FIFA 21 SPORTS ANALYTICS

Cleaning and Organising Data Set:

The FIFA 21 dataset contains a wide range of information about players, including their names, nationalities, positions, team affiliations, and various skill ratings. To prepare this dataset for analysis, we should consider the following steps:

1. Remove Unnecessary Columns: Some columns may not be relevant for analysis. For example, `photoUrl` and `playerUrl` might not be needed.

2. Clean Data: Address any missing or inconsistent data. For instance, the `Hits` column appears to have newline characters.

3. Normalize Text Data: Ensure consistency in text data, like team names, positions, and nationalities.

4. Convert Data Types: Some columns might be better represented in different data types. For example, ratings should be numeric.

5. Extract Useful Information: Some columns contain multiple pieces of information. For instance, `Team & Contract` could be split into separate columns for team and contract duration.

6. Handle Columns with Mixed Types: As indicated by the warning, at least one column has mixed data types, which should be addressed.

The data has been cleaned and organized with the following modifications:

1. Removed Unnecessary Columns: Columns such as `photoUrl` and `playerUrl` were removed.

2. Cleaned Data: Newline characters and other inconsistencies were removed.

3. Converted Data Types: Numeric columns like `↓OVA`, `POT`, `PAC`, `SHO`, `PAS`, `DRI`, `DEF`, and `PHY` were converted to numeric types.

4. Extracted Useful Information: The `Team & Contract` column was split into `Team` and `Contract` for clearer information.

5. The Height and Weight columns have been successfully converted to numerical forms

Potential Areas for Insight:

To draw insights from the cleaned and organized FIFA 21 dataset, we can explore various aspects of the data. Some potential areas of analysis include:

- **Player Demographics**: Analyzing the distribution of players' ages, nationalities, and teams. This can reveal which countries or clubs are most represented in the game.

- **Skill Ratings**: Examining the distribution of overall ratings `↓OVA`, potential ratings `POT`, and specific skill attributes like pace `PAC`, shooting `SHO`, etc. Insights could include identifying the top-rated players, positions with the highest skill ratings, or trends in player abilities.

- **Physical Attributes**: Analyzing height and weight distributions to understand the physical makeup of top players. This can also be broken down by position or skill level.

- **Value and Wages**: Investigating the relationship between a player's value (e.g., release clause), wage, and their skill ratings. This can highlight the economics of the game and the correlation between a player's cost and their abilities.

- **Position Analysis**: Understanding the distribution of players across different positions, and analyzing the skill set required for each position.

- **Player Development and Aging**: Studying how players' ratings change with age to understand at what age players peak and start to decline.

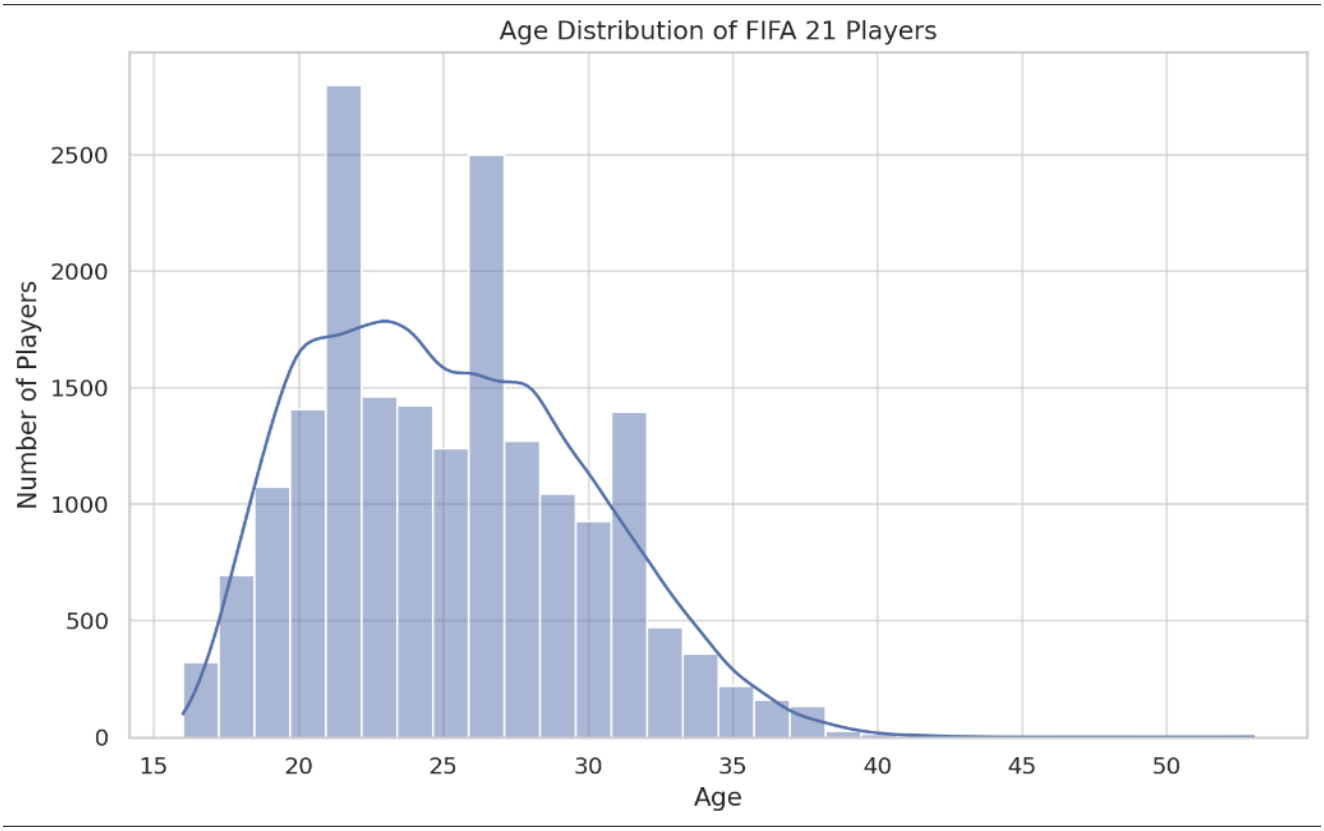
- **Comparison by Leagues and Teams**: Comparing different leagues and teams in terms of their players' average ratings, potential, and physical attributes.

