DATA ANALYSIS: FIFA 2012

INTRODUCTION:

FIFA 12 is a football simulation video game developed by EA Canada and published by Electronic Arts worldwide under the EA Sports label. The game features the UEFA club competitions for the first time, including the UEFA Champions League and UEFA Europa League.

DATASET:

The dataset consists information about attributes of FIFA players. The original dataset contains 89 variables that describe 16,924 players. After cleaning the outliers, the data points were brought down to 14,518. Variables that were included in the study were:

SR. NO	VARIABLE NAME	DESCRIPTION
1.	Age	Age of the player
2.	Strength	Strength of the player
3.	Weights	Weights of the player
4.	Penalties	The Fouls committed resulting in penalties of the Player
5.	Vision	The ability to see potential passes
6.	Ball Control	Having the ability to control the ball in whatsoever situation
7.	Dribbling	Maintaining the possession of the ball across the field
8.	Balance	Balance of the Player
9.	Acceleration	Ability to accelerate quickly for success in their positions
10.	Agility	Agility is how quickly a player can change directions without losing balance
11.	Sprint speed	The ability of a player to run for the possession of the ball.
12.	Reactions	Quick Reaction to various twists and turns in the match
13.	Overall	Overall skills of the player
14.	Potential	It is the ability of a player of how good they can possibly become
15.	Aggression	Aggression is a characteristic that can have many negative as well as positive effects on performance
16.	Stamina	The ability to perform throughout the match

SOFTWARE USED:

R-Studio : Data Cleaning, Data Visualization, Correlation Analysis, Hypothesis Testing

METHODS USED FOR DATA ANALYSIS:

- Histogram
- Cluster Analysis to correlation matrix

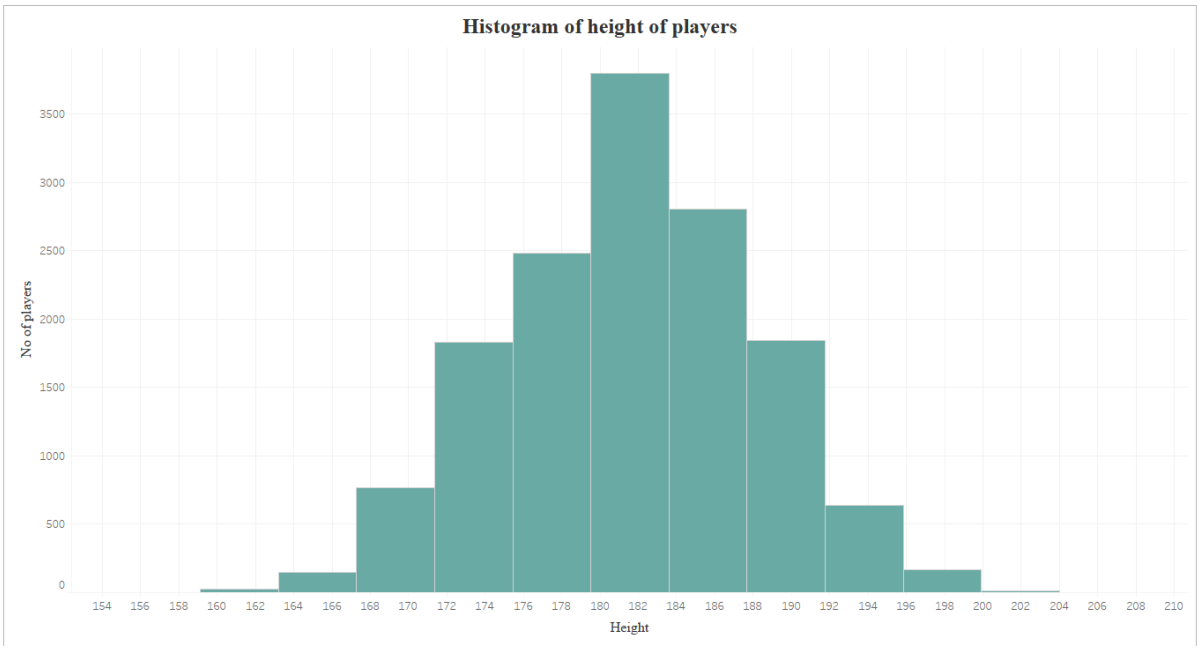
- Scatterplots
- Hypothesis Testing
- Word Cloud

