**FIFA 21 Data Analysis Project**

**Overview**

This project analyzes FIFA 21 player data to gain insights into player attributes, club performances, and financial aspects of football players. The analysis includes data cleaning, transformation, and the creation of a star schema data warehouse for efficient querying and reporting.

**Project Structure**

The project consists of:

Data cleaning and preprocessing (Jupyter Notebook)

Star schema data warehouse implementation

SQL queries for analytical insights

Visualizations of key findings

**Key Features**

Data Cleaning

Handled missing values and data inconsistencies

Converted height and weight measurements to standard units

Parsed financial values (wages, transfer values) from string formats

Standardized club names and player positions

**Data Warehouse**

Implemented a star schema with:

Fact table: PlayerPerformance

Dimension tables: PlayerDim, ClubDim, NationalityDim, PositionDim, TimeDim

Established proper relationships with foreign keys

Loaded data into SQLite database

**Analysis Highlights**

Club performance comparisons by average player ratings

Wage distribution by nationality and position

Player attribute trends over time

Financial efficiency analysis (wage-to-rating ratios)

**Key Findings**

**Top Clubs by Average Rating**:

Bayern Munich leads with an average player rating of 81.48

Real Madrid follows with 79.77

Inter and Napoli complete the top 4

**Wage Distribution**:

English defenders have the highest total wages

Spanish midfielders rank second in wage expenditure

Brazilian midfielders round out the top 5

**Position Analysis**:

Defenders have the highest average rating (65.94)

Midfielders follow closely (65.85)

Goalkeepers have the lowest average rating (64.65)

**Technologies Used**

Python (Pandas, NumPy)

Jupyter Notebook

SQLite

SQLAlchemy

Data visualization libraries (Matplotlib, Seaborn)