

INTRODUCTION

The Olympics games are a major international event with summer and winter sports, where thousands of athletes participate in various competitions. These games occur every four years under their seasonal games. Originally the ancient Olympics games were held in Olympia, Greece.

In this report I have carried out the data analysis of 120 years of Olympics data starting all the way back from 1896 to 2016. The goals of this visualizations are to determine the performances of different athletes and teams over the years across different games. In addition to this, analysis of women's and India's participation, their contribution and performances are also evaluated across different events.

DATASET DESCRIPTION AND SOURCE

The dataset is available in csv format and basically contain the below mentioned data

1. athlete_events.csv file which contains 27116 Rows and 15 columns.
 - ID - Unique number for each athlete
 - Name - Athlete's name
 - Sex - M or F
 - Age – Integer
 - Height – Athlete height in centimetres
 - Weight - Athlete weight in kilograms
 - Team - Athlete Team name
 - NOC – Athlete National Olympic Committee 3-letter code
 - Games - Year and season when the Olympics was held
 - Year – year in which games took place Integer
 - Season – Whether it's a Summer or Winter games
 - City – city which hosted the games
 - Sport – Sport in which the athlete has participated
 - Event – Event under the sport category in which the athlete has participated
 - Medal – Medals won by athlete Gold, Silver, Bronze, or NA

The data was taken from Kaggle and the link is as mentioned below –

PROCESS

To begin with, imported the data into power BI desktop for data cleaning and pre-processing. Promoted the headers, changed the datatypes and replaced the NA values with 0 in the Height and weight columns. Created new columns called Gold, silver, Bronze and Total medals along with a conditional's column (India's participation). Created relevant Measures using DAX Formula and build a report using the tree map, pie chart, line chart, Donut, slicers etc. At the end published the report on the power BI service.

DAX MEASURES USED

SI No	MEASURE NAME	DAX FUNCTION	M-CODE	EXPLANATION
1	Total events	<code>DISTINCTCOUNT()</code>	Total events = <code>DISTINCTCOUNT(olympics_data[Event])</code>	Will count the distinct values in the Event columns
2	Total athletes	<code>COUNT()</code>	Total athletes = <code>COUNT(olympics_data[Name])</code>	Will count the no of cells in the name column
3	Teams participated	<code>DISTINCTCOUNT</code>	Teams participated = <code>DISTINCTCOUNT(olympics_data[Team])</code>	Will count the distinct values in the Teams columns
4.	Male athletes Female athletes	<code>IF()</code> <code>ISBLANK()</code> <code>CALCULATE()</code>	Male athletes = <code>if(isblank(CALCULATE([Total athletes],olympics_data[Sex]="M")),0,CALCULATE([Total athletes],olympics_data[Sex]="M"))</code> Female athletes = <code>if(isblank(CALCULATE([Total athletes],olympics_data[Sex]="F")),0,CALCULATE([Total athletes],olympics_data[Sex]="F"))</code>	<u>If()</u> - will check whether a condition is met . <u>Isblank()</u> -will check if the value we got is blank <u>Calculate()</u> - will evaluate an expression in a context modified by filters
5.	% of Female athletes % of male athletes	<code>DIVIDE()</code>	% of Female athletes = <code>DIVIDE([Female athletes],[Total athletes])</code> % of male athletes = <code>'New Measures'[Male athletes]/[Total athletes]</code>	<u>Divide ()</u> – Will perform division and returns alternate result or blank. ** can use normal division symbol to the same as well as done in %male athletes
6.	Total indian participants	<code>COUNT()</code>	Total indian participants = <code>COUNT(olympics_data[Indian participants])</code>	Will count the no of cells in the name column ** I Have created a column where I have categorised the athletes based on nationality/country and then used the count () to get the indian participation
7.	total indian female participants Total indian male participant	<code>CALCULATE()</code>	total indian female participants = <code>CALCULATE([Total indian participants],olympics_data[Sex]="F")</code> Total indian male participant = <code>CALCULATE([Total indian participants],olympics_data[Sex]="M")</code>	<u>Calculate()</u> - will evaluate an expression in a context modified by filters

INFERENCES AND CONCLUSIONS OF THE OUR GOALS –

1.1 Most successful athlete

Name	Team	Sports	Total medal
Michael Fred Phelps, II	United States	Swimming	28
Larysa Semenivna Latynina (Diriy-)	Soviet Union	Gymnastics	18
Nikolay Yefimovich Andrianov	Soviet Union	Gymnastics	15
Borys Anfiyanovych Shakhlin	Soviet Union	Gymnastics	13
Edoardo Mangiarotti	Italy	Fencing	13
Ole Einar Bjrndalen	Norway	Biathlon	13
Takashi Ono	Japan	Gymnastics	13
Aleksey Yuryevich Nemov	Russia	Gymnastics	12
Birgit Fischer-Schmidt	East Germany	Canoeing	12
Dara Grace Torres (-Hoffman, - Minas)	United States	Swimming	12

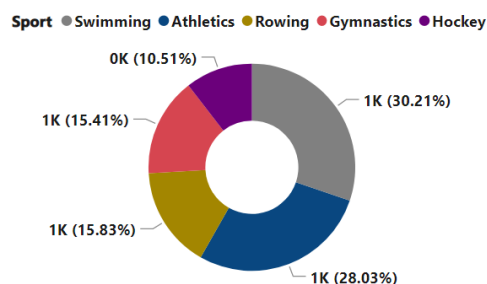
We conclude that **Michael Phelps** from the United States of America is the most decorated athlete to have graced the games as he has won a total of 28 medals in swimming out of which 23 are golds, 3 silver and 2 bronze medals.