OBJECTIVES:

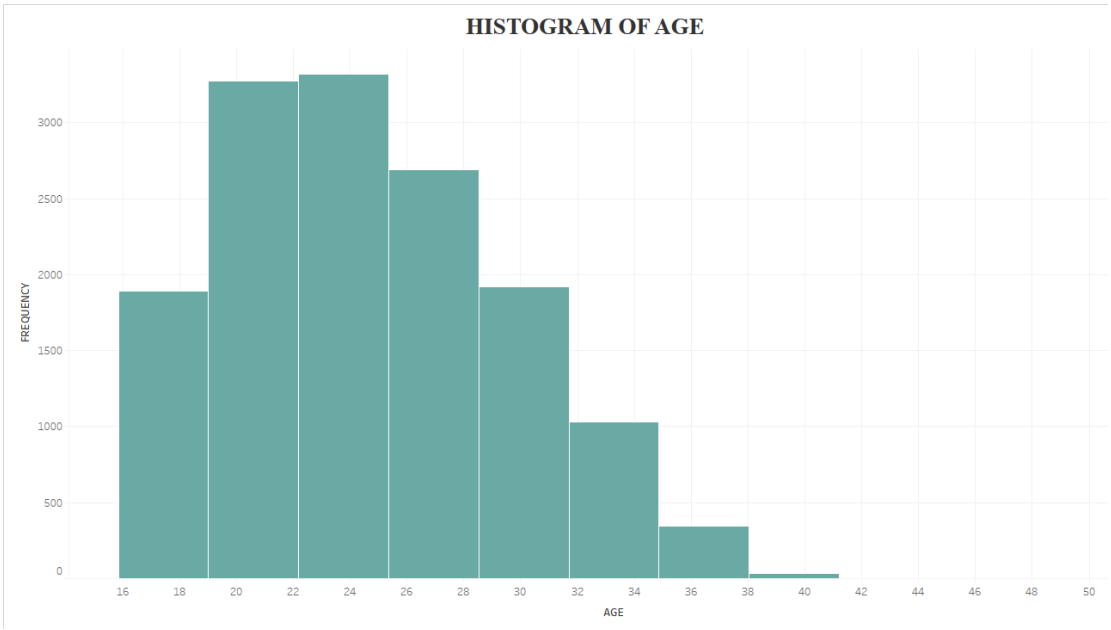
1. Study the impact of different attributes on player's performance.
2. Check whether claims about attributes of player's hold true.

ANALYSIS:

1. **Height**
   Mode range: 180 – 184 cms



2. **Age**: Most of the Fifa players are young.



3. **Weight**

Mode: 70 – 75 kgs

4. **T - test for single population mean**

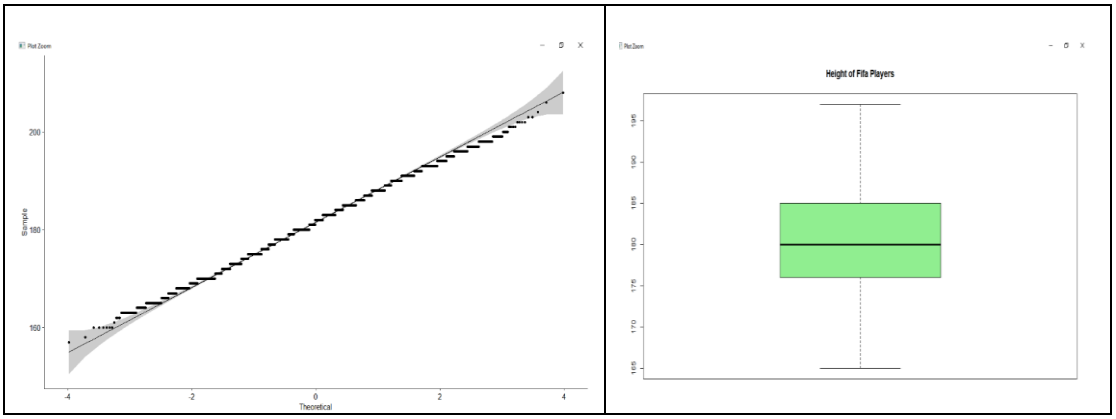
Assumptions:

i. Dependent variable is continuous

ii. Observations are independent in nature.

iii. Dependent variable 'height' comes from a normal population.

iv. dependent variable does not contain any outlier



*All assumptions were met before performing the tests.

Football Observatory claims that the height of a FIFA player on an average is 181.cms

H_o: Average height of FIFA players is 181.7 cms

VS

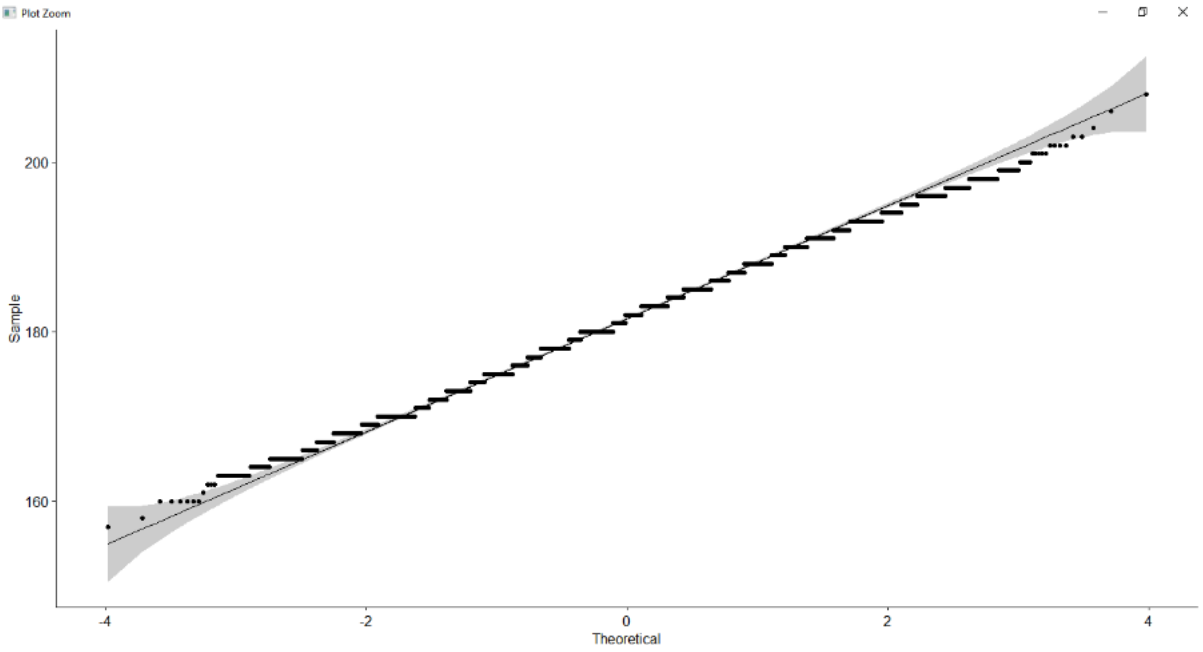
H_1: Average height of FIFA players is NOT 181.7 cms

Results: p-value< 2.2e-16

Level of Significance = 0.05.

Conclusion:
Reject H_0 in favour of H_1 and conclude that the average height of FIFA players is different than 181.7 cm

5. **Test for population variance (Chi- square test for population variance)**
   i. Dependent variable is continuous.
   ii. Dependent variable 'height' comes from a normal population.



