

**Data Analysis on Olympic Games**

**Submitted by**

**M.V.V.N SaiRam**

**Registration No**: 11701975

**Roll no:** A07

**Section**: KM060

**Course Code:** INT 217 (Introduction to Data Management)

**Under the Guidance of**

**Mrs. Savleen Kaur**

**Discipline of CSE**

**Lovely School of Computer Science and Engineering**

**Lovely Professional University, Phagwara**

**DECLARATION**

I, **M.V.V.N SaiRam**, student of Lovely School of Computer Science and Engineering under CSE Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 18-11-2019 Signature:

Registration No:11701975 Name of the student:M.V.V.N SaiRam

**CERTIFICATE**

This is to certify that **M.V.V.N SaiRam** bearing Registration no. **11701975** has completed INT217 project titled, **“Data Analysis of Olympic Games ”** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

**Signature and Name of the Supervisor**

**Designation of the Supervisor**

**School of Computer Science & Engineering**

Lovely Professional University

Phagwara, Punjab.

Date:

**ACKNOWLEDGEMENT**

It is always a pleasure to remind the fine people in the LPU University for their sincere guidance we received to uphold our practical skills in Data Science.

First of all, thanks to our parent for giving encouragement, enthusiasm and invaluable assistance to me. Without all this, we might not be able to complete this properly.

Second, I would like to thanks to Mrs. Savleen Kaur (Our Data Science Professor) for giving us the opportunity to undergo through this interesting project.

A paper is not enough for me too express the support and guidance We received from them almost for all the work I did here.

This project helped me in developing my skills in analyzing such large data on excel and I wouldn’t have been done without the support and teaching of our Professor.

Finally, I apologize all other unnamed who helped me in various ways.

**Table of Content**

1. Introduction
2. Objectives/Scope of the Analysis
3. Source of dataset
4. ETL process
5. Analysis on dataset (for each analysis)
6. Introduction
7. General Description
8. Specific Requirements, functions and formulas
9. Analysis results
10. Visualization
11. List of Analysis with results
12. References
13. Bibliography

**Introduction:**

The modern **Olympic Games** or **Olympics** are leading international sporting events featuring summer and winter sports competitions in which thousands of athletes from around the world participate in a variety of competitions. The Olympic Games are considered the world's foremost sports competition with more than 200 nations participating. The Olympic Games are held every four years, with the Summer and Winter Games alternating by occurring every four years but two years apart.

Their creation was inspired by the ancient Olympic Games, which were held in Olympia, Greece, from the 8th century BC to the 4th century AD. Baron Pierre de Coubertin founded the International Olympic Committee (IOC) in 1894, leading to the first modern Games in Athens in 1896. The IOC is the governing body of the Olympic Movement, with the Olympic Charter defining its structure and authority.

The evolution of the Olympic Movement during the 20th and 21st centuries has resulted in several changes to the Olympic Games. Some of these adjustments include the creation of the Winter Olympic Games for snow and ice sports, the Paralympic Games for athletes with a disability, the Youth Olympic Games for athletes aged 14 to 18, the five Continental games (Pan American, African, Asian, European, and Pacific), and the World Games for sports that are not contested in the Olympic Games. The Deaflympics and Special Olympics are also endorsed by the IOC. The IOC has had to adapt to a variety of economic, political, and technological advancements. The abuse of amateur rules by the Eastern Bloc nations prompted the IOC to shift away from pure amateurism, as envisioned by Coubertin, to allowing participation of professional athletes. The growing importance of mass media created the issue of corporate sponsorship and commercialisation of the Games. World wars led to the cancellation of the 1916, 1940, and 1944 Games. Large boycotts during the Cold War limited participation in the 1980 and 1984 Games.

The Olympic Movement consists of international sports federations (IFs), National Olympic Committees (NOCs), and organising committees for each specific Olympic Games. As the decision-making body, the IOC is responsible for choosing the host city for each Games, and organises and funds the Games according to the Olympic Charter. The IOC also determines the Olympic programme, consisting of the sports to be contested at the Games. There are several Olympic rituals and symbols, such as the Olympic flag and torch, as well as the opening and closing ceremonies. Over 13,000 athletes compete at the Summer and Winter Olympic Games in 33 different sports and nearly 400 events. The first, second, and third-place finishers in each event receive Olympic medals: gold, silver, and bronze, respectively.

