Low Level Design

FIFA WORLD CUP DATA ANALYSIS

Written By	Author 1
Document Version	0.1
Last Revised Date	08/08/2023

Revision Number

Last Date of Revision-

Ayesha Syed Mohammad

Document Control

Date	Version	Description	Author
07/08/2023	1.0	Introduction, Problem Statement	Ayesha
08/08/2023	1.1	Dataset Information, Architecture Description	Ayesha
08/08/2023	1.2	Final Revision	Ayesha

Contents

1. Introduction

1.1 What is Low-Level design document?

The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the FIFA World Cup Data Analysis dashboard. LDD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

1.2 What is Scope?

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

1.3 Project Introduction

The FIFA World Cup, a global football spectacle, captivates the hearts of millions, uniting nations in a fervent pursuit of football glory. While the players and coaches take center stage, an unsung group of analysts toils tirelessly in the background, providing invaluable data to satisfy the curiosity of passionate fans. This project is a tribute to these dedicated analysts, who diligently delve into comprehensive datasets, seeking to unravel the key metrics and factors that influence World Cup triumphs.

In the quest to uncover the secrets behind World Cup victories, we embark on an exciting journey that combines the allure of football with the power of data analysis. With access to a treasure trove of historical World Cup and match result datasets, we aim to shed light on the patterns, trends, and statistics that differentiate champions from contenders.

Through our exploration, we seek to bring recognition to these unsung heroes while deepening our appreciation for the intricacies of football success on the grandest stage. Join us as we celebrate the passion, dedication, and invaluable contributions of the analysts who play a crucial role in shaping the legacy of the FIFA World Cup.

2. Problem Statement

With FIFA being in the blood as many people of the world. You are tasked to tell the story of unsung analysts who put great efforts to provide accurate data to answer every question of fans. The FIFA World Cup is a global football competition contested by the various football-playing nations of the world. It is contested every four years and is the most prestigious and important trophy in the sport of football.

The World Cups dataset show all information about all the World Cups in the history,

while the World Cup Matches dataset shows all the results from the matches contested

as part of the cups.

Find key metrics and factors that influence the World Cup win.

3. Dataset Information

RoundID :- Unique ID assigned to each round played by team in FIFA

MatchID :- Unique ID assigned to each match

Team Initials :- Initials of the teams

Coach name :- name of the coaches

Line-up :- the list of players from a team who are selected to start and play in a particular match

Shirt Number :- Number written on players shirt

Player Name:- Name of the players

Position :- Positions of players in a team

Event :- actions such as goals, fouls, assists, substitutions, and more that occur during a match

Year:- games played in year

Datetime :- Date and time of the game

Stage:- stages of each game played

Stadium :- football stadium

City :- Place where the matches were held

Home Team Name:- the national team that is designated as a home team

Away Team Name:- the foreign team that is designated as a Away team

Home Team Goals:- a national team scored scored

Away Team Goals:- a foreign team scored

Win Conditions:- conditions applied when the games scores are tied.

Attendance:- teams and coaches presences during the matches

Half-time Home Goals:- in half time of match goals scored by national teams

Half-time Away Goals:- in half time of match goals scored by foreign teams.

Referee:- names of the person responsible for officiating and enforcing the rules of a football match during the World Cup tournament

