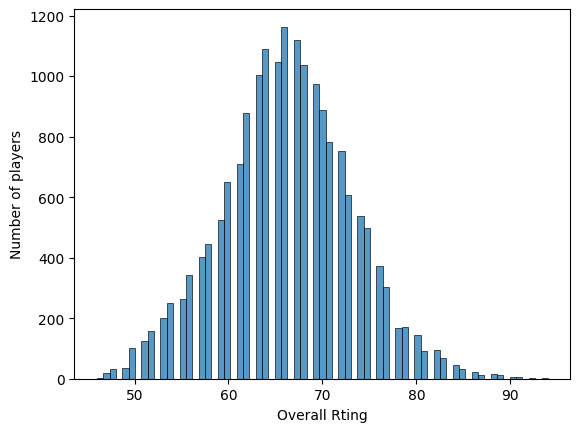
**FIFA Player Data: Cleaning & Analysis**

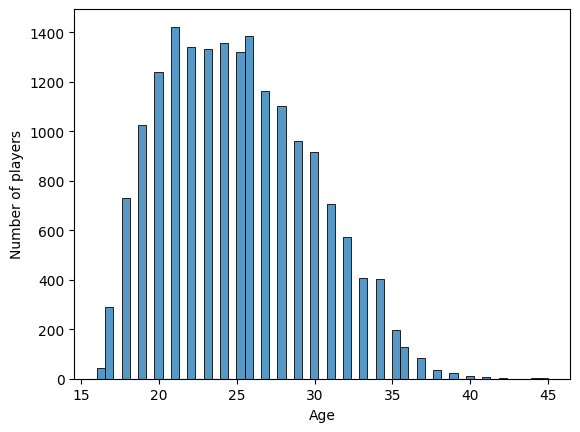
* **Project Objective**

Analyse the FIFA players dataset to extract meaningful insights about player attributes, clubs, and market values. Perform data cleaning and exploratory data analysis (EDA) to understand trends and patterns.

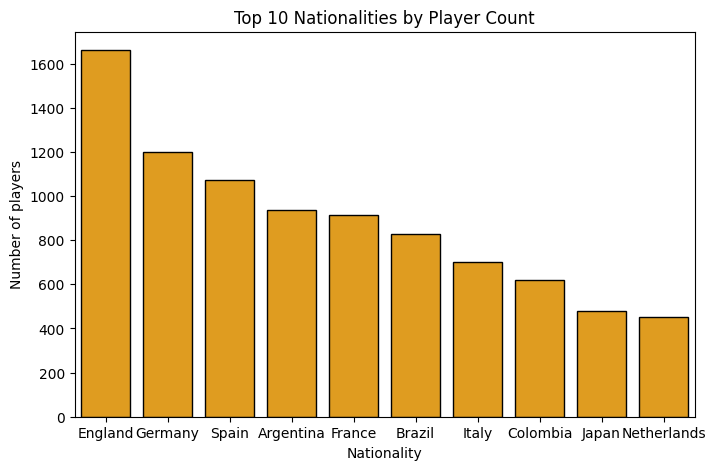
* **Data Description**
* **Rows:** 18,207
* **Columns:** 18
* **Key Columns:**
  + **ID**: Unique player ID
  + **Name**: Player name
  + **Age**: Player’s age
  + **Nationality**: Country of the player
  + **Overall**: Overall rating (1-100)
  + **Potential**: Future potential rating
  + **Club**: Current club
  + **Value**: Player’s market value (in thousands)
  + **Wage**: Weekly wage (in thousands)
  + **Preferred Foot**: Left/Right
  + **International Reputation**: Rating (1-5)
  + **Skill Moves**: Rating (1-5)
  + **Position**: Player’s position
  + **Joined**: Year joined current club
  + **Contract Valid Until**: Contract expiry date
  + **Height & Weight**: Physical stats
  + **Release Clause**: Buy-out clause amount
  + **Libraries Used**
* **pandas** – Data manipulation
* **NumPy** – Numerical operations
* **matplotlib**  – Visualization
* **seaborn** – Visualization
* **Exploratory Data Analysis (EDA)**
* **Exploratory Data Analysis (EDA) is the process of analysing datasets to summarize their main characteristics, often using visual methods. The goal is to understand the structure, detect patterns, find anomalies, check assumptions, and identify relationships among variables before applying any modelling techniques.**
* **Data Cleaning**
* **Checked if there any missing values.**
* **Fill the missing values with (Not Specified)**
* **Checked if there any duplicates present or not**
* **Data Visualization**
* **What is the distribution of players based on their overall rating?**



