# Data Analysis Project Submission Report Template

### 1. Title Page

* **Project Title:** " FIFA World Performance (2022) "
* **Submitted By:**
  + **Student Name: KISHEN RAJ . D**
  + **Roll Number: 2461236**
  + **College Email ID: kishen.raj@btech.christuniversity.in**
* **Course:** e.g., Fundamental Data Analysis
* **Instructor Name: Dhiraja**
* **Institution:** Christ University
* **Date of Submission:** 22/09/2025

### 2. Abstract

This project focuses on analyzing Argentina’s performance in the 2022 FIFA World Cup using player-level data. The primary goal is to evaluate key statistics such as appearances, goals, assists, defensive actions, and overall contributions of individual players to understand how the team achieved success. Microsoft Excel was used as the core tool for data organization, cleaning, and visualization, with sheets dedicated to player stats, goals, duels, appearances, and dashboards. Features such as formulas, sorting, filtering, and charts were applied to uncover insights and present findings in a structured way. The final outcome includes a clear statistical overview of Argentina’s World Cup squad performance, supported by dashboards for easy interpretation. This analysis is useful for identifying player strengths, comparing contributions, and demonstrating how data-driven approaches enhance sports performance evaluation.

### 3. Objectives

The main goals of this project are:

* **Clean and prepare** the raw dataset for accurate analysis and remove inconsistencies, if required.
* **Formulate and answer five key analytical questions** to derive meaningful insights from Argentina’s 2022 FIFA World Cup performance.
* **Design and develop a comprehensive, user-friendly dashboard** that highlights important performance metrics of players and the team.
* **Utilize appropriate charts and graphs** such as bar charts, pie charts, and line graphs to present data visually and enhance understanding.
* **Summarize the findings and highlight their practical implications**, particularly how data-driven insights can support sports performance evaluation and decision-making.

### 4. Scope of the Project

This project is focused exclusively on the cleaning, analysis, and visualization of Argentina’s 2022 FIFA World Cup dataset. The scope is limited to the following boundaries:

* The project emphasizes **data preparation, analysis, and visualization** to extract meaningful insights.
* No external programming languages such as **Python, R, or SQL** are used; all tasks are completed within Microsoft Excel.
* Advanced statistical modeling or predictive analytics is **not included** in the scope.
* The analysis is conducted entirely within a **single Excel file**, ensuring a structured and self-contained workflow.
* Insights are restricted to the **provided dataset only**, without incorporating additional external data sources.

### 5. Tools & Technologies Used

| **Tool / Technology** | **Purpose** |
| --- | --- |
| **Microsoft Excel** | Data manipulation, cleaning, analysis, and dashboard creation |
| **PivotTables** | Summarizing and aggregating player and team statistics for easier interpretation |
| **Charts & Graphs** | Visualizing key metrics and insights through bar charts, pie charts, and line graphs |

### 6. Data Cleaning & Preparation

 **Initial State of the Data:**  
The dataset contained player-level information from Argentina’s 2022 FIFA World Cup squad. It included details such as player name, position, jersey number, club, appearances, goals, assists, and defensive metrics. However, the raw dataset had unnecessary header rows, some empty columns, and inconsistent formatting that required cleaning.

 **Steps Taken to Clean the Data:**

* Removed extra header rows and irrelevant metadata before the actual table began.
* Renamed and aligned column headers for clarity (e.g., “Goals Scored,” “Assists Provided”).
* Checked for and addressed missing or blank values where applicable.
* Verified data types (e.g., dates for “Player DOB,” numeric values for goals, tackles, etc.) and corrected where necessary.
* Removed unnecessary index or placeholder columns.
* Ensured player names and clubs were formatted consistently.

 **New Columns / Features Created:**

* A **“Total Contributions” column** was derived by combining goals and assists to measure overall attacking impact.
* Calculated **average defensive metrics per 90 minutes** for comparisons across players.
* Additional calculated fields (using formulas) were created to support dashboard visualizations and comparative analysis.

### 7. Dashboard Design Strategy

 **Layout and Design:**  
The dashboard was designed to provide a clear and intuitive overview of Argentina’s 2022 FIFA World Cup performance. Key performance indicators such as **total goals, assists, appearances, and defensive actions** were highlighted at the top as summary cards. Below this, charts and tables were arranged in a grid layout to allow easy comparison of player contributions. The design focused on readability, minimal clutter, and quick insights.

 **Choice of Visualizations:**

* **Bar Charts** were used to compare players based on goals, assists, and duels won, making individual contributions easy to interpret.
* **Pie Charts** displayed the distribution of goals and assists among players, highlighting the most influential attackers.
* **Column Charts** were used for appearances to show player participation levels.
* **Line Charts** were applied where trends across matches or cumulative performance needed to be visualized.