Assistant 1 :- referee 2 names

Assistant 2:- referee 3 names

Home Team Initials:- initials or code used to represent the national teams

Away Team Initials:- initials or code used to represent the foreign teams

Country:- place where the world cup held

Winner:- winners of each world cup

Runners-up:- 2nd finalist team of the world cup

Third:- The teams which won 3rd place

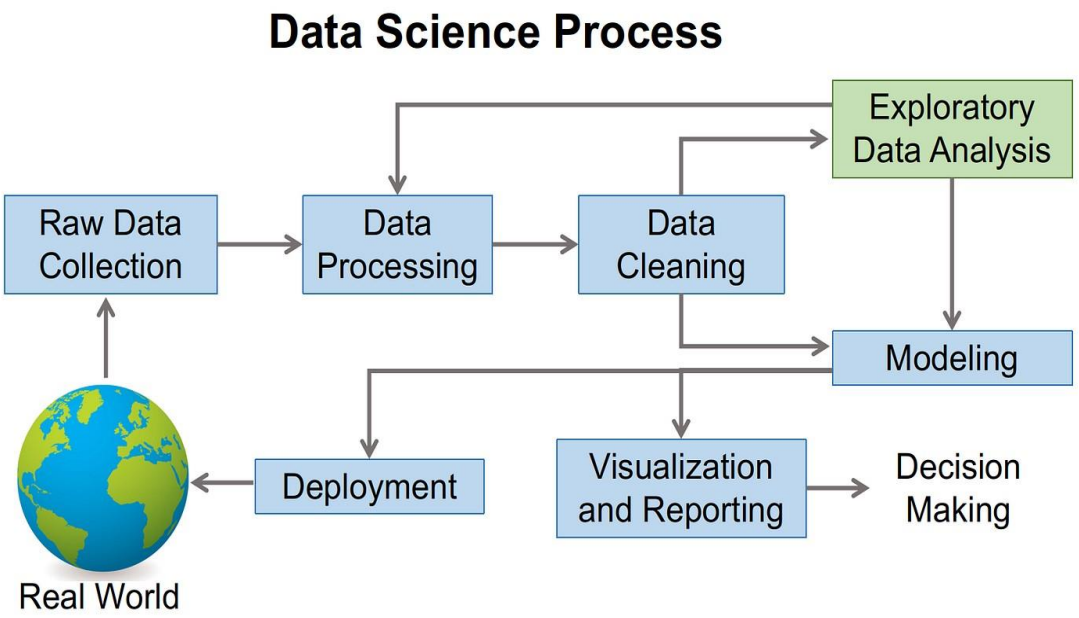
Fourth:- The teams which won 4th place

Goals Scored:- Total Goals scored by each team in world cup every year

Qualified Teams:- teams qualified for each match in world cup

Matches Played:- Total matches played by each team every year

4. Architecture:-



4.1 Architecture Description

1. Raw Data Collection

The Dataset was taken from iNeuron's Provided Project Description Document.

https://drive.google.com/drive/folders/12oHYj0qH2uZD8I13cVDiymTNDYldeJRa?usp=sharing

2. Data Pre-Processing

Before building any model, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict. Model performance depends on the quality of data feeded to the model to train. This Process includes

a) Handling Null/Missing Values

b) Handling Skewed Data

c) Outliers Detection and Removal

3. Data Cleaning

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.

a) Remove duplicate or irrelevant observations

b) Filter unwanted outliers

c) Renaming required attributes

4. Exploratory Data Analysis (EDA)

Exploratory Data Analysis refers to the critical process of performing initial investigations on data to discover patterns, spot anomalies, test hypothesis and to check assumptions with the help of summary statistics and graphical representations.

5. Reporting

Reporting is a most important and underrated skill of a data analytics field. Because being a Data Analyst you should be good in easy and self explanatory report because your model will be used by many stakeholders who are not from technical background. a) High Level Design Document (HLD)

b) Low Level Design Document (LLD)