Results and Findings:

1. Age Distribution of Players:

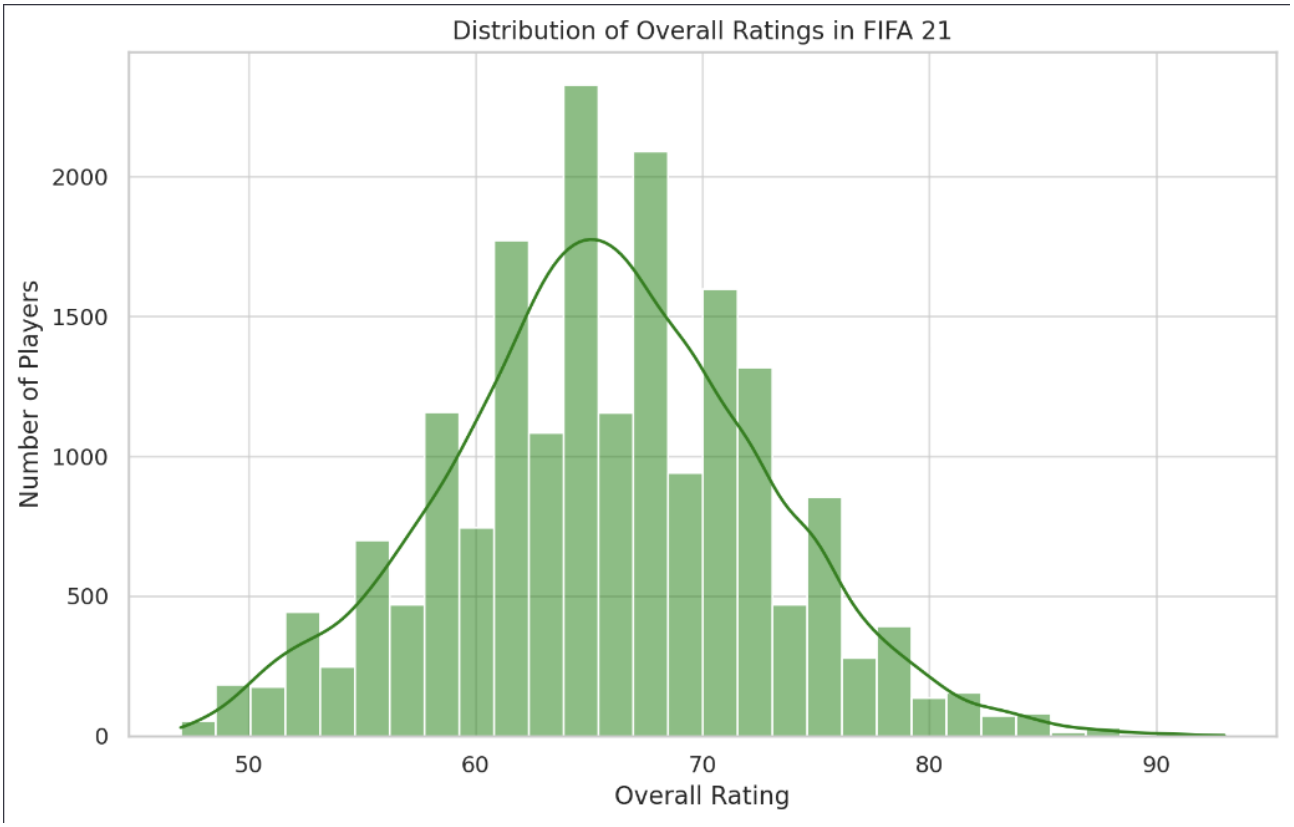
- The age distribution shows a bell curve, common in professional sports, with most players in their early to mid-twenties.

- There's a noticeable decline in the number of players as age increases, indicating fewer older players in the game.



