	5
Polo	1
Racquets	2
Rhythmic Gymnastics	2
Roque	1
Rowing	23
Rugby	1
Rugby Sevens	2
Sailing	37
Shooting	83
Short Track Speed Skating	8
Skeleton	2
Ski Jumping	4
Snowboarding	12
Softball	1
Speed Skating	13
Swimming	54
Synchronized Swimming	3
Table Tennis	6
Taekwondo	8
Tennis	9
Trampolining	2

What is the distribution of events by gender (Men, Women, Mixed)?

SUMMARY ANALYSIS

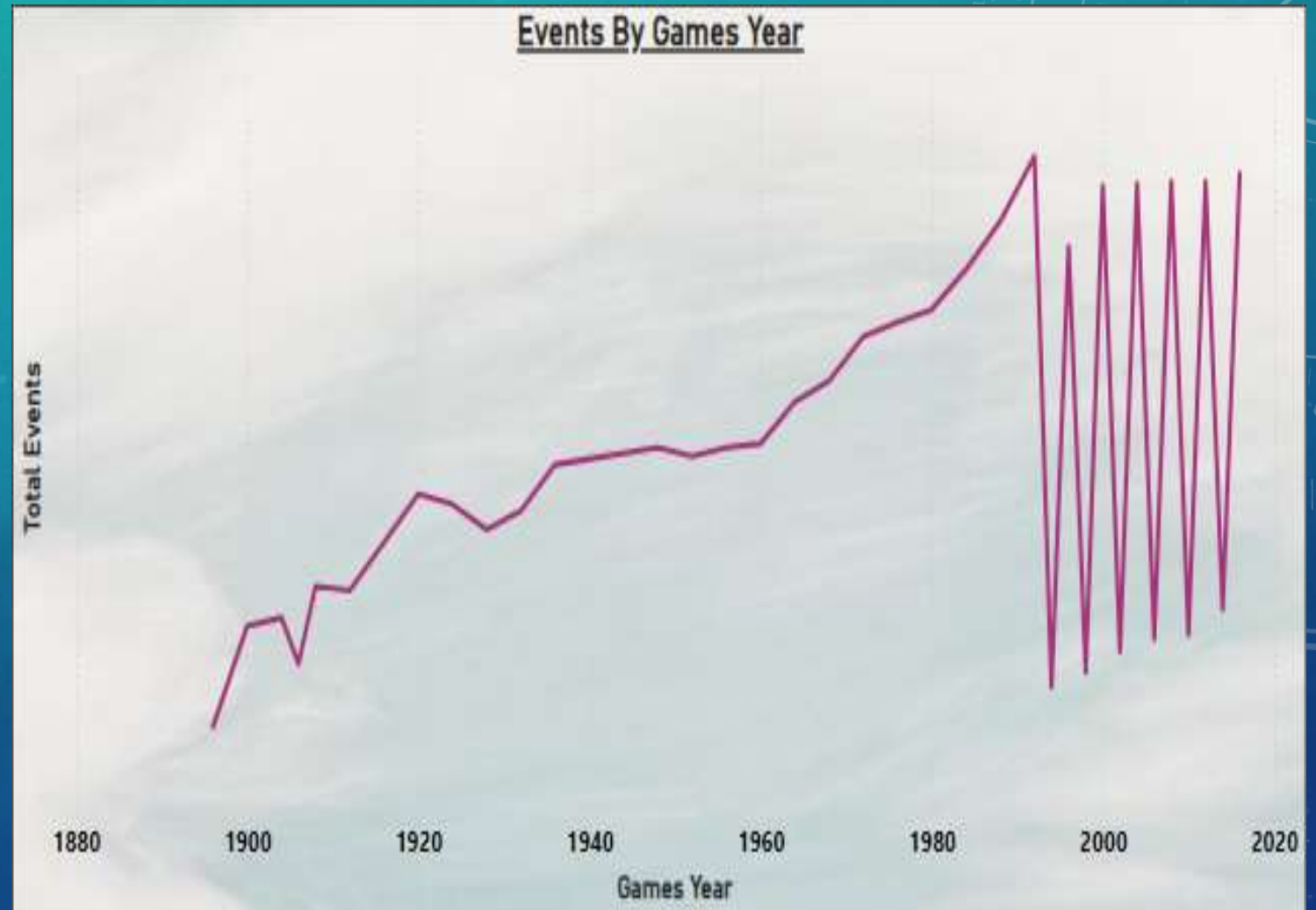
GLEANNING FROM VISUALIZATION MAJORITY OF EVENTS ARE PLAYED BY MEN AND IT'S PROPORTION TO FEMALE COMES OUT TO 67.08 % WHERE AS FEMALE PLAYS 32.92% OF EVENTS IN RESPECTIVE SPORTS.



How has the number of events changed over time?

SUMMARY ANALYSIS

TOTAL NO OF EVENTS FOR PARTICULAR GAMES DEPEND ON THE GAMES SEASON I.E SUMMER OR WINTER AND IN WINTER SEASON NO OF EVENTS ARE QUITE LOW AS FEW NO OF GAMES ARE BEING PLAYED WRT SUMMER SEASON.



PARTICIPANT ANALYSIS

Total Participants

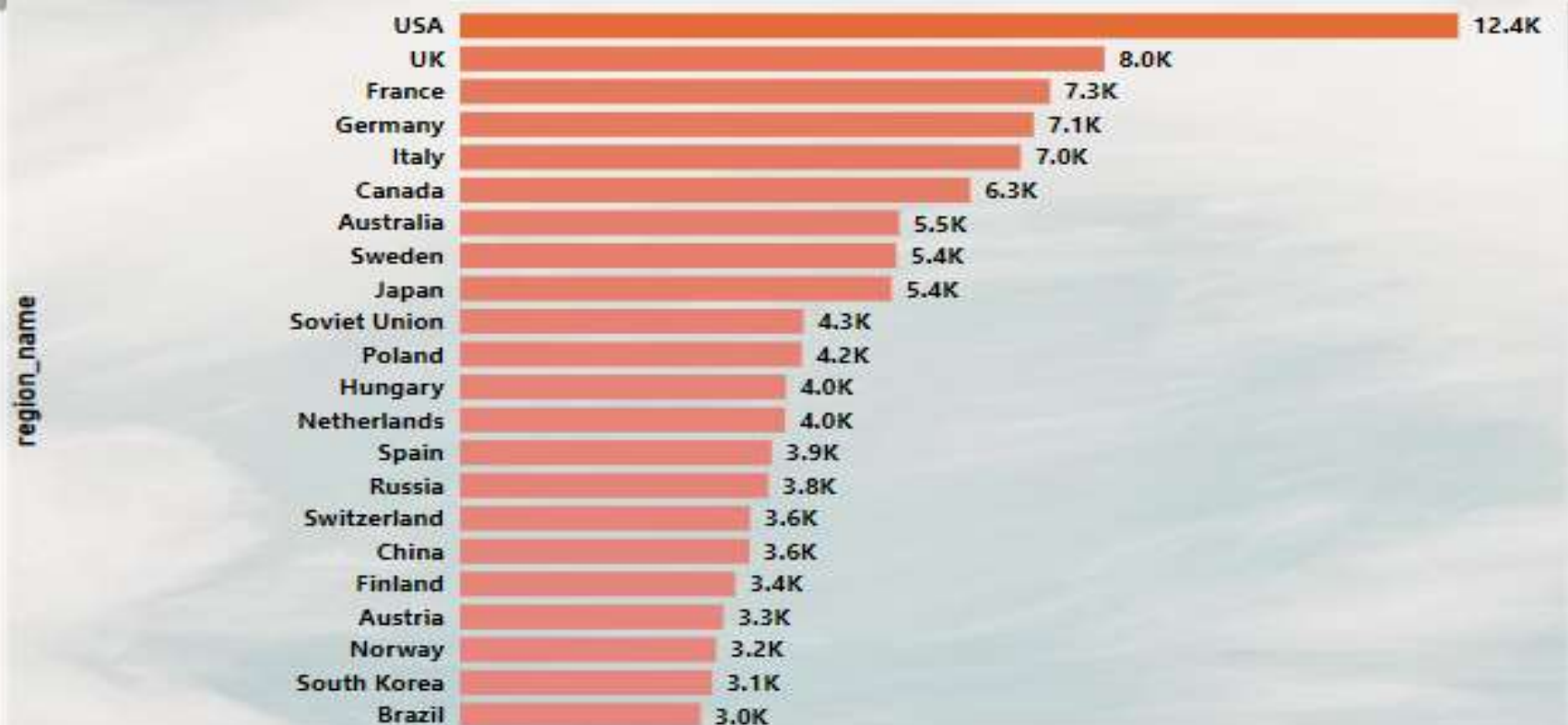
180.25K

Participants By Gender



Countries	Participation Count
USA	12392
UK	7999
France	7323
Germany	7119
Italy	6958
Canada	6332
Australia	5452
Sweden	5416
Japan	5355
Soviet Union	4263
Poland	4241
Hungary	4045
Netherlands	4035
Spain	3867
Russia	3828
Switzerland	3603
China	3594
Finland	3418
Austria	3269
Norway	3180
Total	180252

