

# Asian games and Olympic games data analysis

## CS685 PROJECT REPORT

Group 10

Deepak Raj-21111024-deepakr21@iitk.ac.in  
Dinkar Tewari-21111025-dinkart21@iitk.ac.in  
Divyansh Bisht-21111027-dbisht21@iitk.ac.in  
Rohit Kushwah-21111053-krohit21@iitk.ac.in  
Vikas-21111067-vikas21@iitk.ac.in

Indian Institute of Technology Kanpur (IIT Kanpur)

November 21, 2021

### Abstract

Asian games and Olympic games are very similar to each other in terms of sport events. But only Asian athletes are common in both. There are various insights that can be drawn from these historic sports data. Our work intends to extract this historic data using web scraping technique using BeautifulSoup python package. After data extraction, this work intends to extract various insights from this historic data using data mining techniques. Finally, our work takes an intersection of both Asian games data and Olympic games data and carries out a comparative analysis on this dataset. Visual insights created after data mining process are deployed on a web app for better user interface.

## 1 Problem Statement

Asian games and Olympic games are two of the most famous sport events worldwide. They both occur once in four years. They are scheduled in a way such that there is a two-year gap between the two. The athletes first participate in Asian games. Then after two years they participate in Olympic games as well in general. This cycle goes on. A total of 45 Asian countries participates in both the events and are registered under Olympic Council of Asia. The format and type of events are almost similar in both the events. Yet, there are differences in medal winning patterns in both these events for Asian countries. This work must showcase extensive analysis on player data, country data, gender data, host country data, year wise data, medal data and sport data. The countries taken into account will only be Asian countries. The years taken into account will be after 1951 only as Asian games started from 1951. So, Olympics data before 1951 must be discarded from the study as its intersection with Asian games dataset provides null set.

## 2 Introduction

Our work aims to work on data of Olympic games and Asian games separately first. Then we wish to take their intersection and do a comparative analysis between the two. Our work intends to cover several questions which will be discussed in detail in this report later. After performing analysis on the dataset, we generate output tables. To make it visually appealing, we have created various visualizations of this data using various graphs and plots. For making it easy to visualize data, we have even created a front-end web app using streamlit that creates an overview of all our analysis.

### 3 Dataset

Since our project is to analyze Asian Games and Olympic Games, we needed four datasets, i.e., player-wise and country-wise medal tally of Asian and Olympic Games. There were no Asian Games data available anywhere, so we scrapped the data directly from the website [https://en.wikipedia.org/wiki/Asian\\_Games](https://en.wikipedia.org/wiki/Asian_Games). During scrapping, we had to go to every Asian game year and get the player data and overall sports data. The most challenging job was to check the structure of every year's website, as they were different for a different set of years. So, we divided our work in a way where each person has to scrape data from similar structured websites. Then in each year, we scrapped every player's performance and country's performance in each sport. The next challenge was to find the gender of the players. But only for a few games, gender was specified. So, we had to go through each player page, and from the description, if given, extracted gender. That way, we got around 82% of the player's gender. Even some countries had different names over the year, they were resolved by using NOC which is the same throughout the years. In the end, we had two datasets for Asian Games, "Players\_Medal\_List\_Asian.csv" and "Medal\_List\_Asian.csv".

For the Olympic Games, we got a dataset from [Kaggle dataset](#) which contains Olympics data from 1896 to 2016 of winter and summer Olympics. We needed data only after 1951, the year Asian Games started. Also, we were only analyzing Asian countries' performance in Summer Olympics. So, we considered the Asian countries which were recognized by the National Olympic Committee. We cropped the data accordingly and added 2020 Summer Olympics data manually at the end. We got two datasets for Olympics as well, which are, "Players\_Medal\_List\_Olympics.csv" and "Medal\_List\_Olympics.csv".

The file Players\_Medal\_List\_Asian.csv contains 31020 rows and 10 columns. Each row corresponds to an individual athlete competing in an Asian Game event. The file Players\_Medal\_List\_Olympics.csv contains 3349 rows and 11 columns. Each row corresponds to an individual athlete winning a medal in the Olympic games. The file Medal\_List\_Asian.csv contains 3408 rows and 8 columns. Each row corresponds to the count of medals won by a nation in a sport in a given year in Asian Games. The file Medal\_List\_Olympics.csv contains 838 rows and 8 columns. Each row corresponds to the count of medals won by a nation in a sport in a given year in Olympic Games.

### 4 Methodology

#### Step1 : Data Scraping

Incredible amounts of data exist across all domains of sports. The task is not how to collect the data, but what data should be collected and how to make the best use of it. So, the first part of our problem was to select which data to be collected and decide on how to collect the data. We used the BeautifulSoup library of python to scrape the Asian Games data from Wikipedia. We got Olympics data from Kaggle.

#### Step2 : Data Cleaning

Then, we cleaned all the data in a way that there is no incomplete or inconsistent data. There were many unknowns for Gender which were added. Some columns like Year, Sports were inconsistent, they were fixed. In the end, most of the anomalies were removed. Also, country names were inconsistent as few countries changed their name as Ceylon was changed to Sri Lanka. They were handled according to the NOC of the countries, as the NOC of any country is same throughout the years.

#### Step3 : Data Integration

The Asian game data were collected in parts by all the group members. Now the challenge was to combine all these data without any dependencies. Also, the Olympic data was till 2016. So, it was combined with 2020 Olympics data to get complete datasets ready.

#### Step4 : Data Reduction

As we are comparing Asian and Olympic games, we only considered Summer Olympics after the Asian games started, i.e., from 1952. Also, we were only analyzing the Asian countries. So, we considered the Asian countries which were recognized by National Olympic Committee.

