

FIFA 22 Video Game Character Analysis

Author Contributions

Nan Raghuraman contributed to the Abstract, Background, Aims, Dataset, Description sections

Alex Roginski contributed to the analysis of linear regression and KDE

Veronica Capco contributed to the kernel density estimate

Armando Uribe contributed to the Methods, Results and Discussion sections

Abstract

Our group analyzed player data from the video game FIFA 22. Each playable character in the game is given an overall score determined from evaluations of their skills and physical traits. We aimed to analyze each player's physical traits such as height, weight, and age, to see which of these traits, if any, were prevalent in the dataset. We also investigated the correlation between the players' skills such as shooting, passing, dribbling, and defending, and their overall wage. We found that most FIFA 22 characters have a height between 5'11" and 6'1", a weight between 154 and 187 lbs, and an age between 20 and 30 years. We also found out that there is a quadratic relationship between a player's overall score and their weekly wage.

Introduction

Background

In this project, we are studying characters in the video game series FIFA. We are investigating what predicts the wage of FIFA video game characters. We will be looking at the players' shooting, passing, dribbling, and defending abilities to predict their wage. We will also be looking at the players' physical traits (age, weight, and height) to see if any of these physical traits play any prevalence in the players' statistics. The motivation for collecting this data is to understand which characters (or combination thereof) are best to play the video game well. Although Electronic Arts (EA), the video game company that creates FIFA, makes great effort to make player character traits as realistic as possible, there is an element of fantasy in an effort to make the game more balanced and enjoyable. Using this data, you could attempt to build your own best team while playing the FIFA video games, while also learning some facts about each player, such as their age, height, and the club they play for.

Aims

Originally, our group was under the impression that the in-game statistics were directly modeled off of real-life statistics, so our original aims, although very similar to the ones we ended up choosing, were aimed to predict real-life wages of each player. We quickly realized, that due to the notoriety of certain players (Lionel Messi, Cristiano Ronaldo, etc.), some real-life wages were disproportionately higher than others, something that was not reflected in the game data. We realized that this was due to the fact that players of the video game would not be able to afford to recruit these players, so the game experts at EA determined wages on an adjusted scale that didn't quite reflect the real-life statistics. Therefore, our group shifted our focus from real-life players to the playable video game characters instead. These findings could be used by players of these video games to build the ideal soccer team given the player's funds and access to characters.

With our first aim to use characters' skills to predict wage, we decided to use multiple regression to determine if skill level was linearly correlated with wages of players. We found that there is not a strong linear relationship between player ability and weekly wage. The relationship is actually quadratic.

Our second analysis used Kernel Density Estimation (KDE) to see which of the characters' physical traits, if any, were significant in a character's statistics. We found that player age, height, and weight are all normally distributed. We found that the average player is 165 lbs, 5'11", and 25 years old.

Materials and Methods

Datasets

This data records statistics used to rank soccer players for the video game FIFA 22. The original dataset includes 110 attributes, such as soccer club information, the positions of each player, scores for skills such as dribbling and shooting, as well as more refined skills and traits. We tidied the dataset to only include 14 variables that we will be drawing conclusions from in order to predict each player's in-game weekly wage. There are 19,239 observations in this data. The data was found on <https://sports-statistics.com/soccer/fifa-2022-dataset-csvs>. Every piece of data was recorded directly from the FIFA 22 game as well as the player statistics page from the soFIFA.com website.

Evaluations of the player's skills are determined by experts at EA, as well as outsourced data reviewers who grade on a dimensionless rubric out of 100, or 1-5 stars for the 'international reputation' category. The observational units for the height is in centimeters and the weight is in kilograms. Wage and other monetary values are in euros. Values that measure the player's ability such as shooting, passing, dribbling, defending are measured as dimensionless values that are determined on a rubric by expert soccer analysts at EA. The relevant population is all soccer players who are included in the game's 'Career Mode' setting. The scope of inference from this dataset is broad, as the sampling design seems to encapsulate the population well.