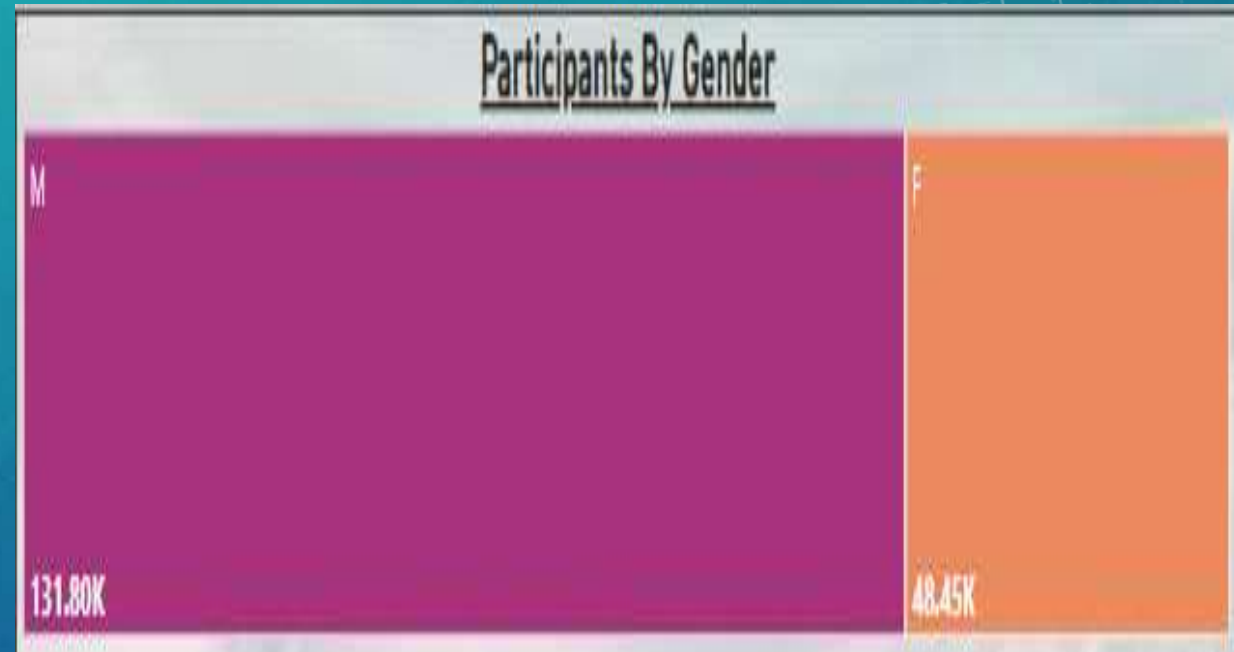
competitor By Region



What is the distribution of participants by gender?

SUMMARY ANALYSIS

TREE MAP VISUAL PROVIDES NO OF PARTICIPANT BY GENDER AND MALE HAVE AROUND 132K AND FEMALE HAVE 48K PARTICIPANT SO BASICALLY MALE ARE 2.5 TIMES OF FEMALE PARTICIPANTS



Which countries have the highest number of participants in the Olympics?

SUMMARY ANALYSIS

GLEANNING FROM THE BAR GRAPH AND TABLE USA HAVE MAX NO OF PARTICIPANTS I.E 12392 AND AFTER THAT UK HAVE 8K.

Countries	Participation Count
USA	12392
UK	7999
France	7323
Germany	7119
Italy	6958
Canada	6332
Australia	5452
Sweden	5416
Japan	5355
Soviet Union	4263
Poland	4241
Hungary	4045
Netherlands	4035
Spain	3867
Russia	3828
Switzerland	3603
China	3594
Finland	3418
Austria	3269
Norway	3180
Total	180252



MEDAL ANALYSIS

Total Medals

39K

Bronze Medal

12.95K

Gold Medal

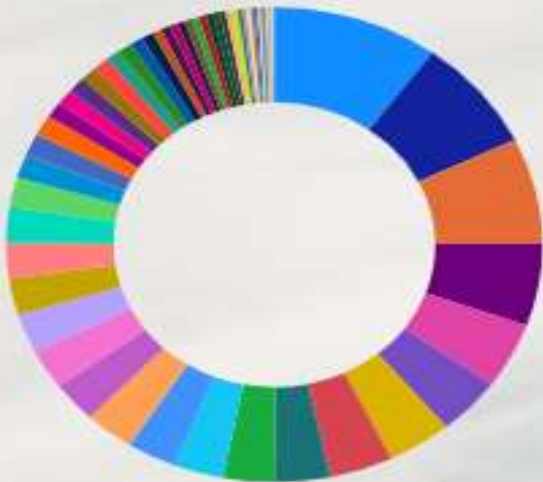
13.11K

Silver Medal

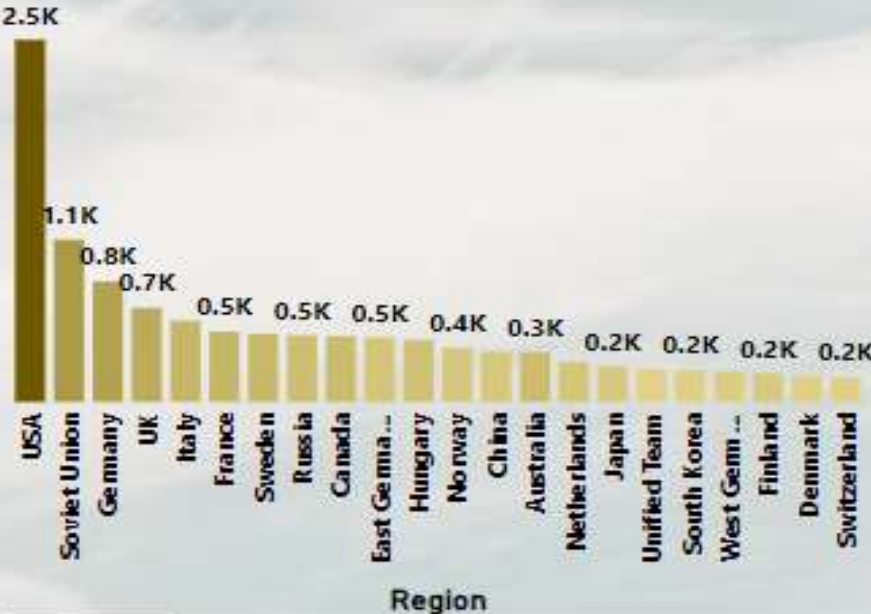
12.77K

Medal Distribution By Sports

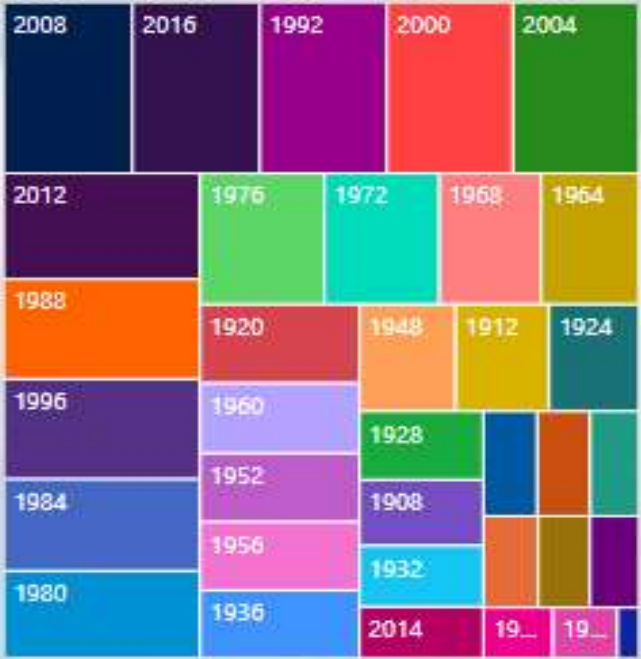
● Athletics ● Swimm... ● Rowing ● Gymna...



