Olympic Games Analysis

By SportStats

Content

Review of Questions to Answer/Hypotheses/Approach

Discuss Technical Challenges

Detail: Entity Relationship Diagram(ERD)

Initial Findings

Deeper Analysis

Hypotheses Results

Section 1:Questions to answer

- The number of athletes, events and nations

    How does this change over time?

    How does the number of women and men athletes change over time?

    Who participated the most Olympic games?Who won the most Medals?

- Height and Weight

    What is the average height & weight of athletes?How does this change over time?

    Do medal winners tend to have an advantage in this?

- Medals

    Which region won the most medals in total?

    In the U.S., how many medals were won by men and women? How does this change over time?

Section 2:Initial Hypotheses

- There are more athletes, countries and events overall;

  There are relatively fewer women athletes

- The average height&weigh will vary in different sports field,and medal winners may have, a stronger advantage in these qualities,for example,in tug-and-war sports,medal winners should be relatively shorter and heavier compared with the average.

- I believe it to be the U.S. or Russia.

Section 3:Data Analysis Approach

- Average metrics for height&weight

- Frequency metrics for medal counts.

Technical challenges

Pandasql is using sqlite, some functions cannot be executed,but we can use other functions to replace that;

There is plenty of missing data in Height&Weight at first glance, we decide to find the patterns in this missing data and try to minimize the impact on our analysis.
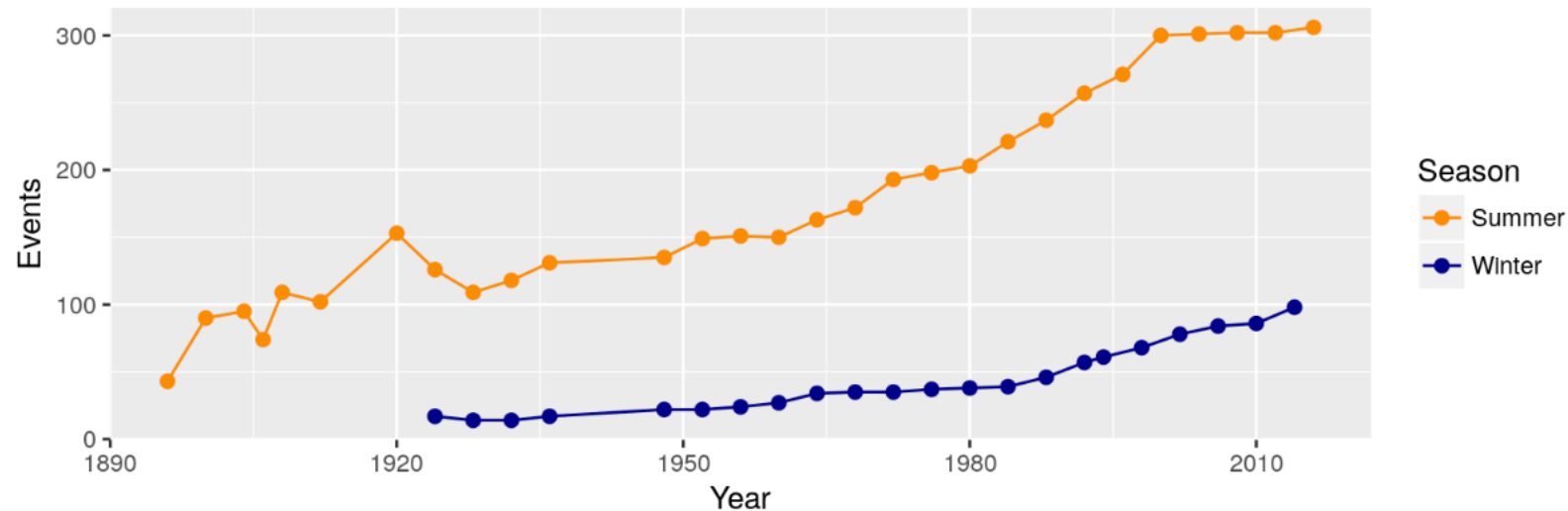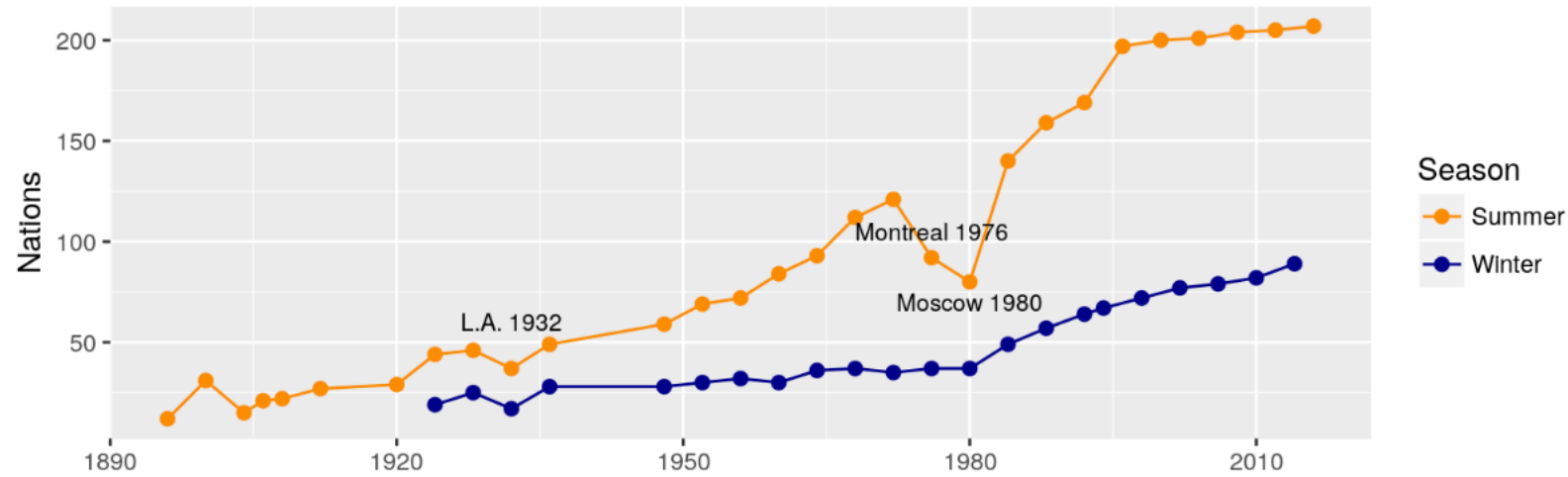
Initial Findings
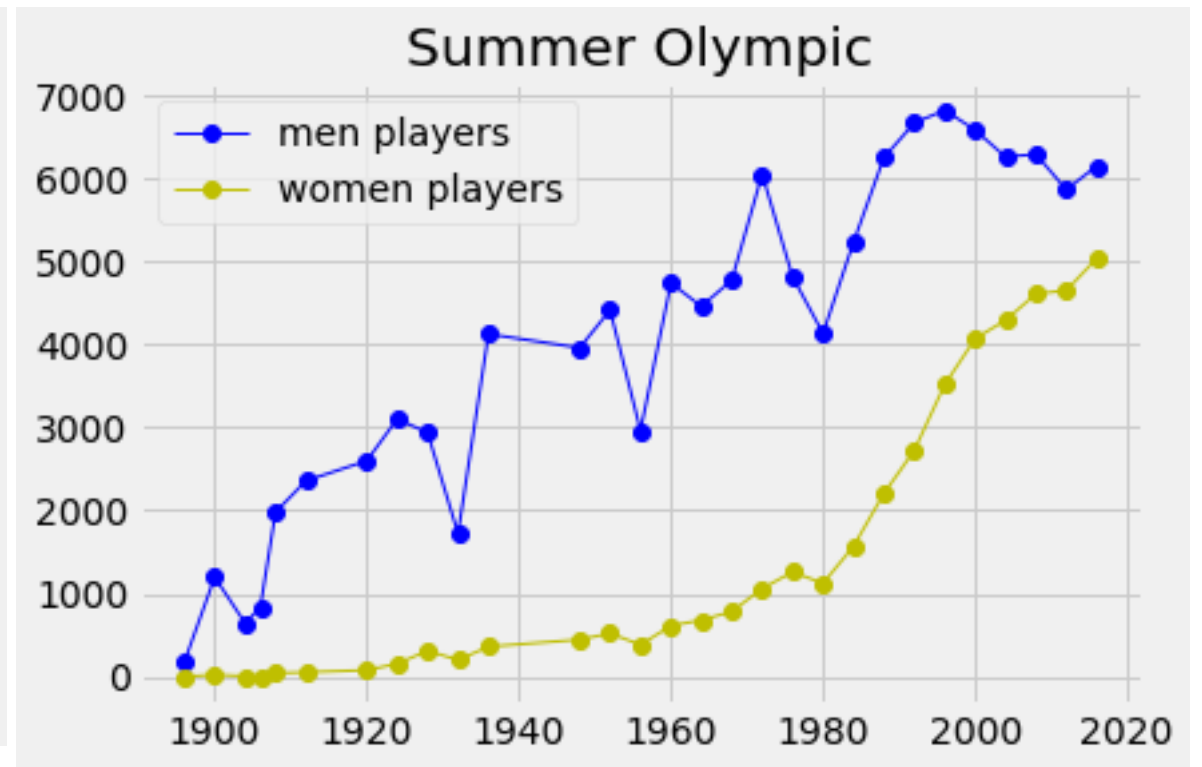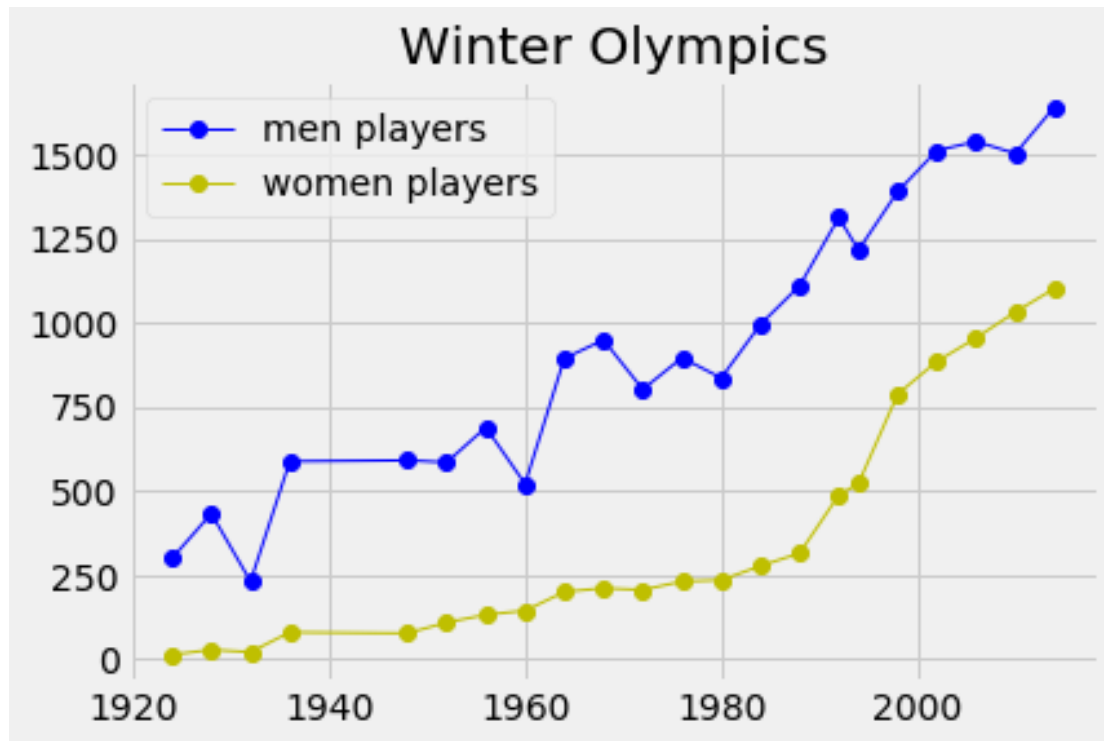
We can see that there are less atheletes in Winter Olympics than Summer Olympics,while both men and women athletes are growing over the years, in general, despite some significant drop in Summer Olmpics in L.A. 1936, Melbourne 1956 and Moscow,1980

