Assignment3: A Statistical Review on The Relationship Between Soccer Players' Age against Player Valuation

Mujaheed Abdulmalik

2023-07-04

BACKGROUND

Football which is also referred to as soccer in north America is one of the oldest sports to exist dating as far back as the 3rd and 4th century BC (History, n.d.). Footballers who are the most important part of the game have seen their individual values skyrocket within the past two decades. Market valuation of footballers can be defined as individualistic amount a club is willing to pay to acquire the services of a the footballer, that is the amount, they would be willing to pay irrespective of the amount the players employer values them at (Herm, S., Callsen-Bracker, H. M., & Kreis, H., 2014). A footballers valuation is very key to a football clubs economic growth as the money acquired from sales helps budget planning (Metelski et al., 2022). Football clubs require information about potential players to determine their qualities and market value and further decide which players to buy or not these information includes age, stats and so on (Bhilawa, & Fahriansyah, R., 2022).

According to Dimitropoulos et. al, (2018) they are many factors that determine the valuation of football players, more established football players have had additional opportunity to create and improve their abilities and capacities through experience. Players' market value rises as they get older and add skills to their human capital. Also,the relationship between age and market value in football players is non-linear. Younger players have lower market value, but as they gain experience and get older, their market valuation increases (Dimitropoulos, P. E., Travlos, A. K., & Panagiotopoulos, S., 2018).

They are other factors that determine like goals per match, competitiveness, contract duration, and the buyer team but these factors have the greatest impact on how much a forward or striker is valued in the market than other positions as that is their main stats. Other factors such as assists, shots, speed, tactics, vision, passes, dribbles, interceptions, and fouls received also contribute to other positions market value and also to a forward but to a lesser extent (Poza., 2020).

The objective of this research is to investigate the statistical difference in valuation between age groups of football/soccer players. The research will provide insights into the approaches used to determine the influencing parameters of players' valuation and help shed light on the player age at which a football club stands to gain more from in terms of sales. Additionally the findings of this analysis will contribute to our understanding of valuation work and provide valuable insights for new fans and football casuals.

