**To**: Managing Director, Shawn Janzen

**From**: Tazvitya Aubrey Chihota Analytics

**Date:** October 14, 2024

**Subject**: Analysis of FIFA Player Ratings: Impact of Age, Potential, Value, and Preferred Foot on Overall Ratings (2014)

**Level of Alpha** = 0.01

**Introduction:** This analysis seeks to assess the relationship between the age, potential, value, and preferred foot of FIFA players on their overall ratings for the year 2014. The objective is to understand how these factors influence players’ overall rating.

**Hypothesis Test**: To discover if there is a relationship Between Player Attributes and Overall Ratings

**Target Population**: All FIFA players who joined in 2014 and whose data is available for analysis.

**Spoiler findings**: The analysis reveals a strong, significant relationship between players' age, potential, value (in millions), and preferred foot on their overall ratings. As these independent variables increase, the overall rating is positively impacted. The regression model explains over 94% of the variability in player ratings, confirming that these variables are strong predictors of overall performance.

|  |  |  |
| --- | --- | --- |
| **Table 1: Preferred Foot Summary** | | |
|  | Freq. | Percent |
| Right | 224 | 78.60 |
| Left | 61 | 21.40 |
| Total | 285 | 100% |

We used a FIFA dataset that includes player attributes in 2014, focusing on the year 2021 with no missing values. Age (IV) represents the players' ages, while Potential (IV) indicates the players' possible future performance. Value (IV) in millions reflects the market value of players, and Preferred Foot (IV) categorizes players based on their dominant foot for play, Overall Rating (DV) the overall performance rating of the player. The Preferred Foot variable is categorical and is treated as a dummy variable in the analysis. It has been converted into a dummy variable format, where the Right foot is used as the reference group. The distribution shows that the **right-footed dominate** the dataset, with nearly 80% of players preferring their right foot.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 2: univariate PQ table for Age, value, overall and Potential** | | | | | | |  |
|  | Freq | Min | Max | Med. | Mean | SD | Skewness |
| Overall | 285 | 54 | 92 | 71 | 70.85 | 6.74 | 0.23 |
| Age | 285 | 19 | 39 | 29 | 28.83 | 4.42 | -0.04 |
| Potential | 285 | 54.00 | 93.00 | 72.00 | 72.80 | 6.37 | 0.28 |
| Value | 285 | 0.00 | 119.50 | 1.50 | 5.97 | 15.05 | 5.02 |

Around **one-fifth of the players** prefer their left foot, which is a smaller but still notable portion of the group.

The 'Value' variable, representing a player's market value, has a mean of 5.97 million euros and a standard deviation of 15.05 million euros, indicating substantial variability. The high skewness of 5.02 suggests a significant right skew, meaning a small number of players have exceptionally high market values, pulling the average up. This is further evidenced by the median value being much lower than the mean, highlighting the influence of these high-value outliers. Most variables exhibit slight skewness and moderate variability, except for 'Value,' which shows significant right skewness and higher variability. This suggests that while characteristics like 'Overall,' 'Age,' and 'Potential' are relatively normally distributed, 'Value' is influenced by a few high outliers, leading to a skewed distribution

**Regression Equation: Overall Rating = -11.828 + 0.554 × (Age) + 0.0.914 × (Potential) + 0.038 × (Value) -0.274 × (Preferred FootLeft) + e**

**Intercept (-11.8276**): This represents the predicted Overall rating when all predictor variables (Age, Potential, Value, Preferred Foot) are equal to zero.

**Age (0.5544):** For each additional year of age, the Overall rating is predicted to an average increase of 0.554 points, assuming all other variables remain constant. The small p-value (< 2e-16) indicates this effect is statistically significant**.**

**Potential (0.9138):** For each one-unit increase in a player's potential rating, the "Overall" rating is predicted to increase by 0.9138 points. This variable has a very strong relationship with the dependent variable, given its high coefficient and significance p <0.01.

**Value (0.0385):** For each unit increase in the player market value, the Overall rating is expected to increase by an average of 0.0385 points. This effect is statistically significant at p < 0.01

**Preferred.FootLeft (-0.2744):** players who prefer the left foot have an average decrease of 0.274 in Overall rating than right-footed players holding other variables constant. However, the p-value (0.272) suggests this effect is not statistically significant, meaning there is no strong evidence that the preferred foot influences a player's overall rating in this model.

