BI

Mini Project

Olympics Games Analysis

2008-2016

Table of Contents

[1 Introduction 3](#_Toc534726757)

[1.1 Setup Checklist for Mini Project 3](#_Toc534726758)

[1.2 Instructions 3](#_Toc534726759)

[2 Problem Statement 4](#_Toc534726760)

[2.1 Objective 4](#_Toc534726761)

[2.2 Abstract of the project 4](#_Toc534726762)

[2.3 Technology used: 4](#_Toc534726763)

[3 Implementation in BI LOT 5](#_Toc534726764)

[3.1 Summary of the Functionality to be built: 5](#_Toc534726765)

[3.2 Guidelines on the functionality to be built: 5](#_Toc534726766)

# Introduction

This document outlines a mini project for the BI LOT. The project is to analyze Olympic games . This document contains the work flow of the system and gives guidelines on how to build the functionality gradually in each of the course modules of the BI LOT.

## Setup Checklist for Mini Project

Minimum System Requirements

* Intel Pentium 90 or higher (P166 recommended)
* Microsoft Windows 2010 or higher
* Memory: 4GB of RAM (64MB or more recommended)
* Internet Explorer 10 or higher
* Oracle 11g / SQL Server.
* Informatica 10.2.0 / SSIS / Data Stage/Abinitio
* Tableau Desktop / SSRS / PowerBi./Qlikview

## Instructions

* The code modules in the mini project should follow all the coding standards.
* Create a directory by your name in drive **<drive>**. In this directory, create a subdirectory **MiniProject**. Store your Project here.
* You can refer to your course material.
* You may also look up the help provided in the BI docs and documentation provided with respective tools.

# Problem Statement

## Objective

Analysis of Olympics Games based on the Olympics data for the year 2008-2016.

## Abstract of the project

This project is aimed to work on past 8 year’s Olympic Games. The system will be used to perform analysis on the games based on the given data and to make decision support with the help of Tableau.

## Technology used:

* Oracle 11G / SQL Server (Database)
* Informatica power center / SSIS / Data Stage/Abinitio(ETL Tool).
* Tableau Desktop / SSRS / PowerBI/Qlikview

# Implementation in BI LOT

## Summary of the Functionality to be built:

The participants need to develop the **Olympics games Analysis** by building the functionality incrementally in each of the course modules of BI LOT using one of the above ETL and reporting tool.

## Guidelines on the functionality to be built:

**Project flow:**

Heterogeneous Sources i.e. operational data (Flat files)

EXTRACT

TRANSFORM

LOAD

Data warehouse (Oracle)

Reports Based On Tableau

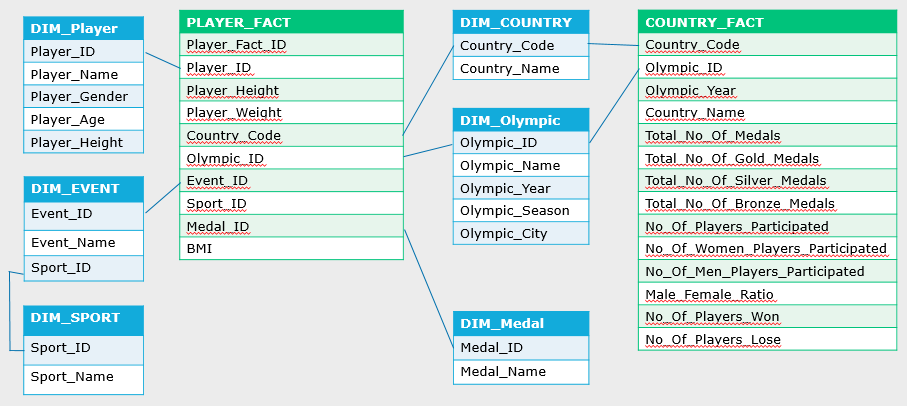
REPORT

**Schema Design:**

The project follows the Fact Constellation schema approach.