Initial Findings (1)
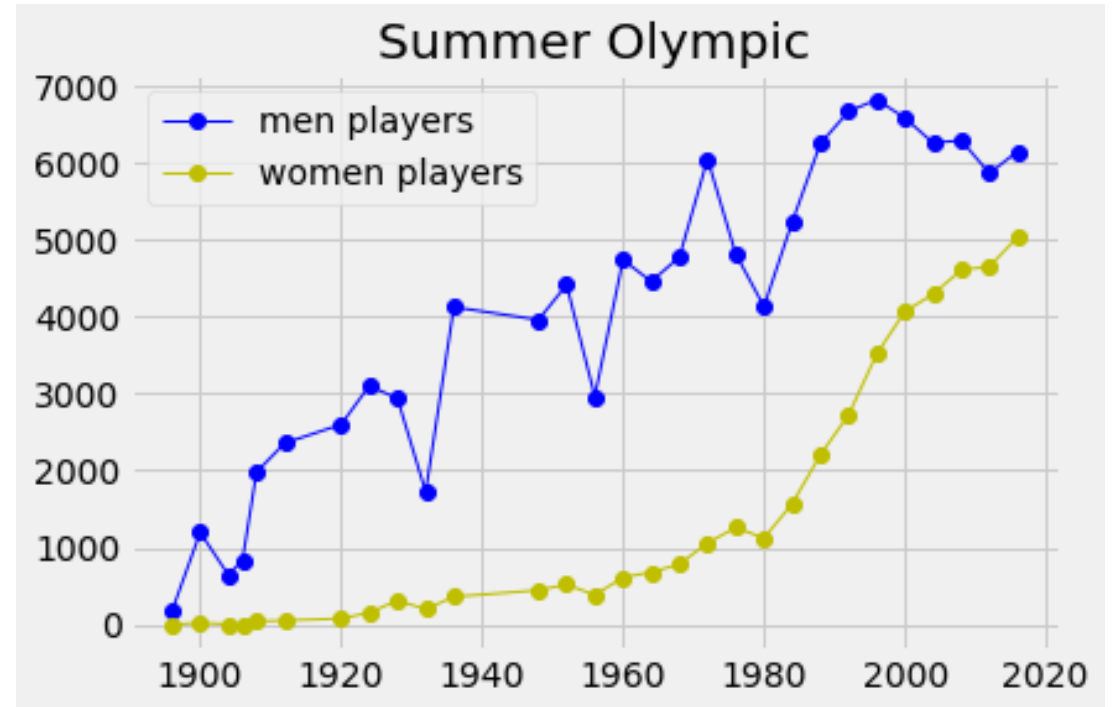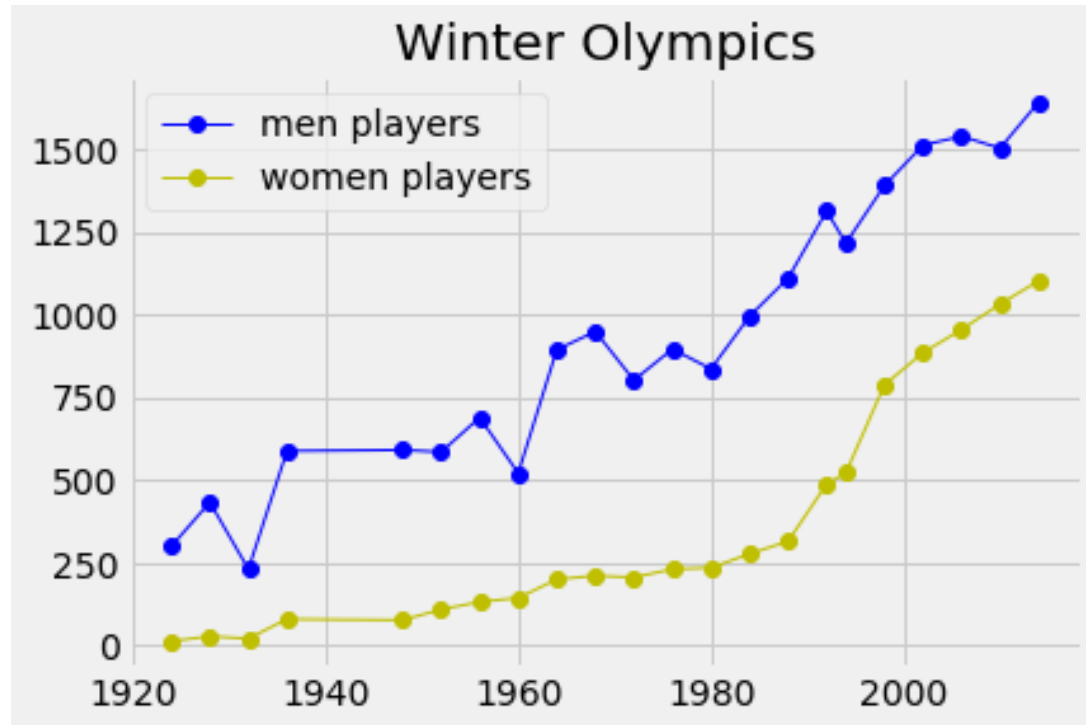
The growth levels off around the year 2000, at least for the Summer Games. The list of events and athletes cannot grow indefinitely, and the Summer Games may have reached a saturation point near the turn of the century, with around 300 events and 10,000 athletes. The Winter Games would seem to have more growing room, but ice and snow sports are not practical or popular in most nations, and that doesn't seem likely to change soon.