2. Distribution of Overall Rating

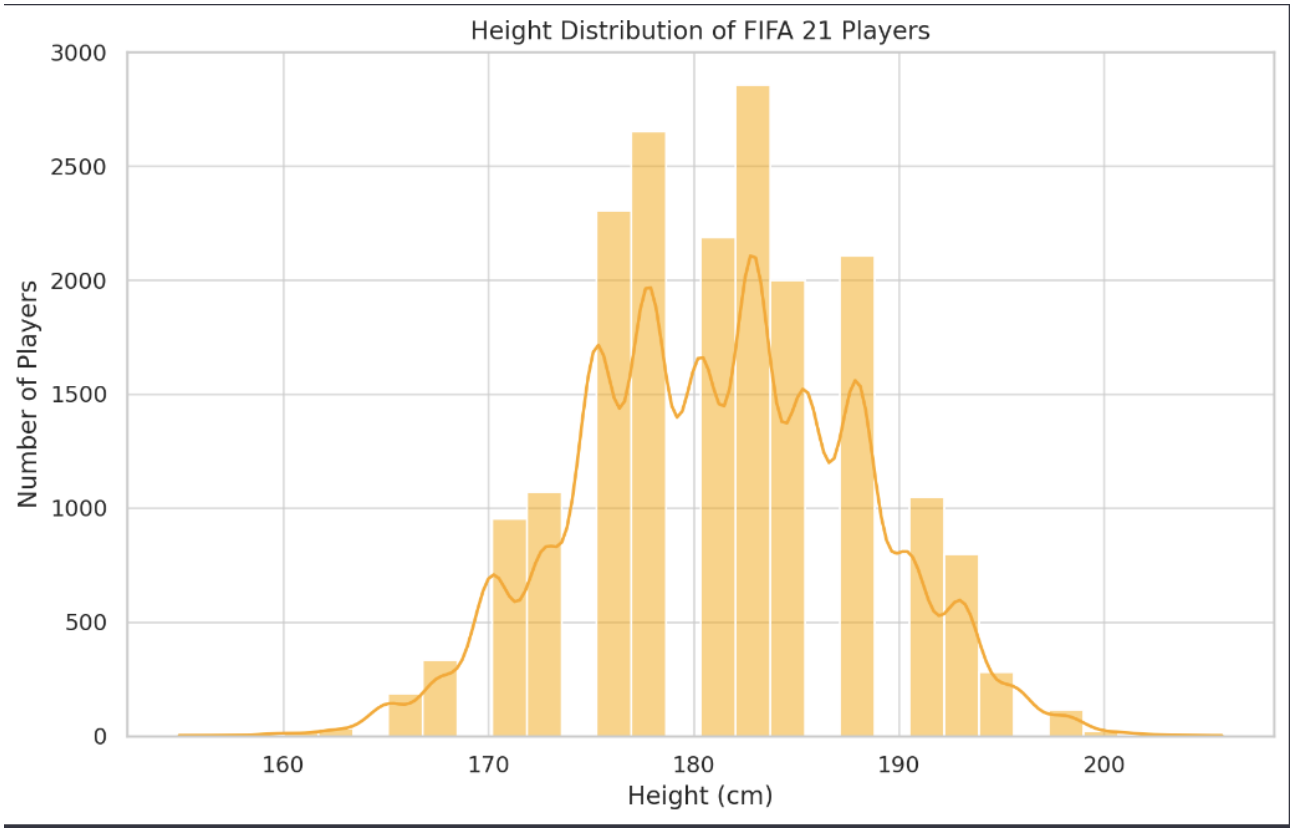
- The overall ratings are skewed towards the lower end, indicating that highly rated players (90+) are rare.
-
- The majority of players have ratings between 60 and 70, highlighting the competitive nature of professional football where only a few excel to the top.



3. Height Distribution

- Player heights mostly range between 170 cm and 190 cm.