A fact constellation has multiple fact tables. It is also known as galaxy schema.

The Dimension tables are Player, Olympics, Event, Sport, Medal and Country. Fact tables are Player Details and Country Details.



**Note:** The datatype/length for the Dimension/Fact table attributes can be changed as required. Additional fields can be added, if required.

* Player Dimension

This dimension contains names of all the player.

This oracle table comprises of the following Fields:

* Player ID
* Player Name
* Player Gender
* Player Age
* Player Height

The Player ID is a running sequence and unique number to be generated for all the player with same Player Name.

As same Player Name could participate in more than one Event at Olympics.

* Olympic Dimension

This dimension contains Olympic year. The Olympic Games are held every four years, with the summer and Winter Games alternating by occurring every four years but two years apart and the oracle table comprises of the following Fields:

* Olympic ID
* Olympic Name
* Olympic Year
* Olympic Season
* Olympic City
* Country Dimension

This dimension contains different countries that participate in Olympics and the oracle table comprises of the following Fields:

* Country Code
* Country Name
* Sport

This dimension contains sport details and the oracle table comprises of the following Fields:

* Sport ID
* Sport Name
* Event

This dimension contains various Event names along with their unique id. Here Sport ID Acts as a foreign key and the oracle table comprises of the following Fields:

* Event ID
* Event Name
* Sport ID
* Medal

This dimension contains medals and the oracle table comprises of the following Fields:

* Medal ID
* Medal Name

The Facts tables are:

* Player Details Fact

This is a Fact table that would contain the all the measuring units related to Player

* Player Fact ID (sequence generated)
* Player ID
* Player Height
* Player Weight
* Country Code
* Olympic ID
* Event ID

Sport ID

* Medal ID
* BMI

This is a Fact table that would contain the all the measuring units related to Country

* Country code (sequence generated)
* Olympic ID
* Olympic Year
* Country Name
* Total No Of Medals
* Total No Of Gold Medals
* Total No Of Silver Medals
* Total No Of Bronze Medals
* No Of Women Players Participated
* No Of Men Players Participated
* No Of Players Participated
* Male Female Ratio
* No of Players Won
* No of Players Lose

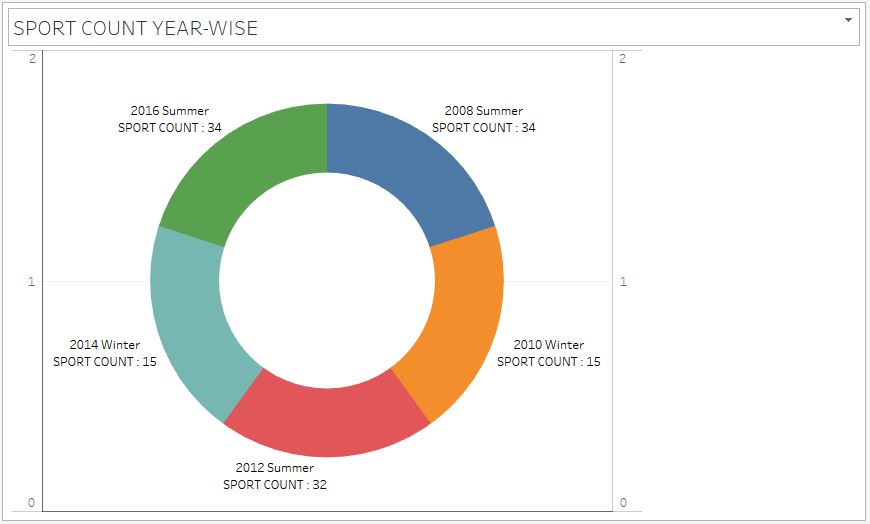
**Data Transformation for data warehouse:**

* Load the Data into Dimension tables using the CSV files provided.
* Load the data into Fact table
  + Points to ponder.
    - Validate the foreign key values before loading.
    - Olympics ID, Player ID, Event ID, Medal ID, Sport ID are the Foreign Keys for this fact table.
    - Fact ID is the primary key for both fact tables.
    - The fact table should store the data pertaining to all the Olympics Games.
    - The measures that do not have data in the CSV file should be given default value NULL in the fact table while loading the data.

