

You can download the data-set from this link:

<https://tinyurl.com/olympicsdata> (from kaggle)

Step 0: Load the csv file using pandas and display the number of records in it. **(30 points)**

Step 1: Modify the data-set by changing NA values in Medal column to string None -- which will mean no medals. (hint: you may use fillna() method). Output how many records in the dataset has a medal associated with it. **(30 points)**

Step 2: Report the top 3 sports in olympics that has the tallest (on average) athletes. Do the same for shortest, heaviest and lightest athletes. **(40 points)**

Step 3: For winter olympics only, report the team with most Gold medals year by year. **(40 points)**

Step 4: For each sport, report which team dominates it -- team with most medals. **(40 points)**

Step 5: For summer olympics only, create a graph displaying the participation of female athletes to olympics year by year. X axis of the graph is year, and the y axis represent the ratio of female athletes to all athletes participated in that year. **(30 points)**

Step 6: Create a graph displaying average athlete age year by year. **(30 points)**

Step 7: We are a small country looking to start participating in Olympics. Based on the data of 2016 olympics only, recommend top 3 sports that is least competitive so we can focus on those. A sport is more competitive if the ratio of athletes participated in that sport is much larger than those that got medals. **(10 points)**

SAMPLE OUTPUT

Step 0

=====

There are 271116 records in the data-set.

Step 1

=====

There are 39783 records in the data-set with medals.

Step 2

=====

Tallest:

Basketball 190.869879
Volleyball 186.994822
Beach Volleyball 186.144954

Shortest:

Gymnastics 162.936020

continues...

Step 3

=====

1924 Great Britain

1928 Canada

1932 Canada

continues...

Step 4

=====

Aeronautics:: Switzerland with 1 medals.

Alpine Skiing:: Austria with 114 medals.

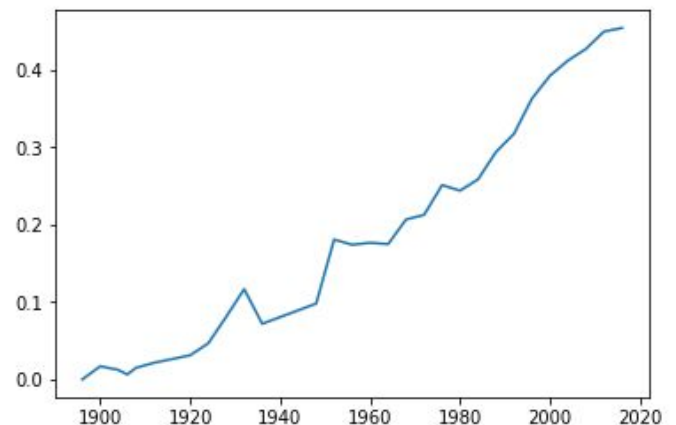
Alpinism:: Great Britain with 12 medals.

Archery:: South Korea with 69 medals.

continues...

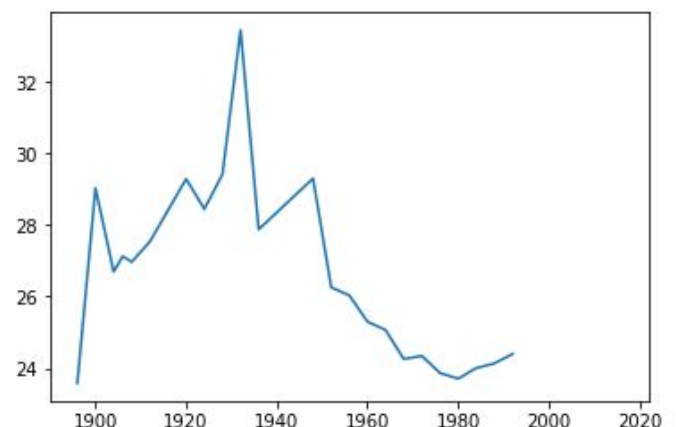
Step 5

=====



Step 6

=====



Step 7

=====

We recommend:

Water Polo

Synchronized Swimming

Rowing