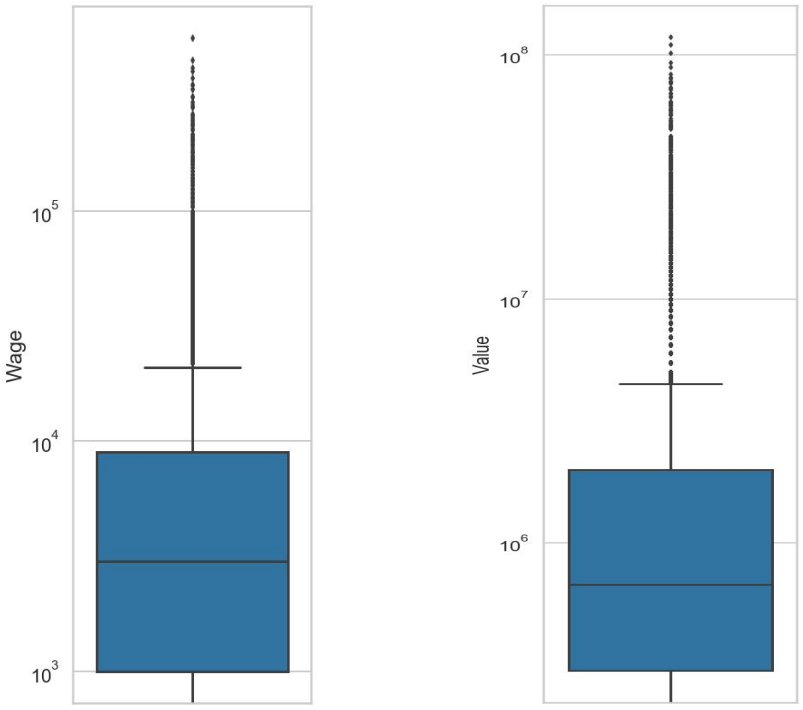
FIFA19 The Story Behind The Player

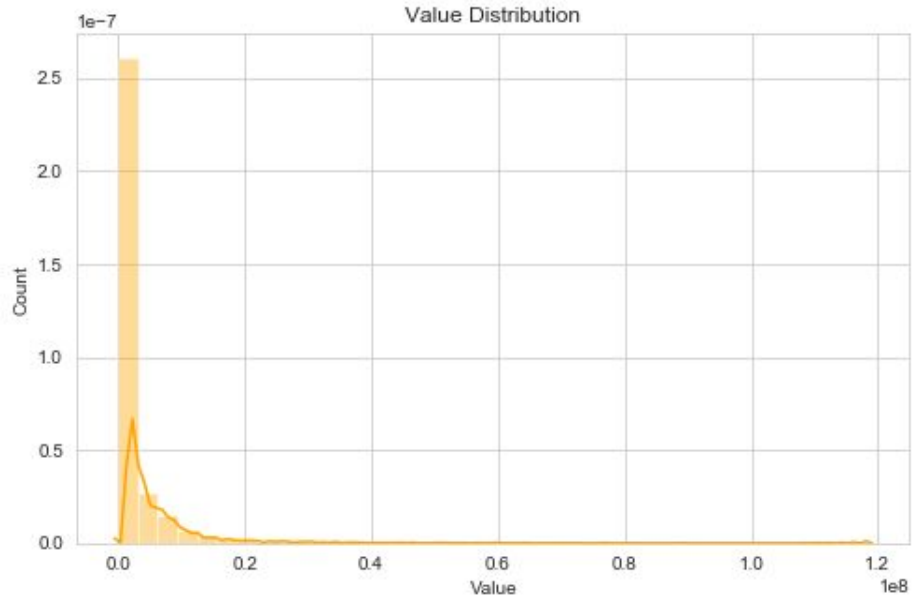
1. *Ask the following questions and look for the answers using code and plots:*
    1. *Can you count something interesting?*
    2. *Can you find trends (e.g. high, low, increasing, decreasing, anomalies)?*
    3. *Can you make a bar plot or a histogram?*
    4. *Can you compare two related quantities?*
    5. *Can you make a scatterplot?*
    6. *Can you make a time-series plot?*
2. *Looking at the plots, what are some insights you can make? Do you see any correlations? Is there a hypothesis you'd like to investigate further? What other questions do the insights lead you to ask?*
3. *Now that you've asked questions, hopefully you've found some interesting insights. Is there a narrative or a way of presenting the insights using text and plots that tells a compelling story? What are some other trends/relationships you think will make the story more complete?*