1.2 Countries with most medals

Team	Gold	Silver	Bronze	Total Medal
United States	2474	1512	1233	5219
Soviet Union	1058	716	677	2451
Germany	679	627	678	1984
Great Britain	519	582	572	1673
France	455	518	577	1550
Italy	535	508	484	1527
Sweden	451	476	507	1434
Australia	342	453	511	1306
Canada	422	413	408	1243
Hungary	432	330	365	1127
Total	7367	6135	6012	19514

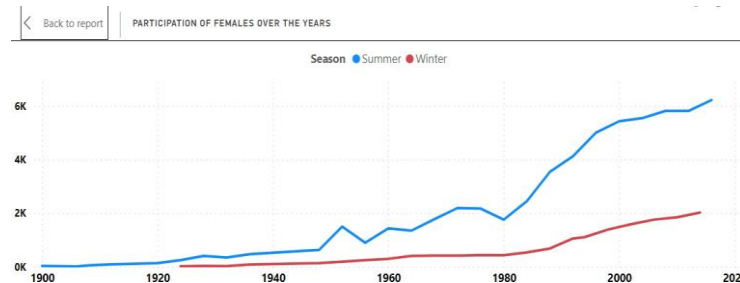
The United States of America have the greatest number of gold medals in the history of the games followed by the Soviet Union and Germany. Out of the 2474 gold medals the female athletes have bagged 830 (33.54%) gold medals which is the highest in the women's category which states the USA has benefited a lot from women participation

1.3 Sports in which most medals are won



Most number of medals are won in swimming 1374, followed by athletics 1275, Rowing 720, Gymnastics -701

2. Females' participation in games over the years



There has been a tremendous increase in the rate of participation by the females in the games since its evolution. The higher the participation of the women's athlete the more, medals that will be brought home. Example of the same is the females from the USA

3. Females' participation Stats

3.1 Most medals by a female athlete

Name	Team	Sports	Total Medal
Larysa Semenivna Latynina (Diriy-)	Soviet Union	Gymnastics	18
Birgit Fischer-Schmidt	East Germany	Canoeing	12
Dara Grace Torres (-Hoffman, - Minas)	United States	Swimming	12
Jennifer Elisabeth "Jenny" Thompson (-Cumpelik)	United States	Swimming	12
Natalie Anne Coughlin (-Hall)	United States	Swimming	12

The highest number of medals bagged by a female athlete over the year is 18 and the record is held by Laryssa latnina from soviet union in gymnastics out of which 9 are gold , 5 are silver and 4 are bronze .

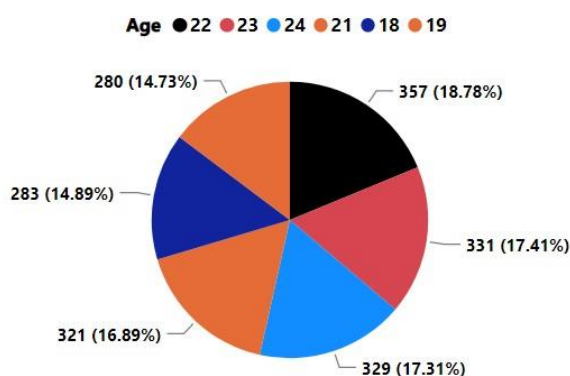
3.2 Most medals by a female athlete – country wise

Team	Gold	Silver	Bronze	Total Medal
United States	830	521	405	1756
China	186	236	178	600
Germany	203	218	236	657
Russia	202	181	150	533
Australia	177	178	188	543
Soviet Union	300	174	193	667
Canada	151	148	206	505
Netherlands	142	146	142	430
East Germany	188	141	104	433
Great Britain	95	125	160	380
Total	2474	2068	1962	6504

The female athletes of the USA have bagged 830 (33.54%) gold medals which is the highest in the women's category followed by Soviet Union, Germany and China.

3.3 Most medals by a female athlete – age wise

FEMALE MEDAL TALLY AGE WISE



Most gold medals are won by the female's athlete when they are 22 years of age, followed by 23, 24, 21 etc.