* Assumptions were met before proceeding with the test
Claim: VARIANCE OF HEIGHT OF FIFA PLAYERS IS 30

H_0: Variance of height of FIFA players is 30
vs
H_1: Variance of height of FIFA players is greater than 30

Results: p-value < 2.2e-16
Level of Significance= 0.05

Conclusion:
Reject H_0 in favour of H_1 and conclude that the variance of height of FIFA players is greater than 30.

6. **To test the significance of single proportion (Chi-Square test)**
   Assumptions:
   1. The sample is a fair representation of the population.
   2. Items are dichotomous
   3. Sample items are independent

   CLAIM: PENN STATE UNIVERSITY CLAIMS THAT 75% OF FIFA PLAYES PREFER THEIR RIGHT FOOT

   H_0: Proportion of fifa players who prefer their right foot is 0.75

VS

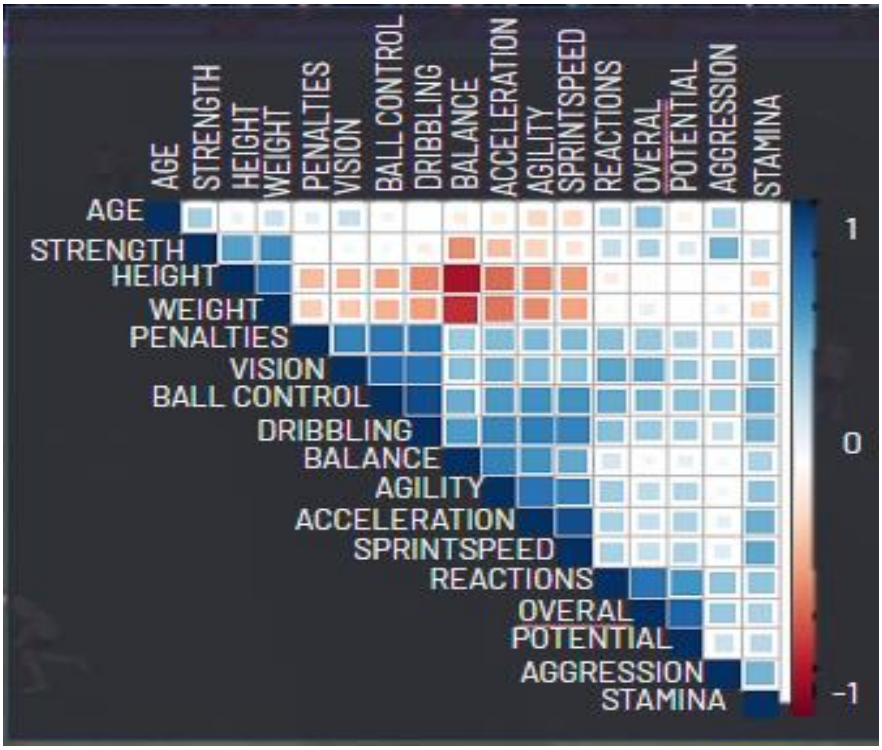
H_1: Proportion of fifa players who prefer their right foot is not 0.75

Results: p-value = 0.2248

Level of significance= 0.05

Hence, we accepted H_0 in favour of H_1 and conclude that PROPORTION OF FIFA PLAYERS WHO PREFER THEIR RIGHT FOOT IS 0.75 AT 5 % LEVEL OF SIGNIFICANCE.