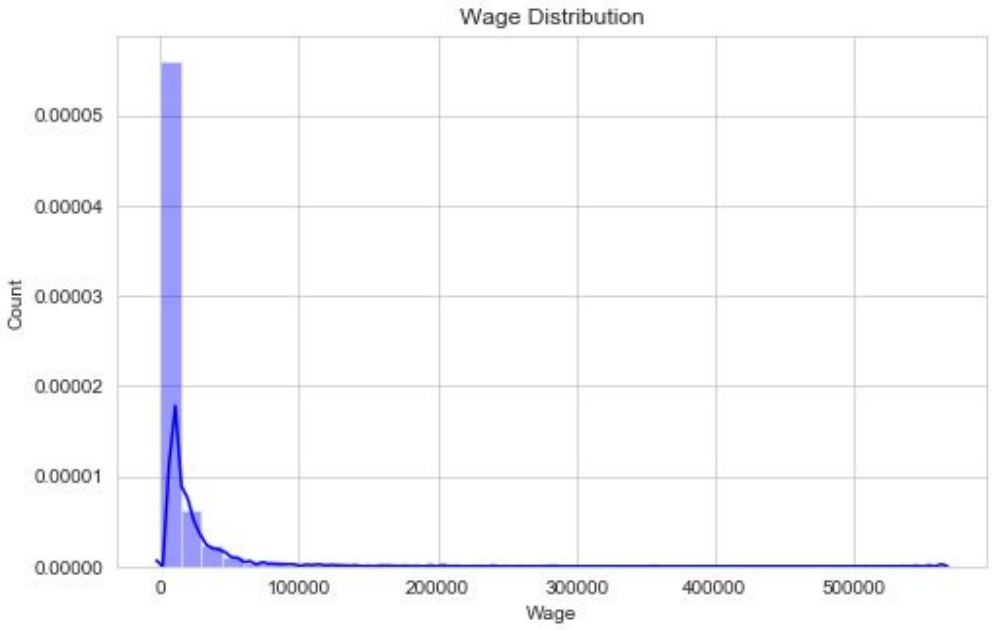
For the Exploratory Visual Analysis (EDA) portion of this project, I will point out noteworthy patterns as well as dive deeper into anomalies such as Messi or Neymar Jr. in attempt to tell the story behind their extraordinary ratings.

I. Wage and Value (Logarithmic scale)

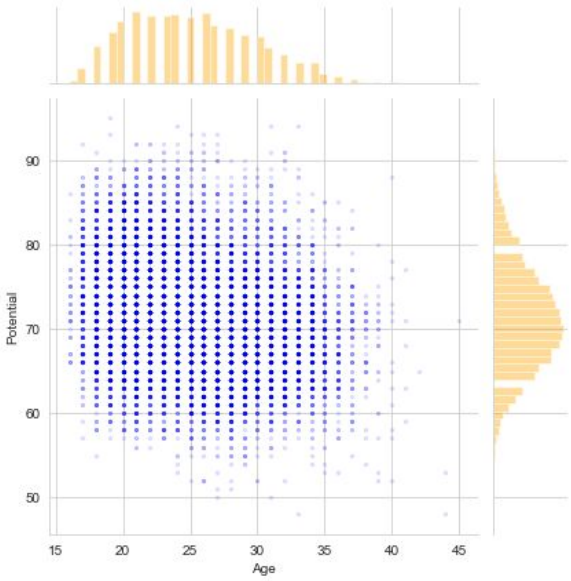


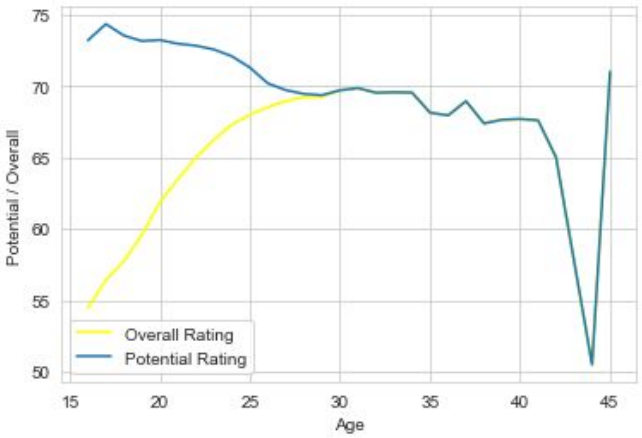
We can see that both wage and value have the same distribution. On average, FIFA19 players are at a specific range; however, there are many outliers. Meaning, there are players whose skills are valued at an exponentially higher rate. A FIFA19 player is either paid “average” or “exceptional”, nothing in between. The 2 outliers to be noted are, Neymar Jr. who is the highest valued player, and Messi who is the highest paid player. Below is their distribution in actual values.