4 To determine the youngest and oldest participant to win a medal

4.1 Youngest athlete to win a medal

Age	Name	Team	First Sport	Gold	Silver	Bronze
10	Dimitrios Loundras	Ethnikos Gymnastikos Syllogos	Gymnastics	0	0	1

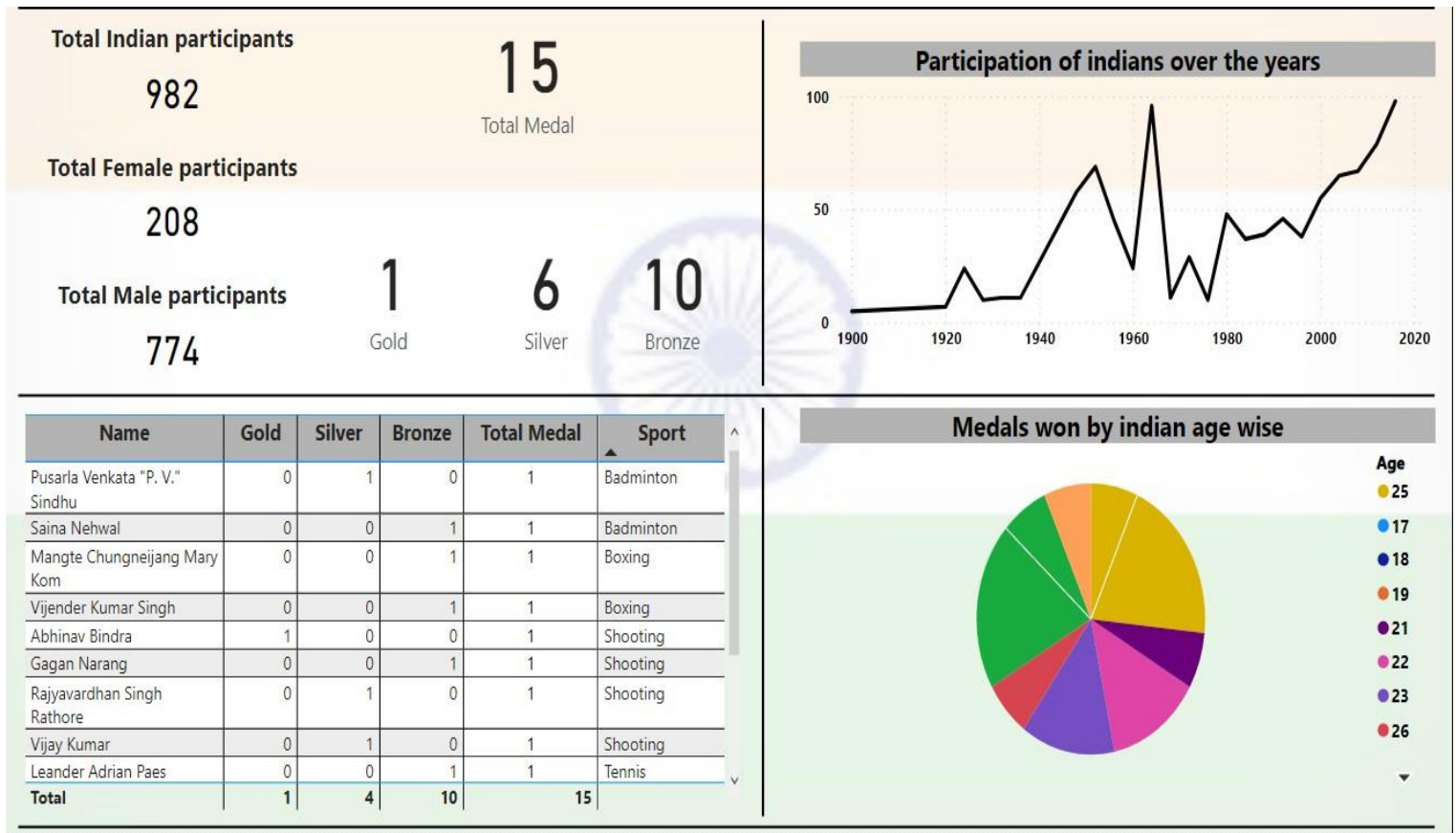
Dimitrios Loundras, is the youngest Olympic winner ever at the age of just 10. He participated in gymnastics in the first-ever modern Olympics in 1896. He took home a bronze medal in the men's team parallel bar event.

4.2 Oldest athlete to win a medal

Age	Name	Total Medal	Sport	Team	Gold	Silver	Bronze
73	Anne Marie Carl-Nielsen (Brodersen-)	1	Art Competitions	Belgium	0	1	0

Anne Marie Carl-Nielsen, is the oldest Olympic winner ever at the age of 73. He participated in art competitions. He took home a silver medal for the Belgium team.

5. India's participation and medal winner



A total of 982 participants have participated in the games. We have a total tally of 15 medals of which there is only 1 gold medal, 6 silver and 10 bronze medals. Amongst these 5 medals are in wrestling, 4 in shooting, 2 in boxing, 2 in badminton and 1 in tennis and 1 in weightlifting. In addition to this we have also won 8 gold medals and 1 silver medal and 2 bronze medals in the men's hockey over the years.

References

- [1] R. H. Griffin, "Kaggle," 15 June 2018. [Online]. Available: <https://www.kaggle.com/heesoo37/120-years-of-Olympic-history-athletes-and-results>. [Accessed 16 December 2018].