### Step5 : Data Transformation

After having all the data, some parts were needed to be changed so that all data aligned with mining goals. Like, one data set has Sports type as "Individual" or "Team", other has "Single" or "Team". Both datasets were transformed into "Single" or "Team". Many such differences were handled. Similar types of columns were named similarly so that it is easier while analyzing all the data.

### Step6 : Data Mining

After having all the data ready, now the challenge was to find interesting patterns within the data and get information out of them. Each member of the team worked on different aspects to get information on the datasets we created. We used python language to analyze the data.

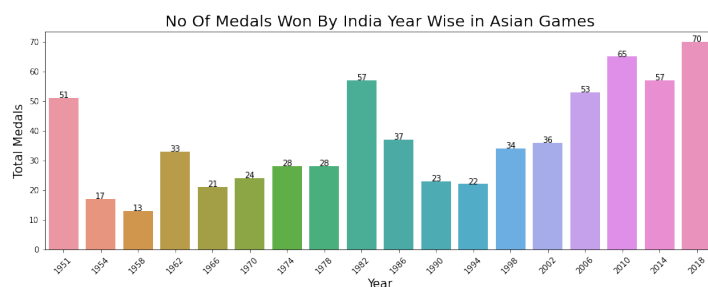
### Step7 : Representing Knowledge in Data Mining

Finally, the task is to represent our analysis. We used the streamlit library of python to represent our whole analysis in a structured way. We used different types of graphs, and tables to represent our analysis. Out of all the information which we got from all the analysis, the most interesting and imported information were all noted in the final report.

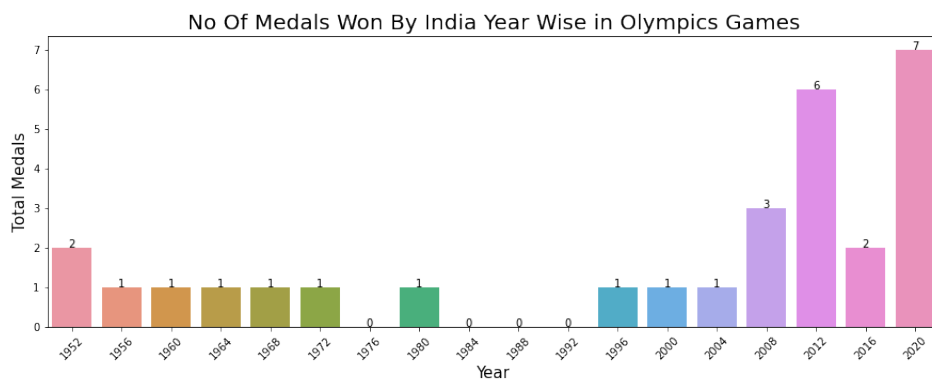
## 5 Results

### 5.1 Medal tally analysis

The medal tally analysis is done for both the Asian Games and the Olympics Games, for every country, year, and sport.

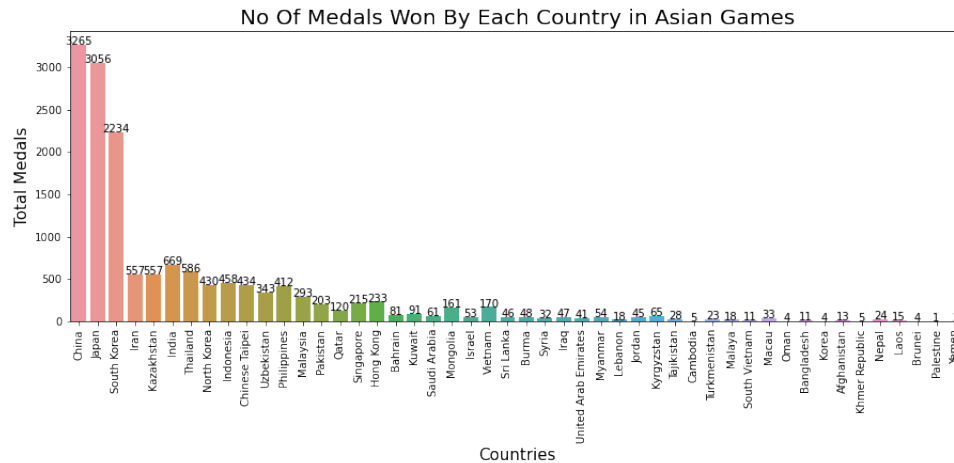


In Asian Games, India has won 669 medals(154 golds, 198 silvers, 317 bronzes). The worst years for India were 1954, 1958, 1966, 1990, and 1994. The best years for India were 2010(65 medals) and 2018(70 medals). India has performed really well in the last 3 Asian games.

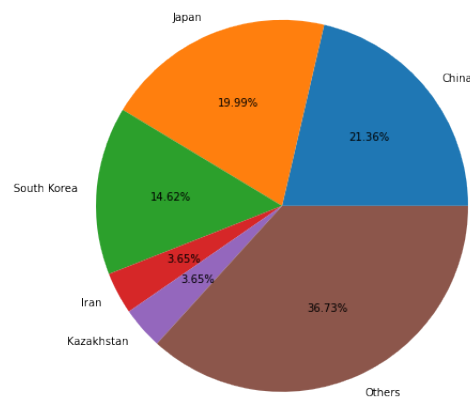


In Olympics games, India has won 29 medals(6 golds, 7 silvers, 16 bronzes). The worst years for India were 1976, 1984, 1988, and 1982 in which India did not win even a single medal. The best years for India were 2012(6 medals) and 2020(7 medals).