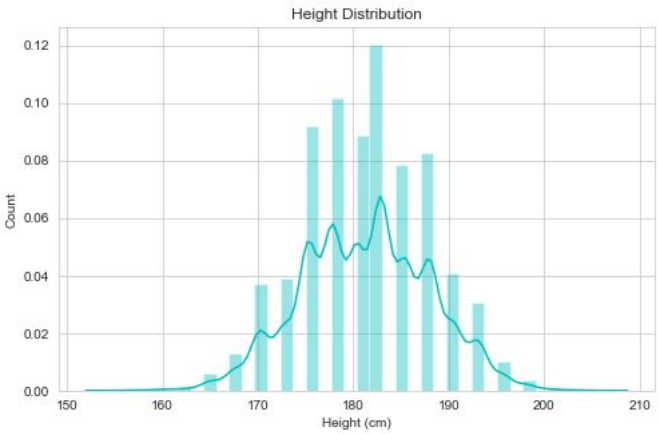
II. Age and Potential

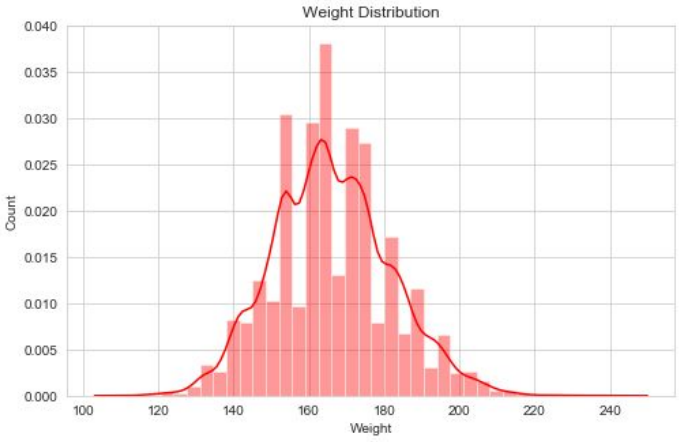




We can see that both age and potential stats of players are mostly uni-modal(~ normally distributed), this is however more true for potential players. Age, on the other hand, has most of its distribution between, 18 to 30 years old. The second graph demonstrates the density of the first graph much more clearly : Player reaches their maximum potential at the age of around 30, this is also where their overall rating is the highest. However, potential tends to fall as you grow old past 30.

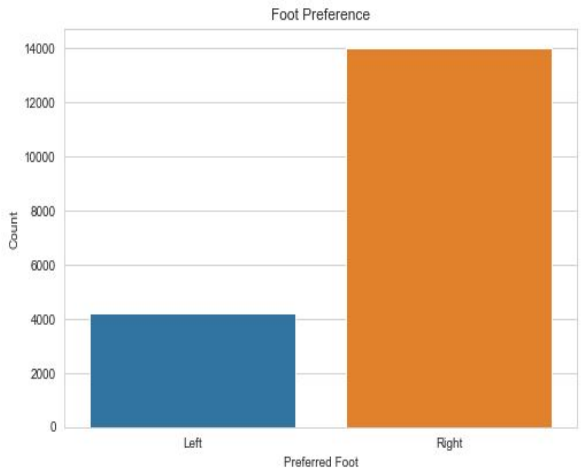
III. Height and Weight Distribution.

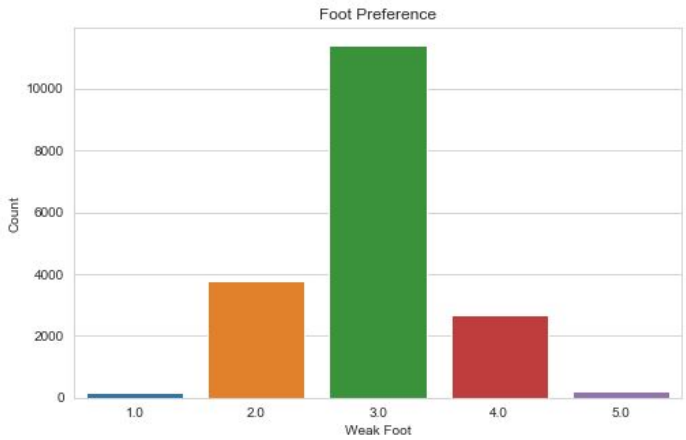




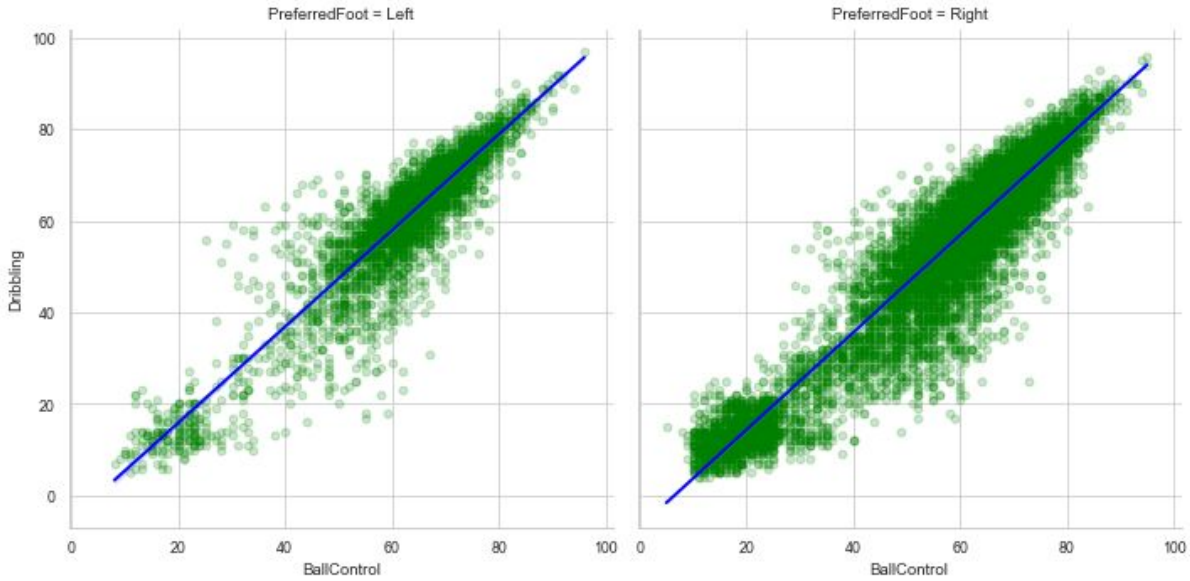
As far as weight and height go, we can see that they are normally distributed, while having certain heights and weights being more common than others. Even though some values of these metrics are more common than others, overall, FIFA19 players have average heights and weights and are very similar to each other.