Gold Medal By Region



Medal Distribution By Year



Afghanistan

Algeria

Armenia

Australia

Azerbaijan

Bahrain

Bronze

Silver

Albania

Argentina

Australasia

Austria

Bahamas

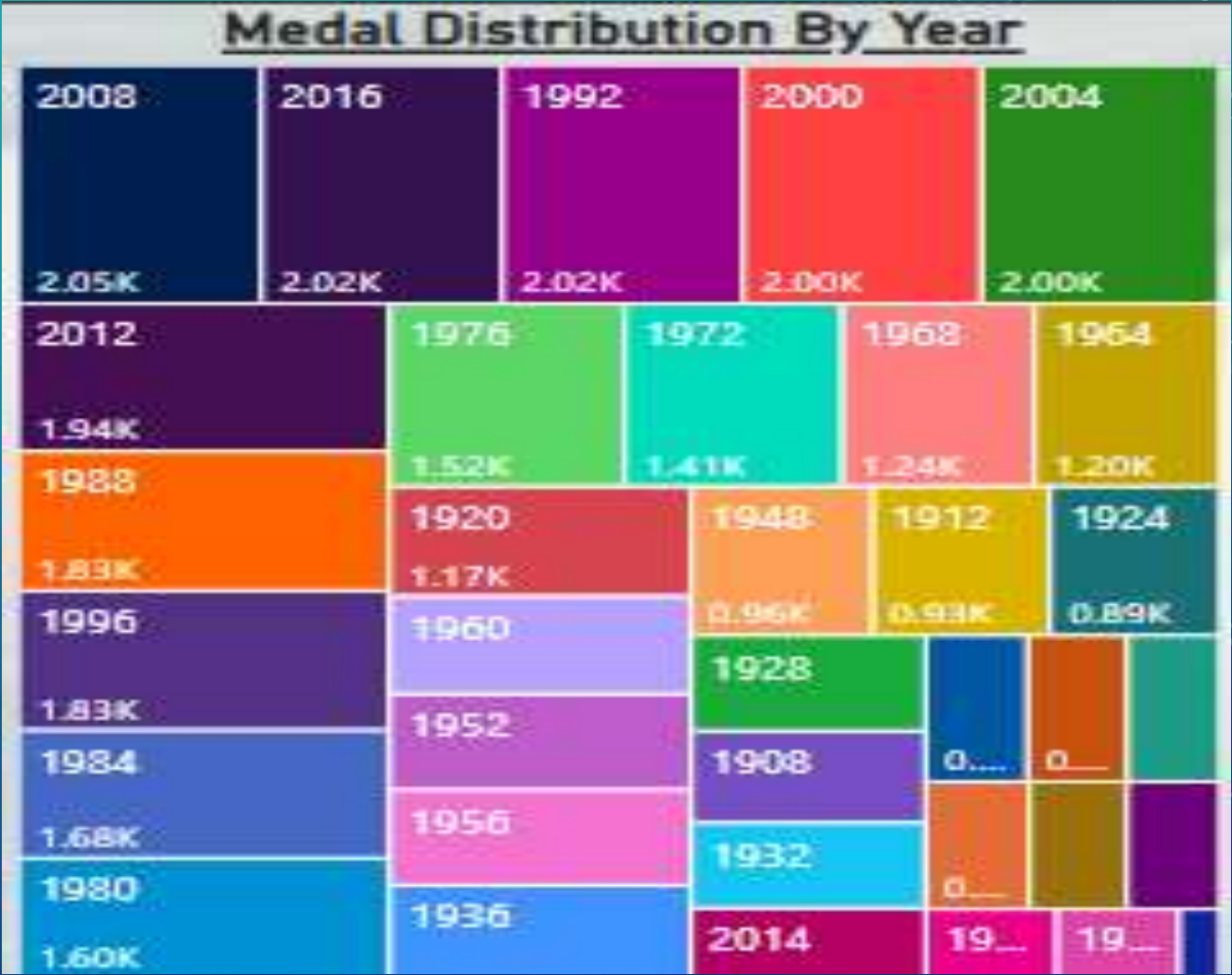
Barbados

Gold

How many medals have been awarded in each Olympics?

SUMMARY ANALYSIS

GLEANNING FROM THE VISUAL
MEDALS DISTRIBUTION IN EACH
OLYMPIC GAME IS AROUND 2K



Which countries have the highest number of gold medals?

SUMMARY ANALYSIS

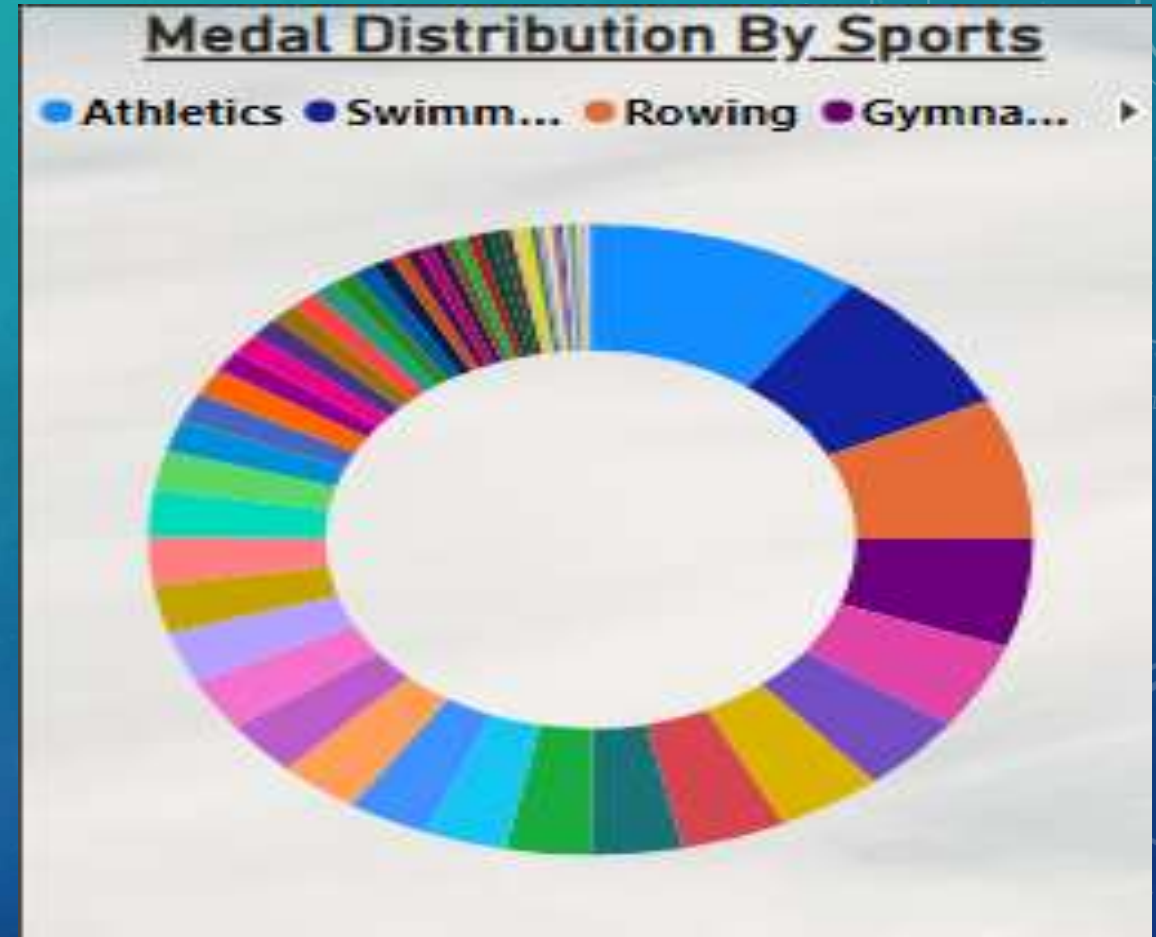
GLEANNING FROM THE VISUALISATION MAX NO OF GOLD MEDAL HAVE BEEN BAGGED BY USA WHICH COUNT TO AROUND 2.5K AND AFTER THAT SOVIET UNION IS ON SECOND PLACE IN GOLD MEDAL TALLY WITH 1.1 K



How does the medal distribution vary across different sports?

SUMMARY ANALYSIS

GLEANNING FROM THE VISUALISATION
ATHLETICS HAVE MAX NO OF MEDAL
AROUND 3.9K AND HAVING SHARE
OF 10% OF TOTAL MEDAL
DISTRIBUTION ON THE OTHER HAND
SWIMMING HAVE AROUND 7.7%
SHARE.