7. **Cluster analysis to Karl Pearson's correlation matrix to obtain the significant results quicker.**



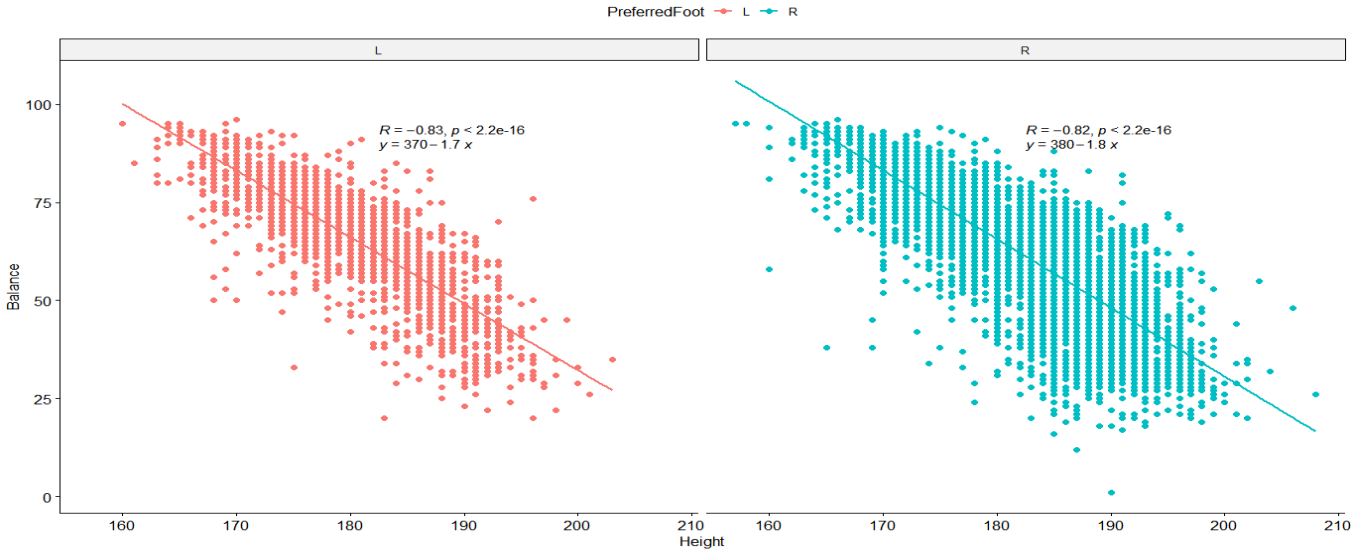
Strong relationship was observed between the following:

- Vision v/s Ball control
- Dribbling v/s Acceleration
- Ball Control v/s Reaction
- Reaction v/s Vision
- Sprint Speed v/s Stamina
- Strength v/s Aggression

8. **Faceted Scatter Plots**

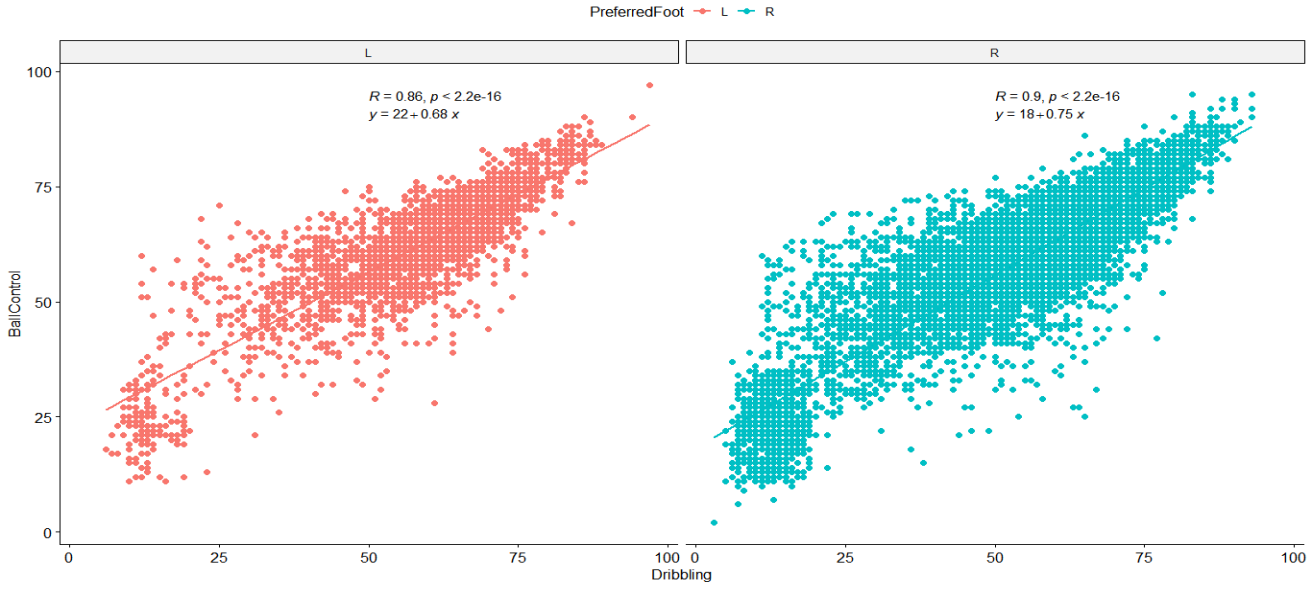
(i) Height vs balance based on preferred foot