IV. Foot Preference

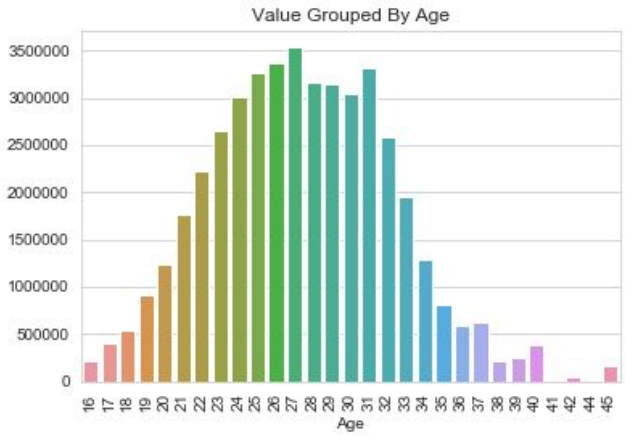


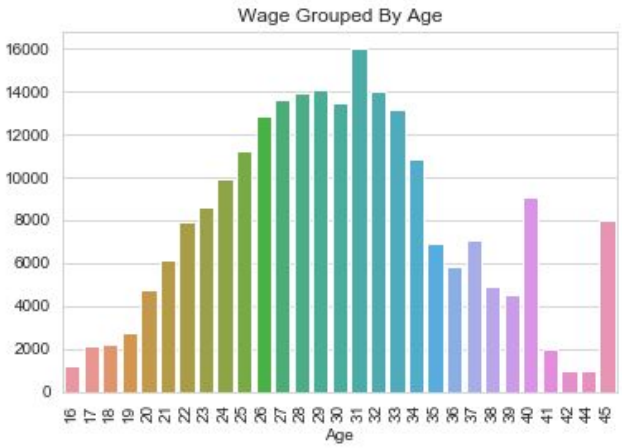


Most FIFA19 players are right footed, however, the second figure shows that most of the players have a 3.0 rating with their weak foot. This means that most players have some professional capacity with both feet. We can also see below that both feet have good ball control on average.



