DATA ANALYTICS - MINI PROJECT

Topic : 120 Years of Olympic History Athletes and Results

Phase 1: Data cleaning in Excel

<mark>Athlete Events Table</mark>

1. Make **Table design**.
2. Remove Duplicates values.
3. Centre-align the data in the following columns: **Id, Age, Height, Weight, Sex, NOC, and Year**.
4. Calculate the averages of **Height**, **Weight**, and **Age**, and format the numbers:

Average of height	175
Average of weight	71
Average of Age	26

Then, use the **Find and Replace** function to replace all **NA** values with the corresponding average values.

5. Correct and replace data of team and their NOC with proper value

<mark>Noc – Region Table</mark>

1. Make **Table design**.
2. Delete Notes Column
3. Unique region and replace with proper NOC by comparing athlete events table

Phase 2: Data visualization in Power BI

1. Dax queries for finding **Total medals, Gold, Silver, Bronze** by creating **New Measure**

```
 Gold Medal Count = COUNTROWS( FILTER(
'athlete_events',  'athlete_events'[Medal] = "Gold" ))
```

```
 Silver Medal Count = COUNTROWS( FILTER(
'athlete_events',  'athlete_events'[Medal] = "silver" ))
```

```
 Bronze Medal Count = COUNTROWS( FILTER(
'athlete_events',  'athlete_events'[Medal] = "Bronze" ))
```

2.Dax queries for finding **Total medals, Gold, Silver, Bronze** for each Athletes with creating **New Column**

```
Gold = IF( athlete_events[Medal]= "Gold",1,0)

Silver = IF( athlete_events[Medal]= "Silver",1,0)

Bronze = IF( athlete_events[Medal]= "Bronze",1,0)

Total Medals =
athlete_events[Gold]+athlete_events[Silver]+athlete_events[Bronze]
```