Initial Findings(2)

The average height and weight centers around 65-85kg,170-185cm,And in most sports fields medal winners tend to be taller and heavier than those who didn't.
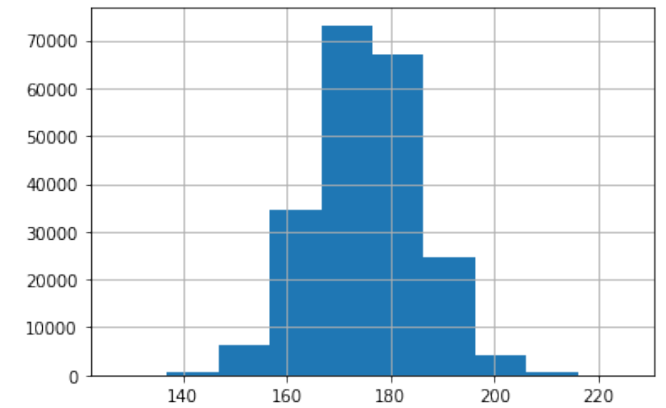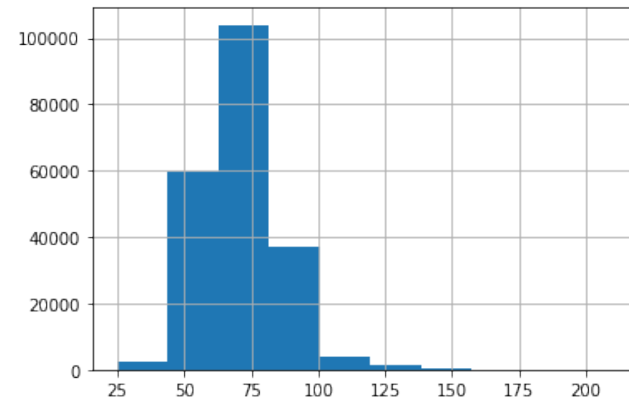

	Sport	winner_counts	loser_counts	Sex	loser_height_avg	winner_height_avg	loser_weight_avg	winner_weight_avg
0	Gymnastics	15130.0	191088	M	168.0	166.0	64.0	62.0
1	Gymnastics	6231.0	66442	F	157.0	157.0	49.0	48.0
2	Athletics	8965.0	59242	M	180.0	181.0	74.0	76.0
3	Swimming	9380.0	49435	M	185.0	188.0	78.0	83.0
4	Shooting	7041.0	41249	M	176.0	176.0	78.0	80.0
5	Swimming	7392.0	37801	F	172.0	175.0	62.0	65.0
6	Cross Country Skiing	3395.0	33970	M	178.0	180.0	72.0	74.0
7	Fencing	7011.0	32352	M	179.0	181.0	76.0	77.0
8	Athletics	4647.0	26051	F	169.0	171.0	60.0	62.0
9	Alpine Skiing	1506.0	25824	M	178.0	180.0	80.0	83.0

Initial Findings (3)

- Who won the most medals?