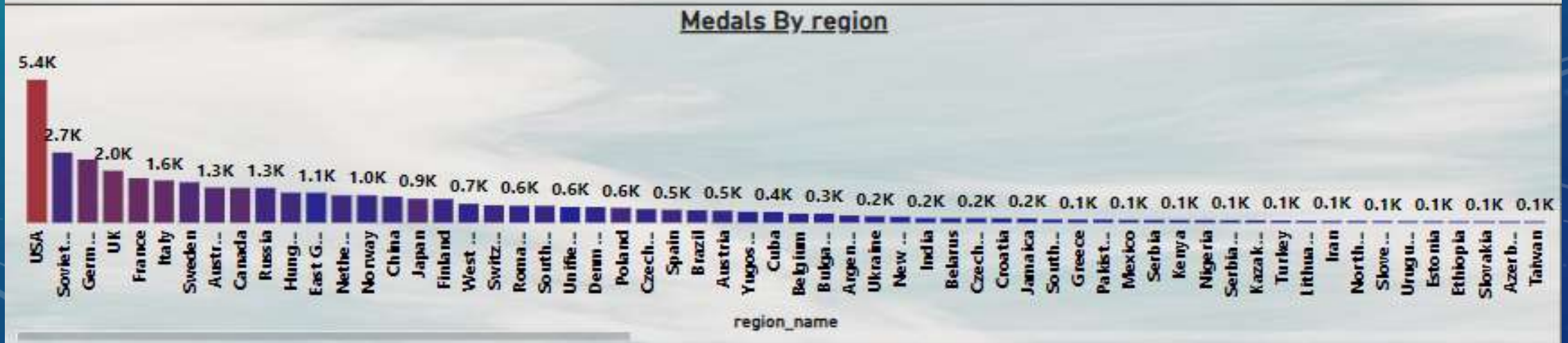
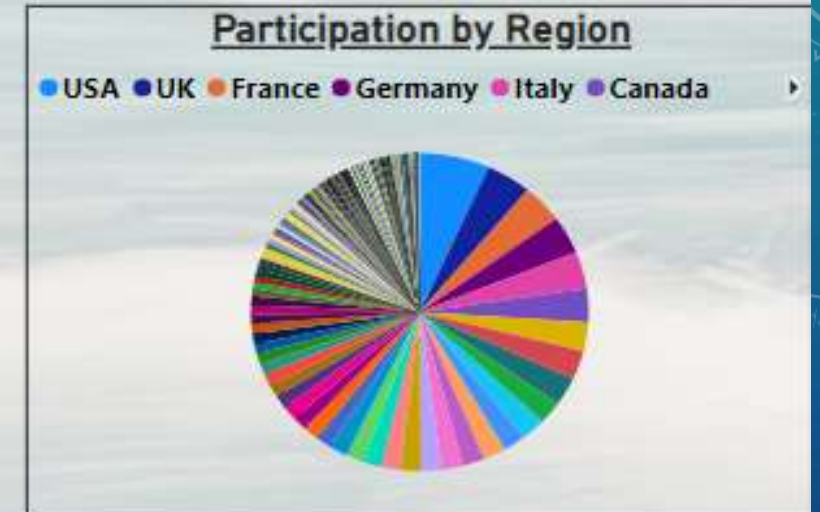


REGION ANALYSIS

Olympic Games	Region Participation
2016	213
2014	91
2012	215
2010	85
2008	212
2006	87
2004	210
2002	85
2000	209
1998	79
1996	206
1994	74
1992	188
1988	182
1984	163
1980	100



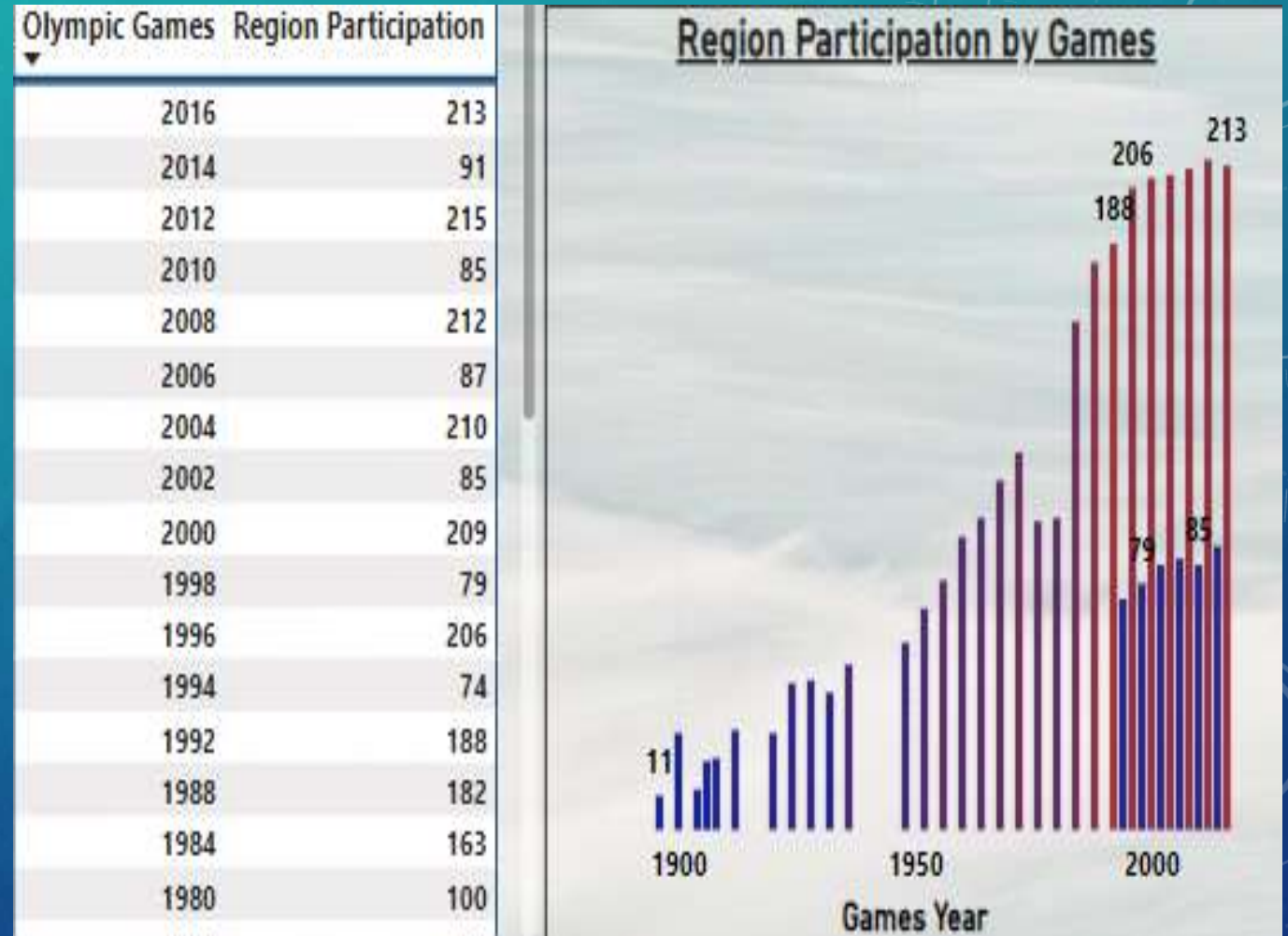
Select all	Afghanistan	Albania	Algeria
------------	-------------	---------	---------



How many regions or NOCs participate in each Olympic Games?

