Project Proposal: FIFA 21 Player Analysis Using Python

Introduction
This project focuses on analyzing player performance and market value using data from FIFA 21. Football is a global sport, and understanding player attributes is critical for team managers, analysts, and fans. Using the FIFA 21 dataset, we aim to explore trends in player skills, identify factors that drive market value, and predict future potential.

Research Questions

1. What attributes most significantly impact a player's overall rating (OVA)?
2. How do physical, technical, and mental attributes influence a player's market value and wage?
3. Are there noticeable trends in player characteristics across positions, such as attackers, defenders, and goalkeepers?
4. Can we predict a player's future potential (POT) or growth based on their current statistics?

Data Description

The dataset contains detailed information on FIFA 21 players, including:

- **Personal Details**: Name, Age, Nationality, Club, Position.
- **Performance Metrics**: Overall rating (OVA), Potential (POT), and Growth.
- **Skill Attributes**: Dribbling, Passing, Finishing, Ball Control, etc.
- **Physical Attributes**: Acceleration, Speed, Strength, and Stamina.
- **Economic Details**: Value, Wage, and Contract Details.

This diverse dataset provides a solid foundation for analysis and machine learning models.

Proposed Analyses

1. **Exploratory Data Analysis (EDA)**:
    - Generate basic statistics to understand player attributes.
    - Visualize player characteristics by position and nationality.
    - Investigate how attributes vary with age or club.
2. **Correlation Study**:
    - Analyze relationships between performance metrics (e.g., OVA, POT) and market value.
    - Examine how wages are influenced by attributes like strength, agility, or skill.

3. **Clustering**:
    - Group players into categories (e.g., speedsters, playmakers) based on their attributes.
4. **Predictive Modeling**:
    - Build a model to predict player market value.
    - Create a second model to estimate growth or future potential based on current stats.
5. **Position-Based Insights**:
    - Compare key metrics for positions such as defenders, midfielders, and attackers.
    - Evaluate skill sets specific to roles like goalkeepers or forwards.

Shiny Dashboard for Python

To make the results accessible and interactive, we will build a Python-based dashboard using Shiny for Python. This tool will:

- Allow users to explore player stats interactively.
- Provide predictions for market value or growth potential based on selected features.
- Show comparison charts and clustering results by position or nationality.

Data Preparation

1. **Cleaning**:
    - Address missing or inconsistent data, such as incomplete wage or value entries.
    - Ensure uniformity in categorical variables like position and nationality.
2. **Feature Engineering**:
    - Create new metrics, such as performance-to-wage ratio.
    - Group players by position and age range for more detailed insights.
3. **Visualization**:
    - Use libraries like seaborn and matplotlib for clear, intuitive visualizations.

Expected Outcomes

1. **Insights**:
    - Identification of key drivers for market value and overall performance.

    - Position-based skill archetypes for players.
    - Trends in growth potential across age groups.
2. **Predictive Models**:
    - A regression model for player market value prediction.
    - A growth potential predictor.
3. **Interactive Dashboard**:
    - A user-friendly interface for exploring player data and predictions.

Deliverables

- **Data**:
    - Raw and cleaned datasets.
- **Code**:
    - Scripts for cleaning, analysis, and modeling.
- **Dashboard**:
    - An interactive web application for exploration and predictions.
- **Final Report**:
    - A comprehensive document summarizing the analysis and findings.
- **Presentation**:
    - A group presentation with highlights and a live dashboard demo.