## Countries performance in Asian Games

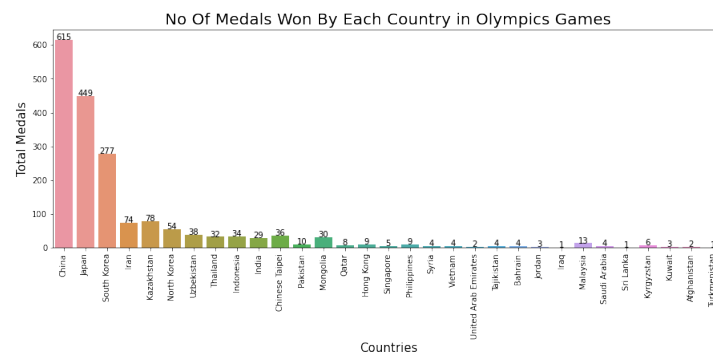


In Asian Games, a total of 15279 medals(4771 golds, 4757 silvers, 5751 bronzes) have been won. The top 5 Performers of Asian games are China, Japan, South Korea, India, and Thailand. India holds 4th position with 669 medals.



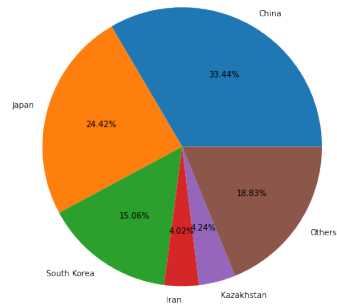
China has 21.36% medals of the total medals. Japan and South Korea's percentage of medals is 19.99% and 14.62%.

## Countries performance in Olympics Games



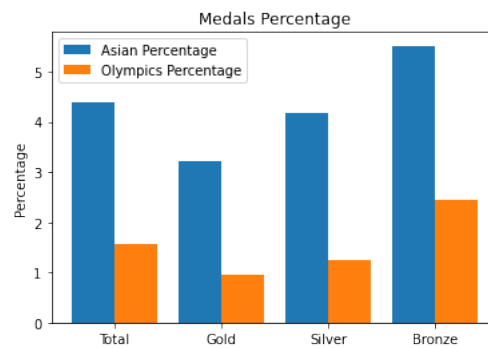
In the Olympics Games, 1839 medals(621 golds, 573 silvers, 655 bronzes) have been won. The top 5 Performers of Asian games are China, Japan, South Korea, Kazakhstan, and Iran. India holds 12th position with 29 medals.

Top performer China has 33.44% of total medals won in Olympic won. Japan has 24.42% and South Korea has 15.06% of total medals won.



A total of 72.92% of medals are won by just three countries - China, Japan, and South Korea. This infers that these three countries clearly dominate Olympic Games.

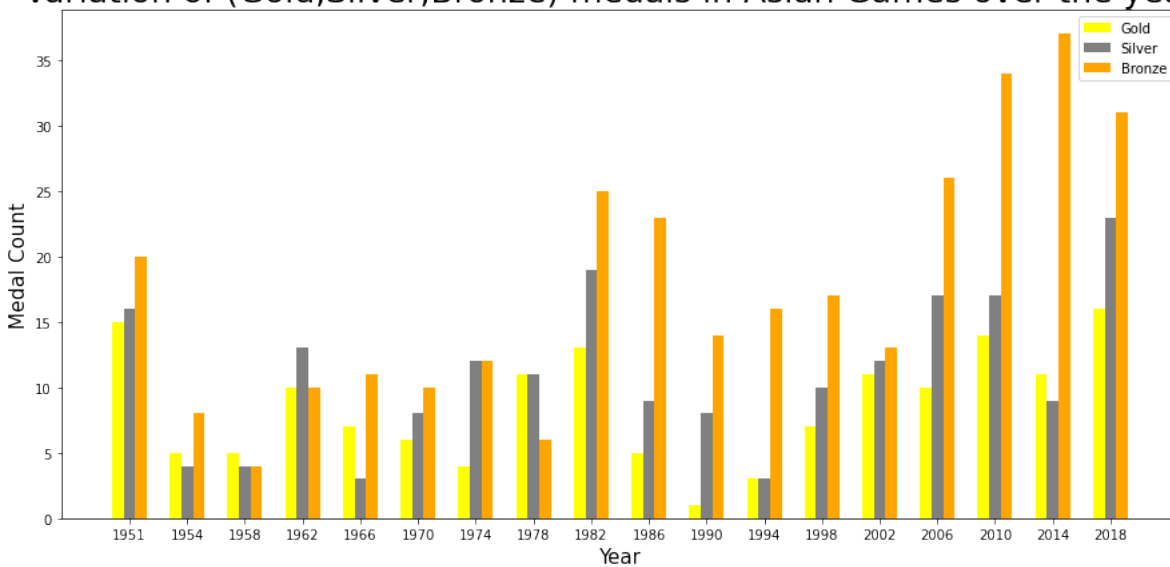
#### Medal percentage comparison of India in Asian Games and Olympics



The percentage of medals won by India in the Olympics is almost half of the percentage of medals won by India in Asian Games.

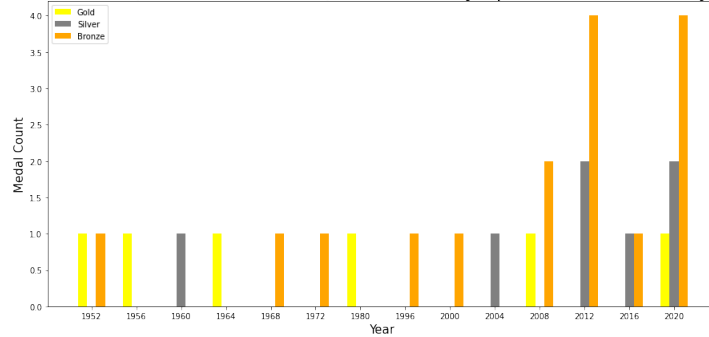
#### Variation of Medals of India over the years

