



جامعة مصر للمعلوماتية  
EGYPT UNIVERSITY  
OF INFORMATICS

Egypt University of Informatics  
Computer and Information Systems  
Data Analysis Course

# Analysis of Age Range 22 to 30 is a Key Factor in Medal-Winning Success.

Submitted by:

- Ibrahim Ahmed[21-101067]
- Samy Mohamed[21-101076]
- Omar Alaa[21-101030]
- Ismail Taher[21-101013]

Under the supervision of:

- Dr Mohamed Taher

## Introduction

Age is often considered a significant factor in athletic performance, with certain age ranges frequently associated with peak physical and mental capabilities. In the context of competitive sports, the age of athletes can influence their endurance, strength, and experience, all of which contribute to their success in international events such as the Olympic Games. This study focuses on Egyptian athletes' performance at the Tokyo 2020 and Paris 2024 Olympic Games, specifically examining whether those within the age range of 22-30 years have a higher medal-winning rate compared to their younger or older counterparts. By analysing this relationship, we aim to determine if this specific age group truly represents the prime of an athlete's career. This will inform future athlete development and training programs.

## Research Question

Do Egyptian athletes aged 22-30 have a higher medal-winning rate compared to those who are younger or older at the Olympic Games?

## Hypothesis

### **Null Hypothesis ( $H_0$ ):**

There is no significant difference between Egyptian athletes aged 22-30 and those who are either younger or older in medal-winning rate.

### **Alternative Hypothesis ( $H_1$ ):**

There is a significant difference between Egyptian athletes aged 22-30 and those who are either younger or older in medal-winning rate.

## Population of Interest:

Egyptian athletes who won medals in Paris 2024 and Tokyo 2020 Olympics.

## Sampling Method:

We selected Egyptian athletes who participated in the most recent Olympic Games, specifically the Tokyo 2020 and Paris 2024 Olympics. The sample includes all Egyptian athletes who won medals across various sports. These athletes are categorised into two age groups: 22-30 and Other (younger than 22 and older than 30). This categorization allows us to compare the medal-winning rates across different age ranges, focusing on whether the 22-30 age group shows a higher success rate compared to other age groups.

## Bias Identification:

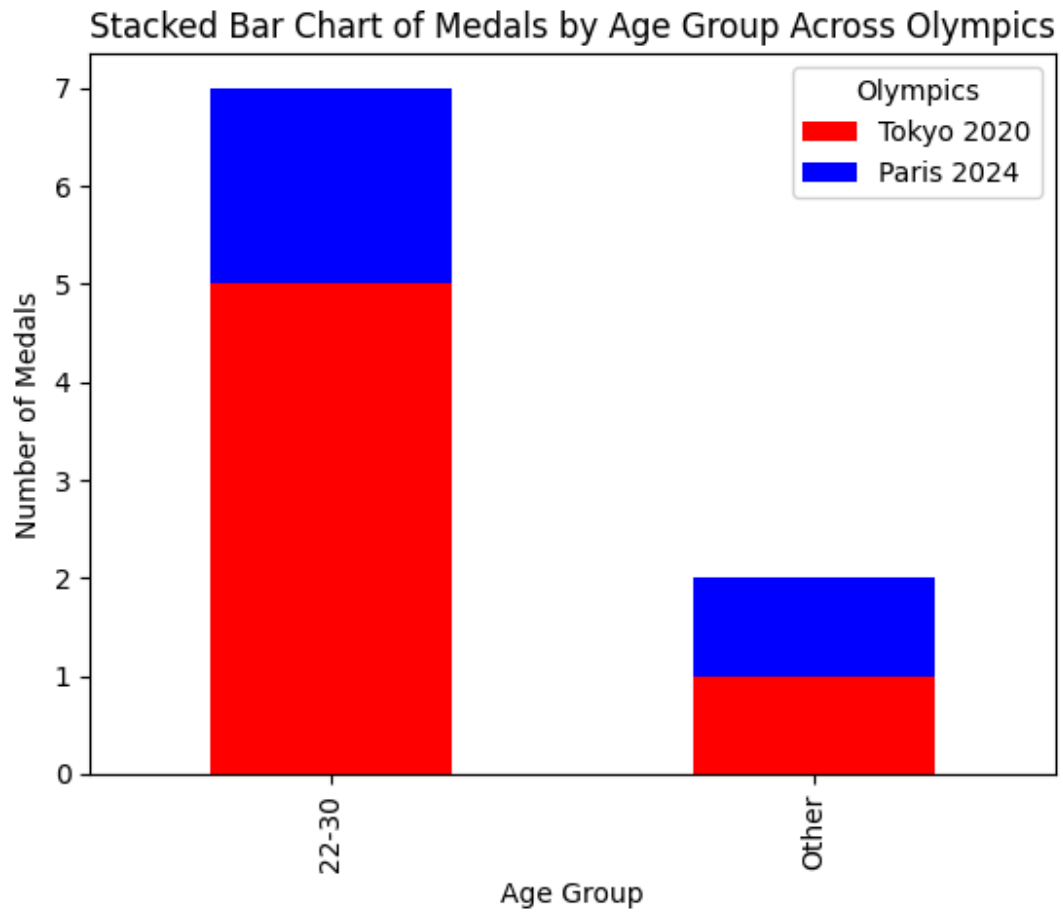
The initial focus was on analysing only those Egyptian athletes who won medals in past Olympics (Tokyo 2020 and Paris 2024). This will introduce bias as it excludes athletes who did not win medals in any of two olympics. So, to address this, the analysis will now include all Egyptian athletes who won medals in all Olympic Games.