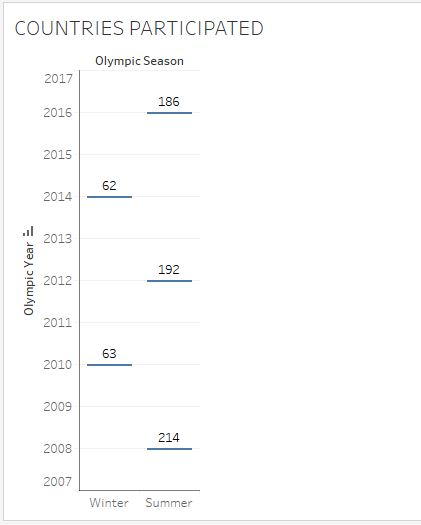
**Create a view in Tableau Designer with the above dimension tables and the fact table.**

**Sample Reports:**

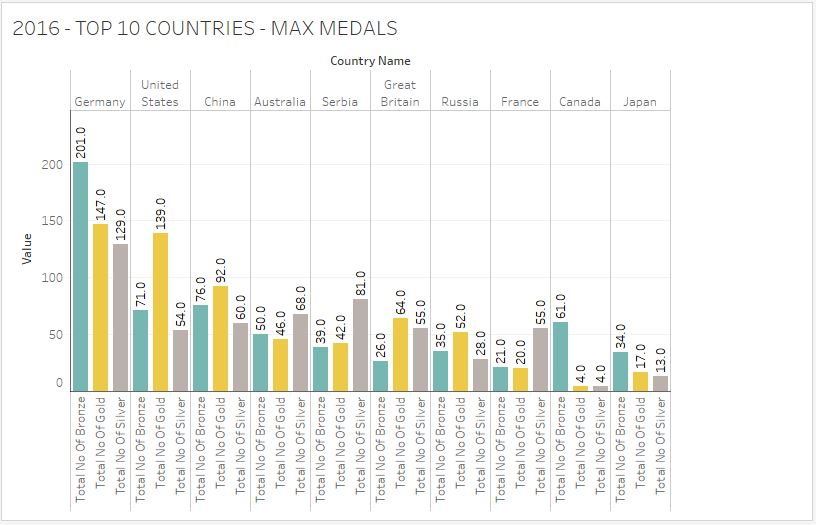
1. Sport Count Year-Wise (Doughnut view).



2.Countries Participated (Gantt view).



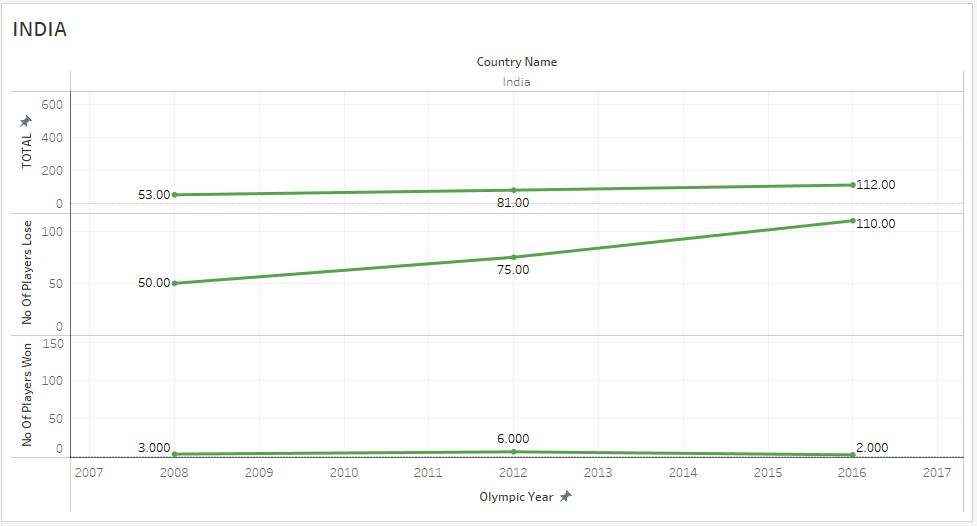
1. 2016- Top 10 Countries of Max Medals (Side-by-Side bars view).