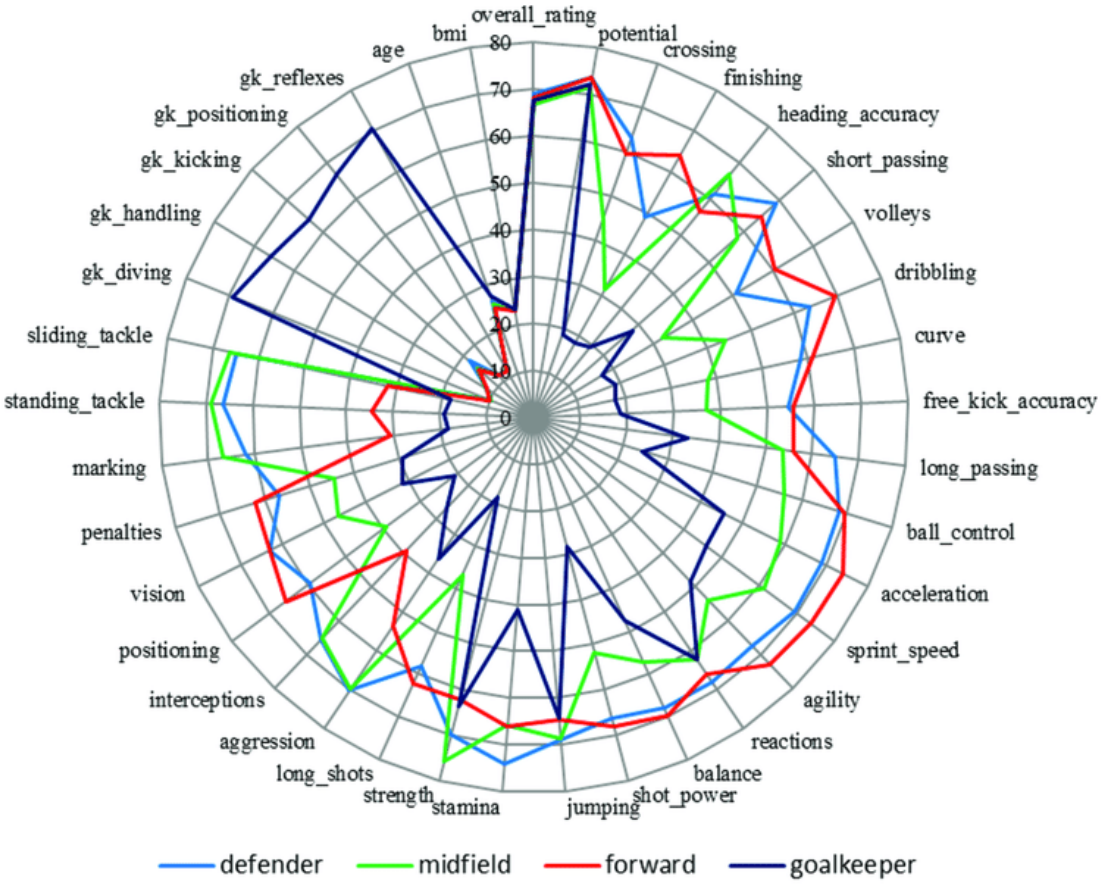

Figure 1: Radar chart of mean values grouped by players'roles(Soto-Valero,César.2017).

DATA DESCRIPTION

The data set used for this analysis is a sample of the 500 most expensive footballers in the world as at 2021, After applying a few data pre-processing to the raw data. i coerced some categories to one for the purpose of this analysis.

Player Positions

This data consists of each players position left-backs, right backs and center backs where all grouped as **defenders** while center midfielders, attacking midfielders and defensive midfielders were all grouped as **midfielders** and lastly all forward and wing positions where grouped under **forwards** with **goalkeepers** remaining unchanged.

Player Age

The age of each of the player were coerced into four groups:

- Young Players which consist of players of ages (17-22).
- Pro Players which consist of players of ages (23-28).
- Peak Players which consist of players of ages (29-32).
- Veteran Players which consist of players of ages (33+).

Player Value

Each of the player are all valued in millions, this valuation was collected based on pricing models on transfermakt community which consist of various key factors like stats and demand (Sanjeet, S. N., 2021).

Player Stats

The main player stats in the data is goals scored and goals assisted. The number of goals a player scores in a match is referred to as a goal score, while the number of assists a player provides in a match is referred to as an assist (Poza., 2020). Both stat was coerced into one group adding by them both together.

GROUP	DESCRIPTION
ID	Unique Identifier
POSITION	Grouping of player by position (Foward, Midfield, Defense or Goalkeeper as F,M,D or G)
AGE-GROUP	Grouping of players based on their ages (Young-Player(17-22), Pro-Player(23-28), Peak Player(29-32) or Veteran Player(33+))
VALUE	Valuation of players measured in millions
GOALS & ASSISTS	A combination of goals scored and goals assisted

Note: The raw dataset preprocessed was extracted from (Sanjeet, S. N., 2021).

METHOD

The two main variables of focus of this research is player value in millions and player age which as stated in the previous section has been coerced into in a group of four. In other to use the right statistical analysis for this research the nature of both variables where considered, as

1. **Player Valuation**
- Continuous variable.
- Dependent.

2. **Player age-groups**
- Categorical variable.
- Independent variable.

Data Analysis

For the purpose of this research an **Independent ANOVA** was to be conducted to find the valuation differences across age groups by comparing its means but a **Kruskal-Wallis test** was performed as non-parametric alternative this was because the Central Limit Theorem (CLT) conditions or assumptions of the Independent ANOVA were not met and eta squared based on the H-statistic was used to check the magnitude of the effect between groups. The participant for this analysis were of different age groups (R Core Team ,2021).

Data Visualization

To picture the consequences of the Kruskal-Wallis test, a violin plot was utilized. These plots showed the appropriation of the reaction variable for each group and assisted with recognizing contrasts between them.