**Scope of analysis:**

This project on Olympic Games Statistics of World provides the overall Statistics details of the Olympic Games held from the year 1868 to 2016.

The project has a dataset with more than 12k rows and 15 columns which are being processed using Excel. These help us to create various graphs for better visualization. The project uses pivot table, pivot chart and some formulas to analyse the data.

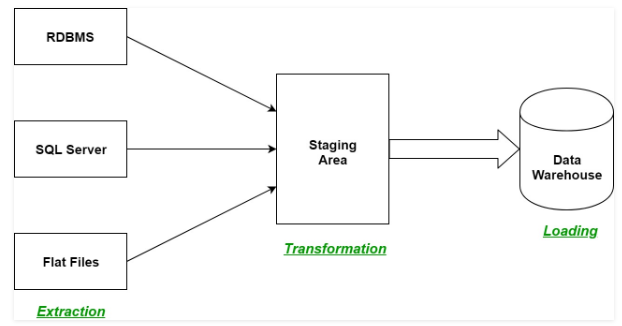
**Source of Dataset Used:**

The data set is being downloaded from the Kaggle Website and has been published by Kaggle. This dataset gives us an in-depth information about Olympic Games.

The link of the dataset is->http://www.kaggle.com/heesoo37/120years-of-olympic-history-atheles-and-results .

**ETL Process:**

ETL is a process in Data Warehousing and it stands for **Extract**, **Transform**and **Load**. It is a process in which an ETL tool extracts the data from various data source systems, transforms it in the staging area and then finally, loads it into the Data Warehouse system.



**Objectives of this project:**

1.Analysis on Number of participants participated in Olympics games Country wise.

2.Analysis on participants categorized by Male and Female gender.

3.Analysis on number of Medals awarded in different sports from year 1868 to 2016

4.Analysis on Medals achieved by different Countries and comparing between them.

5.Analysis on partcipants based on their Age(Age<30,>30 and <50,Age >50).

**Data Analysis:**

**Introduction:**

Data analysis is a process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions, and supporting decision-making. Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, while being used in different business, science, and social science domains.

The process of data analysis:

* Data requirements
* Data collection
* Data processing
* Data cleaning
* Modeling and algorithms

**Specific Requirements, functions and formulas:**

We use different formulas for data cleaning, data processing and modelling and algorithms.

Main function of excel that I have used are

* Power pivot
* Pivot table
* Hyperlink
* Conditional formatting
* Create graph using power pivot
* Create relation between different table
* Function analyser (Slicer)
* Different types of graph
* Different functionality of graph

**Analysis of Dataset:**

**1.Analysis on Number of participants participated in Olympics games Country wise.**

**Introduction:**

This pivot table will help us to visualize the number of participants participated country wise from the data set.

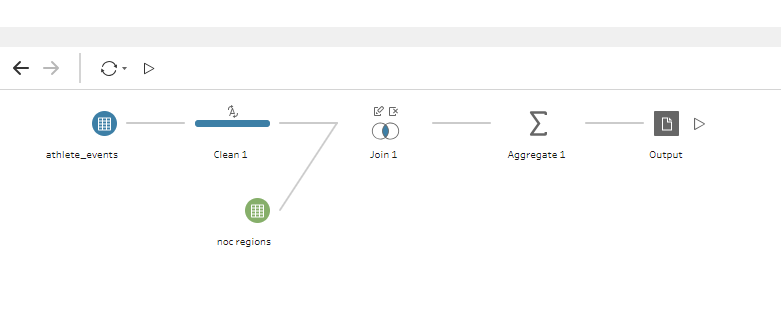
**General Description:**

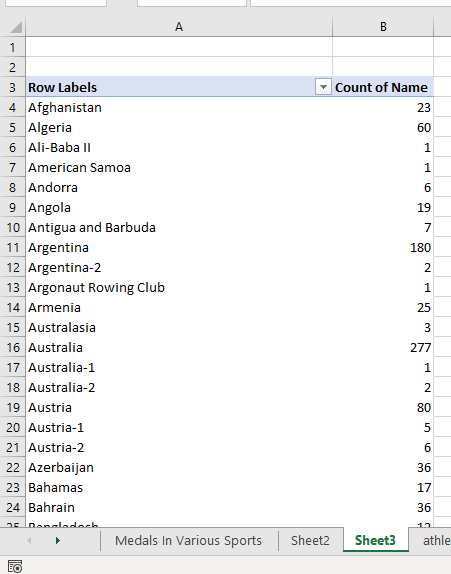
The pivot table used here consists of two columns, they are country name and their respective count of participants.