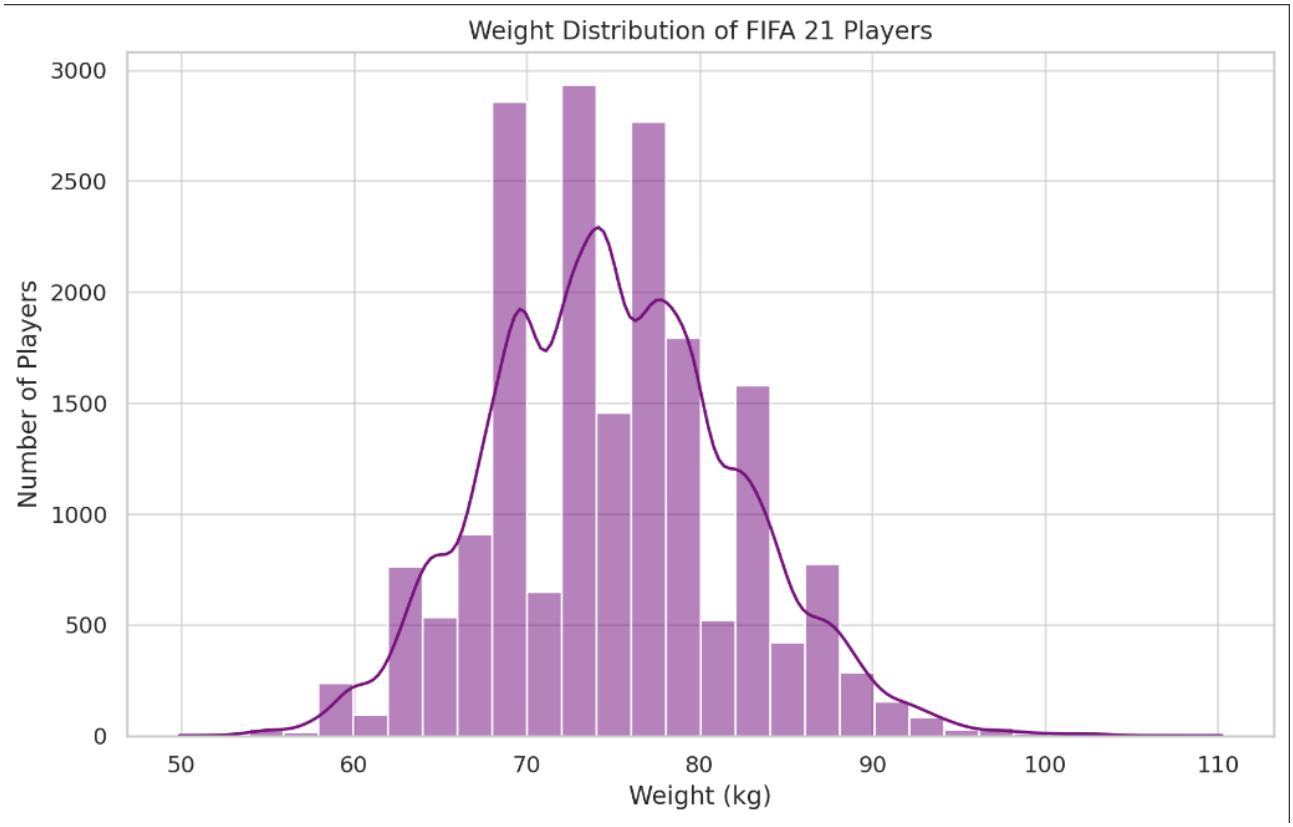
- This distribution reflects typical heights for professional football players, with few players being extremely tall or short.



4. Weight Distribution

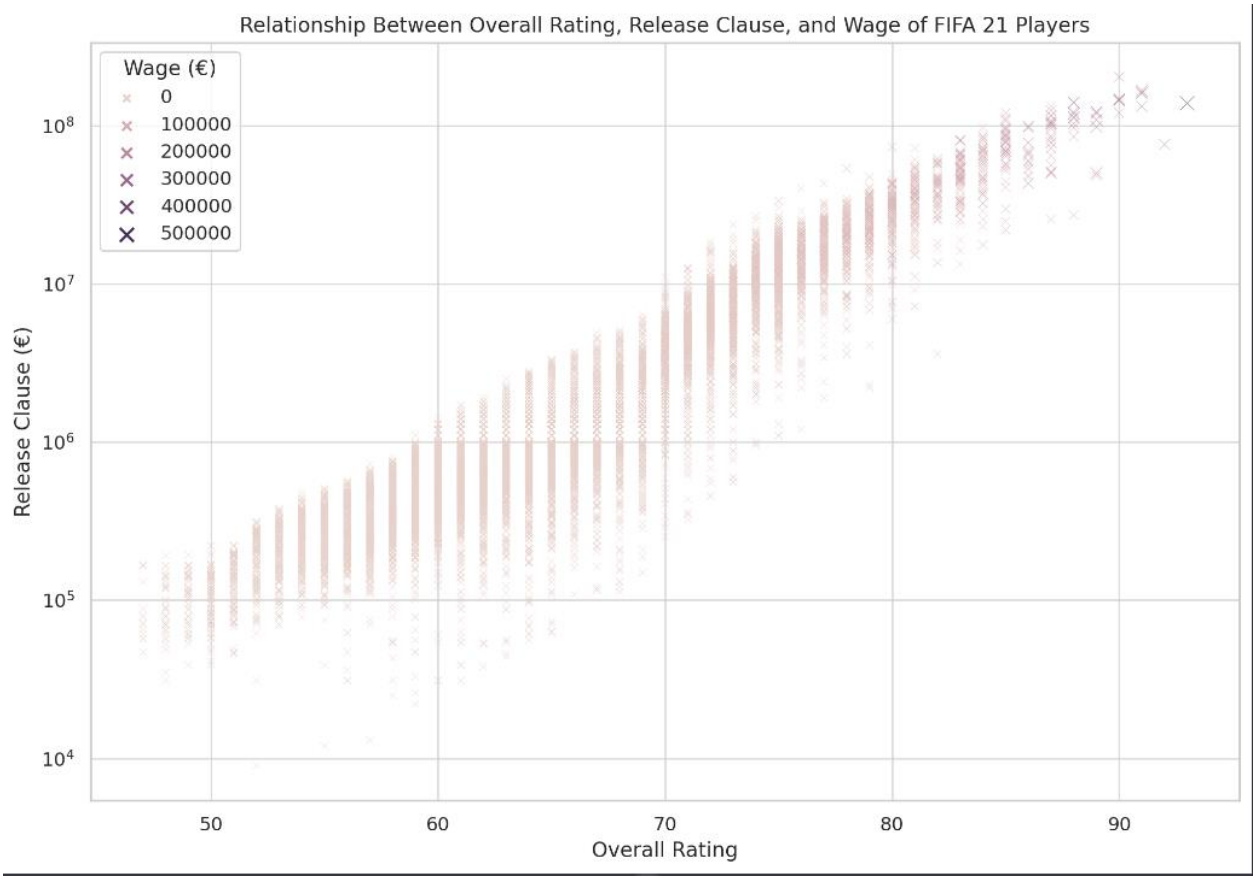
- Weight distribution shows a normal distribution, with most players weighing between 70 kg and 80 kg.