Michael Fred Phelps, II ,swimming athlete from USA,won 8 medals out of events in 2004 Olympics,four years later in Beijing Olympics,he broke the records and won 8 medals out of 8 events.By the time he retired,he already won 22 Olympic gold medals.
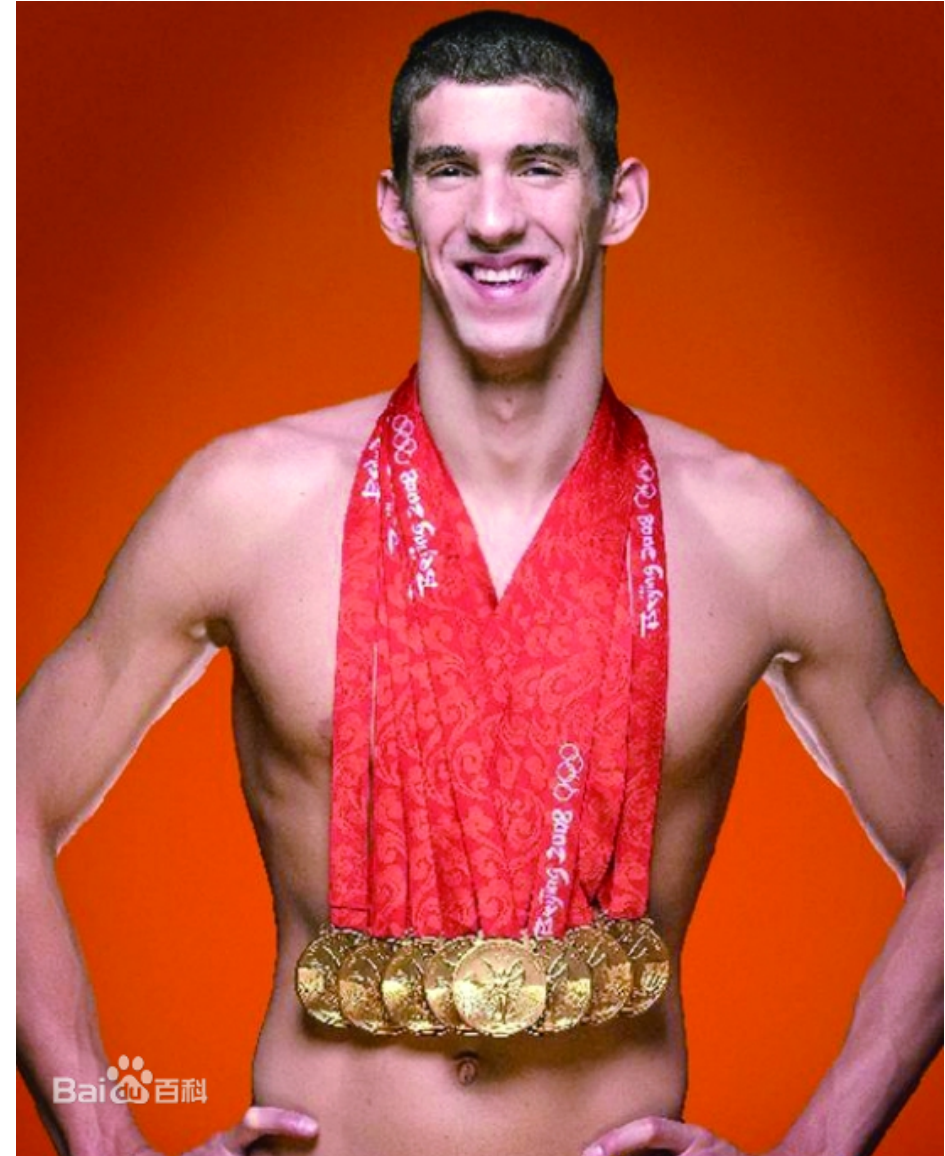
Who attended the most events?

Robert Tait McKenzie from Canada attended 44 events in Art Competition(When it 's still a thing) in 1932,and he won a bronze medal in Art Competitions Mixed Sculpturing, Medals And Reliefs；

Since Art Competition is not an Olympic anymore, the athlete in existing sport fields who attended most Olympic Games is Heikki Ilmari Savolanien from Finland ,who won 8 medals out of 39 events he attended in 5 Olympics between 1928 and 1952.

Initial Findings(4)

Germany has sent most athletes while USA has won most medals.Russia gained the highest medal-winning ratio at 0.34,which means ,roughly, among every three events one athlete attended ,one won a medal.We will look into the US for deeper analysis considering the timescale.

	region	medal_counts	player_counts	ratio
0	Russia	3947	11692	0.337581
1	USA	5637	18853	0.298998
2	Germany	3756	15883	0.236479
3	Pakistan	121	562	0.215302
4	Norway	1033	4960	0.208266