The coefficients for Age, Potential, and Value are statistically significant, with p-values < 0.01, indicating a meaningful increase in overall ratings with increases in these variables. The model and independent variables showed statistical significance (F = 698.1, p < 2.2e-16), illustrating a robust relationship. The R-squared value of 0.93 indicates that the model explains 93% of the variability in overall ratings, confirming a strong fit. For a summary of regression results see **table 3.**

Given the statistical significance of the independent variables, we reject the null hypothesis and confirm a positive relationship between Age, Potential, Value, and Preferred Foot with Overall Rating. The effect sizes suggest substantial real-world implications for player evaluations in FIFA, particularly for players with higher potential and market value Furthermore the calculated power of the regression analysis was found to be approximately 1.0. This indicates that the regression model has a very high probability of detecting an effect, should one exist, suggesting that the sample size is more than adequate for the analysis conducted..

Diagnostic plots revealed several key assumptions were violated such as **Non-linearity**: The Residuals vs Fitted plot suggests some non-linearity.

**Heteroscedasticity**: Both the Residuals vs Fitted and Scale-Location plots show signs of non-constant variance, suggesting heteroscedasticity.

**Non-normality**: The Q-Q plot shows a deviation from normality, especially in the tails.

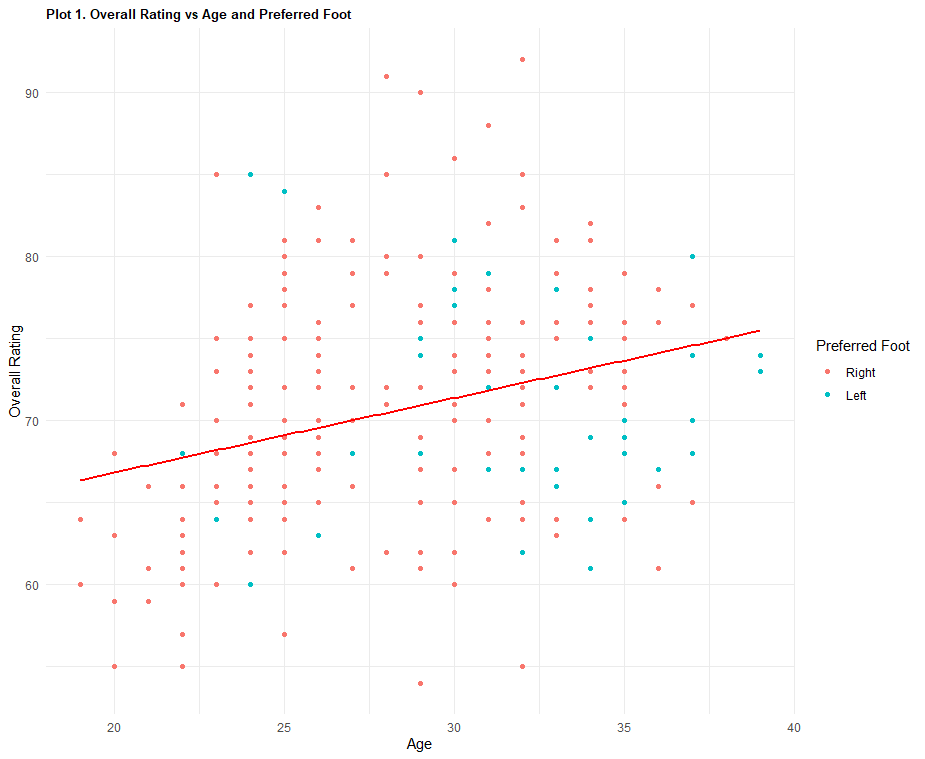
**Influential Points**: The Residuals vs Leverage plot indicates the presence of some influential points.

