|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year | Strength Team 1 | Strength Team 2 | Venue | Toss | Result |
| 2018 | 121 | 119 | 0 | 1 | 0 |
|  |  |  |  |  |  |

Ayush (27/06/2020):

We will have notations related to team 1 always

Eg 1 : Team 1 : MI

Team 2 : CSK

Venue : MI

Toss : CSK

Result : MI

Notations : 0 – Home Team

1 – Foreign Team

When we are calculating the strength of Team 1 and 2 for the match played in IPL Season 2018, we will take stats of players from IPL(Starting year , we have to decide on this) till IPL (2018)

**Strength Calculations**

**Strength of team = Strength of Batsman + Strength of Bowlers + Strength of Fielders**

**Strength of Batsman = Recent Strength(2 years) \* W1 + Rest years Strength \* W2**

**Strength of Bowler = Recent Strength(2 years) \* W1 + Rest years Strength \* W2**

**Recent Strength/ Rest years Strength(Batsman) = W1 \* AVG + W2 \* SR ……..**

**Recent Strength/ Rest years Strength(Bowler) = W1 \* AVG + W2 \* SR ……..**

**W1 , W2 are the weights that we have to decide for our algos**

**Rashid and Vatsal I expect from you to understand this model and try to improve it as much as you can…. The more we explore and read researches and understand them the more we will find loopholes and the chances to improve it…**

**Doubt: While Calculating the Strength for eg Match 24 of season 2018.**

**1 – Should we take stats of player till IPL Season 2018(till 2017 end) or**

**2 - Should we take stats of players till IPL Season 2018 Match 23.**

**I discussed this with Rashid today, Vatsal put your thoughts and analyse as much as you can….**

**Kudos to the team we are doing great work… Keep going guys…**

**Vatsal : 28/06/2020**

**We’ll calculate the weights of each player till the 2019 data which we’ll keep standard formula for batsman and bowlers.**

**Overall form : Weight of player in each year from 2015-2019/5**

**Recent form: Weight of player in 2018 and 2019/2**

**In calculating the team strength, we’ll give more weightage to recent form data and we’ll select the best 11 players.**

**Since teams have changed for 2020, our task would be to create a dataframe which would contain two columns:**

**Overall form, recent form.**

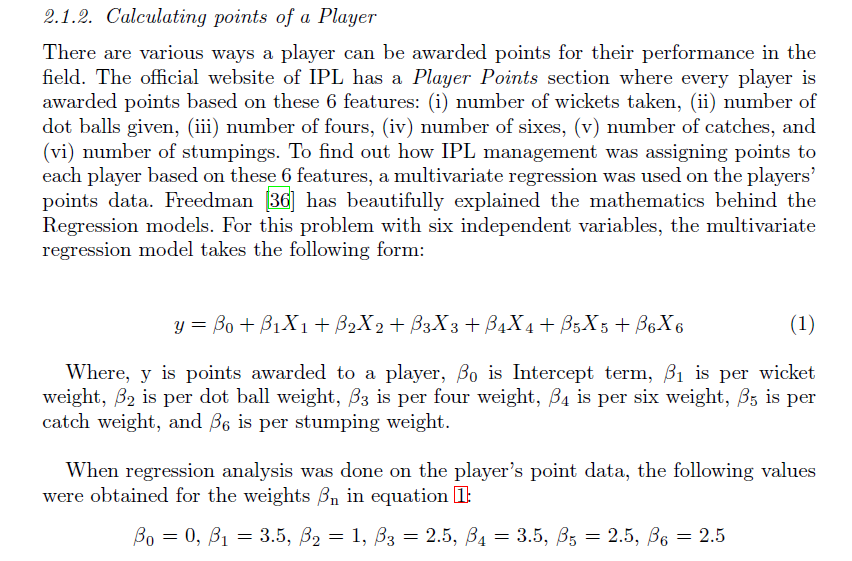
**We’ll select top 11 players from these list. Ig we’ll have to keep them same throughout our series.**

**Using this, we’ll calculate the team strength for 2020 teams.**

**As Ayush explained above, further we’ll be able to calculate the team predictions.**

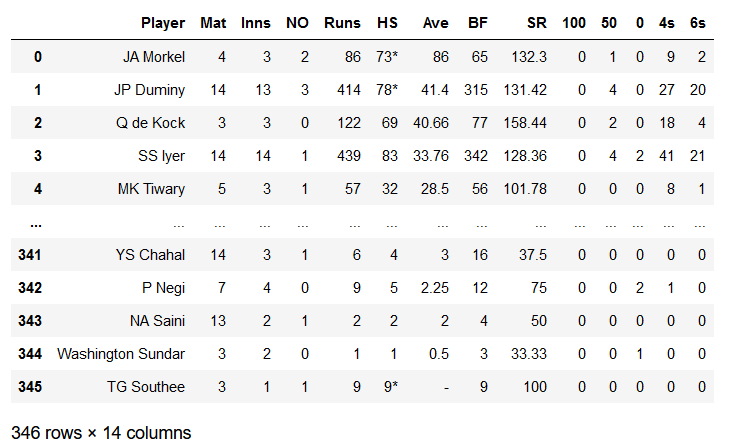
**Biggest Hurdle: How should we calculate weights ?????**

**29/06/2020: Vatsal**



**Here, they haven’t segregated as batsman and bowler points which we are going to do.**

**We’re going to consider different features for batsman to calculate their weights and different for bowlers to calculate their weights. Therefore, our model would be advanced than theirs.**



**They haven’t written what they’ve used as their dependent variable i.e, what they are predicting using their model.**

**I tried predicting runs using different attributes like taking only the features for which we want weights:**

**Average, SR, Innings, 50, 100 to predict runs.**

**The accuracy is always close to 1 for both train and test set.**

**I don’t know where am I going wrong. But there s this two things I found out and confuses me:**

1. **I am getting even negative values of y\_preds i.e, how do I make the model know that runs cannot be -ve. I even applied Standard Scaler.**
2. **The runs are not given for a particular match so I have to change my dependent variable. What should it be then ?**

**(All of this using Multiple Linear Reegression.)**

**I haven’t started bowler’s prediction yet.**