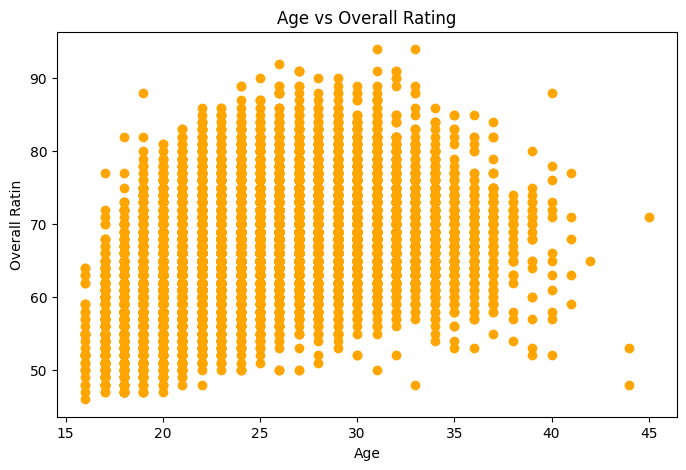
* **The distribution of player ratings is approximately normal (bell-shaped).**
* **Most players have an overall rating between 60 and 70, which is the centre of the distribution.**
* **The peak occurs around 65, meaning the majority of players have an overall rating near this value.**
* **Very few players have extremely low ratings (<50) or very high ratings (>85), indicating these are rare.**
* **The spread shows that average players dominate the dataset, while elite players are limited.**
* **What is the age distribution of FIFA players ?**

****

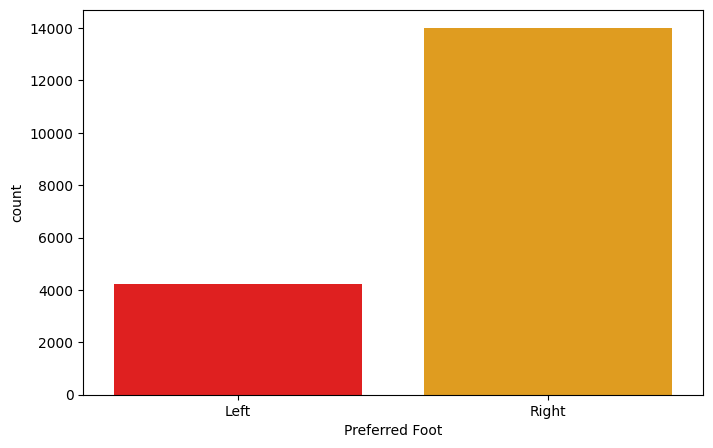
* **The histogram shows the distribution of player ages.**
* **Most players are between 20 and 28 years old, with a peak around 21–24 years.**
* **There are very few players above 35 years old or below 18 years old.**
* **Which 10 nationalities have the highest number of players in the FIFA dataset?**



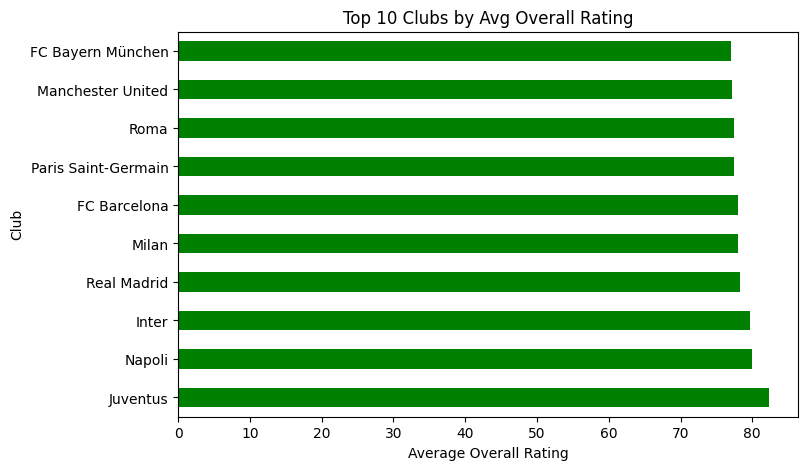
* **England has the most players by a large margin, likely because the English league system (Premier League, Championship, etc.) has many teams and players. European and South American countries dominate the top 10.**
* **What is the relationship between a person's age and their overall rating?**