SUMMARY ANALYSIS

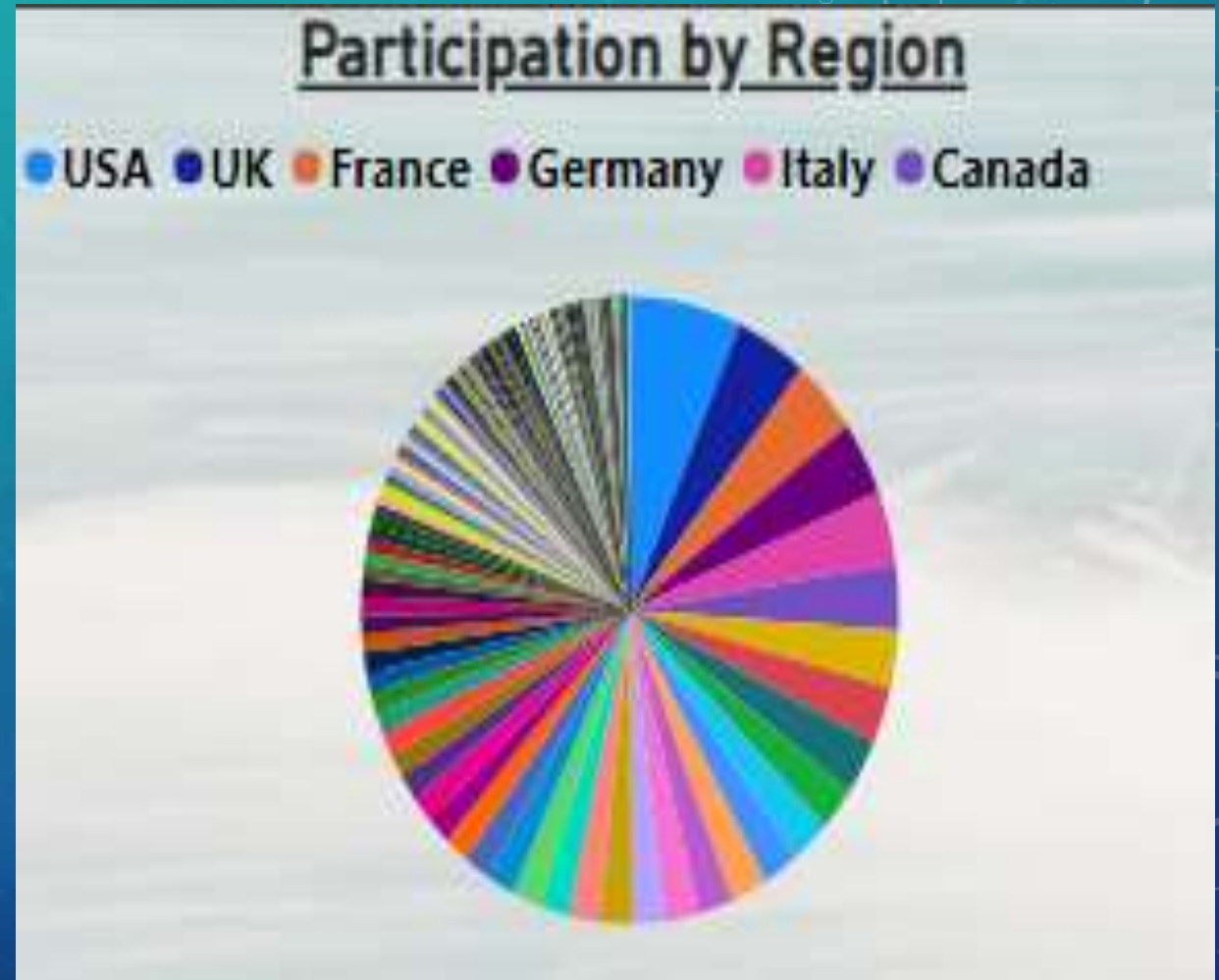
AS OLYMPIC MAIN EVENT TAKE PLACE AFTER EVERY 4 YEAR SO MAX PARTICIPATION BY VARIOUS REGION CAN BE SEEN IN THAT WITH 2016 EVENT HAVING 213 REGION AS PARTICIPANTS



Which regions have the highest number of participants in the Olympics?

SUMMARY ANALYSIS

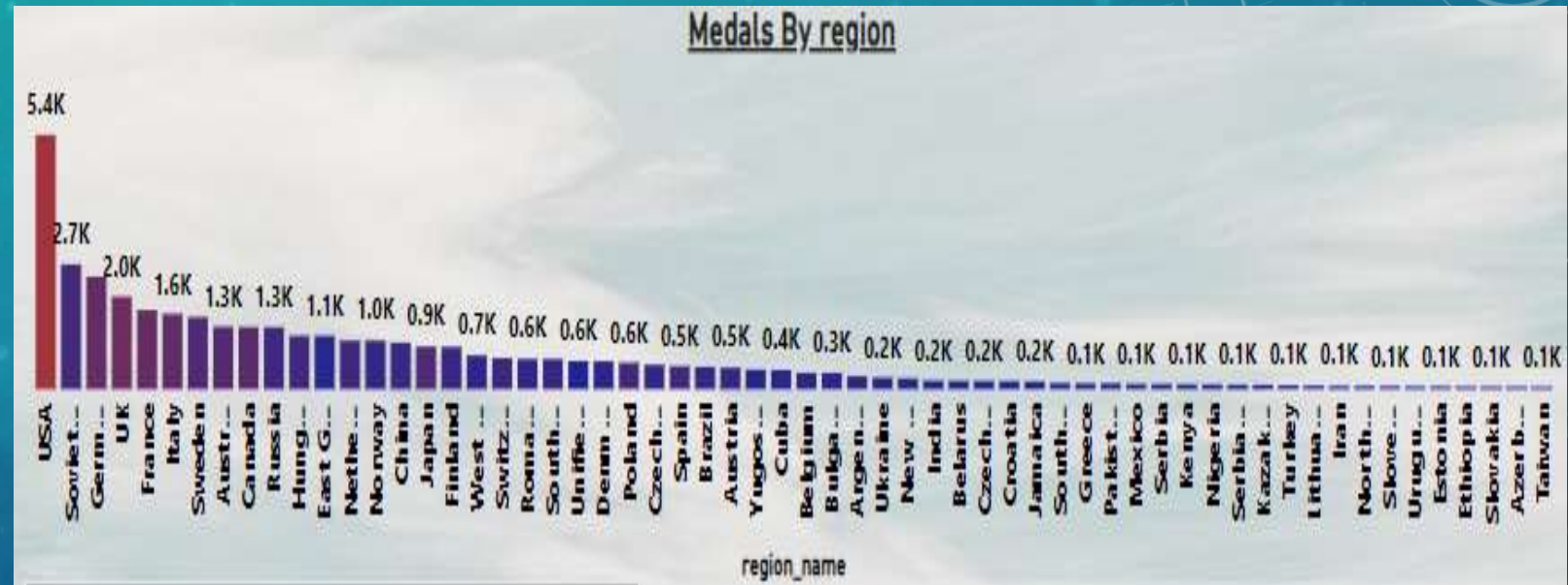
GLEANNING FROM THE VISUAL USA HAVE MAX NO OF PARTICIPANTS AND REGIONS LIKE BENIN,HAITI,CONGO HAVE ONE OF THE LEAST PARTICIPATION IN GAMES.



What is the distribution of medals among different regions?

SUMMARY ANALYSIS

GLEANNING FROM THE VISUAL
USA HAVE MAX NO OF MEDAL
AMONG ALL THE REGION
FOLLOWED BY SOVIET UNION.



EDA QUESTIONS

The background is a gradient of teal and blue, with a subtle pattern of small white dots. On the right side, there is a large, faint circular scale with numbers from 0 to 210. There are also several faint circular lines and arrows scattered across the background, some of which are dashed.

Are there any trends or patterns in the frequency of hosting Olympic Games?

```
select g.games_name,g.games_year,c.city_name
from games as g
join games_city as gct on g.id=gct.games_id
join city as c on gct.city_id=c.id
group by 1,2,3
order by 1,2 desc
```

IT CAN BE ANALYSED FROM THE RETREIVED DATA THAT GAMES ARE OF TWO TYPE SUMMER AND WINTER AND ACC OLYMPIC GAMES THAT TOOK PLACE EARLIER HAD TREND THAT BEFORE 1992 SUMMER AND WINTER USED TO TAKE PLACE IN SAME YEAR BUT AFTER THAT THERE WAS GAP OF 2 YEARS IN BETWEEN.

games_name	games_year	city_name
1896	1896	Athina
1900	1900	Paris
1904	1904	St. Louis
1906	1906	Athina
1908	1908	London
1912	1912	Stockholm
1920	1920	Antwerpen
1924	1924	Paris
1924 Winter	1924	Chamonix
1928	1928	Amsterda
1928 Winter	1928	Sankt Moritz
1932	1932	Los
1932 Winter	1932	Lake
1936	1936	Berlin
		Garmisch-

How has the duration of Olympic Games changed over time?

```
WITH CTE AS (
  SELECT
    games_year,
    LAG(games_year) OVER
    (ORDER BY games_year) AS
    previous_game_year
  FROM
    games
)
SELECT
  games_year,
  previous_game_year, (games_year -
  previous_game_year) AS
  years_between_games
FROM CTE
WHERE
  previous_game_year IS NOT NULL
ORDER BY 1,2 desc
```

games_year	previous_games_year	years_between_games
1900	1896	4
1904	1900	4
1906	1904	2
1908	1906	2
1912	1908	4
1920	1912	8
1924	1920	4
1924	1924	0
1928	1924	4
1928	1928	0
1932	1928	4
1932	1932	0
1936	1932	4
1936	1936	0
1948	1936	12
1948	1948	0
1952	1948	4
1952	1952	0
1956	1952	4
1956	1956	0
1960	1956	4
1960	1960	0
1964	1960	4
1964	1964	0
1968	1964	4
1968	1968	0
1972	1968	4
1972	1972	0
1976	1972	4
1976	1976	0
1980	1976	4

AS DISCUSSED OLYMPIC GAMES HAVE 2 SEASONS AND BEFORE THAT DURATION BETWEEN 2 SEASON WAS NOT THERE AND OVERALL O.GAMES WERE ORGANISED IN 4 YEARS BUT AFTER 1992 WITH SOME EXCEPTIONS BUT AFTER 1992 THERE IS DEF TREND OF 2 YEAR GAP.