Name	Variable description	Type	Units of measurement
long_name	full name of player	string	none
nationality_name	nationality of player	string	none
age	age of player	Numeric	Calendar year
height_cm	height of player	Numeric	centimeter
body_type	body type of player	string	none
weight_kg	weight of player	Numeric	kilogram
overall	overall rating of player based on in game stats	Numeric	dimensionless
wage_eur	in-game per-week wage of player	Numeric	euro
international_reputation	international reputation of player determined by EA experts	Numeric	dimensionless
shooting	shooting ability of player ranked by EA experts	Numeric	dimensionless
passing	passing ability of player ranked by EA experts	Numeric	dimensionless
dribbling	dribbling ability of player ranked by EA experts	Numeric	dimensionless
defending	defending ability of player ranked by EA experts	Numeric	dimensionless

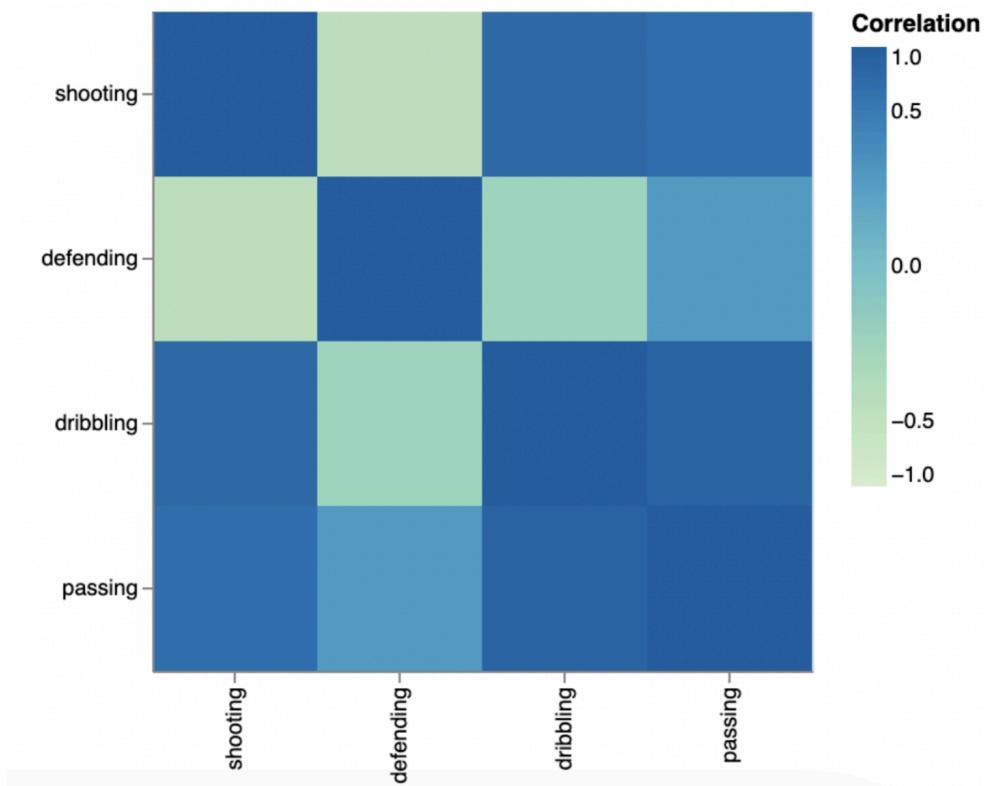
Methods

With respect to the first question regarding the potential correlation between a player's four physical variables of shooting, passing, dribbling, and defending and a player's wages, we implement a heat map to represent an initial visual correlation between the factors. Then,

multiple linear regression of the four physical variables is calculated to show deviation from the actual wages, followed by a calculation of R^2 and sum of squares of the residuals to check the validity of fit to the data. Lastly, a polynomial function is used in comparison to the results of linear regression. Both the linear regression and polynomial function are transposed as lines over a scatter plot of *Overall Score vs Player Weekly Wage* as a visual. In the second question, the kernel density estimate (KDE) is calculated using age, weight, and height respectively, in order to determine the distribution of each. From all three of these estimates, the trait with the highest prevalence is produced alongside with the average FIFA player's age, weight, and height.