1. Germany Win-Lose Year wise Analysis (Lines continuous view).



1. India Win-Lose Year wise Analysis (Lines Continuous view).



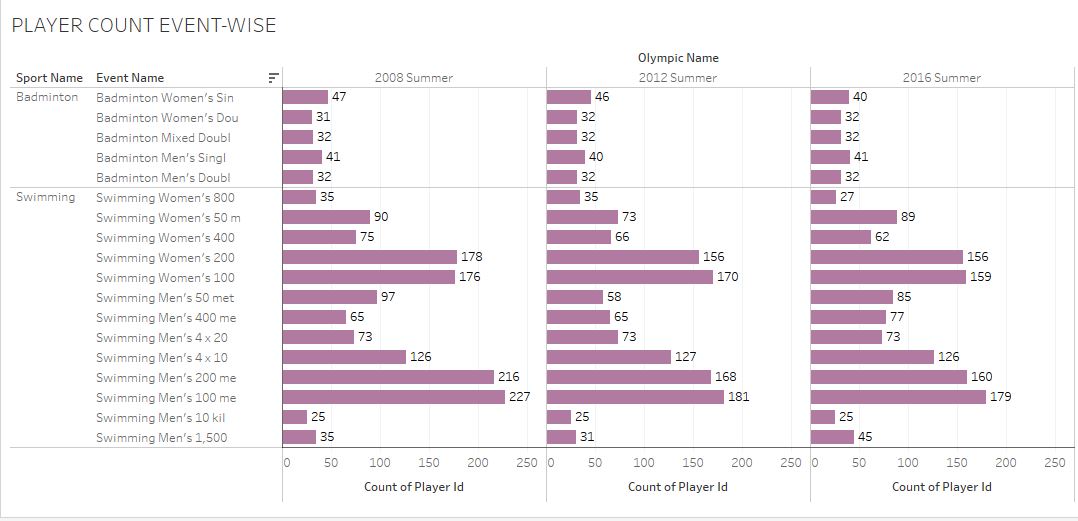
1. Male to Female Ratio Year- wise (Horizontal bars view).



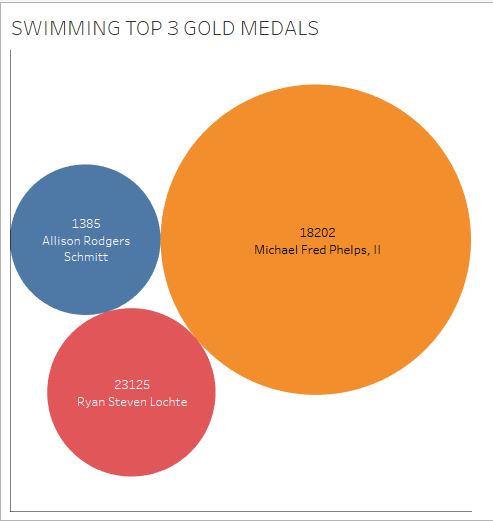
1. Top 10 Countries with max medal year-wise (Text tables view).



1. Player Count Event-Wise (Horizontal Side-by-side bars view).



1. Top 3 Medals in Swimming (Packed Bubbles view).



1. Top 10 Players Max Medals all 5 years (Tree Map view).