#### Variation of (Gold,Silver,Bronze) medals in Asian Games over the years



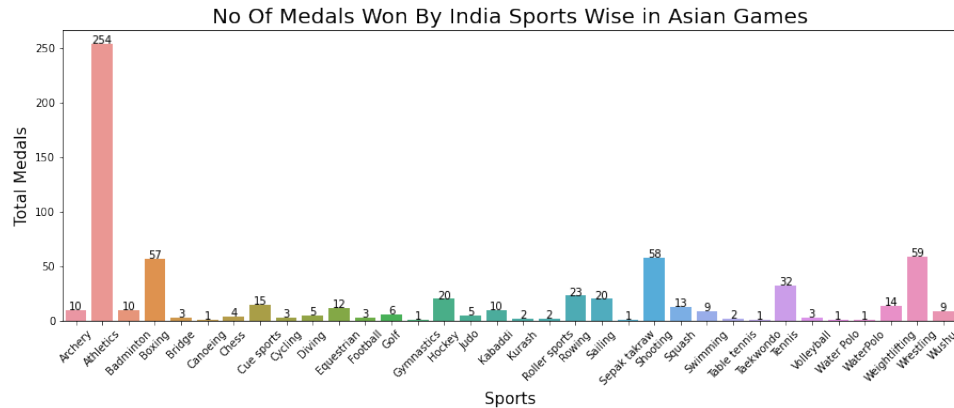
In the years 1951 and 2018, India has more gold medals than other years. In 2014 India had a maximum number of bronze medals, and in 2018, a maximum number of silver medals.

Variation of (Gold,Silver,Bronze) medals in Olympic Games over the years



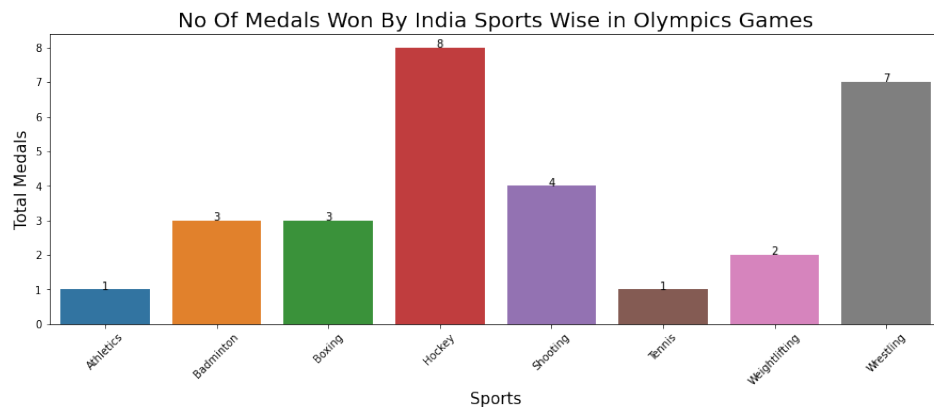
In the years 2012 and 2020, India has won a maximum number of silver and bronze medals as compared to other years. There are some years in which India did not get any gold medals.

### Sport-wise medals of India in Asian Games



The top sport for India is Athletics in Asian Games. Out of 669 medals, 254 medals are won by India in athletics alone. The other three best sports of India are Boxing, Shooting, and Wrestling. The sports India is very bad at are Canoeing, Gymnastics, Sepak takraw, Taekwondo, and Water polo.

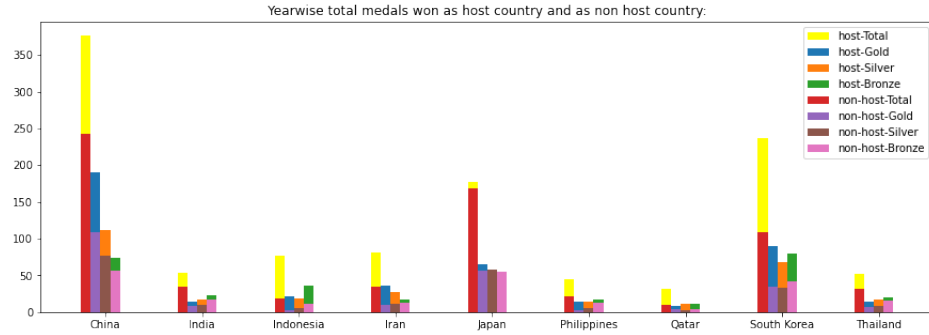
### Sport-wise medals of India in Olympics Games



To date, India has won medals in only eight sports in Olympics. The top sports for India are Hockey(7 medals), Shooting(4 medals), and Wrestling(7 medals). Although the best sport for India in Asian Games is athletics, India has performed very poorly in the same sport in Olympics.

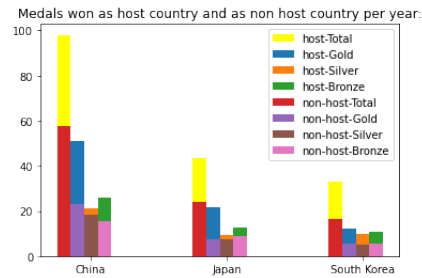
## 5.2 Medals won as a host country vs non host country

### Asian Games



Asian games have been hosted by 9 countries so far. Thailand hosted Asian games a maximum number of times which is 4 times. South Korea has hosted 3 times, India, China, Indonesia, Iran, Japan have hosted 2 times, Iran, Philippines, and Qatar have hosted Asian games for 1 time only. The total number of medals won by China, Indonesia, and South Korea is almost doubled when these countries played on home ground. From the above graph, we can infer that the number of total medals and gold medals are highly affected by playing as the host nation. Silver and Bronze are little affected but still, they are slightly more than medals when won as a non-host country. For Japan only, the performance as host nation and away nation is almost the same.