Deeper Analysis (1)
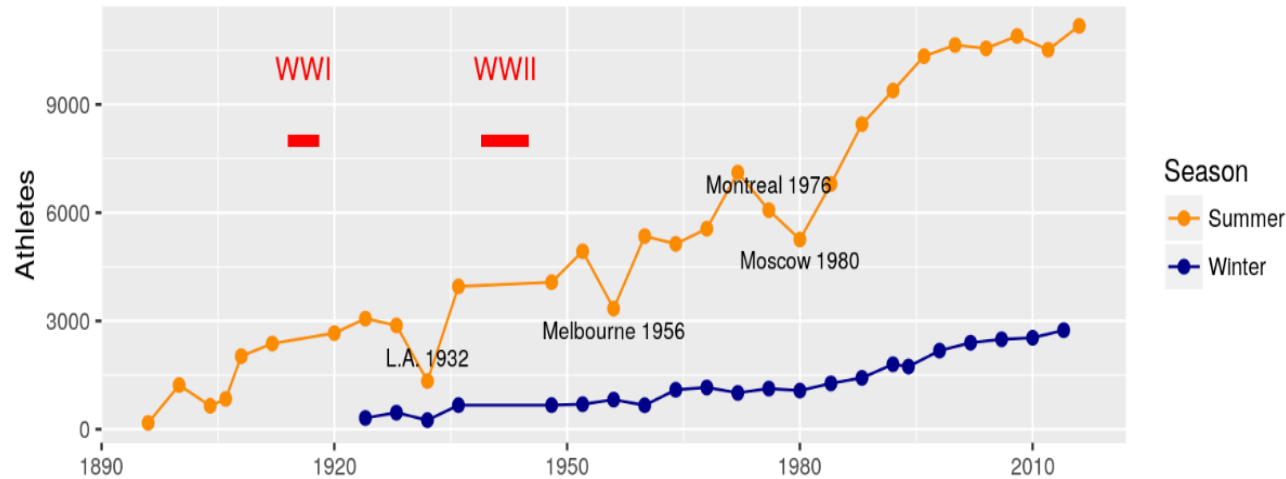
L.A., 1932: Attendance dipped because these Games occurred in the midst of the Great Depression and in a remote location, such that many athletes were <u>unable</u> to afford the trip to the Olympics.

Melbourne, 1956: Attendance dipped due to several boycotts: Iraq, Egypt, and Lebanon did not participate due to the involvement of France and Britain in the <u>Suez Crisis</u>; the Netherlands, Spain, Switzerland, and Cambodia did not participate due to the Soviet Union's beat down of the <u>Hungarian Revolution</u>; and China did not participate in protest of the IOC's <u>recognition</u> of Taiwan.

Montreal, 1976: Attendance dipped because <u>25 nations</u>, mostly African, boycotted the Games in response to apartheid policies in South Africa. Attendance at the 1980 Winter Olympics in Lake Placid wasn't affected much since African nations have a limited presence at the Winter Games.

Moscow, 1980: Attendance dipped because <u>66 nations</u>, including the U.S., boycotted the Games in response to the Soviet invasion of Afghanistan.

Deeper Analysis (2)

We can see there is a stunning increase in data completeness after 1960, so I decide to use data after 1960 for analysis.


