FIFA

WORLD CUP

(1930 - 2022)

22

TOURNAMENTS

DEPI FINAL PROJECT REPORT

FIFA World Cup Analysis (1930 – 2022) Report

Digital Egypt Pioneers Initiative – DEPI

Data Analyst Specialist Track

Project Name: FIFA World Cup Analysis (1930 – 2022)

Team Members:

- **Sameh Mahmoud Mahmoud**
- **Mohammed Helmy Fadel**
- **Kamilya Abdelnaser Ahmed**

Special Thanks to our Great Instructor Amira Shahin for her invaluable guidance throughout this project.

1. Data Collection

To begin the analysis of the World Cup, data was sourced from a variety of platforms, including Kagle and Maven, ensuring a comprehensive and diverse dataset. These sources provided valuable information, such as:

- Player statistics
- Team records
- Match results
- Historical World Cup data

2. Data Cleaning

The raw dataset contained missing values and irrelevant columns. The cleaning process involved:

- **Handling null values**: Missing data was either removed or imputed based on context.
- **Removing irrelevant columns**: Unnecessary columns were dropped to streamline the analysis.
- **Standardizing data**: Ensured consistency in names, dates, and numeric formats to avoid discrepancies.

3. Data Preparation

Once the data was cleaned, it was prepared for analysis.

- **Standardizing player names**: Ensured uniformity in player names across datasets to avoid duplicates or mismatches.
- **Inserting Country flags:** Added flags for each player's country, enhancing visual clarity in analysis, and making the analysis more engaging.

4. Data Analysis

The following analysis techniques were applied to extract insights:

- **KPI exploration**: Key Performance Indicators (KPIs) were created to track important metrics like goals scored, win ratios, and team performance.
- **Survey insights**: Sentiment analysis was conducted on teams, players, and events to gauge public opinion.
- **Custom data tables**: Additional metrics were calculated to provide deeper insights into performance.

5. Connection Modeling

Relationship modeling was essential to making the data interactive.

- Connecting teams and years: Established relationships between teams, players, and years to allow for dynamic analysis across different time periods and match comparisons.

6. Dashboard Creation

A custom dashboard was built using Power BI:

- **Research**: Reviewed existing World Cup dashboards to identify best practices and effective visualization techniques.
- **Power BI implementation**: Interactive charts, graphs, and maps were used to visualize the data.
- **Custom measures**: Calculated statistics such as goal averages, team rankings, and player performance metrics to enhance analysis.

7. Visualization & Page Layout

Effective visualizations were chosen to present the data:

- **Visual choices**: Bar charts, line graphs, and interactive maps were used to convey insights clearly.
- **Page structure**: The dashboard was divided into different pages, each focusing on specific aspects, such as:
    - Player performance
    - Team progress
    - Tournament history

8. Overview & Navigation

The dashboard begins with an overview page, which presents key World Cup information. Users can navigate easily between different analysis pages for in-depth exploration.

9. Final Report & Recommendations

Each dashboard section includes a brief report summarizing key findings.
Recommendations and forecasts are provided based on the analysis:

- **Team recommendations**: Based on data, certain teams may need to adjust their strategies. For example, Argentina tends to play aggressively, as reflected in their high number of red cards received during tournaments.
- **Future performance forecast**: Using historical data, predictions were made about future tournaments. For instance, Brazil is expected to maintain a strong presence in the upcoming World Cups.