Project: Data Analysis Using Excel

It is an ongoing project that I am doing to apply my skills and knowledge about Data Analysis using Excel. I keep on adding to it as I learn. For this Project, I have chosen the data of 120 years of Olympic history: athletes and results. That is the data of all the athletes who participated in Olympics from Athens 1896 – Rio 2016. (source:- https://www.kaggle.com/heesoo37/120-years-of-olympic-history-athletes-and-results).

This project has two files:
  1) Athletes Events
  2) NOC region.

The NOC region file consist of following field:
Region: Name of the country.
NOC: Three letter abbreviation for the name of the country.

The Athletes Events file consist of following fields:

ID: A serial number to organize the data.

Name: Name of the Athlete.

Sex: Gender of the Athlete.

Age: Age of the Athlete.

Height: Height of the Athlete.

Weight: Weight of the Athlete.

Team: The name of the Team they are playing for.

NOC: Three letter abbreviation of the country.

Games: The Year and the season, when the Olympic was held.

Year: The Year when the Olympic was held.

Season: The Season when the Olympic was held (Summer or Winter).

City: Name of the City where the Olympic was held.

Sport: The Sport in which the Athlete participated (Skating, Judo, Basketball etc.,)

Event: The particular event in which the Athlete participated (Speed Skating Men,

Speed Skating 500m, Speed Skating 1000m etc.,)

Report

Although the data was in .csv format but, I first converted the data in .txt format (tab delimited) and then converted the .txt format to Excel file (.xlsx format). This conversion has nothing to do with the data analysis but, the first thing that I learned about the Data Analysis using Excel was to convert the files into Excel file for Analysis. This conversion was just to start from what I have learned first.

Analyze by Country v/s Medals won

This Pivot table and Pivot chart shows the tally of medals won during the first ever Olympics held at Athens in 1896. The Pivot tables gives us the information about the numbers of medals (Gold, Silver, Bronze) won by each country and the Pivot Chart displays the same information in chart form.

For this analysis, I have done a little extra work. As mentioned in the beginning of the document, there are two files: 1) noc region and athletes events.

Athletes events has field called "noc", which is three letter abbreviations use for the name of the country.

Using the Excel formula "VLOOKUP", I have created an extra field in Athletes events file named as "Country". This extra field gives us the exact name of the country and hence is helpful for analysis.