Are there any notable events or occurrences associated with specific Olympic Games?

	games_year	total_events	total_participants
SELECT g.games_year,count(distinct e.id) as total_events,count(gc.id) as total_participants FROM games as g join games_competitor as gc on g.id=gc.games_id join competitor_event as ce on gc.id=ce.competitor_id join event as e on ce.event_id=e.id group by 1 order by 1,2,3 desc	1896	42	217
	1900	90	1143
	1904	94	1020
	1906	72	989
	1908	109	2437
	1912	107	3875
	1920	153	3437
	1924	148	4540
	1928	136	4590
	1932	145	2964
	1936	167	7156
	1948	175	6257
	1952	171	3026
	1956	175	5758
	1960	177	8964
	1964	197	9376
	1968	207	10315
	1972	228	11809
	1976	235	10423
	1980	241	8745
	1984	260	11339
	1988	283	14530
	1992	314	16335
	1994	61	3153
	1996	271	13750
	1998	68	3600
	2000	300	13815
	2002	78	4105
	2004	301	13440
	2006	84	4381
	2008	302	13597
	2010	86	4398

WITH THE TIME AND GREAT CALIBRE ONE
CAN SEE INCREASE IN NUMBER OF
EVENTS AND PARTICIPANTS IN O.GAMES



Are there any emerging sports that have been recently added to the Olympics?

```
SELECT s.sport_name, MIN(g.games_year)
      olympic_entry_year
FROM games as g
join games_competitor as gc on g.id=gc.games_id
join competitor_event as ce on
      gc.id=ce.competitor_id
join event as e on ce.event_id=e.id
join sport as s on e.sport_id=s.id
      group by 1
      order by 2 desc
      limit 5
```

sport_name	olympic_entry_year
Rugby Sevens	2016
Triathlon	2000
Trampolining	2000
Taekwondo	2000
Snowboarding	1998

ACCORDING TO THE DATA RUGBY SEVENS IS THE LATEST ENTRY INTO THE O.GAMES 2016 AND AFTER THAT IT WAS TRIATHLON ,TRAMPOLINING,TAEKWONDO WHICH WAS ADDED IN 2000

How has the popularity of certain sports changed over the years?

```
SELECT s.sport_name, MIN(g.games_year) as
olympic_entry_year, MAX(g.games_year) as last_year
FROM games as g
join games_competitor as gc on g.id=gc.games_id
join competitor_event as ce on gc.id=ce.competitor_id
join event as e on ce.event_id=e.id
join sport as s on e.sport_id=s.id
group by 1
Having last_year < 2016
order by 3
```

```
SELECT s.sport_name, COUNT(ce.competitor_id) AS
participant_count
FROM sport s
JOIN event e ON s.id = e.sport_id
JOIN competitor_event ce ON s.id = ce.event_id
JOIN games_competitor gc ON ce.competitor_id = gc.id
JOIN person_region pr ON gc.person_id = pr.person_id
JOIN noc_region nr ON pr.region_id = nr.id
GROUP BY 1
order by 2 DESC
```

THIS QUESTION HAVE BEEN TAKEN IN 2
WAYS ONE CAN FIND THE POPULARITY OF
SPORT BY WEATHER IT IS STILL PLAYED
OR NOT AND IN ANOTHER WAY BY
FINDING THE NO OF PARTICIPANT THAT
GAME CATER.

sports_name	olympic_entry_year	last_year
Basque Pelota	1900	1900
Cricket	1900	1900
Croquet	1900	1900
Roque	1904	1904
Jeu De Paume	1908	1908
Lacrosse	1904	1908
Motorboating	1908	1908
Racquets	1908	1908
Tug-Of-War	1900	1920

sports_name	participation_count
Athletics	38301
Gymnastics	25354
Swimming	23039
Shooting	11162
Fencing	10335
Cycling	10312
Rowing	10250
Cross Country Skiing	9539
Alpine Skiing	8877

Are there any sports that have a higher number of events for one gender compared to others?

```
WITH gender_event_count AS (
    SELECT
        s.sport_name, p.gender, COUNT(e.id) AS
        event_count
    FROM sport s
    JOIN event e ON s.id = e.sport_id
    JOIN competitor_event ce ON e.id =
        ce.event_id
    JOIN games_competitor gc ON
        ce.competitor_id = gc.id
    JOIN person p ON gc.person_id = p.id
    GROUP BY 1,2
),
sport_event_totals AS (
    SELECT s.sport_name,
    COUNT(e.id) AS total_event_count
    FROM sport s
    JOIN event e ON s.id = e.sport_id
    JOIN competitor_event ce ON e.id =
        ce.event_id
```

FROM THE RETRIEVED DATA SPORTS SUCH
AS BADMINTON, RHYTHMIC
SWIMMING, SOFTBALL, FIGURE SKATING
AND SUNC SWIMMING HAVE MORE
PARTICIPANT FROM FEMALE.

sports_name	gender	event_cour	total_even	event_%
Aeronautics	M	1	1	100
Alpine Skiing	F	3387	8712	38.88
Alpine Skiing	M	5325	8712	61.12
Alpinism	F	1	16	6.25
Alpinism	M	15	16	93.75
Archery	F	1005	2249	44.69
Archery	M	1244	2249	55.31
Art	F	301	3071	9.8
Art	M	2770	3071	90.2
Athletics	F	11571	37361	30.97
Athletics	M	25790	37361	69.03
Badminton	F	740	1457	50.79
Badminton	M	717	1457	49.21
Baseball	M	891	891	100
Basketball	F	1256	4334	28.98
Basketball	M	3078	4334	71.02
Basque Pelota	M	2	2	100
Beach	F	276	563	49.02
Beach	M	287	563	50.98
Biathlon	F	1863	4887	38.12
Biathlon	M	3024	4887	61.88
Bobsleigh	F	143	2966	4.82
Bobsleigh	M	2823	2966	95.18
Boxing	F	72	5720	1.26
Boxing	M	5648	5720	98.74
Canoeing	F	1376	6125	22.47
Canoeing	M	4749	6125	77.53
Cricket	M	18	18	100

Are there any new events that have been introduced in recent editions of the Olympics?

```
SELECT e.event_name,min(g.games_year) latest_event
      FROM event e
      JOIN competitor_event ce ON e.id = ce.event_id
      JOIN games_competitor gc ON ce.competitor_id = gc.id
      JOIN games g ON gc.games_id = g.id
      group by 1
      having latest_event > 2014
      order by 2 desc
```

Event_name

Rugby Sevens Men's Rugby Sevens

Rugby Sevens Women's Rugby Sevens

Sailing Women's Skiff

Wrestling Women's Featherweight, Freestyle

Wrestling Women's Light-Heavyweight, Freestyle

THESE 5 EVENTS HAVE BEEN ADDED AFTER 2014 WHICH ARE RELATED TO
RUGBY,SAILING AND WRESTLING.

Are there any events that have been discontinued or removed from the Olympics?

```
WITH CTE AS (SELECT e.event_name, MIN(g.games_year) AS
              A, MAX(g.games_year) AS B
              FROM event e
              JOIN competitor_event ce ON e.id = ce.event_id
              JOIN games_competitor gc ON ce.competitor_id = gc.id
              JOIN games g ON gc.games_id = g.id
              group by 1
              HAVING B < 2016
              order by 2,3 desc)
SELECT event_name from CTE
```

DIFFERENT SPORTS HAVE VARIOUS EVENTS AND EVENTS KEEPS ON WITH DEMAND OF TIME SO TILL 2016 451 EVENTS HAVE BEEN DISCONTINUED TILL NOW.