****

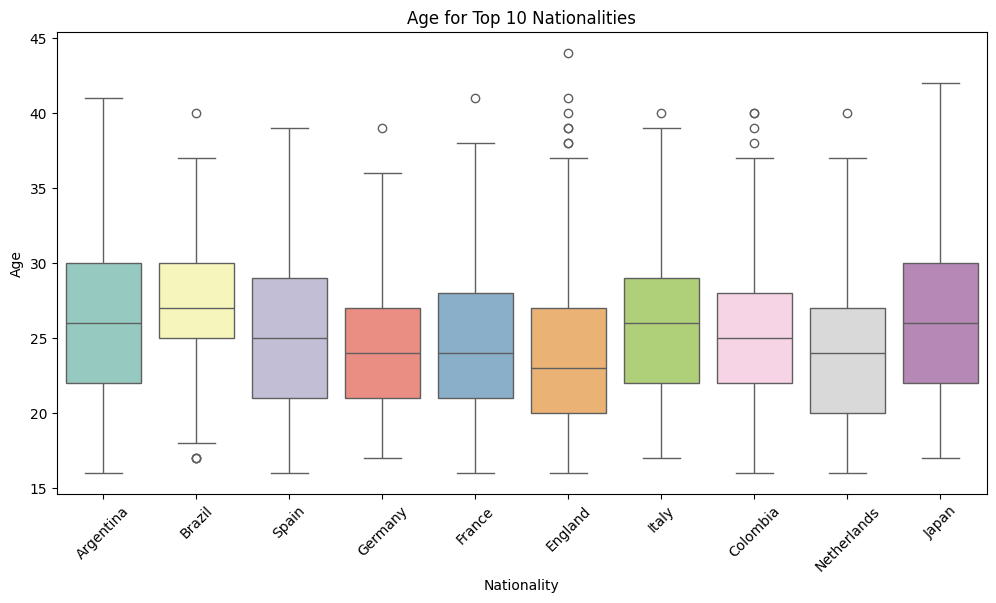
* **Age and Overall Rating have a non-linear relationship.**
* **The highest overall ratings appear to be concentrated in the late 20s to early 30s age range. This is where the scatter plot shows a dense cluster of points with high values on the y-axis.**
* **Ratings tend to be more spread out for younger and older individuals, with a noticeable decline in the number of high ratings as age increases beyond the early 30s.**
* **What is the distribution of FIFA players based on their preferred foot (Left vs. Right)?**

****

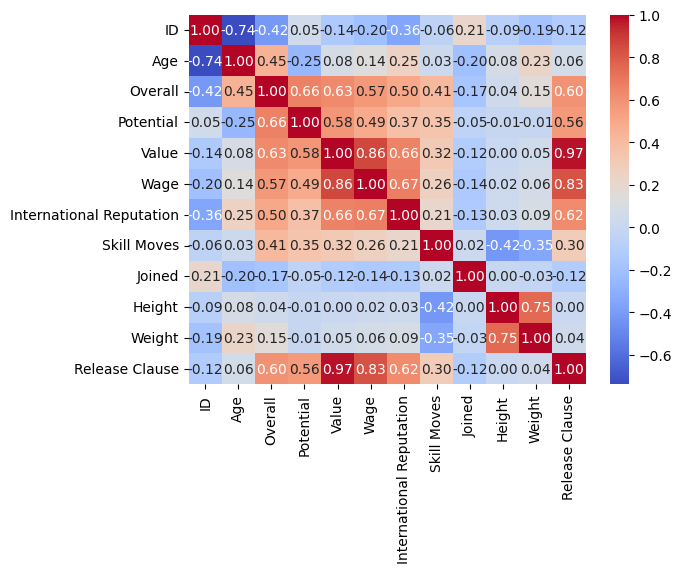
* **The bar chart shows that the majority of players prefer their Right foot, with around 14,000 players using it. In contrast, about 4,000 players prefer their Left foot.  
  This indicates that right-footed players dominate FIFA datasets, while left-footed players form a much smaller group.**
* **Which are the top 10 football clubs by average overall player rating in the dataset?**