### Olympics Games



Similarly for Olympic Games, from the above graph, we can infer that the number of total medals and gold medals are highly affected by playing as the host nation. Silver and Bronze are little affected but, they are still more than medals won as a non-host country. Other Asian countries have never been host in Olympic games from 1950 onwards. So their data is filtered out in this analysis by default.

## 5.3 Player-wise analysis

### Top athletes for India in Asian Games

	Player_Name	Gold	Silver	Bronze	Total
0	Leander Paes	5	0	3	8
1	Milkha Singh	5	0	0	5
2	P. T. Usha	4	7	0	11
3	Jaspal Rana	4	2	2	8
4	Parduman Singh Brar	3	1	1	5

Leander Paes is the most successful player for India in Asian games who plays tennis. He has won 5 gold medals and 3 bronze medals for India. P.T. Usha holds the record of winning the maximum medals(11 medals) for India in the Asian Games.

### Top athletes in Asian Games

	Player_Name	Gender	Country	Gold	Silver	Bronze	Total
0	Wang Yifu	M	China	12	7	3	22
1	Zhang Li	F	China	10	1	0	11
2	Yoshimi Nishigawa	F	Japan	10	0	0	10
3	Sun Yang	M	China	9	5	0	14
4	So Gil-san	Unknown	North Korea	9	2	0	11

Wang Yifu is a male Chinese pistol shooter and, in terms of Asian medals, one of the most successful sport shooters of all time. He has won a total of 22 medals for china.

### Top athletes for India in athletics in Asian games

	Player_Name	Gold	Silver	Bronze	Total
0	P. T. Usha	4	7	0	11
1	Milkha Singh	4	0	0	4
2	Parduman Singh Brar	3	1	1	5
3	Manjeet Kaur	3	1	0	4
4	M. R. Poovamma	3	0	1	4

### Top athletes in athletics in Asian Games

	Player_Name	Gender	Country	Gold	Silver	Bronze	Total
0	Anat Ratanapol	M	Thailand	6	1	1	8
1	Shigenobu Murofushi	M	Japan	5	1	0	6
2	Mohamed Suleiman	M	Qatar	5	0	1	6
3	Maryam Yusuf Jamal	F	Bahrain	5	0	0	5
4	Toyoko Yoshino	F	Japan	5	0	0	5

### Top athletes for India in Olympics Games

Udham Singh Kular was an Indian Hockey Player from Punjab. He shares the distinction of being the only Indian player to win four Olympic medals.



	Name	Gold	Silver	Bronze	Total
0	Udham Singh Kular	3	1	0	4
1	Leslie Walter Claudius	2	1	0	3
2	Shankar Pillay Laxman	2	1	0	3
3	Balbir Singh Dosanjh, Sr.	2	0	0	2
4	Govind Perumal	2	0	0	2

#### Top athletes in Olympics Games

Takashi Ono is the player who has won maximum medals(13 medals) in gymnastics. All the top 5 athletes belong to japan.

	Name	Gender	Country	Gold	Silver	Bronze	Total
0	Sawao Kato	M	Japan	8	3	1	12
1	Akinori Nakayama	M	Japan	6	2	2	10
2	Takashi Ono	M	Japan	5	4	4	13
3	Yukio Endo	M	Japan	5	2	0	7
4	Mitsuo Tsukahara	M	Japan	5	1	3	9

#### Top athletes for India in athletics in Olympics games

Neeraj chopra is the only player who won medal for india in athletics.

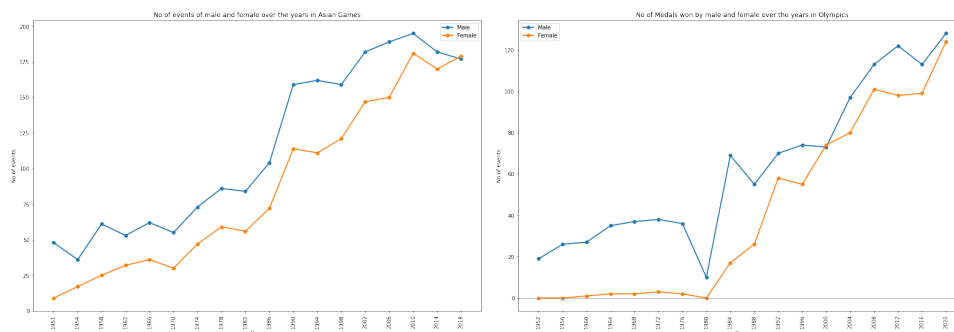
#### Top athletes in athletics in Olympics games

Top top Asian athletes overall for athletics in Olympics games are epresented in the table given.

	Name	Gender	Country	Gold	Silver	Bronze	Total
0	Mutaz Essa Barshim	M	Qatar	1	1	1	3
1	Wang Junxia	F	China	1	1	0	2
2	Gong Lijiao	F	China	1	0	1	2
3	Koji Alexander Murofushi	M	Japan	1	0	1	2
4	Liu Hong	F	China	1	0	1	2

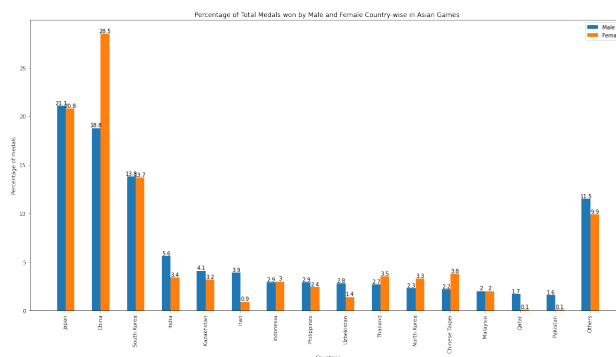
## 5.4 Male vs Female Analysis

Does the increase in no of events for males and females in Asian games increase medals in Olympics?



