Statistical Analysis of Olympics Dataset

Now we have a concrete understanding of the dataset. We now know what our dataset looks like. Now we will move toward the statistical analysis of this dataset. We will answer these questions:

1. We have an Age column that contains some null values. Rather than deleting them, fill these missing values by mean after grouping them by games.

2. What is the probability distribution "of the age" of all the players who participated in the swimming competition between the years 2000 - 2016.

3. Is the overall distribution of Gold Medals tally in India, positive or negative skewed?

4. What is the standard deviation/error of the distribution of gold medals when the UK is hosting the Olympic games and when not hosting the games? (When games are hosted in the UK, the city is London.)

5. Standardise the age of all the participants participating in the Shooting Competition and calculate this standardised data's mean and standard deviation.

6. Is age positively correlated or negatively correlated with winning the medal?

7. From the dataset, Calculate the IQR for the Age column.