## Survey Questions/Collected Data/Dataset:

The analysis utilises two datasets that represent the performance of Egyptian athletes at the Tokyo 2020 and Paris 2024 Olympic Games. Each dataset contains detailed information about the athletes, including their demographics and the medals they won. The datasets are crucial for understanding how age impacts athletic success, specifically focusing on the medal-winning rates of athletes within the 22-30 age range compared to those who are younger or older.

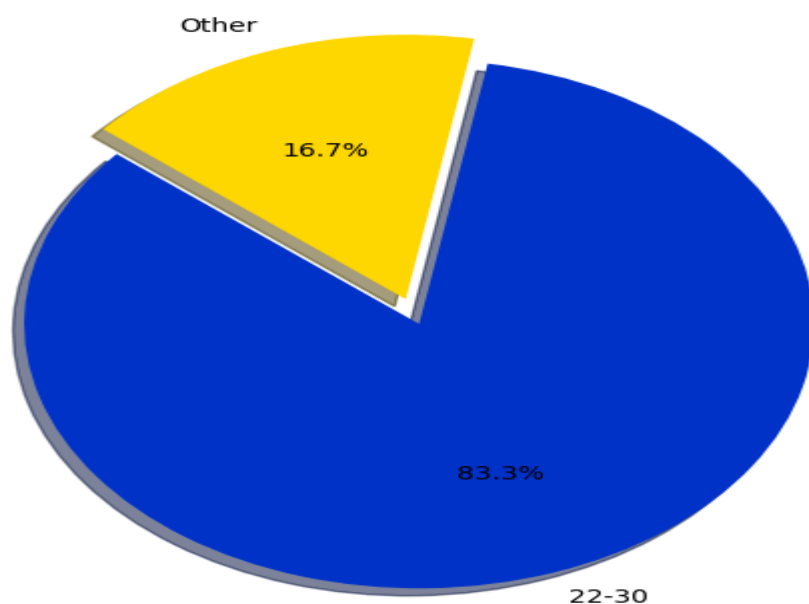
The datasets are filtered to include only Egyptian athletes, as the focus of the analysis is on this specific group. The athletes are then categorised into two groups based on their age: those within the 22-30 age range and Other (younger than 22 and older than 30).