RESULT

The current study sought to determine whether or not player age had an effect on valuation. 500 study participants were sampled from the general football market (109 were young players, 317 were footballers just who were at a pro stage, 69 were players at their peak and 5 players were veterans closer to retirement). The sample contained extreme outliers which were left in as they were important to the analysis, also the sample failed to demonstrate normality by group,and Levene's test demonstrated homogeneity of variance. The data failed the assumptions therefore a Kruskal-wallis test was conducted as a non-parametric alternative. The mean valuation for the Young Player group was **32.67 (SD = 21.53)**, the mean score for the Pro Player group was **30.4 (SD = 15.67)**,the mean score for the Peak Player group was 34.25 (SD = 18.39) and the mean score for the Veteran Player group was **41.4 (SD = 22.37)**.

A Kruskal-wallis test showed that the mean difference in **Player Valuation** score across **Player Age-groups** was not statistically significant, $\boldsymbol{F(2,55) = 4.45,\ p = 0.22,\ \eta^2 = 0.0029}$. The results suggest that there is no significant difference in player valuation score across player age-groups. According to Kruskal-wallis eta squared based on the H-statistic the magnitude of effect between groups is small.

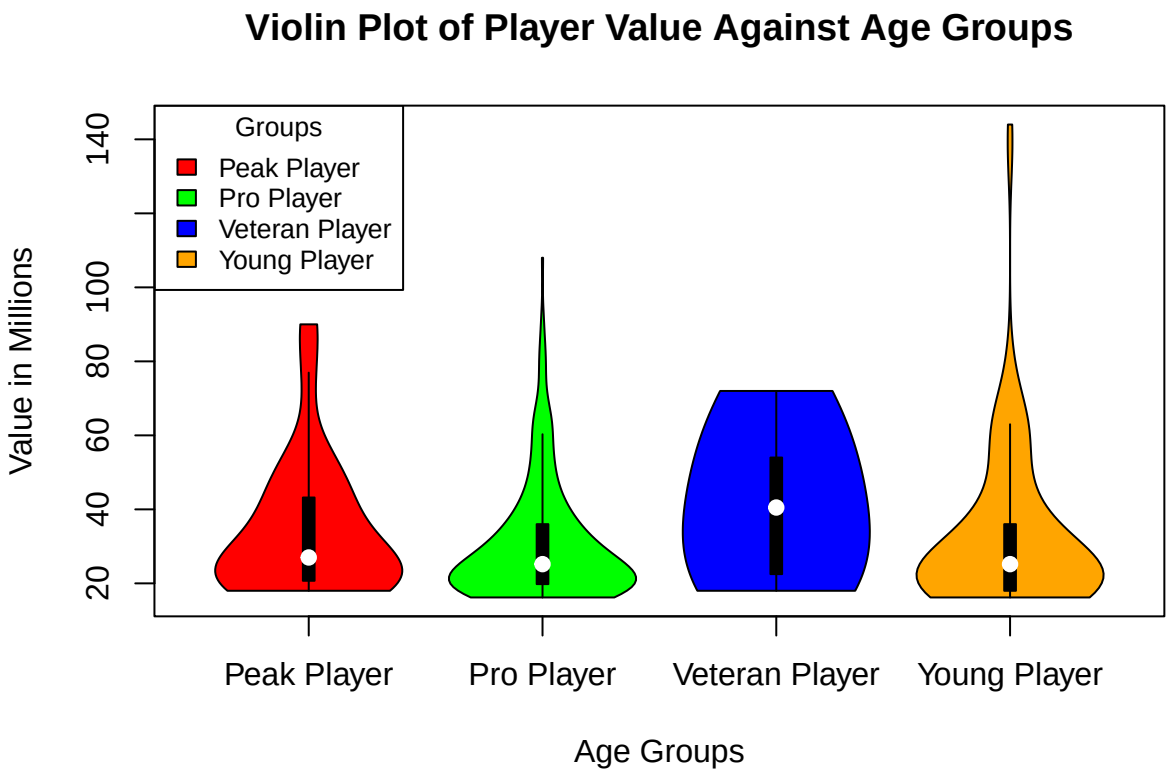

Figure 2: A Violinplot of All Age-Groups .