Event_name

Shooting Men's Free Rifle, Three Positions, 300 metres

Gymnastics Men's Rope Climbing

Cycling Men's 100 kilometres

Cycling Men's 333 metres Time Trial

Weightlifting Men's Unlimited, One Hand

Weightlifting Men's Unlimited, Two Hands

Fencing Men's Foil, Masters, Individual

Cycling Men's 10,000 metres

Cycling Men's 12-Hours Race

Gymnastics Men's Horizontal Bar, Teams

Gymnastics Men's Parallel Bars, Teams

Shooting Men's Free Pistol, 30 metres

Shooting Men's Military Pistol, 25 metres

Shooting Men's Military Rifle, 200 metres

Shooting Men's Muzzle-Loading Pistol, 25 metres

Swimming Men's 1,200 metres Freestyle

Swimming Men's 500 metres Freestyle

Wrestling Men's Unlimited Class, Greco-Roman

Cycling Men's Points Race

Rowing Men's Coxed Fours

Rowing Men's Coxed Pairs

Are there any notable trends in the height and weight of participants over time?

```
SELECT
g.games_year,round(avg(height),2)
as avg_ht,round(avg(weight),2) as
avg_wt
FROM person as p
join games_competitor as gc on
p.id=gc.person_id
join games as g on gc.games_id=g.id
group by 1
order by 1 desc
```

games_yea	avg_ht	avg_wt
2016	174.07	70.62
2014	174.42	68.7
2012	174.43	70.36
2010	175.27	72
2008	174.84	71.65
2006	174.75	71.92
2004	176.5	72.48
2002	173.92	72.12
2000	175.29	71.8
1998	171.18	70.99
1996	153.45	63.45
1994	165.24	68.53
1992	143.4	59.32
1988	167.18	68.68
1984	168.88	68.89
1980	167.36	69
1976	164.71	67.6
1972	172.56	70.39
1968	172.73	70.7
1964	162.87	66.58
1960	160.32	64.88
1956	79.49	33
1952	43.98	18.34
1948	33.55	13.77
1936	30.44	10.9
1932	68.75	13.4

WITH CHANGING PHYSICAL DEMAND OF GAME AVERAGE HEIGHT AND AVERAGE WEIGHT IS AROUND 170 AND 70.IT SHOW A PERSISTENT EFFORT SHOWN BY ATHLETES TO KEEP THERE MEASURES IN REQ CATEGORY.

Are there any dominant countries or regions in specific sports or events?

<pre> with medals_by_region_sport as (SELECT nr.region_name,s.sport_name,count(ce.medal_id) as total_medal from noc_region as nr join person_region as pr on nr.id=pr.region_id join games_competitor as gc on pr.person_id=gc.person_id join competitor_event as ce on gc.id=ce.competitor_id join event as e on ce.event_id=e.id join sport as s on e.sport_id=s.id WHERE ce.medal_id <> 4 group by 1,2 order by 2,3 desc), rank_by_sports_acc_to_region as (select region_name,sport_name,total_medal,rank() over(partition by sport_name order by total_medal desc) as XYZ from medals_by_region_sport) select region_name,sport_name,total_medal from rank_by_sports_acc_to_region where XYZ= 1 order by 2 </pre>	region_name	sport_name	total_medal	VAROIOUS REGIONS ARE SORTED IN MANNER WHERE THEY HAVE WON MAX MEDAL IN PATICULAR SPORT.LIKE USA HAVE MAX MEDAL IN BASKETBALL,ATHLETICS AND SWIMMING.CHINA IS BEST IN BADMINTON AND SO ON.
	Switzerland	Aeronautic	1	
	Austria	Alpine	114	
	UK	Alpinism	11	
	South Korea	Archery	69	
	Germany	Art Competitio ns	26	
	USA	Athletics	1035	
	China	Badminton	64	
	Cuba	Baseball	112	
	USA	Basketball	331	
	Spain	Basque Pelota	2	
	Brazil	Beach Volleyball	26	
	Germany	Biathlon	85	
	Switzerland	Bobsleigh	92	
	USA	Boxing	105	
	Germany	Canoeing	184	
	UK	Cricket	17	
	France	Croquet	8	
	Norway	Cross Country Skiing	164	
	Canada	Curling	48	
	UK	Cycling	172	
	USA	Diving	138	
	Germany	Equestriani sm	179	
	Italy	Fencing	358	

What factors contribute to the success or performance of participants from different countries?

```
SELECT
  nr.region_name, round(avg(gc.age), 2) as
  avg_age, count(ce.medal_id) as total_medal
  from noc_region as nr
join person_region as pr on nr.id=pr.region_id
join games_competitor as gc on
  pr.person_id=gc.person_id
join competitor_event as ce on
  gc.id=ce.competitor_id
WHERE ce.medal_id <> 4
  group by 1
  order by 3 desc
```

region_name	avg_age	medal
USA	24.91	5414
Soviet	25.55	2658
Germany	26.41	2395
UK	27.87	1971
France	27.97	1686
Italy	26.94	1603
Sweden	28.05	1535
Australia	24.83	1333
Canada	25.82	1324
Russia	25.43	1321
Hungary	26.5	1140
East	23.64	1137
Norway	27.92	1034
Netherlands	26.49	1034
China	23.21	991
Japan	24.52	909
Finland	27.39	901
West		
Germany	26.47	719
Switzerland	28.66	662
Romania	24.22	644
South Korea	23.43	642
Unified	25.21	600
Denmark	26.8	597
Poland	26.19	567
Czechoslov		
akia	25.9	521
Spain	26.17	493
Brazil	25.93	476

SUCCESS AND FAILURE OF ATHLETES DEPENDS ON MANY QUALITATIVE AND QUANTITATIVE FACTORS AND ACC TO THE DATA PROVIDED AGE IS APPROPRIATE FACTOR WHICH SHOW MID 20'S IS BEST TIME FOR ATHLETES AS THEY ARE IN THERE PRIME AND AFTER/BEFORE THERE IS NOT SO MUCH SUCCESS.THEREFORE COUNTRIES IN THIS GROUP PERFORM

Are there any countries that consistently perform well in multiple Olympic editions?

with medals_by_year as (SELECT nr.region_name,g.games_year,count(ce.medal_id) as total_medal from noc_region as nr join person_region as pr on nr.id=pr.region_id join games_competitor as gc on pr.person_id=gc.person_id join games as g on gc.games_id=g.id join competitor_event as ce on gc.id=ce.competitor_id where ce.medal_id <> 4 group by 1,2 order by 2,3 desc), average_medal_won as (select region_name,COUNT(DISTINCT games_year) AS editions, SUM(total_medal) AS total_medals, AVG(total_medal) AS avg_medals_per_edition from medals_by_year group by 1), consistent_country as (select region_name,editions,total_medals,avg_medals_per_edition from average_medal_won where editions > 5 and avg_medals_per_edition > (select avg(avg_medals_per_edition) from average_medal_won)) select region_name,editions,total_medals, avg_medals_per_edition FROM consistent_country	country	editions	medals	avg_medal	This approach focuses on identifying not just countries with a large total medal count but those that have maintained consistent high performance across many editions. The filtering by average medals per edition ensures that only the most consistent performers are highlighted.					
	Serbia	8	107	13.4						
	Croatia	13	170	13.1						
	Belarus	14	178	12.7						
	Ukraine	14	245	17.5						
	Unified Team	14	600	42.9						
	Cuba	15	410	27.3						
	China	15	991	66.1						
	East Germany	16	1137	71.1						
	Romania	18	644	35.8						
	Argentina	19	275	14.5						
	Bulgaria	19	348	18.3						
	Brazil	19	476	25.1						
	Spain	19	493	25.9						
	West Germany	19	719	37.8						
	Yugoslavia	20	416	20.8						
	Russia	20	1321	66.1						
	Soviet Union	20	2658	132.9						
	South Korea	22	642	29.2						
	Poland	25	567	22.7						
	Czechoslovakia	26	521	20.0						
	Japan	28	909	32.5						
	Hungary	28	1140	40.7						
	Denmark	29	597	20.6						
	Finland	32	901	28.2						
	Netherlands	32	1034	32.3						
	Germany	32	2395	74.8						
	Austria	33	458	13.9						
	Sweden	33	1535	46.5						
	Italy	33	1603	48.6						
	Switzerland	34	662	19.5						
	Norway	34	1034	30.4						
	Canada	34	1324	38.9						

Are there any sports or events that have a higher number of medalists from a specific region?

```

with MEDALS_BY_SPORT_REGION AS
(SELECT
  nr.region_name,
  s.sport_name,
  COUNT(cc.medal_id) AS total_medal
  FROM noc_region AS nr
  JOIN person_region AS pr ON nr.id = pr.region_id
  JOIN games_competitor AS gc ON pr.person_id = gc.person_id
  JOIN competitor_event AS ce ON gc.id = ce.competitor_id
  JOIN event AS e ON ce.event_id = e.id
  JOIN sport AS s ON e.sport_id = s.id
  WHERE ce.medal_id <> 4
  GROUP BY 1,2),

MAX_MEDAL_BY_SPORT AS
(select sport_name,
  MAX(total_medal) as max_medals
from MEDALS_BY_SPORT_REGION
  group by 1),

dominant_region as
(select
  MBSR.region_name, MBSR.sport_name, MBSR.total_medal
from MEDALS_BY_SPORT_REGION as MBSR
  JOIN MAX_MEDAL_BY_SPORT AS MMB on
  MBSR.sport_name=MMB.sport_name and
  MBSR.total_medal=MMB.max_medals
  GROUP BY 1,2)

select region_name,
  sport_name,
  total_medal
FROM dominant_region
ORDER BY 2,3 desc

```

region_name	sport_name	total_medal
Switzerland	Aeronaautics	1
Austria	Alpine Skiing	114
UK	Alpinism	11
South Korea	Archery	63
Germany	Art Competitions	26
USA	Athletics	1035
China	Badminton	64
Cuba	Baseball	112
USA	Basketball	331
Spain	Basque Pelota	2
Brazil	Beach Volleyball	26
Germany	Biathlon	85
Switzerland	Bobsleigh	92
USA	Boxing	105
Germany	Canoeing	184
UK	Cricket	17
France	Croquet	8
Norway	Cross Country	164
Canada	Curling	48
UK	Cycling	172
USA	Diving	138
Germany	Equestrianism	173
Italy	Fencing	358
USA	Figure Skating	61
Brazil	Football	136
USA	Freestyle Skiing	21
USA	Golf	36
Soviet Union	Gymnastics	290
Soviet Union	Handball	97
Netherlands	Hockey	255
Canada	Ice Hockey	347
UK	Jeu De Paume	2
Japan	Judo	84
Canada	Lacrosse	23

This query identifies which region has won the most medals in each sport. It first calculates the total medals by region and sport, then finds the maximum medal count for each sport, and finally selects the region(s) that achieved that maximum. The results show the top-performing regions in each sport. LIKE BOXING HAVE 105 MEDAL

What are some notable instances of unexpected or surprising medal wins?