It has an indirect linear relationship, i.e., balance decreases with increase in height.



(ii) Dribbling vs Ball control based on preferred foot

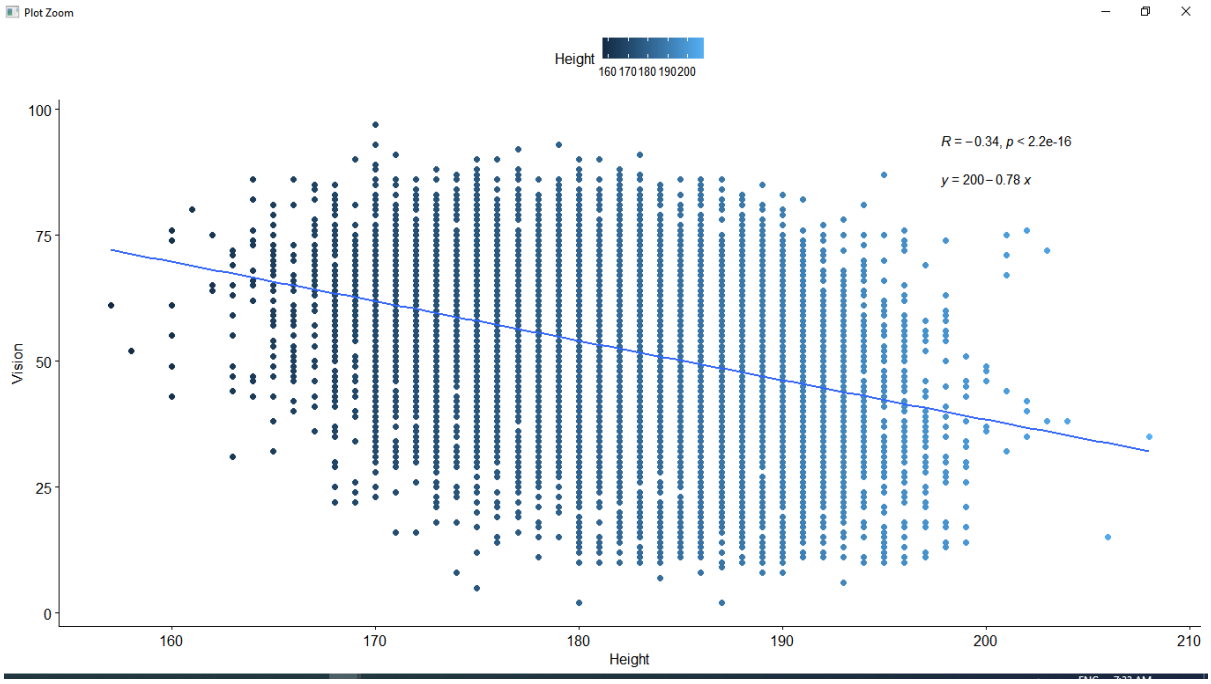
It has an direct linear relationship, i.e., ball control and dribbling skills are positive related.