****

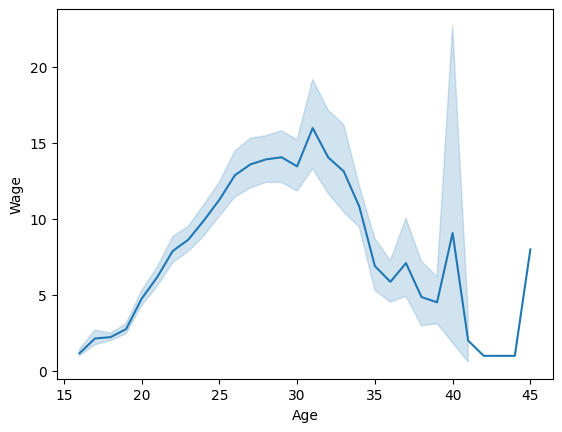
* **Juventus has the highest average rating, followed closely by Napoli and Inter.**
* **Other top clubs include Real Madrid, Milan, FC Barcelona, Paris Saint-Germain, Roma, Manchester United, and FC Bayern München.**
* **Overall, these clubs maintain an average player rating between 78–82, indicating a high concentration of skilled players.**
* **What is the age distribution of players among the top 10 nationalities in the dataset?**

****

* **Most players across all countries fall within the 20 to 30 age range, with medians typically around 24–26 years.**
* **Some outliers exist in older age groups (35–40+ years), representing veteran players.**
* **The spread (interquartile range) is quite similar across most nationalities, suggesting a consistent age profile in professional football.**
* **England and the Netherlands appear to have slightly younger distributions compared to Argentina and Japan, which show a wider spread including older players**
* **What are the correlations among key numeric features (e.g., Overall, Potential, Value, Wage, Age, Height, Weight) in the FIFA player dataset?**

****

* **Overall vs Potential → Strong positive correlation (0.66): Higher-rated players usually also have higher potential.**
* **Value vs Overall (0.63) and Value vs Potential (0.58) → Player market value depends heavily on both overall rating and potential.**
* **Wage vs Overall (0.67) → Wages are strongly linked to player overall performance.**
* **Release Clause vs Value (0.97) → fAlmost perfect correlation, since release clauses are typically derived from market value.**
* **International Reputation vs Wage (0.67) → More reputed players earn higher wages.**
* **Height vs Weight (0.75) → Expected strong positive correlation, as taller players tend to weigh more.**
* **Age vs Overall (-0.25) → Slight negative relationship, suggesting younger players may sometimes have higher ratings, while aging lowers overall.**
* **Weak or no correlation between physical traits (Height/Weight) and skills (Overall, Potential, Skill Moves).**
* **How does a football player’s wage vary with age?**

****

* **Ages 16–20 → Wages start very low and gradually increase.**
* **Ages 20–30 → Wages rise steadily, peaking around 28–30 years old, which represents players’ prime years.**
* **Ages 30–35 → Wages begin to decline, as players move past their peak performance.**
* **Ages 35–40+ → Wages drop sharply, with only a few exceptional cases earning high wages (outliers, likely superstar veterans).**
* **Conclusion**

**The FIFA dataset (18,207 players) highlights that most players are average-rated (60–70), with peak performance and wages in the late 20s to early 30s. Football is dominated by young athletes (20–28 years), mainly right-footed, with England contributing the most players due to its large league system. Elite clubs such as Juventus, Real Madrid, Barcelona, PSG, and Bayern Munich host the highest-rated talent. Player value, wages, and reputation are strongly correlated, with market value and release clauses almost identical. Overall, football economics and performance are tightly linked to age, peak years, and club prestige.**