<pre> SELECT nr.region_name, s.sport_name, COUNT(cc.medal_id) AS total_medal FROM noc_region AS nr JOIN person_region AS pr ON nr.id = pr.region_id JOIN games_competitor AS gc ON pr.person_id = gc.person_id JOIN competitor_event AS ce ON gc.id = ce.competitor_id JOIN event AS e ON ce.event_id = e.id JOIN sport AS s ON e.sport_id = s.id where ce.medal_id <> 4 and ce.medal_id=1 group by 1,2 having total_medal = 1 </pre>	Result shown by above query have region_name,sport_name and total medal won by that specific region.These medal count only have gold medal in it as this show a suprising win by that region that in gold medal winning sport		
	region_name	sport_name	total_medal
	Uzbekistan	Athletics	1
	Jordan	Tackwond	1
	Italy	Volleyball	1
	Kuwait	Shooting	1
	United Arab Emirates	Shooting	1
	Austria	Triathlon	1
	Lithuania	Modern Pentathlon	1
	Kazakhstan	Swimming	1
	Soviet Union	Ski	1
	Latvia	Figure Skating	1
	Canada	Diving	1
	Italy	Skeleton	1
	India	Shooting	1
	Yugoslavia	Swimming	1
	Poland	Gymnastics	1
	Bulgaria	Swimming short	1
	Australia	Track Speed Skating	1
	Australia	Snowboarding	1
	Poland	Speed Skating	1
	Australia	Tackwond	1
	Turkey	Athletics	1
	Mexico	Diving	1
	New Zealand	Triathlon	1

Are there any regions that have experienced significant growth or decline in Olympic participation?

```

with participation_by_decade as
(SELECT
  nr.region_name,
  FLOOR(g.games_year / 10) * 10 AS decade,
  COUNT(DISTINCT ce.competitor_id) AS total_participants
FROM noc_region AS nr
JOIN person_region AS pr ON nr.id = pr.region_id
JOIN games_competitor AS gc ON pr.person_id = gc.person_id
JOIN competitor_event AS ce ON gc.id = ce.competitor_id
join games as g on gc.games_id=g.id
GROUP BY 1,2),

participation_trend AS (
  SELECT
    region_name,
    decade,
    total_participants,
    LAG(total_participants, 1) OVER (PARTITION BY
      region_name ORDER BY decade) AS
      previous_decade_participants,
    total_participants - LAG(total_participants, 1) OVER
      (PARTITION BY region_name ORDER BY decade) AS
      change_in_participation
  FROM
    participation_by_decade
  where decade > 1990
),

trend_analysis AS (
  SELECT

```

region_name	decade	total_participants	previous_decade_participants	change_in_participation	percent_change
Fiji	2010	62	21	41	195.2381
Montenegro	2010	71	25	46	184
Djibouti	2010	11	4	7	175
United Arab Emirates	2010	40	17	23	135.29412
Cook Islands	2010	17	8	9	112.5
Brunei	2010	6	3	3	100
Gabon	2010	27	14	13	92.85714
Marshall Islands	2010	9	5	4	80
Serbia	2010	239	134	105	78.35827
Virgin Islands, British	2010	7	4	3	75
Bahrain	2010	46	28	18	64.28571
Timor-Leste	2010	6	4	2	50
Colombia	2010	245	166	79	47.59036
Ecuador	2010	74	52	22	42.30769
Eritrea	2010	24	17	7	41.17647
Tuvalu	2010	4	3	1	33.33333
Togo	2010	13	10	3	30
Vietnam	2010	40	31	9	29.03226
Turks and Caicos	2010	10	8	2	25
Estonia	2010	133	167	-34	-20.35928
Saint Vincent	2010	7	9	-2	-22.22222
Syria	2010	17	22	-5	-22.72727
Trinidad and Tobago	2010	53	69	-16	-23.18841
Poland	2010	549	715	-166	-23.21676
Guam	2010	13	17	-4	-23.52941
Portugal	2010	170	223	-53	-23.76682
Chile	2010	86	113	-27	-23.89381
Belize	2010	6	8	-2	-25
Maldives	2010	9	12	-3	-25
Chile	2010	86	113	-27	-23.89381
Belize	2010	6	8	-2	-25
Maldives	2010	9	12	-3	-25
Czech Republic	2010	415	557	-142	-25.49372
Georgia	2010	87	117	-30	-25.64103
Japan	2010	826	1118	-292	-26.11807
Micronesia	2010	11	15	-4	-26.66667
San Marino	2010	11	15	-4	-26.66667
Croatia	2010	225	307	-82	-26.7101
Venezuela	2010	154	211	-57	-27.01422
American Samoa	2010	8	11	-3	-27.27273
Democratic Republic of the Congo	2010	8	11	-3	-27.27273
Morocco	2010	115	159	-44	-27.67296
Liechtenstein	2010	13	18	-5	-27.77778
Bermuda	2010	18	25	-7	-28
Haiti	2010	15	21	-6	-28.57143
Palau	2010	10	14	-4	-28.57143
USA	2010	1511	2121	-610	-28.76002
Israel	2010	93	131	-38	-29.00763
Jordan	2010	17	24	-7	-29.16667
Jamaica	2010	106	150	-44	-29.33333
Lithuania	2010	145	206	-61	-29.61165
South Africa	2010	265	377	-112	-29.70822
Puerto Rico	2010	66	94	-28	-29.78723
Sierra Leone	2010	7	10	-3	-30
Yemen	2010	7	10	-3	-30
Bahamas	2010	50	72	-22	-30.55556
El Salvador	2010	18	26	-8	-30.76923
Turkmenistan	2010	20	29	-9	-31.03448

```

    trend_analysis AS (
      SELECT
        region_name,
        decade,
        total_participants,
        previous_decade_participants,
        change_in_participation,
        CASE WHEN previous_decade_participants IS NOT NULL
          AND previous_decade_participants > 0 THEN
          (change_in_participation * 100.0) /
            previous_decade_participants
        ELSE NULL
        END AS percent_change
      FROM participation_trend
    )

    SELECT
      region_name,
      decade,
      total_participants,
      previous_decade_participants,
      change_in_participation,
      percent_change
    FROM
      trend_analysis
    WHERE
      ABS(percent_change) > 20
    ORDER BY
      percent_change DESC

```

Are there any regions that have had a notable impact on the overall medal tally?

```
with total_medal as
  (SELECT
    nr.region_name,
    COUNT(ce.medal_id) AS total_medal
  FROM noc_region AS nr
  JOIN person_region AS pr ON nr.id = pr.region_id
  JOIN games_competitor AS gc ON pr.person_id = gc.person_id
  JOIN competitor_event AS ce ON gc.id = ce.competitor_id
  where ce.medal_id <> 4
  group by 1 )
  select
    region_name,total_medal,
    round((total_medal*100)/(select sum(total_medal) from total_medal),2)
    as percentage
  from total_medal
  group by 1,2
  having percentage > 4
  order by 3 desc
```

region_name	total_medal	percentage
USA	5414	13.36
Soviet Union	2658	6.56
Germany	2395	5.91
UK	1971	4.86
France	1686	4.16

MAX MEDAL BAGGED BY REGION HAVE
USA ON TOP ACCOUNTING FOR AROUND
13.36 % MEDALS AND FOLLOWED BY
SOVIET UNION.THIS SHOWS THE TEAM
EFFORT SHOWN BY VARIOUS COUNTRIES
TO ACHIEVE LAURELS IN FIELD OF
SPORTS



THANK YOU