**Member:**

1. Aswadi Abdul Rahman (WQD180082)
2. Lee Kwan Li (WQD180019)
3. Chai Kun Ting (WQD180040)
4. Zulkanain Hasan (WQD180031)

**Predict the best team for Fantasy Football**

group assignment 2 – wqd7011 Numerical optimization

2019

Contents

[**1.** **Introduction** 2](#_Toc10208464)

[**2.** **Dataset** 3](#_Toc10208465)

[**3.** **Objective Function** 3](#_Toc10208466)

[**4.** **Constraints** 3](#_Toc10208467)

[**5.** **Python** 5](#_Toc10208468)

[**6.** **Justification** 7](#_Toc10208469)

[**7.** **Conclusion** 8](#_Toc10208470)

1. **Introduction**

Fantasy football is a game in which the participants serve as the general managers of virtual professional football teams. The competitors choose their team rosters by participating in a draft in which all players of a real football league are available. Points are based on the actual performances of the players in the real-world competition.

Once you register yourself for this you are provided with virtual money which you need to spend on selecting 15 players from the 20 teams who are in the Premier League.



*Figure 1: Fantasy football player’s selection GUI*

1. **Dataset**

Source from <https://www.kaggle.com/delayedkarma/fantasy-premier-league-20182019/downloads/fantasy-premier-league-20182019.zip/6>

There will be 21 columns/variable and around 500 rows (depend on which week we choose) inside the dataset. For simplicity we will just choose 8 columns and week 0 (FPL\_2018\_19\_Wk0.csv) for this assignment. Below are the first 10 samples and some of the important column.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No | Name | Team | Position | Cost | Assists | Yellow\_cards | Minutes | Points |
| 1 | Adam Smith | BOU | DEF | 45 | 3 | 6 | 2067 | 56 |
| 2 | Adrian | WHU | GKP | 45 | 0 | 2 | 1710 | 72 |
| 3 | Aguero | MCI | FWD | 110 | 6 | 2 | 1960 | 169 |
| 4 | Ake | BOU | DEF | 50 | 3 | 5 | 3352 | 102 |
| 5 | Albrighton | LEI | MID | 55 | 8 | 5 | 2533 | 107 |
| 6 | Alderweireld | TOT | DEF | 60 | 0 | 3 | 1177 | 43 |
| 7 | Alexander-Arnold | LIV | DEF | 50 | 2 | 3 | 1573 | 83 |
| 8 | Alisson | LIV | GKP | 55 | 0 | 0 | 0 | 0 |
| 9 | Alli | TOT | MID | 90 | 13 | 7 | 2957 | 175 |
| 10 | Alonso | CHE | DEF | 65 | 2 | 6 | 2855 | 165 |

1. **Objective Function**

Given a budget (total cost) 1200 and 15 players for each team constraints, we want to maximize the point earn by using Constrain Optimization (Linear Optimization) algorithm. The 15 players should be divided into below criteria: (2GKP, 5DEF, 5MID, 3FWD)

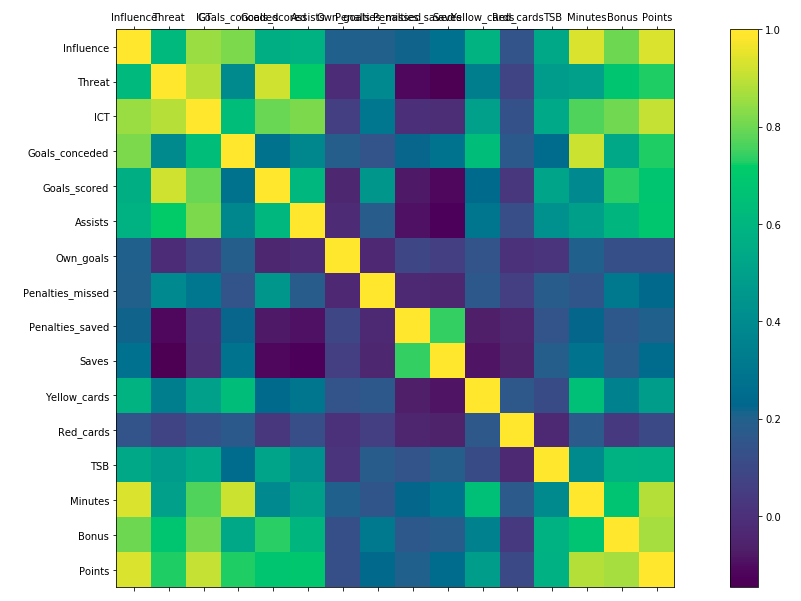
* 2 x Goal Keeper position (GKP)
* 5 x Defender position (DEF)
* 5 x Mid field position (MID)
* 3 x Forward position (FWD)