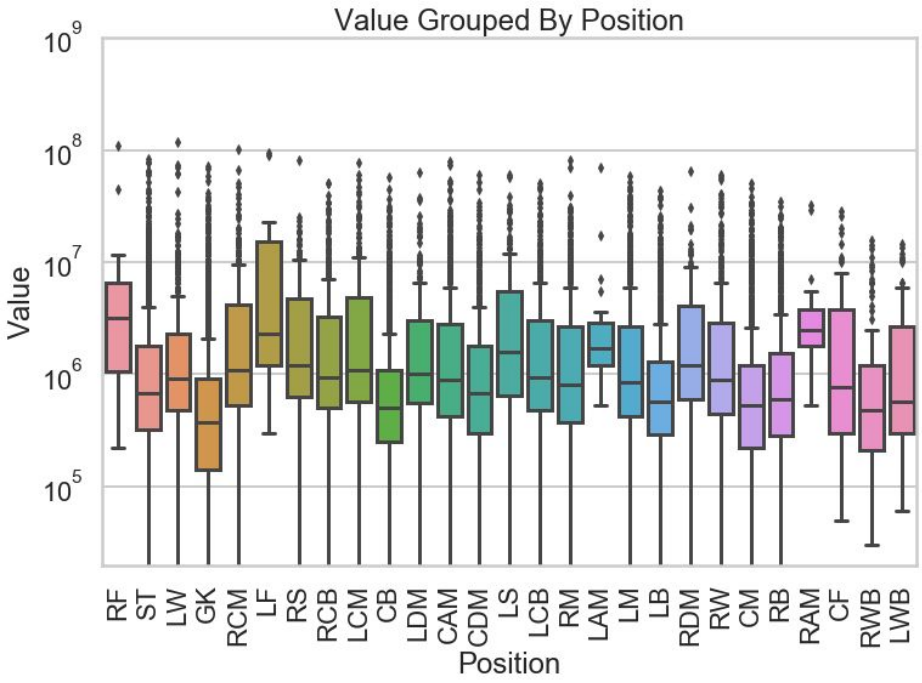
V. Back to Value/Wage and Age





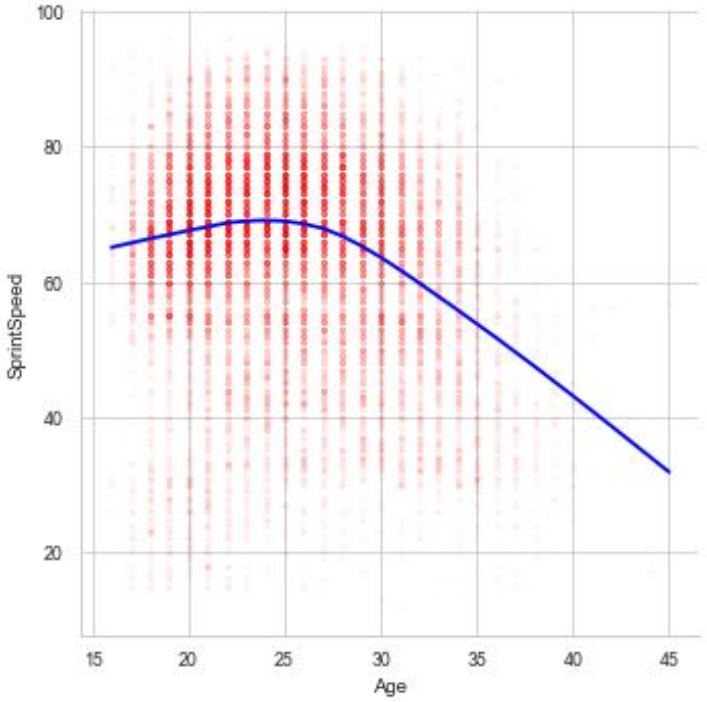
As we mentioned earlier, players reach their peak potential at the age of 30. We we can see in the figure " Wage Grouped by Age" that players are paid the most at the age of 31, after which this value slowly goes down. There are spikes at the age of 40 and 45, these are likely for goalkeepers because the average age span of goalkeepers is much higher than of field players. Looking at the "Value Grouped by Age", we can see that players are valued very high, before they reach their full potential. One, could say that the age of 27 is a player's "golden year" on average. As a result, they are paid higher even when they reach the age of 30. Next, let's look at player values based on their positions.

VI. Value Grouped by Position



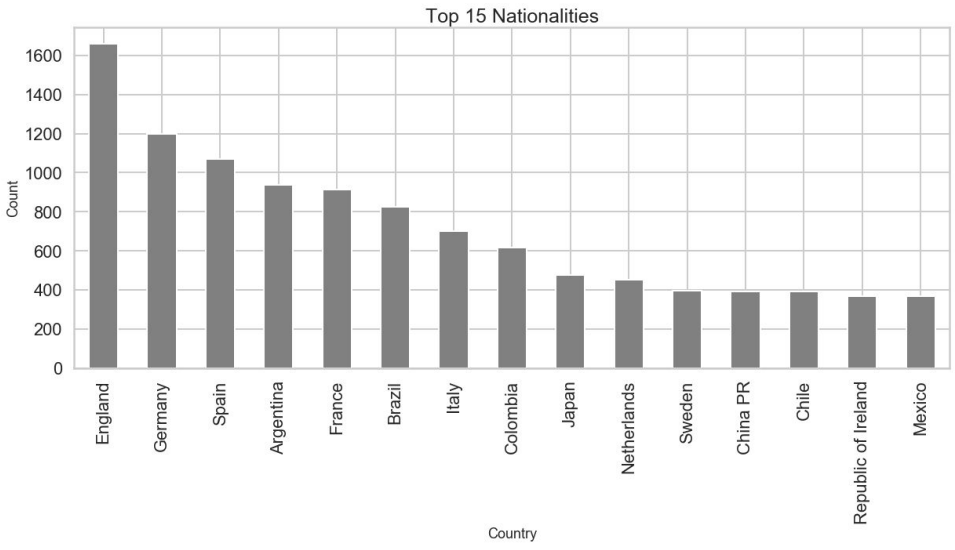
We can see that Right Forward (RF) and Left Forward(LF) positions have the most “high end” players who are valued the highest. We can also see that even though on average goalkeepers are valued much less than field players, there are many outliers who surpass the average field player in value. Unlike other forward positions, Center Forward (CF) has much less deviation in value, they are mostly on the higher end of the spectrum.