Height&Weight Data Completeness porportion

Average Height — Average Weight
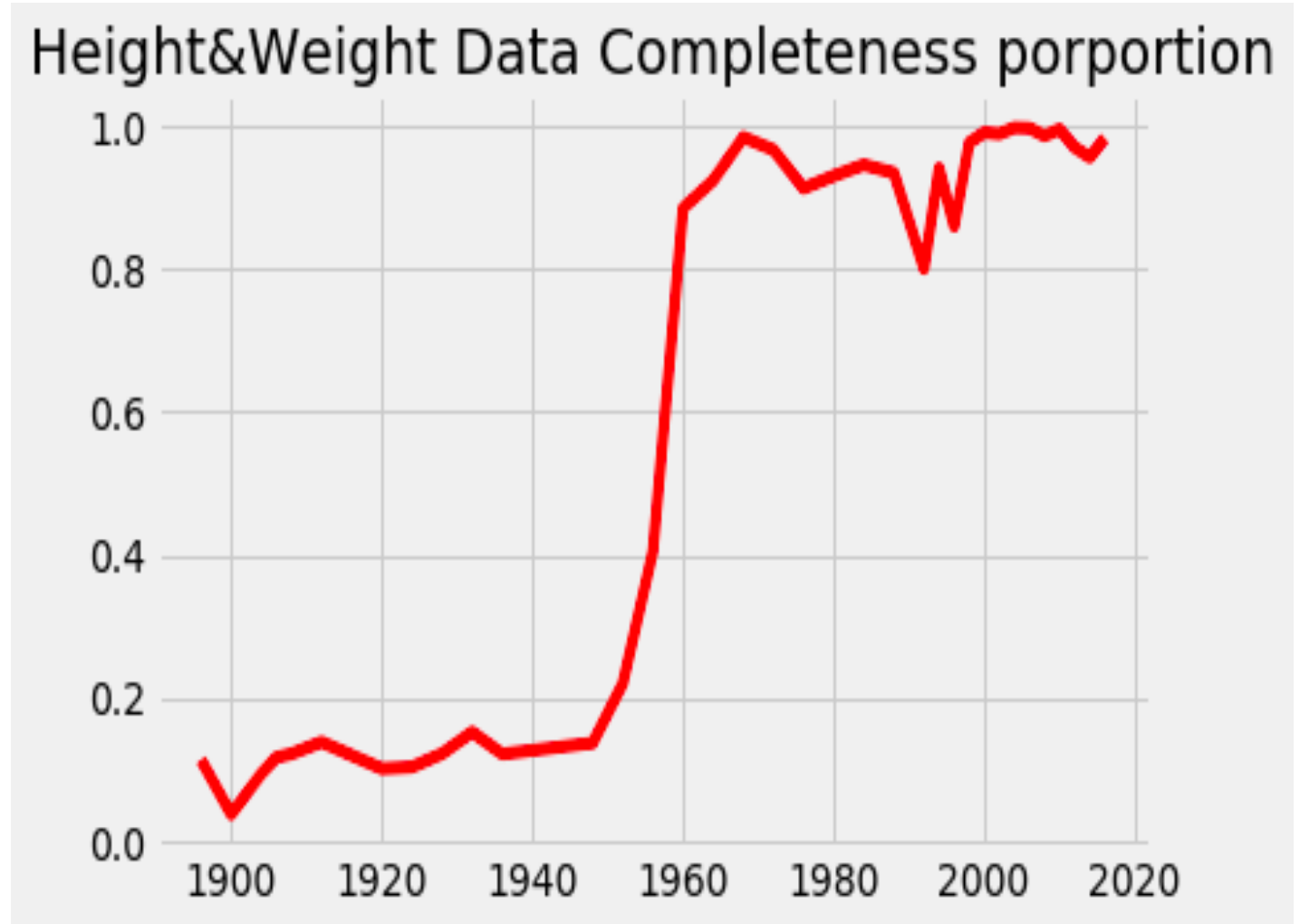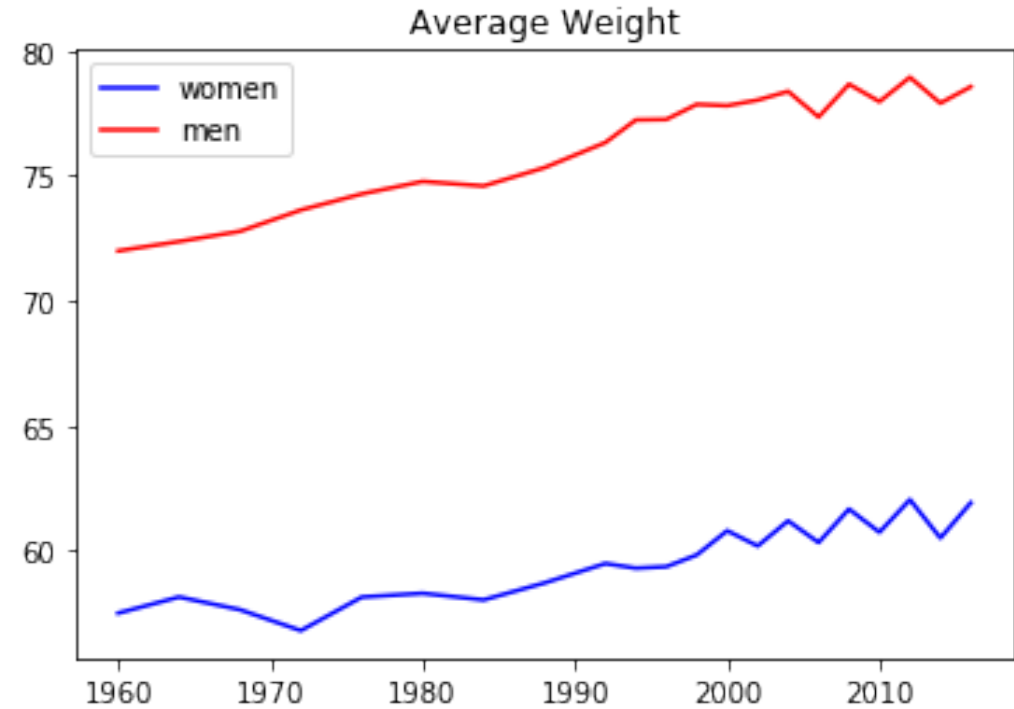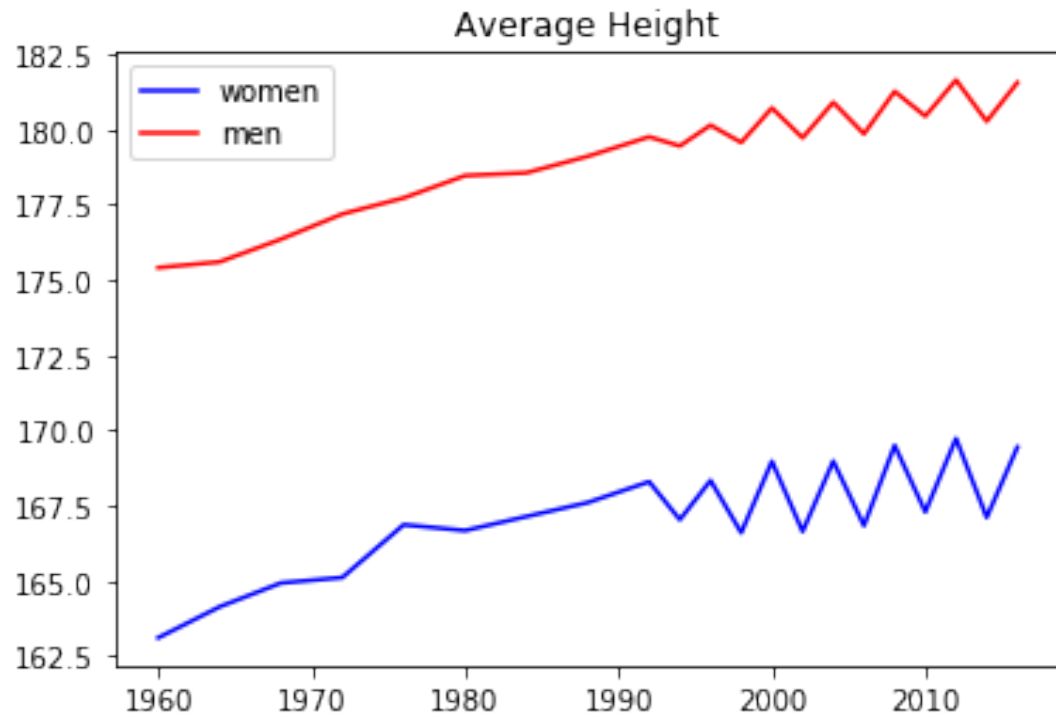
The motto for the Olympics is "Citius, Altius, Fortius", which means "Faster, Higher, Stronger" in Latin. Indeed, as illustrated by the long history of record-breaking performances at the Olympics, athletes at every Olympics seem to be taller and stronger than the one before.