**Specific Requirements:**

Data cleaning in Tableau prep is performed as well as pivot tables and slicers are used to form relationship.

**Analysis Result**:





**Visualization:**

**2.** **Analysis on participants categorized by Male and Female gender.**

**Introduction:**

This pivot table will help us to visualize the data of participants Gender wise(Male and Female)

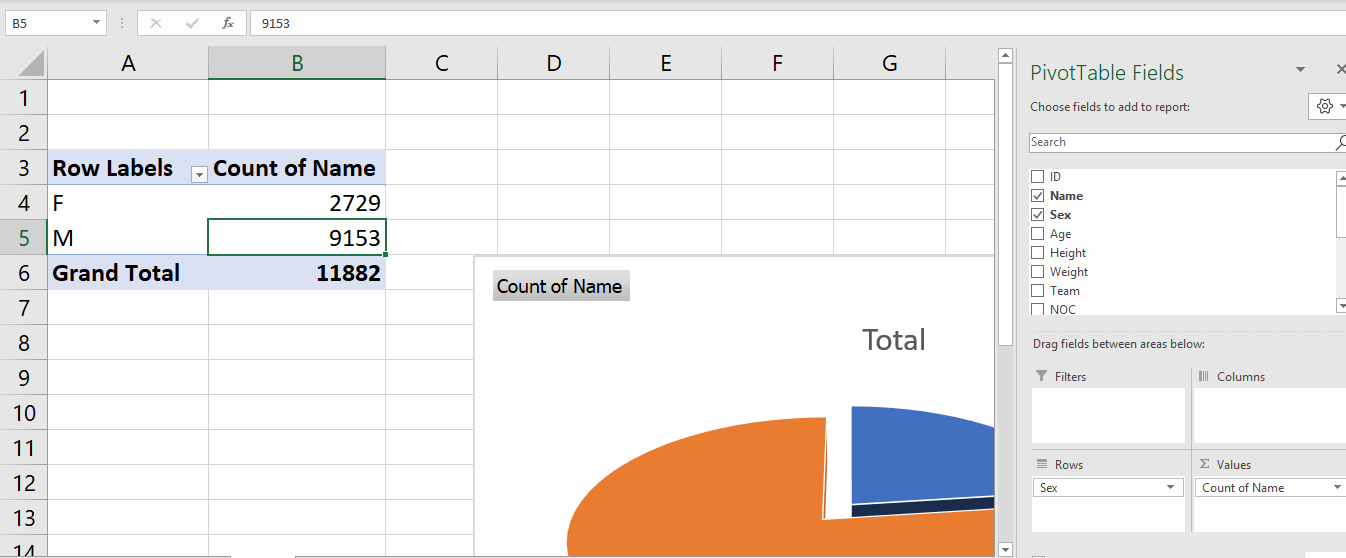
**General Description:**

The pivot table consists of two columns, they are count of Name of participants and Gender .

**Specific Requirements:**

Data cleaning is performed as well as pivot tables and slicers are used to form relationship.

**Analysis Result:**



**Visualization:**

**3.** **Analysis on number of Medals awarded in different sports from year 1868 to 2016**

**Introduction:**

This pivot table will help us to visualize the data based on Season and Count of the Medals awarded in each sport so far.