The first graph shows the number of events for male and female athletes over the years in Asian games. The second graph shows the number of medals won by Asian male and female athletes in Olympics from 1952(After the Asian games have started). With the increase in games for males and females in Asian games, there is an increase in the number of medals won in the Olympics by males and females. Also, the gap between medals won by males and females in the Olympics decreases as the gap between the number of games in Asian games decreases.

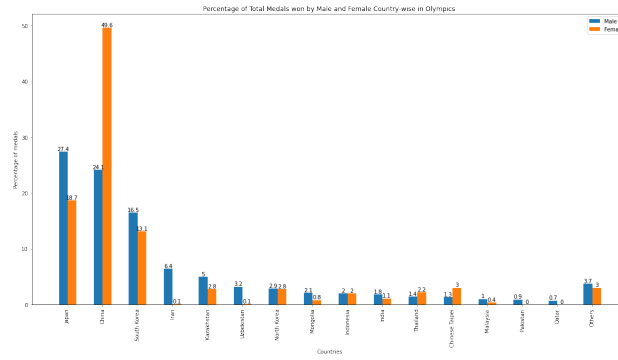
### Medals distribution for male and female in Asian Games



Out of all medals distributed to males, 21.1% medals are won by Japanese male athletes, 18.8% are won by china male athletes, 13.8% medals are won by South Korean male athletes, 5.6% medals are won by Indian male athletes, and the remaining by other countries male athletes. Out of all medals distributed to females, 20.8% medals are won by Japanese female athletes, 28.5% are won by china female athletes, 13.7% medals are won by South Korean female athletes, 3.4% medals are won by Indian female athletes, and the remaining by other countries female athletes. Japan and South Korea male athletes and female athletes have almost similar winning percentages. However, Chinese female athletes winning percentage is very much greater than Chinese male athletes.

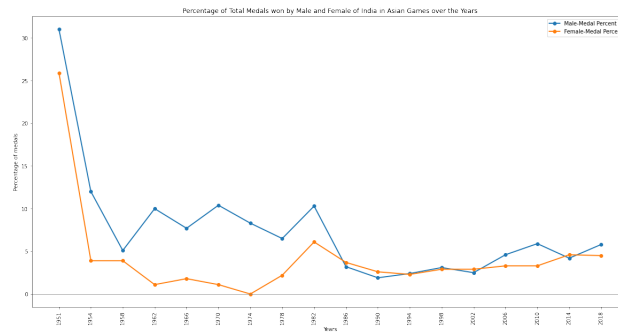
### Medals distribution for male and female in Olympics Games

Out of all medals distributed to males, 27.4% medals are won by Japanese male athletes, 24.1% are won by china male athletes, 16.5% medals are won by South Korean male athletes, 1.8% medals are won by Indian male athletes, and the remaining by other countries male athletes. Out of all medals distributed to females, 18.7% medals are won by Japanese female athletes, 49.6% are won by china female athletes, 13.1% medals are won by South Korean female athletes, 1.1% medals are won by Indian female athletes, and the remaining by other countries female athletes. Chinese female athletes winning percentage is very much greater than Chinese male athletes. Chinese female athletes have won almost 50% of total medals won by female athletes. This infers that female athletes clearly dominate



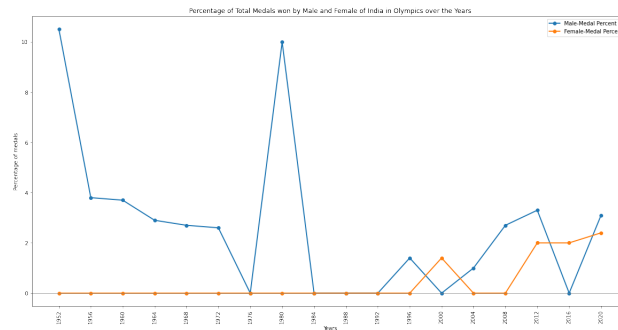
the other country female athletes. Japanese male athletes have won most male medals in Asian games, and they also won most male athletes' medals in the Olympics. Similarly, Chinese female athletes have won most female medals in the Asian Games, and they also have won the most female medals in Olympics. So we can clearly see that the countries with the best performance of male and female athletes in Asian games also performs well in Olympics.

### Male and Female athletes of India in Asian Games



From 1951 to 1986, female athletes have won fewer medals than male athletes. But from 1990 onwards, female athletes have performed really well and even outperformed male athletes in some years like 1990, 2002.

### Male and female athletes of India in Olympics



No female athlete has won any medal till 1996. The performance of female athletes has improved from the year 2000 in both the Asian and Olympics Games.

## 5.5 Comparison between Asian game and immediate Olympics

### Indian Athletes who won medal in Asian game and immediate Olympics

	Player	Year_Asian	Year_Olympic	Sport	Total_Asian	Total_Olympic
2	Vijay Kumar	2010	2012	Shooting	2	1
0	Karnam Malleswari	1998	2000	Weightlifting	1	1
1	Gagan Narang	2010	2012	Shooting	1	1
3	P. V. Sindhu	2018	2020	Badminton	1	1
4	Neeraj Chopra	2018	2020	Athletics	1	1
5	Bajrang Punia	2018	2020	Wrestling	1	1

Only Neeraj Chopra won the same type of medal i.e. Gold in Asian Games 2018 and its immediate Olympics 2020. There is no other player who won the same medal in Asian Games and its immediate Olympics.