VII. Sprint Speed vs Age

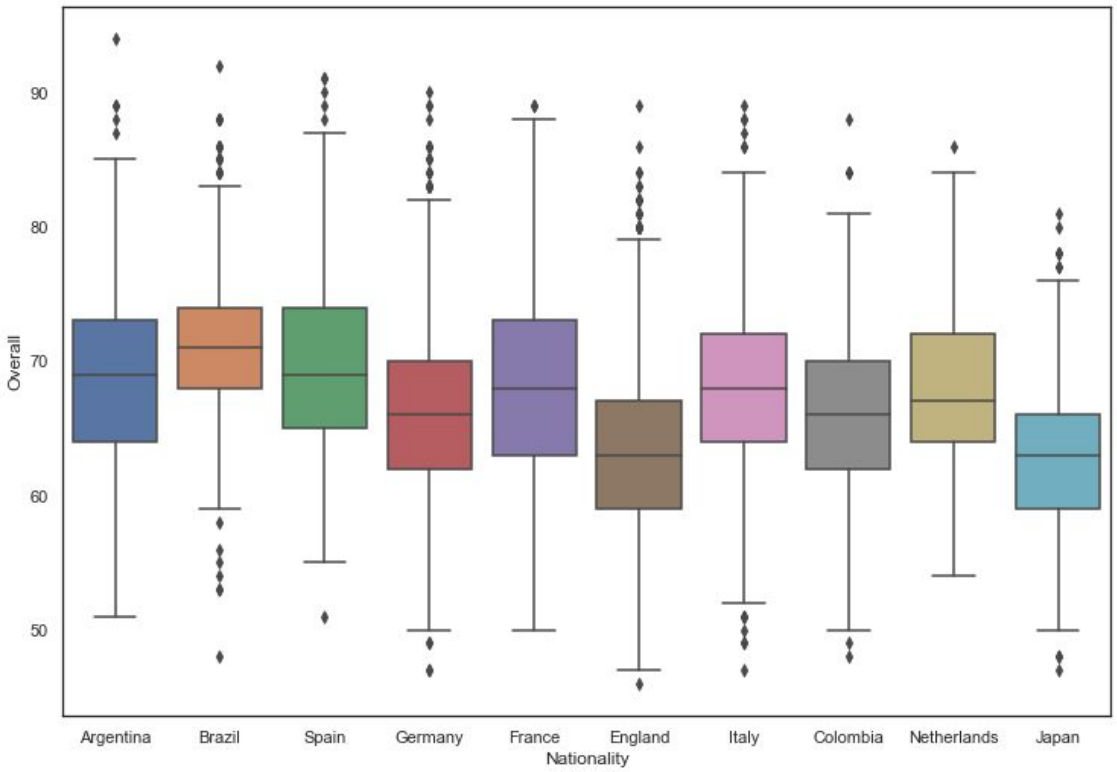


Sprint speed fall as players age

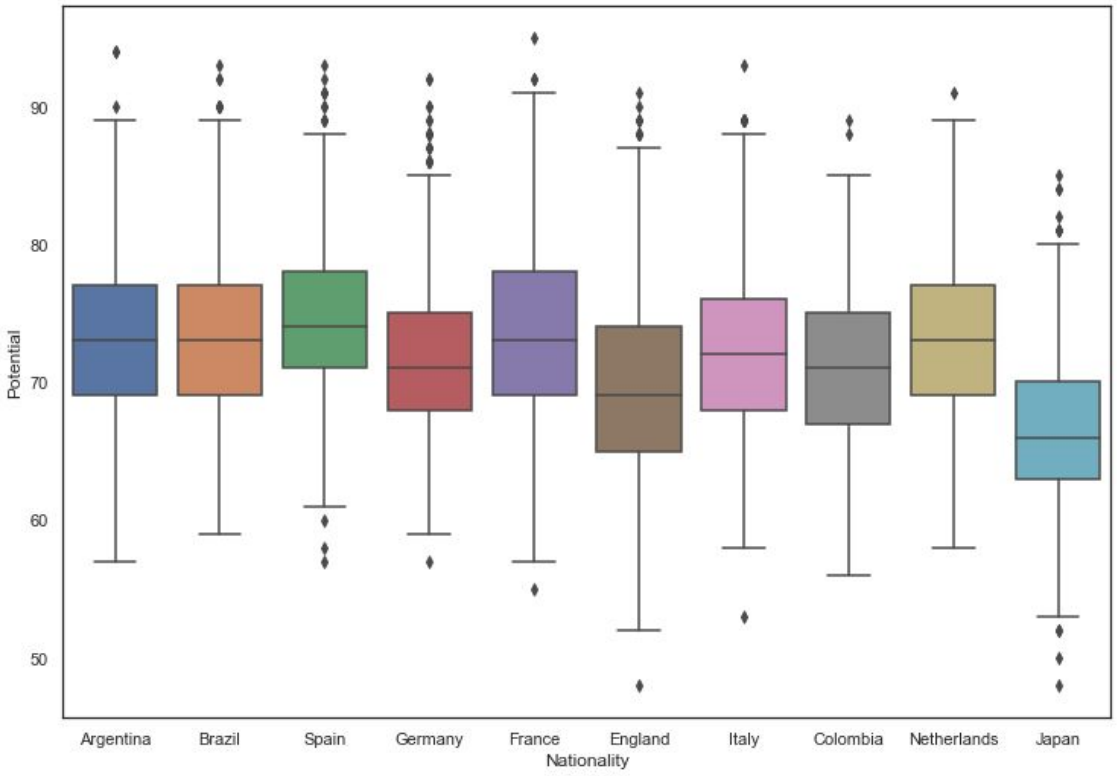
VIII. Top 15 Nationalities



It is clear that the sport is dominated by European nationalities followed by South American ones. Let's take a look at the overall scores for the top 10 Nationalities.