## Analysis:

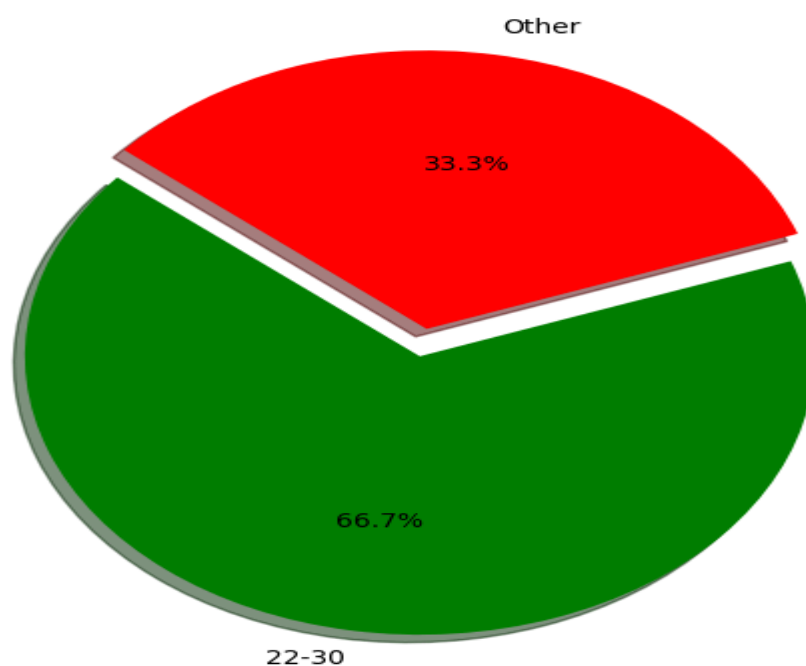


In this Stacked Bar chart we compare the number of medals won by two groups: 22 to 30 and Other (age less than 22 or greater than 30) in Tokyo 2020 and Paris 2024 Olympics. x-axis represents age groups, y-axis represents number of medals won by each group, while colours are for olympic events where the red colour is for Tokyo 2020 while blue is for the Paris 2024 olympics. The chart shows that the age group 22-30 has a higher medal count compared to the Other age group in both Tokyo 2020 and Paris 2024.