Results

Heat Map

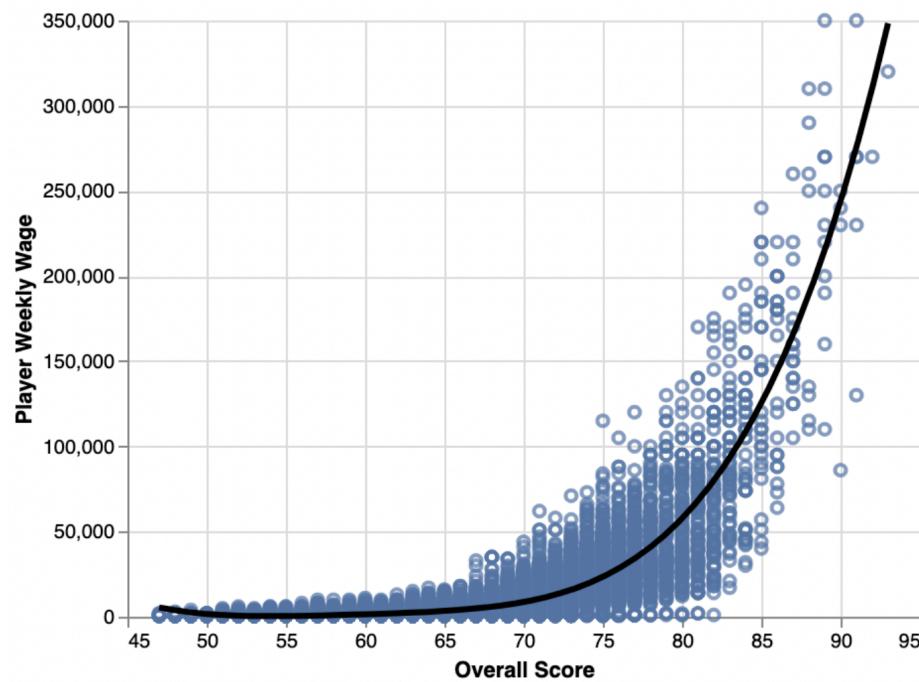


There appears to be a negative correlation between shooting and defending, which makes sense since players tend to focus their abilities on attacking or defending, but not both at the same time. There is a positive correlation between shooting and dribbling, as well as passing and dribbling, since these abilities go hand in hand for attackers. This makes sense because both these abilities are essential in an attacker. Furthermore, there seems to be little correlation

between passing and defending. This makes sense because a defender does not need to pass excellently to fulfill their role.

Linear Regression

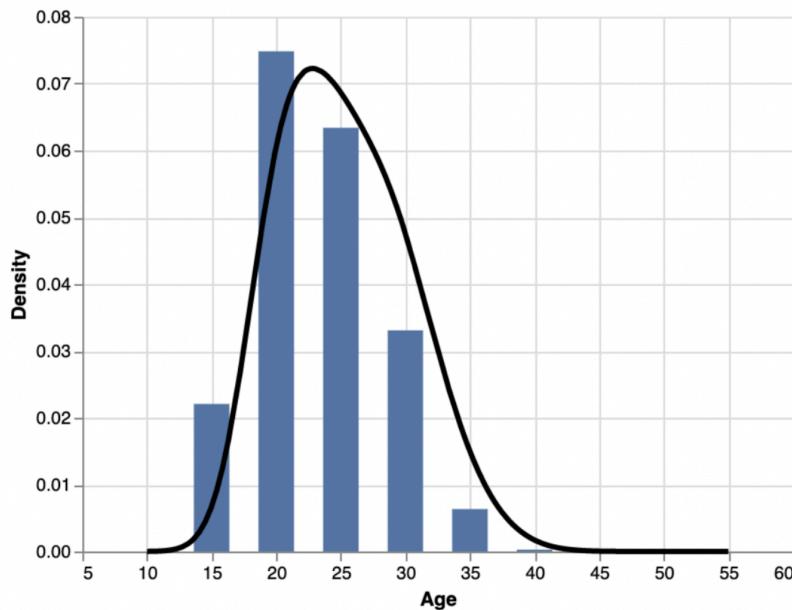
In this investigation, we are trying to predict a player's weekly wage using their scores in the following abilities: shooting, passing, dribbling, and defending. We will perform a multiple linear regression using these four variables. We can look at the R^2 value to see how well the regression fits the data. We get a R^2 value of 0.24, which indicates a poor linear fit on the data. We can also look at the sum of the squared residuals. We get a value of 5.09×10^{12} , which is a huge residual value. This is because there is not a linear relationship between the four abilities and weekly wage. Can we fit a polynomial instead?



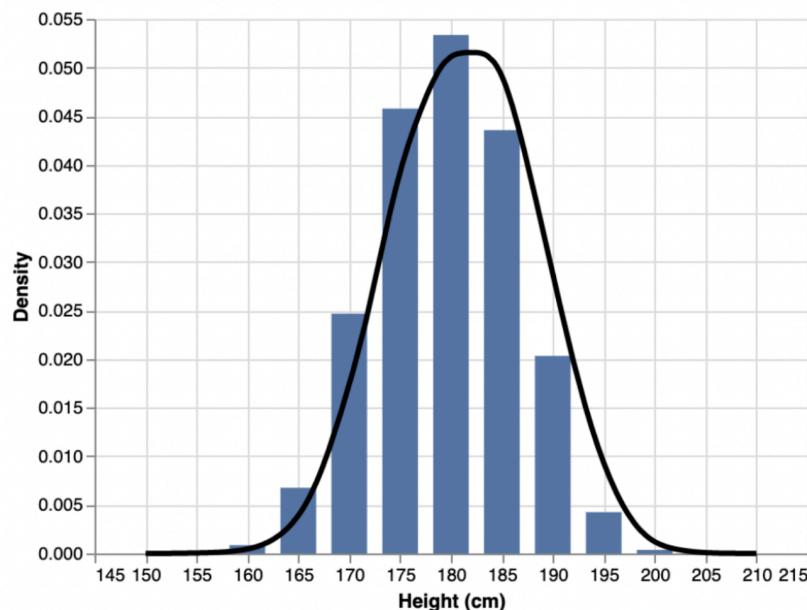
The fit is much better now. Another method would be to fit an exponential function instead of a polynomial function.

