FIFA WORLD CUP ANALYSIS

PROJECT DETAILS :

Technology - Business Intelligence
Domain - Sports
Difficulty - Advance
Language - Exploratory Data Analysis using Python

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.

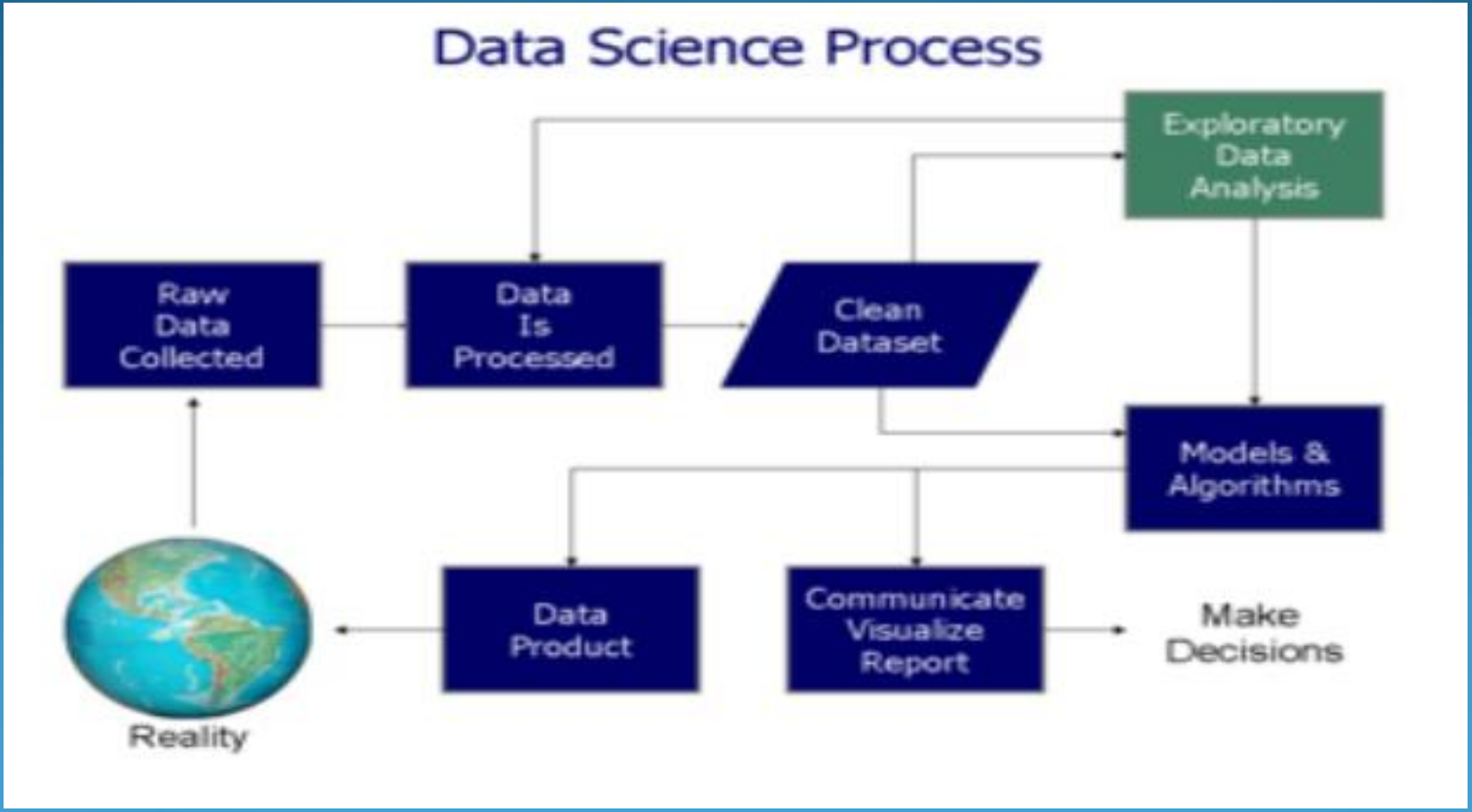
- Do your own research and come up with your findings.

OBJECTIVE :

- To get useful insights from the FIFA WORLD CUP dataset and visualize to give some interesting figures to the football fans .

- This project is based on **FIFA World Cup Analysis** to get insights about the Matches, patterns about players, Patterns about cup holders, Home Ground Advantage and many more things**

ARCHITECTURE :



BENEFITS :

- Most Number of World Cup Winning Title
- Number of Goal Per Country
- Attendance, Number of Teams, Goals, and Matches per Cup
- Goals Per Team Per World Cup
- Matches With Highest Number Of Attendance
- Stadium with Highest Average Attendance
- Which countries had won the cup ?
- Number of goal per country
- Match outcome by home and away teams

VISUALIZATION PROCESS : 1

Collecting Data

- The first and most important step of data visualization is gathering data in large amounts. Only after we have substantial data, we can apply data visualization techniques on the collected data and get some helpful insights from it.

VISUALIZATION PROCESS : 2

Clean Your Data

- Data cleaning is an essential step to perform before creating a visualization. A bunch of data out of a large dataset that has inappropriate, empty, or false values may lead to adding erroneous visuals with anomalies in it.
- The output received from a data cleaning process is usually a dataset that is free of errors and anomalies etc. which gives much more accuracy when data is processed. Data cleaning is pretty much dependent on the dataset domain that you're working with.

VISUALIZATION PROCESS : 3

Prepare Data

- To prepare the data before sending it further for visualization is to determine the type of graph, chart, or any other visualizations you need to create and the supporting library you will be integrating for it. After the chart is finalized it may be necessary to transform the data as per requirements.
- Data preparation tasks include finding data columns that help make some decisions out of it, giving some meaningful insights about data, grouping data, creating aggregate values for groups, combining variables to create new columns, etc.

VISUALIZATION PROCESS : 4

Choose A Chart Type

- Before choosing a visual chart or graph, it is important to understand your audience and then choose a chart or graph accordingly which will best communicate the message.
- Choosing a chart totally depends on what findings you need to convey to your audience.

VISUALIZATION PROCESS : 5

Visualize Data

- In the final step, you'll have the required data you need to create visualizations. Now you can apply all your visualizations skills on the prepared data and represent the data in charts or graphs with meaningful insights.

QUESTION / ANSWER :

Q1) What is the source of data?
ANS - This dataset is publicly available for research. Available in the form of
Jovian link –
pdf –
GitHub Link –
Tableau Link –
Title = (FIFA World Cup Analysis)

Q 2) What was the type of data?
ANS - The data was contained in the zip folder in csv format.

Q 3) What is the complete flow you followed in this Project?
ANS - Refer slides for better Understanding.

Q 4) What steps should I follow to get insights from the data?

Step1 –Download the data and store it in a location in your PC.
Step2 –Open .ipynbsolution notebook, enter the path and run all cells.

QUESTION / ANSWERS :

Q5) Can I use different IDE to open Solution.ipynb?
ANS - Yes, this file is compatible to run in any IDE that supports python.

Q 6) What were the libraries that you used in Python?
ANS - I used Pandas, NumPy, Matplotlib , Seaborn libraries in Python.

Q 7) Do I have to download and import all these libraries?
ANS - These libraries are downloaded in build, So you don't need to download separately. For importing these libraries, the code is already written.

Q 8) How much time will it take to run all cells in all cells and get insights?
ANS - It will usually take less than a minute to run and execute the solution file.