Deeper Analysis (3)

This is interesting.Men athletes gained a huge advantage over women between 1900 and 1980,but the difference starts to narrow down from 1980 and becomes trivial,in recent Olympic Games, we can see medals won by women outnumbered men. Considering men athletes still outnumbered women athletes by thousand, we can say American women are doing a better and better job.

In general, USA are gaining more and more medals, which is largely resulted by the increasing number of medals gained by women
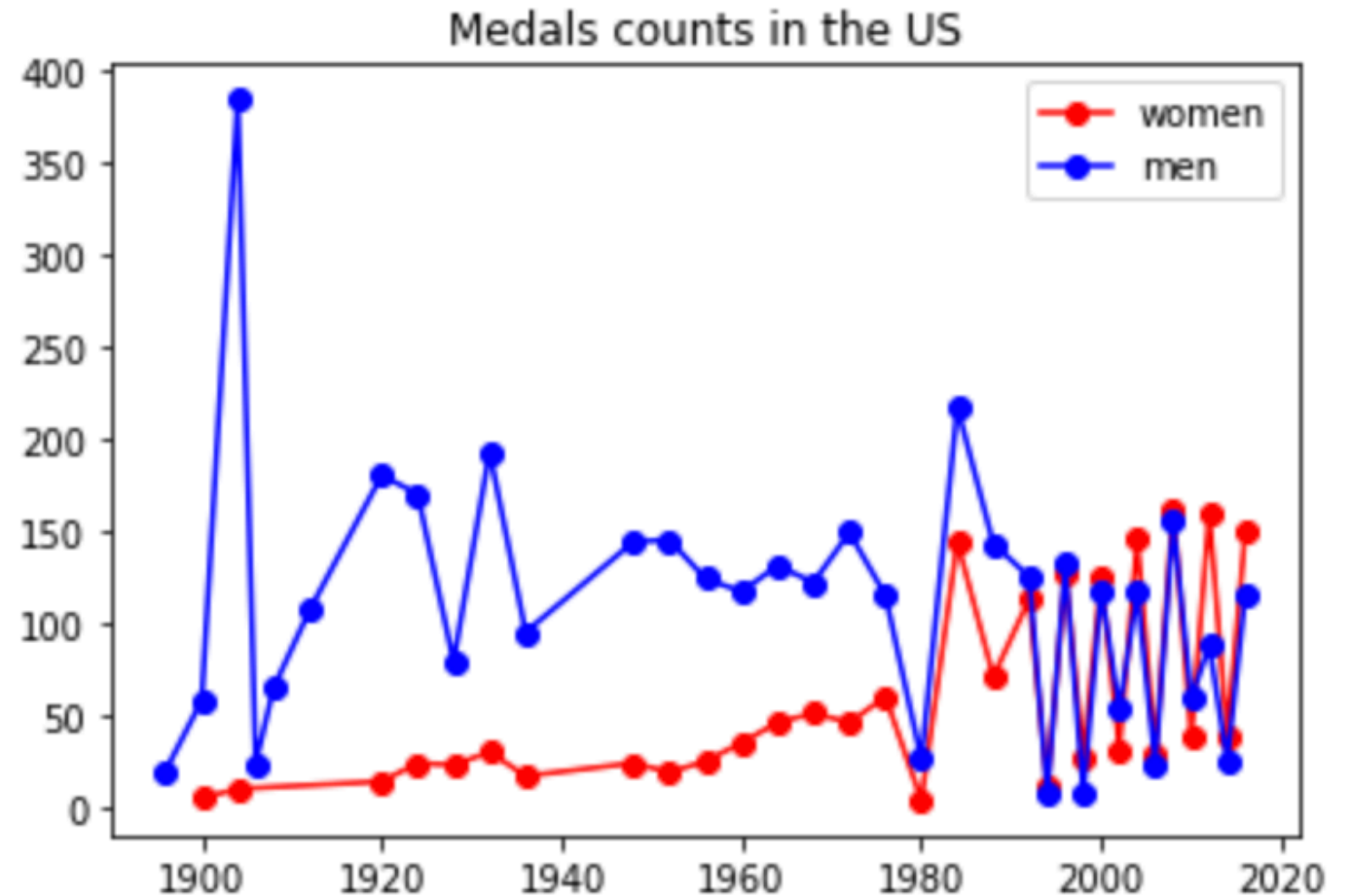
Hypotheses Results

- There are more and more athletes, countries and events;Women athletes are less than men athletes but the gap is narrowing down.

- Athletes are becoming heavier and taller in general,medal winners tend to be stronger too.

- Michael Fred Phelps, II has won the most medals while Heikki Ilmari Savolanien has attended the most Olmypic events in terms of existing Olympic sports fields.

- USA gained the most medals,while women started to gain more medals than men since 2008 Summer Olympic.

Recommendations for news report

- Report about the Legend Phelps
- Women in Olympics

Thank you!