The weekly wage of FIFA characters has a quadratic distribution. We see that most players have a weekly wage between 0 and 5,000 euros. Fewer players have a weekly wage between 5,000 and 30,000 euros, and very few players have a weekly wage more than that. This suggests that most FIFA characters receive a weekly wage in the several thousands range while very few players make tens or hundreds of thousands of dollars per week.

KDE

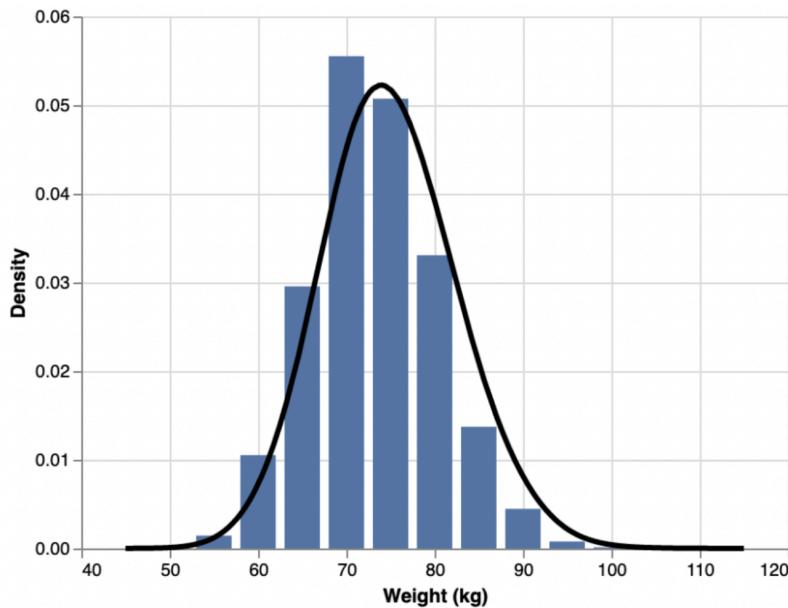


Looking at the kernel density estimate, we see that the distribution of Age has a somewhat bell curve-like shape. Most players in FIFA 22 are between the ages of 20 and 30, while players younger than 20 and older than 30 are more rare.



Like the Age distribution, we see that the height distribution of players in FIFA appears to be distributed in a bell curve. The players center around a height between 180-185 cm, which is

equivalent to 5'11" to 6'1". There is a significant minority of players less than 175 cm (5'9") and greater than 190 cm (6'3").



Like the other distributions we've seen, we find that weight of FIFA 22 players is normally distributed. Most players have a weight between 70-85 kilograms (154-187 lbs). Like the player's height and age, their weight statistics are also applicable to real life. We know that most players have a weight between 154 and 187 pounds with a significant minority lighter and heavier than that range. Our analysis suggests that the average FIFA22 character has a height between 5'11" and 6'11", a weight between 154 and 187 pounds, and an age between 20-30.

Discussion

In this project, we found that there is a quadratic relationship between a player's overall playing ability and their score. We speculate that the relationship is quadratic because the video game designers who created the characters designed the game that way. Also, it is true that in many sports the best players get paid the most, so this conclusion makes sense. One important caveat of our interpretation is that our conclusions about weekly wage only apply to the video game FIFA 22, not real life. A more refined question to investigate would be, "do the overall scores of players in the FIFA 22 video game correlate with the actual soccer players that they represent?" In the second part of our project, we looked at the prevalence of player's ages, heights and weights. We found that the most prevalent heights for soccer players are between 5'11" and 6'1" and between 154 and 187 pounds. These ranges of height and weight seem practical because professional soccer players would want to have a balance between being strong and

fast, as being taller or heavier may not offer an advantage at all. In terms of age, the most prevalent age being between 20 and 30 years old makes sense because those are the years where soccer players are both relatively young and highly experienced. In a further investigation, we would like to investigate the question, “do physical traits like age, height, and weight correlate with a player’s overall skill?” Another further investigation we could explore would be how the average height and weight of professional soccer players compare to the best soccer players worldwide such as Messi.