### **Final Report: Analysis of FIFA 21 Player Dataset**

### **Introduction**

#### **Research Question:**

The central research question for this analysis is:

* **"What physical attributes and skills correlate most strongly with the market value of a soccer player?"**

Understanding the relationship between physical attributes, skill ratings, and a player's market value is crucial for clubs, scouts, and analysts. These insights can guide talent identification, improve player evaluation methods, and optimize decision-making in player acquisitions. In football, a player's value is influenced by a variety of factors, such as age, skills, performance, and reputation. This research aims to explore these correlations, particularly focusing on the factors that drive market value.

#### **Context of the Dataset:**

The **FIFA 21 Player Dataset** provides detailed information on soccer players from the FIFA 21 video game, reflecting their physical attributes, skills, and market values. This dataset includes over 18,000 players, with attributes such as:

* **Skills**: Ratings for key abilities like shooting, passing, dribbling, etc.
* **Physical Attributes**: Data on physical capabilities such as pace, strength, balance, and stamina.
* **Market Value**: The estimated transfer market value of players in euros.
* **Positions, Clubs, and Nationality**: The players on-field position, club, and nationality, among other metadata.

This dataset offers a snapshot of player performance and value at the time of **FIFA 21's release**, which makes it an excellent resource for analyzing key drivers of a player's market value.

### **Data Preprocessing and Exploratory Data Analysis (EDA)**

#### **Data Cleaning:**

To ensure accurate analysis, the dataset was cleaned and preprocessed. The following steps were taken:

1. **Missing Data Handling**: Rows with missing values in essential columns such as market\_value, age, and overall were removed to prevent skewing the analysis. This ensured that only complete and reliable data was used.
2. **Conversion of Market Value**: The market value of players (given in millions, thousands, or billions of euros) was converted to a consistent numeric format for easier analysis. Values with suffixes like ‘M’ (million) and ‘K’ (thousand) were standardized to their numeric equivalents.
3. **Categorical Variables**: For the purpose of analysis, categorical variables such as player position were handled by extracting the main position from multi-position data, and the data was encoded accordingly.

#### **Exploratory Data Analysis (EDA):**

During the EDA phase, we explored key patterns in the dataset, including distributions, correlations, and relationships between the players' attributes and market value. The following insights were gathered from the visualizations:

1. **Market Value Distribution**: The distribution of market values was highly skewed, with most players having lower market values. Only a small proportion of players (superstars like Messi and Ronaldo) had exceptionally high market values, which is typical in professional football.
2. **Correlations**: A **correlation heatmap** revealed that **physical attributes** such as **pace** and **dribbling** showed moderate positive correlations with **market value**. Other attributes like **shooting** and **passing** also had notable correlations, but not as strong as the aforementioned skills.
3. **Position and Market Value**: A **boxplot** showed that attacking players (e.g., forwards, wingers) generally had higher market values than defensive players (e.g., defenders, goalkeepers), which suggests that goal-scoring and attacking contributions are more highly valued in the market.
4. **Pace and Dribbling**: Scatter plots showed that players with higher **pace** and **dribbling** ratings tended to have higher market values. However, the correlation was not perfect, indicating that other factors like reputation and club performance also influence market value.

### **Methodology and Analysis**

#### **Statistical Analysis:**

Several statistical methods were used to identify relationships between attributes and market value. Key steps included:

* **Correlation Analysis**: Pearson's correlation coefficient was used to identify linear relationships between numerical variables like pace, shooting, passing, dribbling, and market value. This helped identify the most influential attributes in determining a player's worth.
* **Regression Analysis**: A linear regression model was employed to predict a player's market value based on their physical and skill ratings. Key predictors such as **pace**, **dribbling**, **shooting**, and **passing** were found to have moderate predictive power.

#### **Visualization:**

Various visualizations were used to better understand the data, including:

* **Histograms and Box Plots**: These visualized the distribution of player market values and helped identify outliers and trends.
* **Scatter Plots**: These plots provided insights into the relationships between individual attributes (e.g., pace, dribbling) and market value, confirming the importance of these attributes in player valuation.
* **Radar Charts**: A radar chart was used to compare key skills (dribbling, shooting, passing, etc.) between players, helping to highlight the strengths and weaknesses of top players (e.g., Messi vs Ronaldo).

### **Findings and Interpretation**

Based on the analysis, the following findings were made:

1. **Strong Correlations Between Market Value and Skills**: Attributes such as **pace**, **dribbling**, and **shooting** showed the most significant correlations with market value. These physical and technical attributes are crucial in determining a player's worth in the transfer market, especially for attacking players.
2. **Attacking Positions Have Higher Market Value**: The analysis confirmed that players in attacking positions, such as forwards and wingers, generally have higher market values compared to defenders and goalkeepers. This aligns with real-world football economics, where goal-scoring and playmaking abilities are more highly prized.
3. **Age and Experience Factor**: Although not directly analyzed in detail, age plays an important role in market value. Younger players with high potential and skill ratings tend to have higher market values, reflecting the desire for long-term investment.
4. **Physical Attributes Matter, But Not Alone**: While **pace** and **dribbling** correlate positively with market value, they are not the sole determinants. Other factors, such as a player's overall reputation, consistency, and position, contribute significantly to a player's worth.

### **Conclusion and Recommendations**

#### **Conclusion:**

The analysis of the FIFA 21 Player Dataset has provided clear insights into the key factors that determine a player's market value. While physical attributes like **pace** and **dribbling** are strong predictors, it is clear that a combination of **technical skills**, **position**, and **market visibility** drive a player's worth. Players in attacking roles with higher dribbling and pace are generally valued higher than defenders and goalkeepers.

#### **Recommendations:**

1. **For Scouting and Player Acquisition**:
   * Focus on acquiring **attacking players** with strong **dribbling** and **pace**. These attributes appear to correlate strongly with market value and are crucial for success in modern football.
   * **Versatility** is also important. Players who can play in multiple attacking roles or adapt to different team strategies tend to be more valuable.
2. **For Team Management**:
   * If your team needs to improve its **attack**, focus on signing players who are strong in **dribbling** and **shooting**, as these players are likely to be high-impact contributors.
   * For **defensive roles**, while the market value may be lower, selecting players with strong **tackling** and **positioning** skills remains crucial for building a balanced squad.
3. **For Player Development**:
   * Emphasize improving technical attributes such as **dribbling** and **passing** for young players. These skills are highly correlated with a player’s market value and long-term career success.
   * Ensure that players develop **physical attributes** like **pace** and **strength**, particularly for attacking roles where these attributes provide an edge.

Messi > Ronaldo