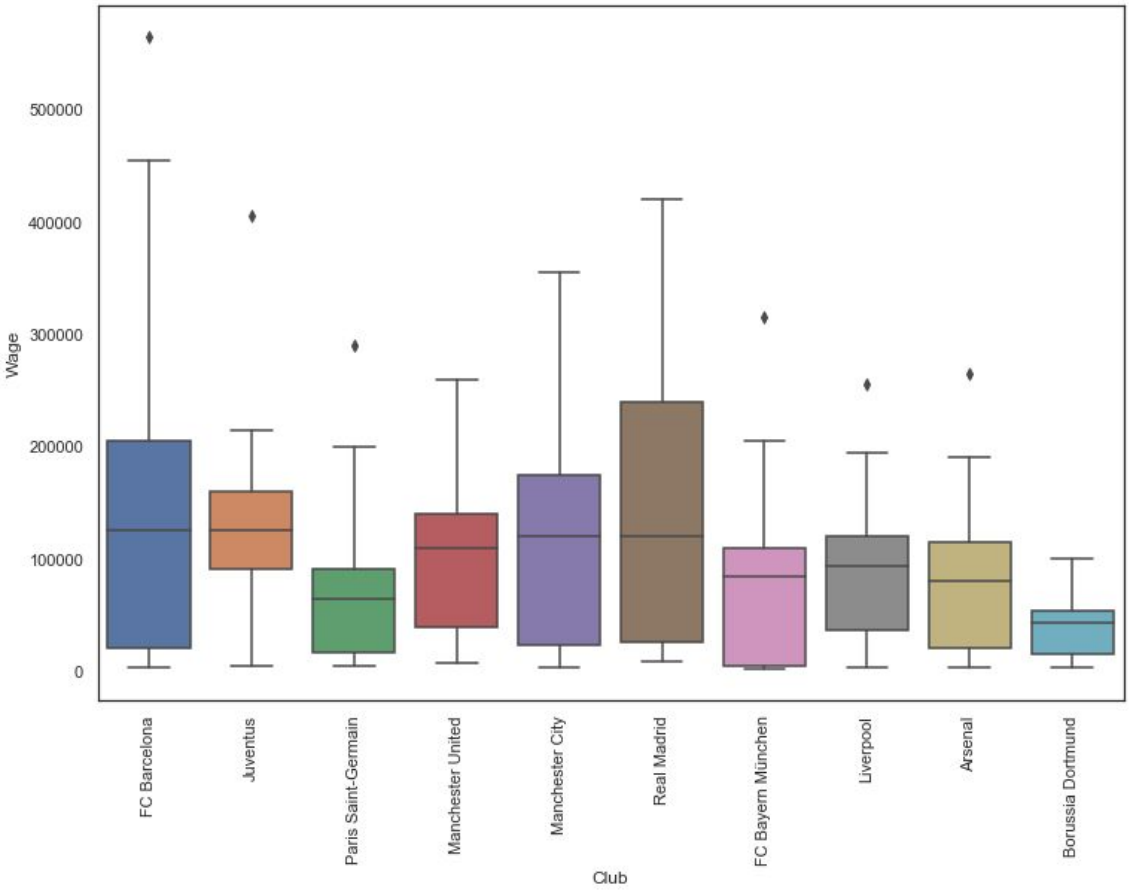


We can see that the highest over score belongs to the South American countries -- Argentina and Brazil -- followed by the European countries. Their spread is relatively the same even though we can see that Brazilian players have a tighter spread: meaning their players on average are above the rest in terms of overall score.