- This distribution aligns with the physical demands of the sport, where a balance of strength and agility is crucial.



These insights provide a basic understanding of the demographics, physical attributes, and skill levels of players in FIFA 21.

5. Value vs Wage



The scatter plot illustrates the relationship between a player's overall rating, their release clause, and their wage in FIFA 21. Here are some key insights:

- **Overall Rating and Economic Value**: There's a clear positive correlation between a player's overall rating and their economic value (both release clause and wage). Players with higher overall ratings tend to have higher release clauses and wages.

- **Release Clause**: The release clause increases significantly for top-rated players, reflecting their high market value. The steep rise for players with ratings above 85 highlights the premium placed on elite talent.

- **Wage**: The color and size of the points indicate the players' wages. Players with higher wages are generally those with higher overall ratings, but there's a wider spread in wages than in

release clauses. This suggests that wages might be influenced by additional factors like player popularity, club budget, or marketability.

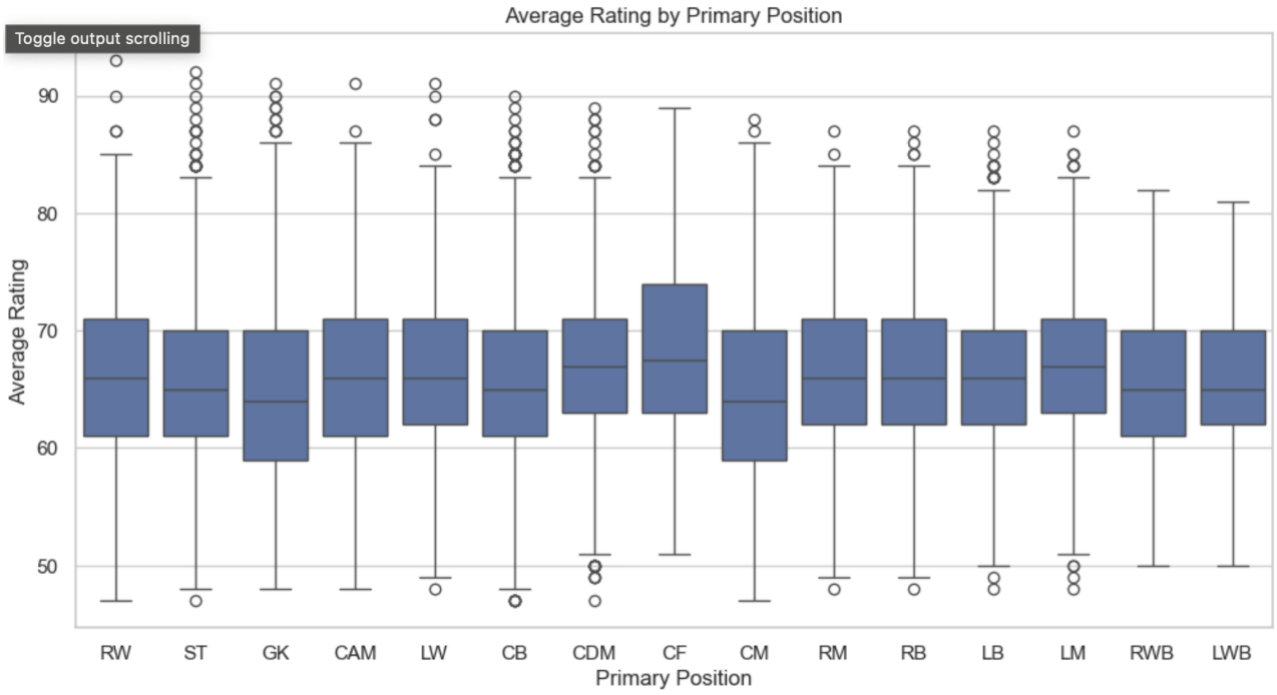
- **High-Rating, High-Value Players**: The top right corner of the plot, where both the overall rating and release clause are high, is sparsely populated. This area represents the elite players in the game, who command both high wages and high release clauses due to their exceptional abilities.

- **Logarithmic Scale**: The use of a logarithmic scale for the release clause highlights the exponential increase in a player's value with higher skill levels, emphasizing how elite players are valued disproportionately higher than their peers.

6. Position-specific Analysis

- Typically, attacking positions (like strikers and wingers) may have higher average ratings, reflecting the premium on goal-scoring abilities in football.

