
ANALYSIS OF WOMEN PARTICIPATION IN OLYMPICS



Jose Moti

ABOUT OUR DATASET

- Link for the dataset → ([Link](#))
- Contains 120 years of Olympics participant data.
- Features → Name, age, height, weight, sex, country, medal won, event, participated year..
- Total 271k Olympic participants.
- 9k nulls in age, 60k in height, 62k in weights, no other NULL values
- Nulls in medal shows that the participant didn't win a medal.
- Consists of last 29 summer Olympics and last 22 winter olympics.

HYPOTHESES

- Hypothesis 1

The women participation in Olympics increased over years.

- Hypothesis 2

The women participation increases in both Winter and Summer Olympics with year. The women participation is highly correlated to each other.

- Hypothesis 3

The women participation differs a lot according to the countries.

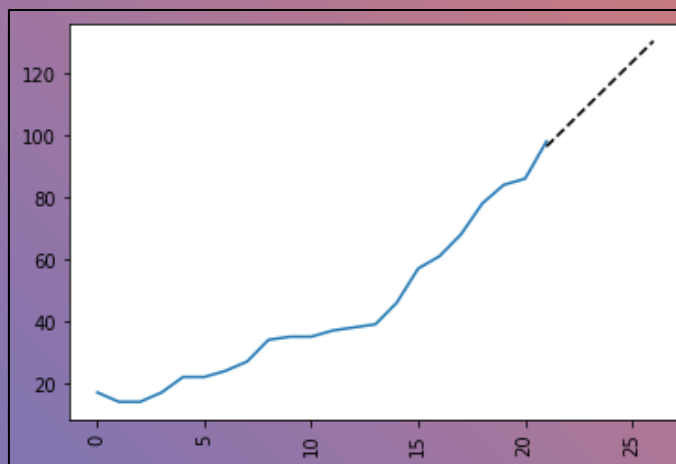
ALL HYPOTHESE CAME OUT TO BE TRUE AFTER ANALYSIS

HYPOTHESIS 1 TEST

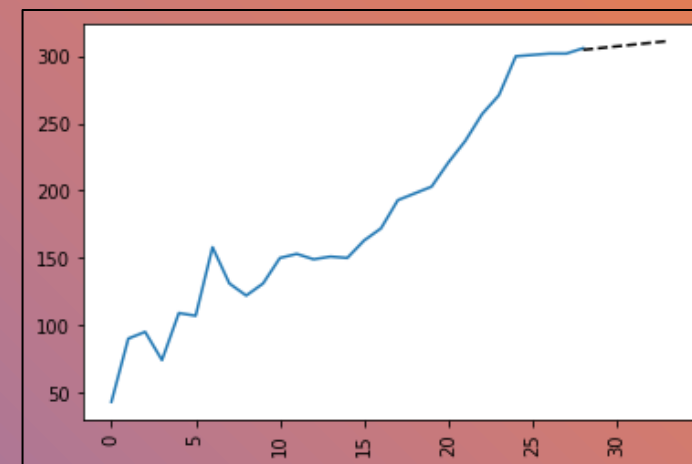
- Women participation increased from 0% in 1896 summer Olympics to 45.26% in 2016 summer Olympics.
- Similarly, the percentage of women in winter Olympics increased from 3.6% to 41.3%.

Extrapolated to predict the number of women events in the upcoming years.

Number of events for summer Olympics seems to be saturated while there is still potential to improve for winter Olympics.