CONCLUSION

The objective of this research was to investigate the statistical difference in valuation between age groups of football/soccer players. Previous researches showed that they football teams or organizations need better insight on footballer and on field life to help them decide on the acquiring a player thus making them need to properly valuate the player. They also shed light on various factors that determines the market valuation of a footballer which includes age, years of experience, key stats like goals and assist, the position in which they play and injury records. Additionally this research was made shed light on the player age at which a football club stood to gain the most in terms of sales and provided insights into the methods used to determine the influencing parameters of players' valuation. The outcomes were to add more to how new fans and football casuals might interpret valuation.

The study sampled 500 of the most valued footballers as of 2021, the sample consisted of footballers of different age groups and positions. A Kruskal-Wallis test was conducted as a non-parametric alternative to Independent ANOVA. The findings suggest that there is no significant difference in player valuation score across player age-groups and there is only a small difference between the valuation of each group thus translating to the fact that age as a stand alone factor of market value is not reliable. To conclude the aim of this research was inconclusive as the alternative hypothesis was not proven.

Future research could explore other factors that determine player valuation, like exploring age with other key stats like goals per match, clean sheets, defensive actions, injury record, competitiveness, contract duration, and the buyer team. Also, future research can be conducted using current player valuation as the financial market of the world in 2023 is not the same as two years back. Additionally, further research could investigate the relationship between player age and market value in other sports.

REFERENCES

1. Metelski, A., Kornakov, K., Sharda, R., & Chłapowski, J. (2022). How the value of football players influences a team's chances of victory - a euro 2020 example. *Journal of Physical Education and Sport, 22(1)*, 167-173. (https://doi.org/10.7752/jpes.2022.01021)

2. Herm, S., Callsen-Bracker, H. M., & Kreis, H. (2014). When the crowd evaluates soccer players' market values: Accuracy and evaluation attributes of an online community. *Management Review, 17(4)*, 484-492.

3. Dimitropoulos, P. E., Travlos, A. K., & Panagiotopoulos, S. (2018). Migration and Football Player Market Value: Evidence from Greece. *International Journal of Sport Finance, 13(2)*, 119+. (https://link-gale-com.proxy1.lib.trentu.ca/apps/doc/A653502720/AONE?u=ocul_thomas&sid=bookmark-AONE&xid=5204bb14)

4. Poza. (2020). A Conceptual Model to Measure Football Player's Market Value. A Proposal by means of an Analytic Hierarchy Process. [Un modelo conceptual para medir el valor de mercado de los futbolistas. Una propuesta a través de un proceso analítico jerárquico]. *Revista Internacional de Ciencias Del Deporte, 16(59)*, 24–42. (https://doi.org/10.5232/ricyde2020.05903)

5. History, F. (n.d.). *The History of Football(Soccer).* Retrieved from FootballHistory.org: (https://www.footballhistory.org/)

6. Bhilawa, & Fahriansyah, R. (2022). THE INFLUENCE OF PERFORMANCE, AGE, AND NATIONALITY ON THE MARKET VALUE OF FOOTBALL PLAYERS. *Assets (Madiun, Online), 11(1)*, 1–9. (https://doi.org/10.25273/jap.v11i1.8422)

7. Sanjeet, S. N.(2021). *Most Expensive Footballers 2021* [Data set]. (https://www.kaggle.com/datasets/sanjeetsinghnaik/most-expensive-footballers-2021?resource=download)

8. R Core Team (2021). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. (https://www.R-project.org/).

9. Henning, C. (n.d.). *AMOD-5210 Workshop 1: Introduction to Basic and Inferential Statistics in R.* Trent Blackboard. Retrieved July 10, 2023, from (https://trentu.blackboard.com/webapps/blackboard/execute/content/file?cmd=view&content_id=_1551876_1&course_id=_50501_1)

10. Soto-Valero, César. (2017). A Gaussian mixture clustering model for characterizing football players using the EA Sports' FIFA video game system. RICYDE. *Revista internacional de ciencias del deporte. 13.* 244-259. 10.5232/ricyde2017.04904.