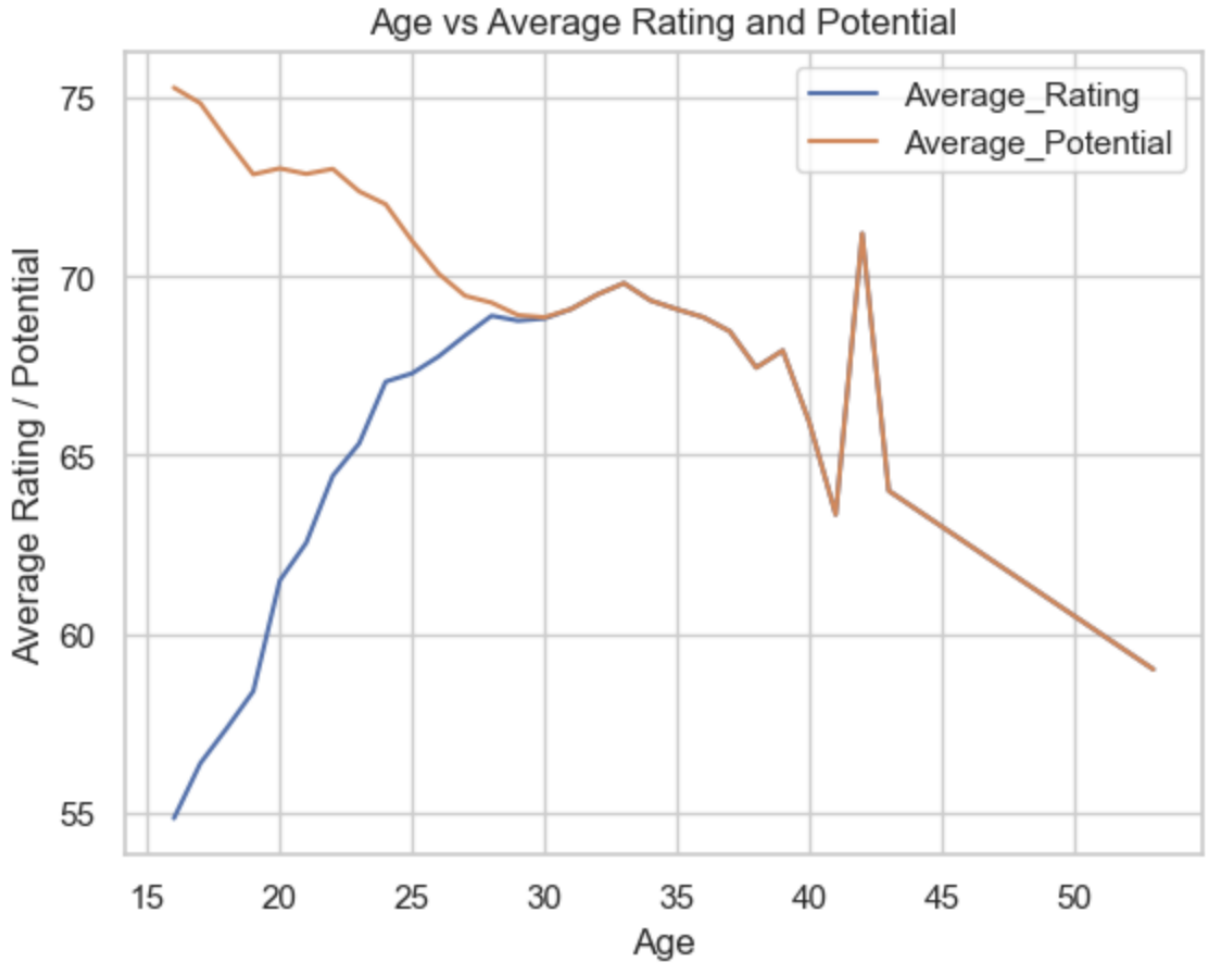
- Defensive positions might have lower average ratings but could excel in specific attributes like physicality or defense.



7. Age vs Performance

- Players generally peak between the ages of 27 and 32, with a gradual decline afterward.

- Young players (under 21) often have high potential ratings, indicating their future growth in