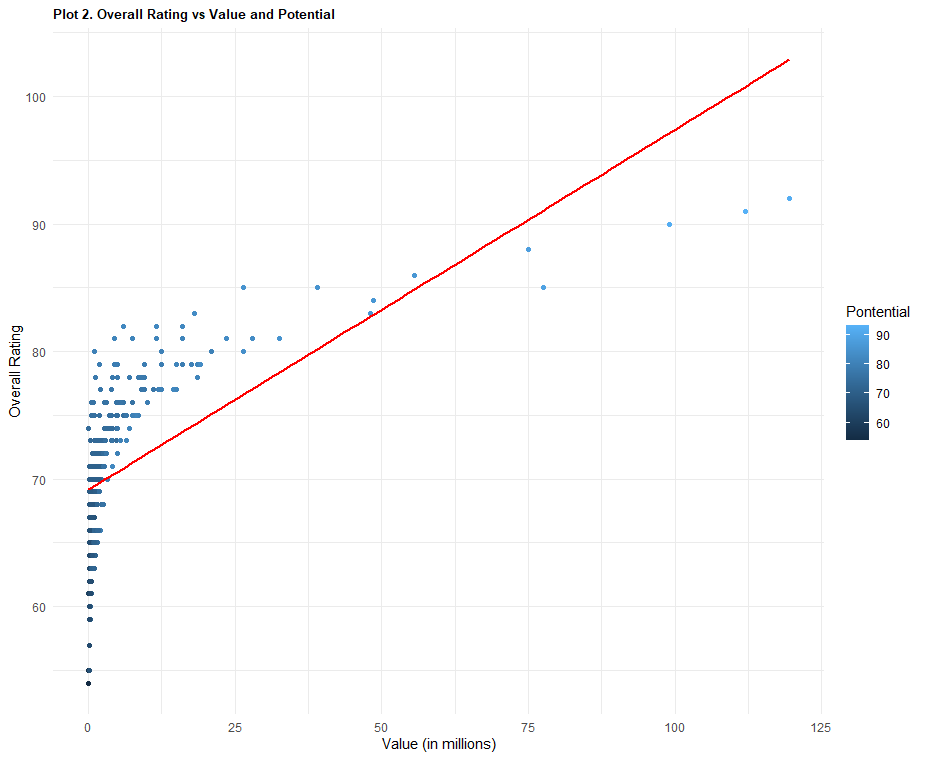
Potential Weakness: One potential limitation is the assumption that all players of the same age or potential have identical impacts on overall ratings, without accounting for variations in individual performance. The analysis indicates potential violations of normality, suggesting our results could be influenced by outliers. Despite this, the key takeaway is that increasing player potential and market value significantly enhances overall ratings. Depicted by **PQ** **Plot 3**.

Age has a slight positive relationship with Overall Rating, while Potential shows a strong positive relationship and is a significant predictor. Value has a moderate positive effect, especially for higher-value players, but Preferred Foot does not significantly impact Overall Rating. In conclusion, Potential and Value are key predictors, with Age having a weaker influence and Preferred Foot showing little effect as shown by **PQ** **Plot 1 and 2.**

Future analyses should incorporate additional attributes, such as player positions and historical performance data, for deeper insights. We recommend using robust regression methods to account for outliers and ensure the accuracy of the results.

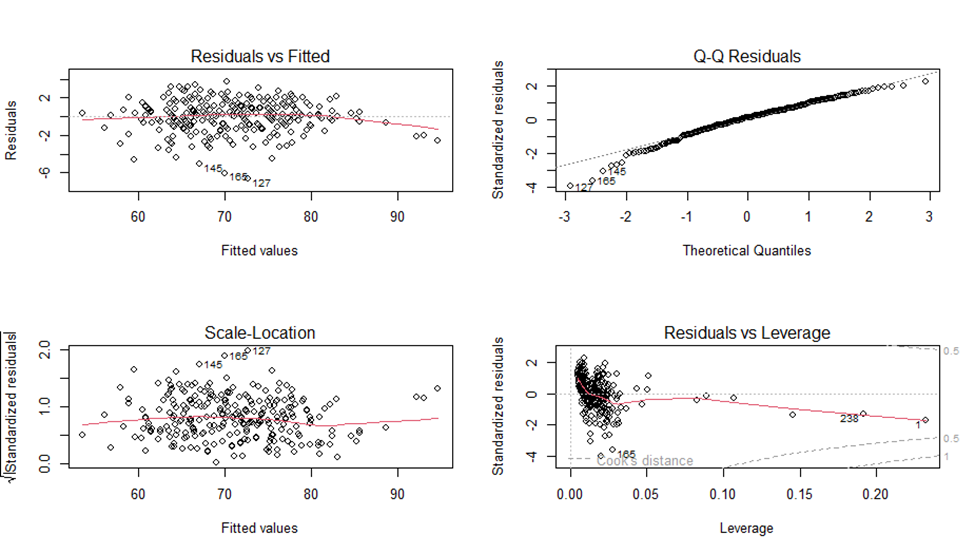
Recommendations: To improve player evaluations and scouting strategies, team managers should consider the influence of age, potential, value, and preferred foot on overall player ratings. Additionally, teams could benefit from assessing successful training and development programs that enhance these attributes, ultimately improving player performance. In conclusion, our analysis demonstrates a significant positive relationship between player attributes and overall ratings in FIFA, highlighting the importance of these factors in player evaluation and market strategies.

****



**PLOT 3**

**A screenshot of a computer

Description automatically generated****TABLE 3**