**Valuable players (who won medals in both Asian and Olympics)**

	Name	Gender_Olympic	Sport_Olympic	Total_Asian	Total_Olympic
	P.V. Sindhu	F	Badminton	2	2
	Abhinav Bindra	M	Shooting	3	1
	Vijay Kumar	M	Shooting	3	1
	Saina Nehwal	F	Badminton	2	1
	Gagan Narang	M	Shooting	2	1
	Rajyavardhan Singh Rathore	M	Shooting	2	1
	Karnam Malleswari	F	Weightlifting	2	1
	Bajrang Punia	M	Wrestling	2	1
	Yogeshwar Dutt	M	Wrestling	2	1
	Neeraj Chopra	M	Athletics	1	1

## 5.6 Country with maximum sport winning percentage: Asian

	Year	Country	Total_Sports	Sports_Won	Sports_Lost	Percentage_Won	Percentage_Lost
1	1951	India	8	7	1	87.500000	12.500000
8	1954	Japan	10	9	1	90.000000	10.000000
21	1958	Japan	15	13	2	86.666667	13.333333
37	1962	Japan	16	13	3	81.250000	18.750000
52	1966	Japan	16	15	1	93.750000	6.250000
68	1970	Japan	14	13	1	92.857143	7.142857
84	1974	Japan	18	16	2	88.888889	11.111111
103	1978	Japan	21	19	2	90.476190	9.523810
124	1982	Japan	23	21	2	91.304348	8.695652
145	1986	South Korea	27	27	0	100.000000	0.000000
168	1990	Japan	29	25	4	86.206897	13.793103
171	1990	China	29	25	4	86.206897	13.793103
195	1994	Japan	33	32	1	96.969697	3.030303
226	1998	Japan	41	34	7	82.926829	17.073171
256	2002	South Korea	42	38	4	90.476190	9.523810
298	2006	Japan	42	36	6	85.714286	14.285714
334	2010	China	48	42	6	87.500000	12.500000
369	2014	South Korea	42	37	5	88.095238	11.904762
409	2018	Japan	51	41	10	80.392157	19.607843

In 1951, India had the most sport-winning percentage. Out of 8 sports, it won medals in 7 sports. South Korea is the only country that has a 100% sport-winning percentage. It has won medals in every sport it participated in the year 1986.

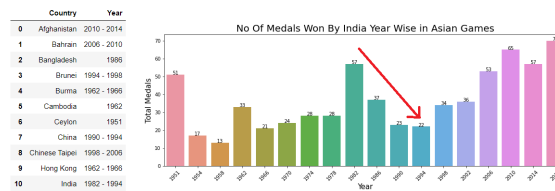
### Olympics

In the Olympics, Japan has the maximum sport winning percentage(90%). In 1964, it won medals in 9 sports out of 10 sports.

	Year	Country	Total_Sports	Sports_Won	Sports_Lost	Percentage_Won	Percentage_Lost
1	1952	Japan	6	3	3	50.000000	50.000000
5	1956	Japan	6	3	3	50.000000	50.000000
10	1960	Japan	8	6	2	75.000000	25.000000
16	1964	Japan	10	9	1	90.000000	10.000000
23	1968	Japan	8	7	1	87.500000	12.500000
29	1972	North Korea	9	5	4	55.555556	44.444444
30	1972	Japan	9	5	4	55.555556	44.444444
36	1976	Japan	8	6	2	75.000000	25.000000
43	1980	North Korea	5	3	2	60.000000	40.000000
46	1984	China	16	10	6	62.500000	37.500000
54	1988	South Korea	16	10	6	62.500000	37.500000
62	1992	China	20	15	5	75.000000	25.000000
75	1996	China	24	15	9	62.500000	37.500000
92	2000	China	20	12	8	60.000000	40.000000
108	2004	China	24	20	4	83.333333	16.666667
123	2008	China	28	25	3	89.285714	10.714286
141	2012	China	23	20	3	86.956522	13.043478
165	2016	China	23	20	3	86.956522	13.043478
184	2020	Japan	30	20	10	66.666667	33.333333

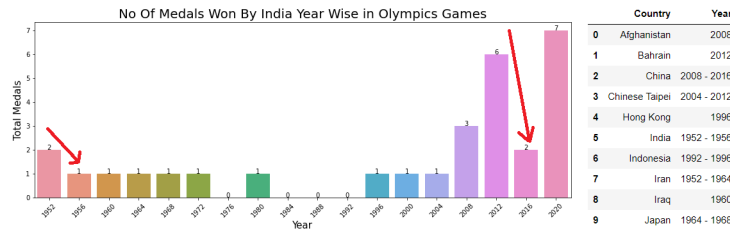
## 5.7 Recession and Booming Period

The recession period is the period in which the medals of a particular country keep on decreasing. The booming period is the period in which the medals of a particular country keep on increasing. **Recession period in Asian Games**



The recession period of India in Asian games is 1982 to 1994, India won 52 medals in 1982, 37 medals in 1986, 23 medals in 1990, and 22 medals in 1994.

### Recession period in Olympic Games



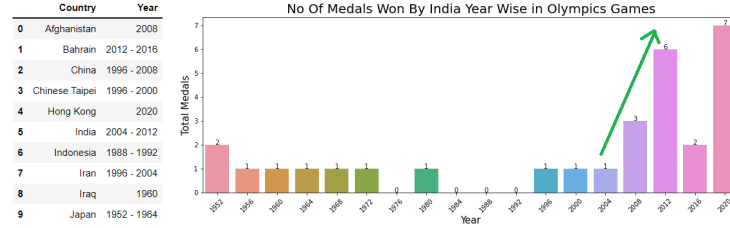
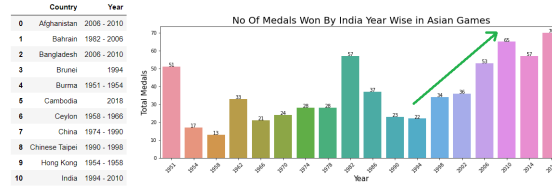
The recession period of India in the Olympics Games is from 1952 to 1956, India won 2 medals in 1952 and 1 medal in 1956.