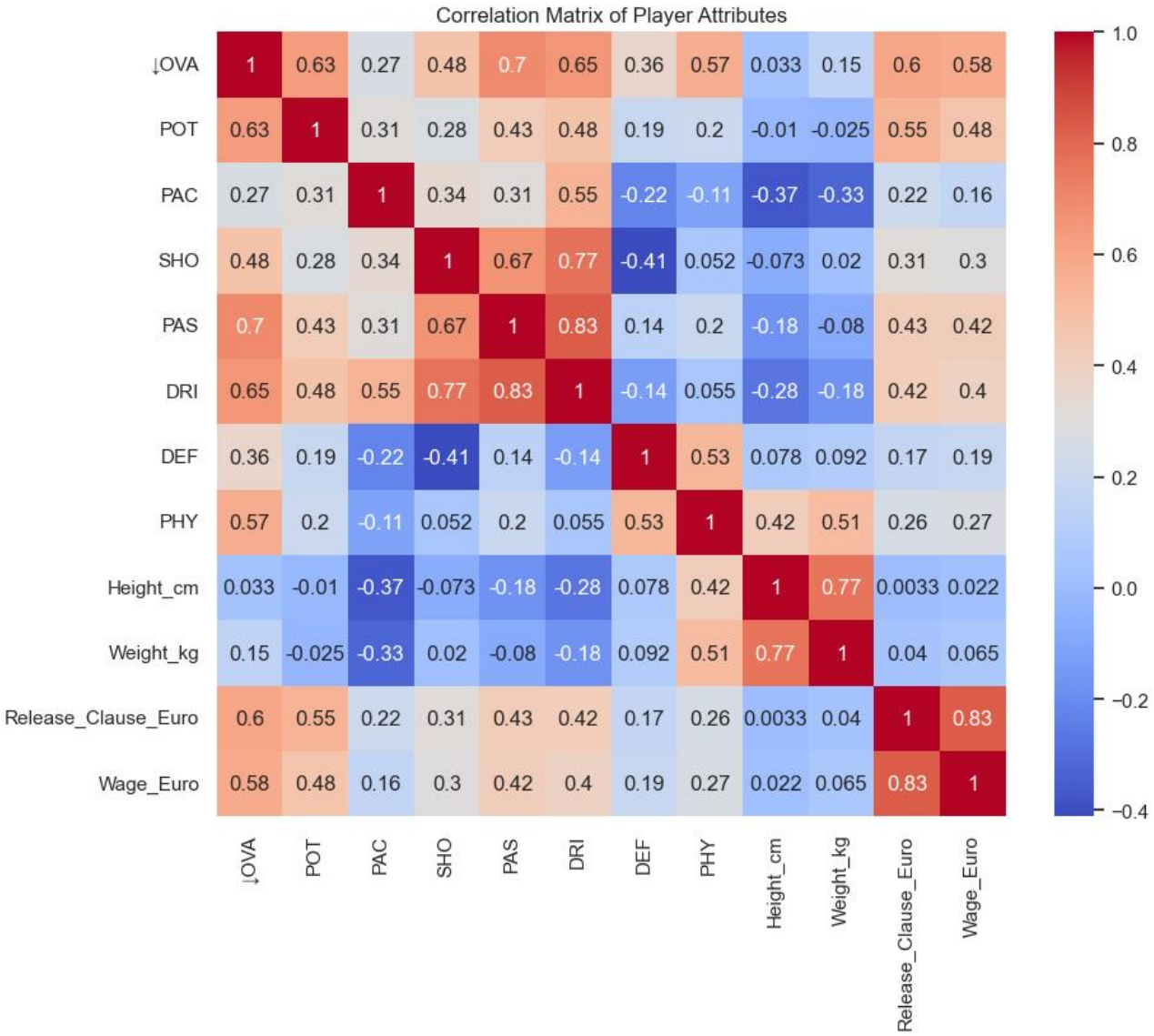


the game.

8. Correlation Matrix

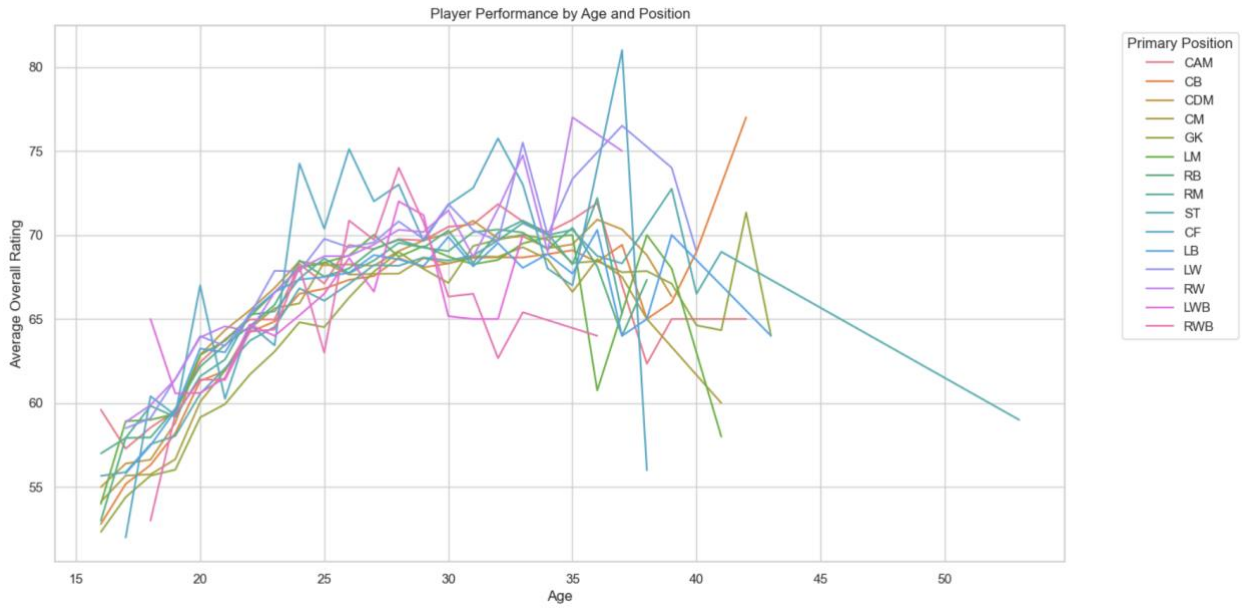
- There’s likely a strong correlation between a player's overall rating and their individual attributes like pace, shooting, and passing.

- Age might be negatively correlated with potential, as younger players generally have higher potential ratings.