c) Architecture

d) Wireframe

e) Detailed Project Report

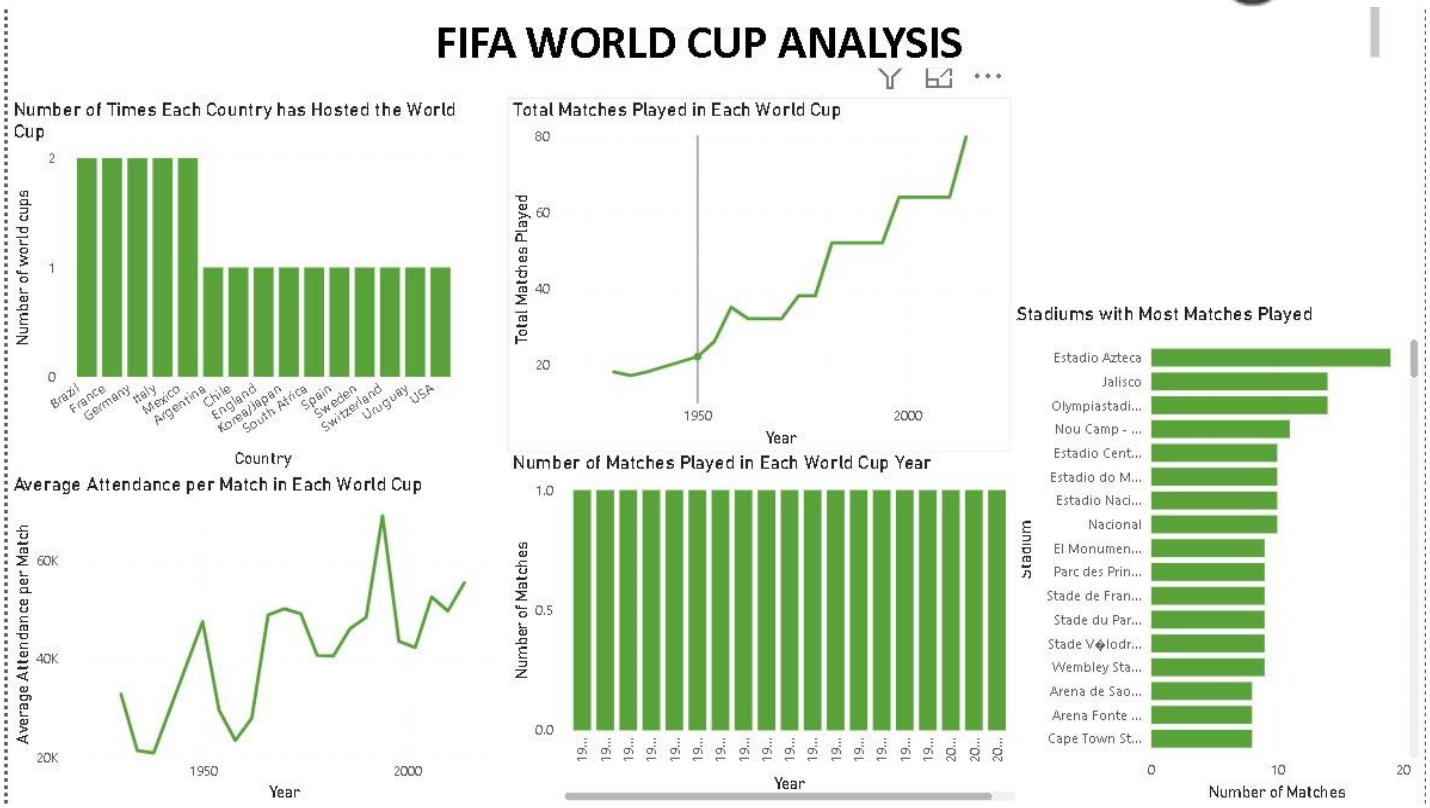
f) Power Point Presentation

6. Modelling

Data Modelling is the process of analysing the data objects and their relationship to the other objects. It is used to analyse the data requirements that are required for the business processes. The data models are created for the data to be stored in a database. The Data Model's main focus is on what data is needed and how we have to organize data rather than what operations we have to perform.

7. Deployment

We created a Power BI Dashboard


FIFA WORLD CUP ANALYSIS
Number of Times Each Country has Hosted the World Cup
Number of world cups
Brazil
France
Germany
Italy
Mexico
Argentina
Chile
England
Korea/Japan
South Africa
Spain
Sweden
Switzerland
Uruguay
USA
Country
Total Matches Played in Each World Cup
Total Matches Played
Year
Stadiums with Most Matches Played
Estadio Azteca
Jalisco
Olympiastadi...
Nou Camp - ...
Estadio Cent...
Estadio do M...
Estadio Naci...
Nacional
El Monumen...
Parc des Prin...
Stade de Fran...
Stade du Par...
Stade V◆lodr...
Wembley Sta...
Arena de Sao...
Arena Fonte ...
Cape Town St...
Stadium
Number of Matches
Average Attendance per Match in Each World Cup
Average Attendance per Match
Year
Number of Matches Played in Each World Cup Year
Number of Matches
Year