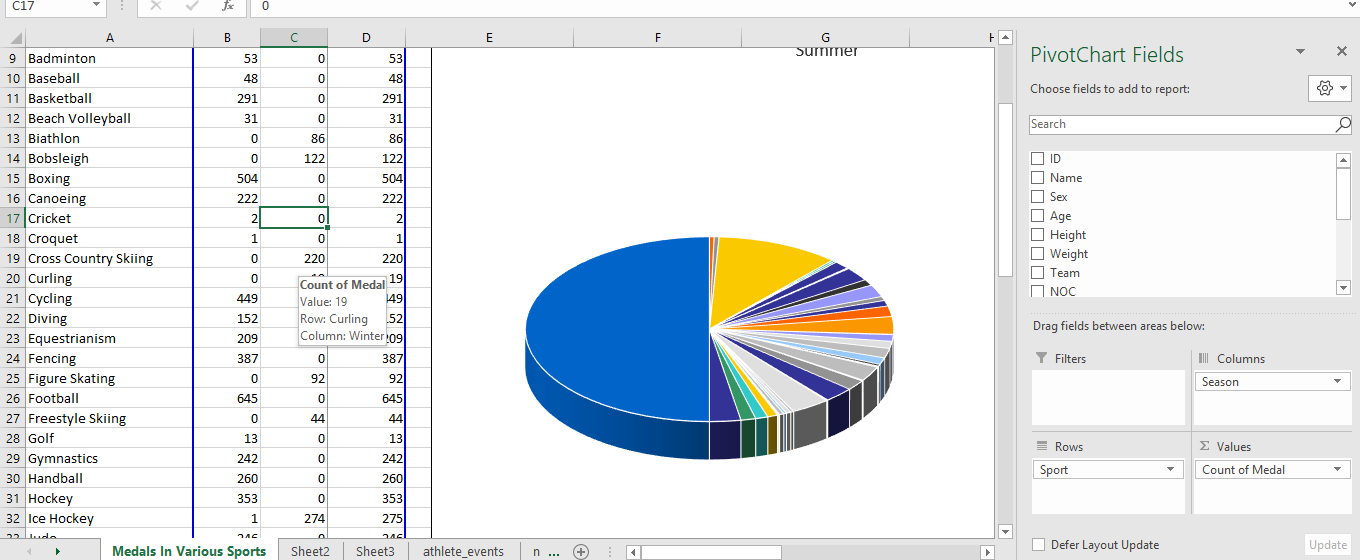
**General Description:**

The pivot table consists of the Game name ,Season (Winter and Summer) and the number of medals arranged in 4 columns.

**Specific Requirements:**

Data cleaning is performed as well as pivot tables and Pie chart is used to visualise the data.

**Analysis Result:**



**Visualization:**

**4.** **Analysis on Medals achieved by different Countries and comparing between them**

**Introduction:** This pivot table will help us to visualize the data based on the number of Gold,Sliver and Bronze medals achieved by countries so far.

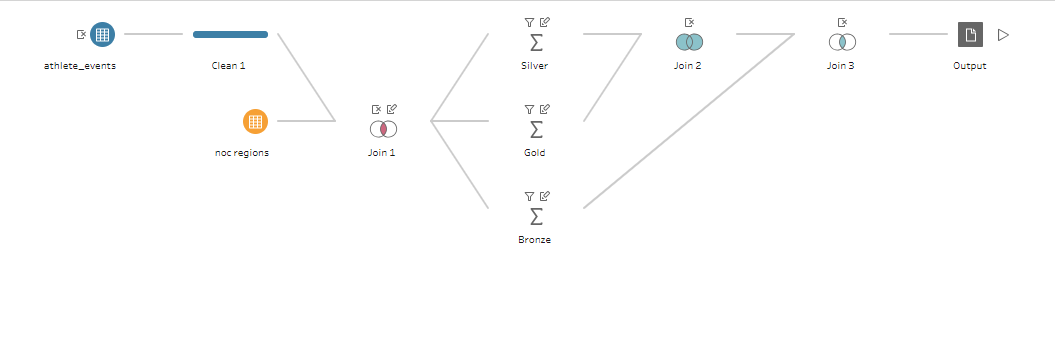
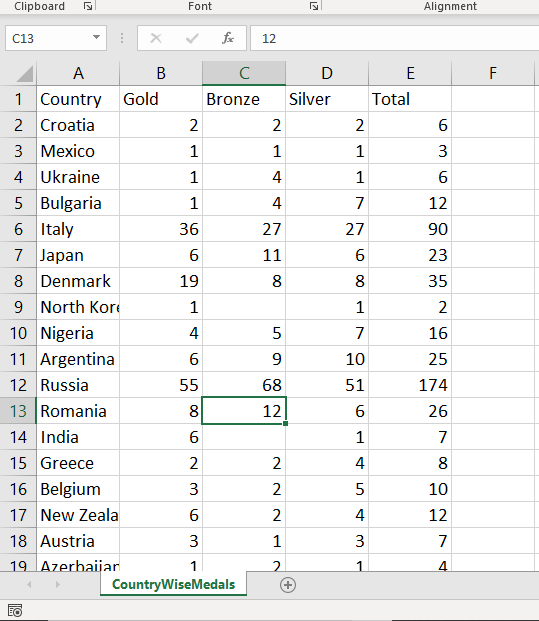
**General Description:**