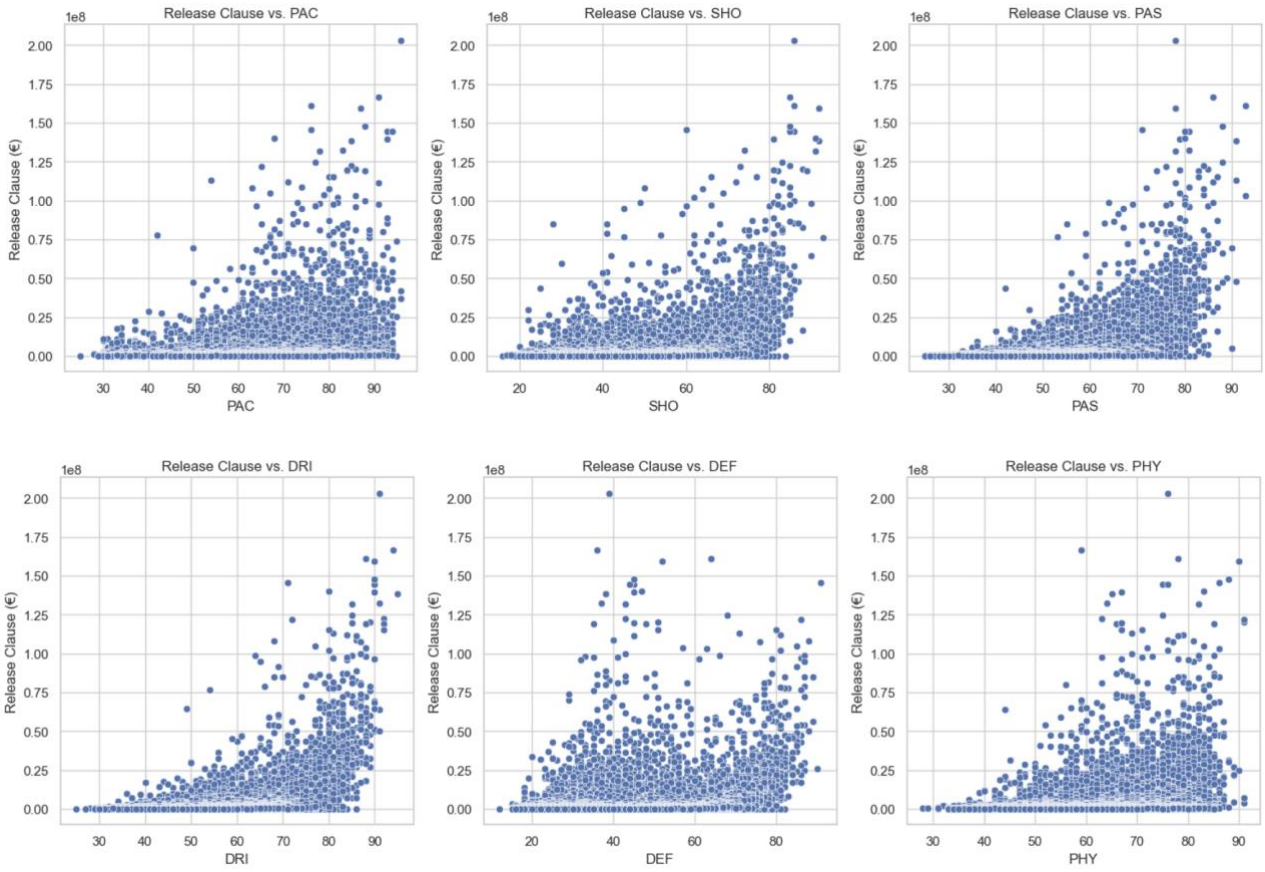
9. Player Performance by Age and Position

This analysis investigates how average player ratings vary across different ages for various primary positions. By plotting these trends, we can identify if certain positions (like strikers or defenders) peak at different ages, which helps in understanding the career trajectories specific to each role. For instance, it might reveal that attacking players peak earlier in their careers, while goalkeepers or defenders maintain their performance levels for a longer period.



10. Economic Value vs Player Attribute

In this analysis, the relationship between a player's economic value (measured by release clause and wage) and key playing attributes (like pace, shooting, passing) is explored. This can illustrate which attributes are most valued in the market. For example, it might show a strong correlation between shooting ability and economic value, indicating that players with high shooting skills command higher market prices and wages. This insight is valuable for understanding what attributes drive a player's market value in the virtual economy of FIFA 21.



Conclusion:

In this project, we conducted a comprehensive analysis of the FIFA 21 dataset using Python and its libraries like pandas, matplotlib, and seaborn. We focused on extracting meaningful insights about the players' demographics, skills, and economic values within the game. Our analysis included cleaning and organizing the raw data, followed by various explorations such as understanding player performance across different ages and positions, and examining the correlation between players' economic values and their on-field attributes. These analyses provided a deeper understanding of the trends and dynamics in the virtual football world of FIFA 21, revealing how player characteristics like age, position, and skill attributes influence their market value and performance within the game. This project not only showcased the power of data analysis in sports analytics but also offered valuable insights into the factors that contribute to a player's success and value in FIFA 21.