Medal Distribution by Age Group - Tokyo 2020

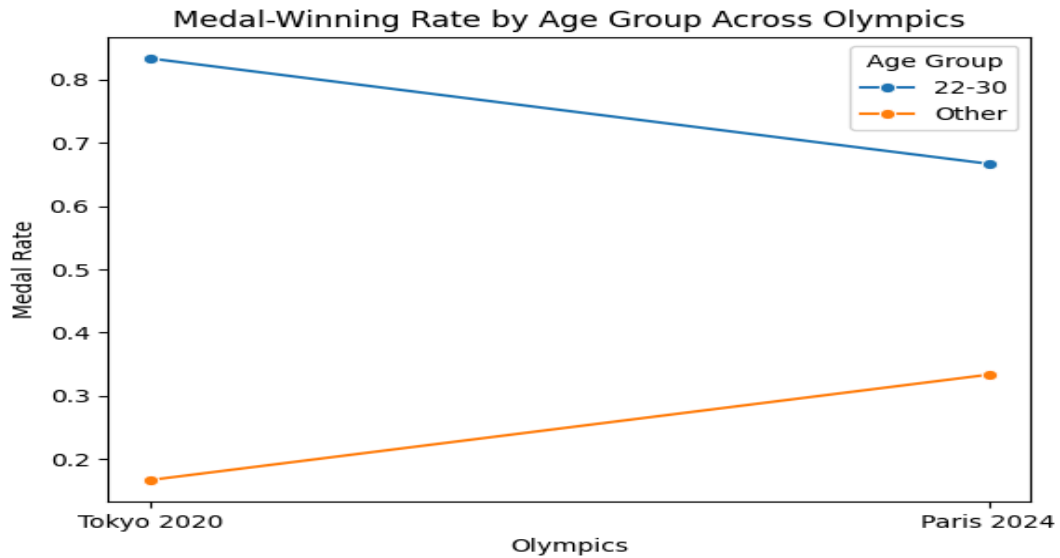


Medal Distribution by Age Group - Paris 2024

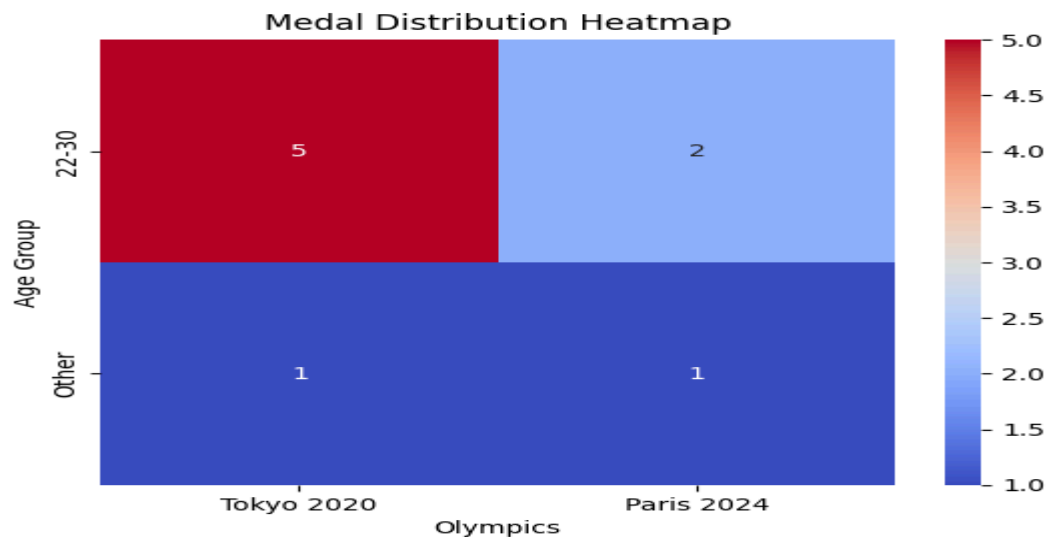


---

The pie chart shows the proportion of medals won by Egyptian athletes, divided into two age groups: **22-30** and **other**. The pie chart is divided into two slices, with each slice representing the proportion of medals won by each age group. The **green slice** corresponds to the "22-30" age group while The **red slice** corresponds to the "Other" age group. The chart displays the percentage of total medals won by each age group



In this line chart, we compare the medal-winning rate between two age groups: 22 to 30 and Other (ages less than 22 or greater than 30) in the Tokyo 2020 and Paris 2024 Olympics. x-axis represents the two olympic events, while the y-axis shows the medal-winning rate for each group. We have two lines, one is blue and the other is orange, the blue line represents the age group from 22 to 30 while the orange line represents the age group Other. The chart shows that the age group 22-30 has a higher medal-winning rate in both Olympic events, although their rate decreases slightly from Tokyo 2020 to Paris 2024.



In this heatmap, we compare the distribution of medals won by two groups, 22 to 30 and Other (age less than 22 or greater than 30), in the Tokyo 2020 and Paris 2024 Olympics. The x-axis represents the olympic events, while the y-axis represents the age groups. The colour intensity indicates the number of medals won, with becoming darker giving more medals. The heatmap shows that the age group 22-30 won more medals in both Olympics compared to the age group Other.

## Hypothesis Testing Steps

- Step 1: Define null and alternative hypothesis
  - We mentioned above
- Step 2: Choose the appropriate test
  - T-test
- Step 3: Calculate the p-value
  - p-value = 0.053
- Step 4: Determine the statistical significance
  - Since p-value is greater than alpha (0.05). Therefore we accept the null hypothesis and there is no significant difference between Egyptian athletes aged 22-30 and those who are either younger or older in medal-winning rate.

## Conclusion

Our analysis proved that there is no significant difference between Egyptian athletes aged 22-30 and those who are either younger or older in medal-winning rate. Where Egyptian athletes within age 22-30 have a higher medal-winning rate compared to those who are younger or older.

## Any potential issues

1. For the Paris 2024 dataset, I chose the medals dataset because it contains the data I need to analyse Egyptian athletes medal wins. However, the problem was that it didn't have an age column. To address this, I used the athletes dataset, but it was also missing the age column. But, it includes each athlete's birthday. Using Python's datetime library, I calculated each athlete's age based on the start date of the Olympics. After calculating the athletes' ages, I added this age column to the medals dataset by performing a join operation based on the code column, which is present in both datasets.
2. For the Tokyo 2020 dataset, I selected the results dataset because it is very similar to the medals dataset for Paris 2024. The issue with the results dataset was that it had empty rows in the Medals column for athletes who didn't win any medals. To solve this, I removed these rows. After cleaning, the dataset became comparable to the Paris 2024 medals dataset.