It is done using Aggregates functions and joins in Tableau Prep.

**Specific Requirements:**

Data cleaning is performed using aggregates and other functions in Talbeau Prep as well as pivot tables and bar graph is used to visualise the data.

**Analysis Result:**



**Visualization:**

**5.** **Analysis on partcipants based on their Age(Age<30,>30 and <50,Age >50).**

**Introduction:**

The data analysis is done based on age.Age is divided into 3 categories namely less than 30,age between 30 and 50 and age greater than 50.

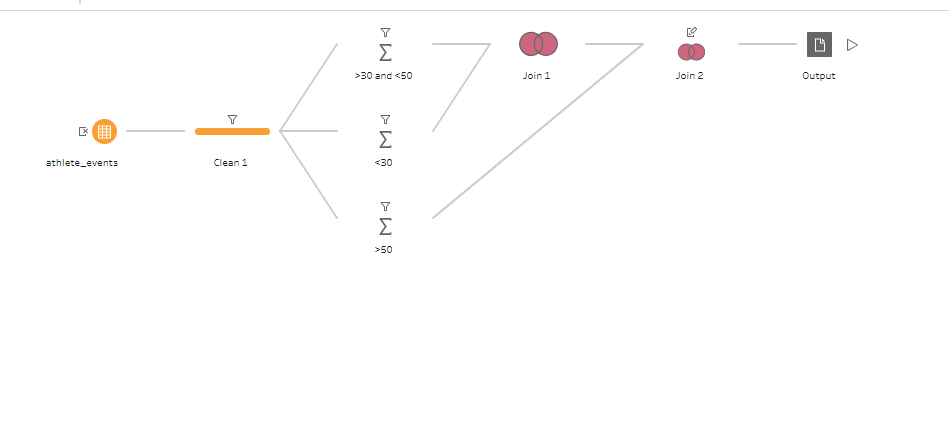
**General Description:**

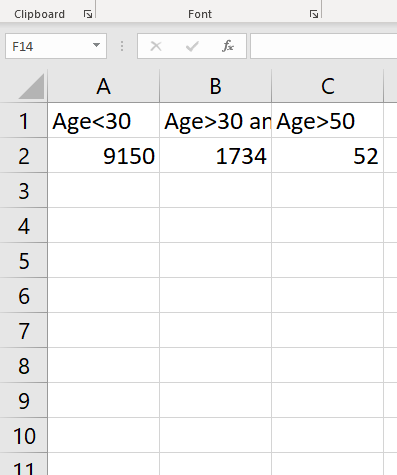
Data handling in done in tabeau prep.It used some aggregate functios and Some joins to handle data.

**Specific Requirements:**

Data cleaning is performed as well as pivot tables and Doughnut chart is used to visualise the data.

**Analysis Result:**





**Visualization:**

**REFERENCES:**

-Tutorials Point

LINK: <https://www.tutorialspoint.com/excel/index.htm>

-Java Point

LINK: <https://www.javatpoint.com/excel-tutorial>

YOUTUBE.COM

-Tutorials Point (India) Pvt. Ltd. Channel

**LINK:** <https://www.youtube.com/channel/UCVLbzhxVTiTLiVKeGV7WEBg/channels>

- MyOnlineTrainingHub

**LINK:** <https://www.youtube.com/user/MyOnlineTrainingHub/videos>

**THANK YOU**