9. Scatter Plots

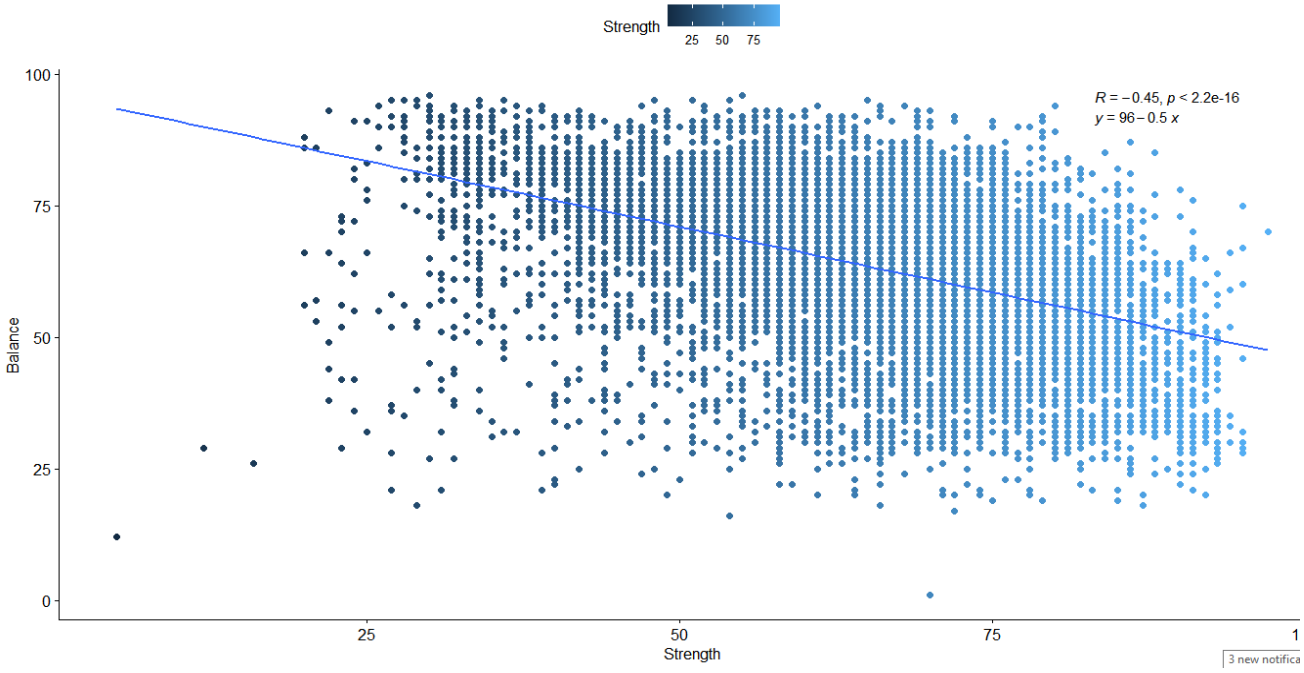
(i) Height vs Vision

Surprisingly enough, taller players have shown lesser vision(ability to see potential passes) than shorter players.



(ii) Strength vs Balance

The strength of a FIFA player shows a negative linear relationship with the balance.



10. WORD CLOUD

The attributes like strength, speed, tackling, playmaking and many other is what defines a football player. To get know about the relationship between these attributes are necessary to understand the ability and role of a specific player in his position.

CONCLUSIONS:

1. Stamina, Reaction, Vision, Ball control, Balance of height and weight are the key attributes in FIFA Players.
2. The average height of the players turned out to be lesser than 181.7 cm
3. 75% of the FIFA players prefer to use their right foot.
4. Variance of Avg Height of Players is greater than 30.
5. There is an unusual negative linear relationship between Height and Vision.