### Booming period in Asian Games

The booming period of India in the Asian Games is from 1994 to 2010. India has performed really well in this period. India won 22 medals in 1994, 34 medals in 1998, 36 medals in 2002, 53 medals in 2006, and 65 medals in 2010.

### Booming period in Olympic Games

The booming period of India in the Olympics Games is from 2004 to 2004. India won 1 medal in 2004, 3 medals in 2008, and 6 medals in 2012. The intersection in the booming period of India in Asian Games and Olympics is 2004 to 2008. This somewhat infers that whenever any country performs better in Asian games, it is very likely to perform better in the Olympics Games in the same period.



## 6 Conclusion

Asian Games and Olympics are some of the most important games which are played once every four years. Athletes prepare for 4 long years to represent their countries and earn medals for their country. Some fail, some achieve great feats, but there is lots of hard work by every athlete over the years. Countries invest large amounts of money and efforts to win lots of medals. Here in our project, we are looking into Asian Games trying to see if there are similar trends in the Olympics too. Some of the findings are listed below:

- Top countries of Asian Games, like China, Japan, South Korea, are also top Asian countries in Olympics, that too in the same order.
- Increasing number of events in the Asian Games increases the chances of Asian countries medal winning chances.
- Countries where females performed better than males in Asian games, also performed better in Olympics. Same for males.
- China's women dominate in Asian Games, they are also to most medal winners in Olympics for Asian nations.
- China's women dominate in Asian Games, they are also to most medal winners in Olympics for Asian nations.
- China's women dominate in Asian Games, they are also to most medal winners in Olympics for Asian nations.
- The chances of winning a medal for a player in the Olympics increase when he starts winning medals in the Asian Games. E.g., Neeraj Chopra, Bajrang Punia are Indian athletes who first won medals in Asian games, then in Olympics. There are many such players. The Asian games winning players are more valuable for any country.
- Sports, where a country is good at Olympics, is also very good as Asian Games. But, some countries are good at some sports in Asian games but that doesn't make them good at the Olympics. E.g., India is good at Athletics in Asian games, but their performance at athletics in the Olympics is not so good.

These all facts indicate that Asian Games performance directly influences Olympic Games performance for any country. So, a country must give their best in Asian Games, to increase their chances of winning in the Olympics as well. Winning in the Olympics is the ultimate aim for any country, but Asian countries have the advantage to first do their best in Asian games, which will prepare them for the Olympics.

## 7 Discussion and Future Work

Our work has tried to cover as many aspects as possible for the comparative analysis between Asian Games and Olympic Games. We found lots of helpful information that might be highly beneficial

for the Sports authorities of various nations for better planning for the upcoming events. Athletes, coaches, organizing committee of these games will be highly benefitted from our analysis. This work can further be improved a lot. We plan to carry this project forward and work on it to make it better for the players, coaches, and authorities.

Our future plan includes obtaining the height and weights of players too and trying to gather some insights from that data. We will even perform some predictive analytics to predict the exact number of medals that a country might win based on its past performance in the Asian Games and Olympic Games. We wish to expand this analysis even further on Winter Olympics and Winter Asian games. We plan to make some improvements to our user interface as well to make it more compact and user-friendly. With all these improvements planned, we wish to sincerely look forward to improving this work to its full extent.

## 8 Individual Contributions

### Deepak Raj (21111024)

- Web scraping of asian games table from wikipedia (1998, 1982, 1978, 1974, 1966, 1962, 1958)
- Data integration and file management
- Jupyter Files
  - Q13 - Sport percentage analysis of countries
  - Q14 - Finding Recession and Booming period of countries
  - Q15 - Finding best year and best sport of countries
- Streamlit work
- Project Report

### Dinkar Tewari (21111025)

- Web scraping of asian games table from wikipedia (1998, 1982, 1978, 1974, 1966, 1962, 1958)
- Jupyter Files
  - Q7 - Male vs Female analysis
  - Q8 - Individual vs Team analysis
- Streamlit work
- Project Report
- Presentation

### Divyansh Bisht (21111027)

- Web scraping of asian games table from wikipedia (1951 - 1954)
- Gathering files of Olympic database
- Jupyter Files
  - Q4 - Host vs Non host country analysis
  - Q5 - Finding top athletes of countries
  - Q6 - Finding top athletes of counties sport wise
- Streamlit work
- Project Report

### Rohit Kushwah (21111053)

- Web scraping of asian games table from wikipedia (2002 - 2018)
- Jupyter Files

- Q9 - Asian and its immediate Olympic analysis
- Q10 - Medal-winning/ not winning analysis
- Q11 - Detailed India analysis
- Q12 - Finding valuable players(who won medals in both asian games and olympics)
- Streamlit work
- Presentation

### **Vikas (21111067)**

- Web scraping of asian games table from wikipedia (1994, 1990, 1986, 1970)
- Jupyter Files
  - Q1 - Medal tally analysis of countries
  - Q2 - Medal Percentage analysis of countries
  - Q3 - Sport-wise medal analysis
- Streamlit work
- Presentation
- Readme and other files

Every member of our group has contributed equally to this project.

## **9 Web app link**

We created a web app using streamlit to generate better and visual insights to the user. The link of our web app is given below:

<https://share.streamlit.io/divyansh009/dmprojectgrp10/main/main.py>

## **References**

We used two different websites to get the datasets

- [https://en.wikipedia.org/wiki/Asian\\_Games](https://en.wikipedia.org/wiki/Asian_Games)
- <https://www.kaggle.com/heesoo37/120-years-of-olympic-history-athletes-and-results>