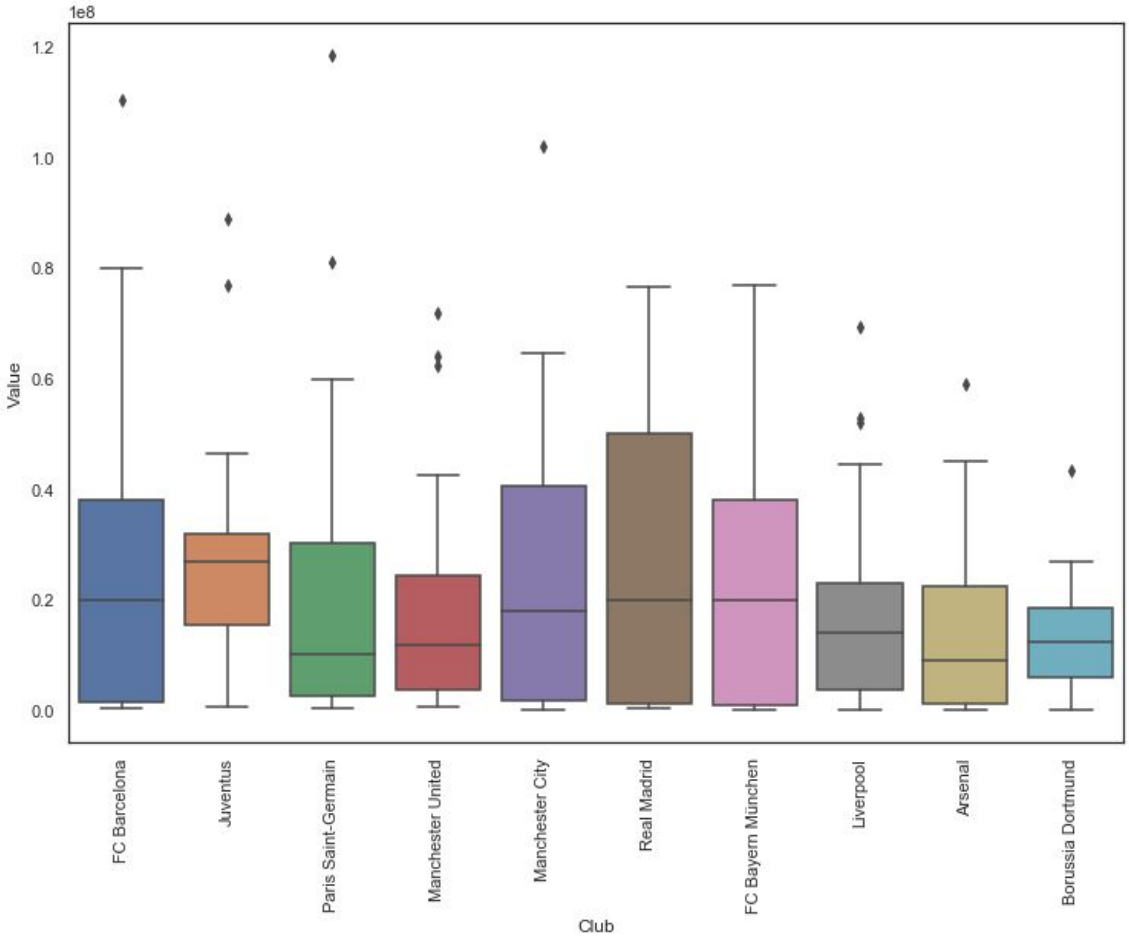


Player potential score tells a similar story. Countries with top potential scores are Spain, Netherlands, Argentina and Brazil. All of these scores are evenly spread. We can also see the many outliers that most of these countries demonstrate.

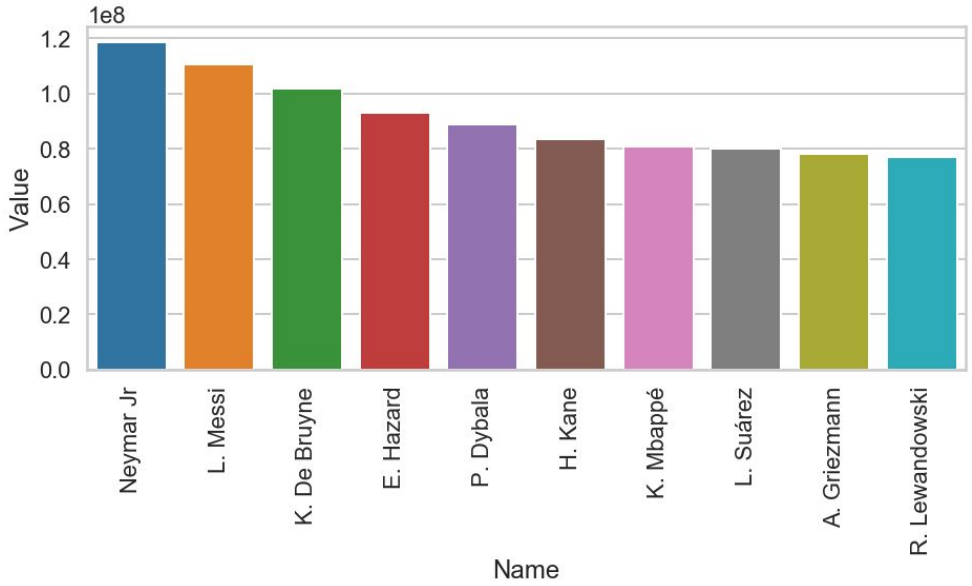
IX. Top Clubs Wages and Values



As we observe the Wages for the top 10 clubs, we can see that all of these clubs are European. However, they are split between different “leagues” such as Champions League(English clubs: Manchester) and the Spanish League (Barcelona, Madrid). This does not mean that only European nationalities are in the top 10, because different nationalities can play for different clubs during the normal season. We can see that the highest paid players are in Real Madrid and Barcelona, followed by Juventus that seems to value a smaller “range” of Wages that are still high on average. We can check whether highest valued players are also in the same clubs.



This was expected as players who are paid the highest “Wages” have a direct correlation with their “Value”. One detail to be noted is that on average, Juventus players are of the highest value. It is likely that the highest rated players are all in one of these clubs with perhaps one or two exceptions.



X. Attributes that may be good predictors of player's Value.