Winter Olympics

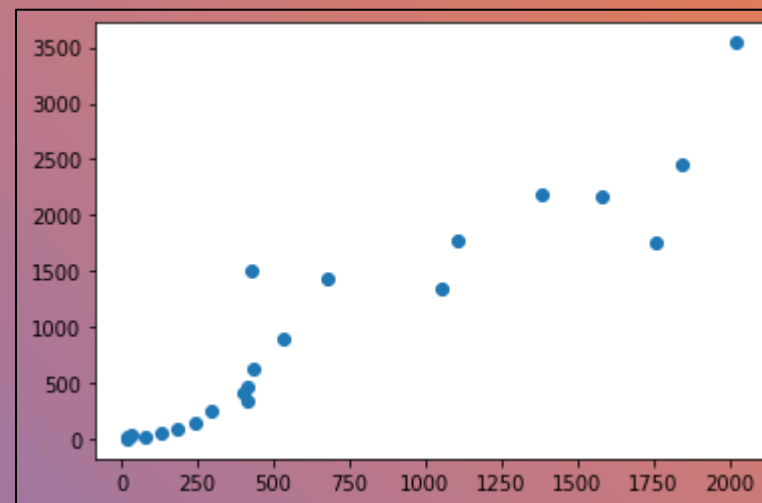


Summer Olympics

HYPOTHESIS 2 TEST

- The scatter plot between number of woman participants in winter and summer Olympics is plotted in the order of their occurrences
- Clear positive association from scatter plot. Increase from 0 to 2000 for winter Olympics while 0 to 3500 increase for summer Olympics

Pearson correlation coefficient $\rightarrow 0.94$



HYPOTHESIS 3 TEST

- Took 2016 Summer Olympics data for this analysis. Grouped the women participants based on country.
- Took the countries which had at least 20 participants.

Countries with high women percentage → Cameroon 17/22 (0.77), Senegal 16/6 (0.72), Zimbabwe 22/31 (0.70), Angola 18/26 (0.69)

Countries with low women percentage → Iraq 0/20 (0.00), Honduras 1/22 (0.04), Qatar 2/40 (0.04), Iran 11/56 (0.16), Algeria 13/61 (0.17)

Countries with high number of participants and perfect ratio → USA 360/719 (0.51), Brazil 259/583 (0.44), Germany 246/536 (0.45).

HYPOTHESIS 3 TEST (CONTD)

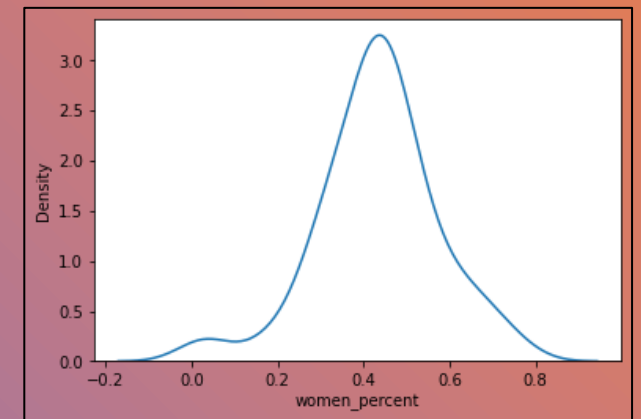
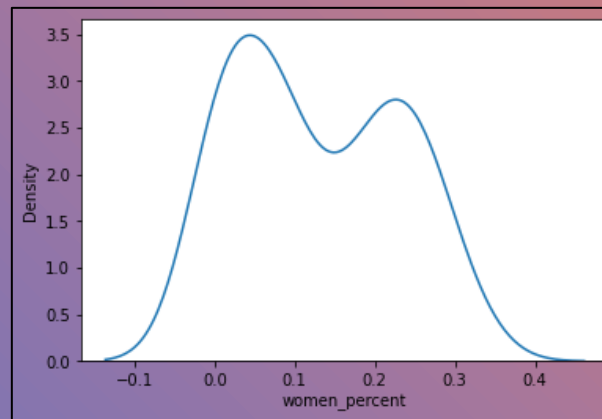
- We plotted the distribution of number of woman participants by country for both 1964 and 2016 summer Olympics.
- We used 2 metrics in order to quantify the spread of distribution (mean and standard deviation). High standard deviation so correct hypothesis

1964

Mean \rightarrow 0.12, STD \rightarrow 0.10

2016

Mean \rightarrow 0.43, STD \rightarrow 0.16



CONCLUSION

- Even though there has been a general increase in women participants some countries still fall behind in ensuring this.
- Does this directly correlate with gender equality in these countries? Maybe or maybe not.

Wut to do?

- Nothing much.. If any of you are citizens of these countries try to think of why women are less represented from your country in the Olympics. (Political reasons?.. Social reasons?.. Religious reasons?..)
- Try to organize more events in these less represented countries and motivate more towards these games.



**WE CANNOT SUCCEED WHEN HALF OF US ARE
HELD BACK**

MALALA YOUSAFZAI