 **Interactive Elements:**

* **Slicers and filters** were integrated to allow users to explore data by player, position, or club.
* Dropdown filters were provided to focus on specific metrics (e.g., attacking vs. defensive stats).
* Interactive PivotCharts ensured dynamic updates across visualizations when selections were changed.

### 8. Questions & Solutions

 **Question 1: Who was the top goal scorer for Argentina in the 2022 FIFA World Cup?**

* **Analysis:** Sorted the dataset by Goals Scored column in descending order.
* **Solution:** Lionel Messi scored the highest number of goals, making him the top scorer.

 **Question 2: Which player provided the most assists?**

* **Analysis:** Filtered and compared values in the Assists Provided column.
* **Solution:** Ángel Di María contributed the highest assists, playing a key role in Argentina’s attack.

 **Question 3: Which defender had the strongest defensive stats?**

* **Analysis:** Compared players in DF position using Tackles per 90 Min and Interceptions per 90 Min.
* **Solution:** Nicolás Otamendi had the best overall defensive performance, with high tackles and interceptions.

 **Question 4: Which player made the most appearances in the tournament?**

* **Analysis:** Checked the Appearances column to identify players with maximum matches played.
* **Solution:** Goalkeeper Emiliano Martínez and Lionel Messi had the highest appearances, featuring in all matches.

 **Question 5: What was the overall attacking contribution of players (goals + assists)?**

* **Analysis:** Created a new calculated column (Total Contributions = Goals + Assists) and ranked players.
* **Solution:** Lionel Messi led overall contributions, followed by Julián Álvarez, showing their central role in Argentina’s success.

### 9. Challenges Faced & Solutions

| **Challenge** | **Solution** |
| --- | --- |
| **Challenge 1: Difficulty in handling missing or blank values** | Used Excel’s **Find and Replace** and conditional formatting to identify missing data, then replaced them with “N/A” or logical estimates to maintain consistency. |
| **Challenge 2: Choosing the right chart type for effective visualization** | Experimented with **bar, line, and pie charts**, and finalized the most appropriate type based on whether the metric required trend analysis, distribution, or comparison. |
| **Challenge 3: Dataset was not in a tidy format for PivotTables** | Applied **Text to Columns**, removed unnecessary headers, and reformatted the data into a structured table suitable for PivotTable analysis. |
| **Challenge 4: Ensuring dashboard readability and avoiding clutter** | Designed the layout with a **clean grid structure**, grouped related metrics together, and used slicers/filters for interactivity instead of overloading visuals. |
| **Challenge 5: Comparing both attacking and defensive performance fairly** | Created **calculated columns** (e.g., Total Contributions = Goals + Assists, Avg. Defensive Actions) to allow balanced comparisons across positions. |

### 10. Outcome

 **Key Insights Gained:**  
The analysis provided clear insights into Argentina’s 2022 FIFA World Cup performance. It highlighted top scorers like Lionel Messi and Julián Álvarez, key playmakers such as Ángel Di María, and defensive leaders including Nicolás Otamendi. The dataset also revealed patterns in appearances, attacking contributions, and defensive strengths across positions.

 **Usefulness of the Dashboard:**  
The dashboard served as an interactive and user-friendly tool to summarize large volumes of player statistics. By using charts, PivotTables, and slicers, it allowed quick comparisons, easy trend spotting, and focused analysis of both individual and team contributions.

 **Skills Learned and Enhanced:**  
This project strengthened skills in **data cleaning, transformation, and visualization** using Microsoft Excel. It improved proficiency in PivotTables, chart selection, dashboard design, and applying formulas for calculated metrics. Beyond technical skills, it also enhanced the ability to interpret sports performance data and present insights in a clear, visual format.

### 11. Screenshots of Final Output

*1. Extract & Display → I can load the Dashboard and Sort and Filters / combo sheets from your Excel file, then render their contents here so you can see the final output tables/figures.*

*2. Recreate as Images → I can generate clean charts/figures (e.g., goals, appearances, averages) based on the data, so you’ll have “screenshot-style” images of the outputs that look like a dashboard.*

### 12. Conclusion

This mini-project significantly enhanced my data analysis skills using Microsoft Excel. I practiced essential steps such as data cleaning, transformation, and visualization, which are critical for extracting meaningful insights from raw data. Working with a real-world dataset like the FIFA World Cup statistics allowed me to see how analytical techniques can be applied beyond theory.

Through creating dashboards and answering specific performance-related questions, I learned how to structure data in a way that supports clear decision-making. This hands-on experience not only improved my technical proficiency in Excel but also deepened my appreciation for how data-driven insights can be used to solve business problems, evaluate performance, and support strategic decisions.

### 13. References

FIFA World Cup 2022 Official Statistics – https://www.fifa.com

Microsoft Excel Documentation – https://support.microsoft.com/excel

Data Visualization in Excel Tutorials – Excel Campus, YouTube Channel

Chandoo.org – Excel Dashboards & Data Analysis Resources