By representing each player/row into different variable we can derive point maximizing using below mathematical equation

56\*x0 + 72\*x1 + 169\*x2 + 102\*x3 + … Points\*xn

1. **Constraints**

Before we chose a value for each constraint, we try to find it’s correlation using a Pearson correlation. Example in below figure we can see Minutes highly positive correlate (indicate in greenish yellow) to the Points.



*Figure 2: Pearson Correlation Visualisation*

Below are all the constraints that we need to use while trying to find the maximum points.

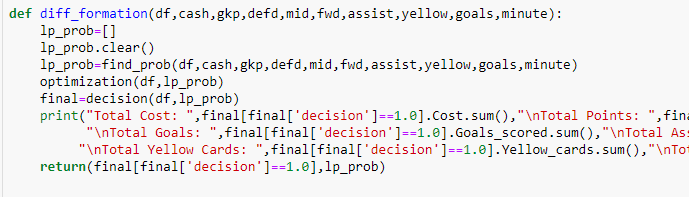
1. Cash constrain: 45\*x0 + 45\*x1 + 110\*x2 + 50\*x3 + … Cost\*xn <= 1200
2. GKP Player Position: x1 + x30 + x38 + x48 + … xn (GKP position ONLY) = 2
3. DEF Player Position: x0 + x3 + x5 + x6 + … xn (DEF position ONLY) = 5
4. MID Player Position: x4 + x8 + x10 + x11 + … xn (MID position ONLY) = 5
5. FWD Player Position: x2 + x12 + x15 + x18 + … xn (FWD position ONLY) = 3
6. Total assist constraint: 3\*x0 + 6\*x2 + 3\*x3 + … Assist\*xn >= 90
7. Total yellow card constraint: 6\*x0 + 2\*x1 + 2\*x2 + 5\*x3 + … Yellow Card\*xn <= 20
8. Total goals scored constraint: x0 + 21\*x2 + 2\*x3 + … Goals\_scored\*xn >= 150
9. Total minutes played constraint: 2067\*x0 + 1710\*x1 + 1960\*x2 + 3352\*x3 + … Minutes played\*xn >= 44100

All the total value (example total goals >= 150) was defined using normal statistic such as mean to indicate that we expect the players that we get is better than average players.

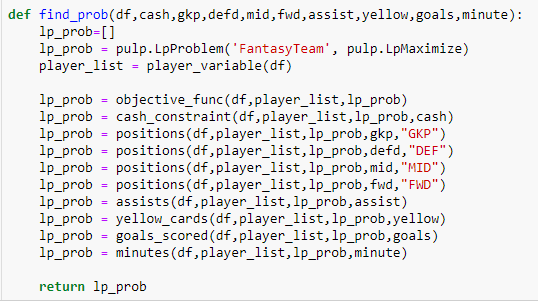
1. **Python**

The main code start initiated from this code

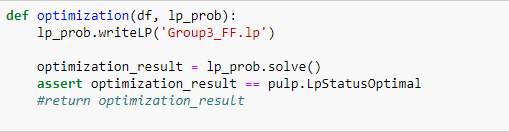
*final, lp\_prob = diff\_formation(df,1200,2,5,5,3,90,20,150,44100)*



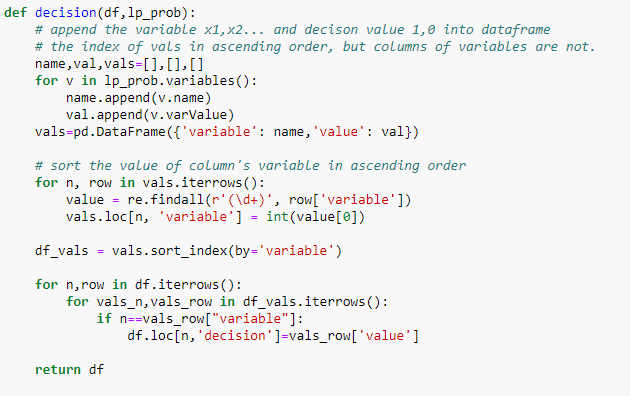
Then after wards the code will call above funtion and it will start with creating mathematical equations from each constraint.



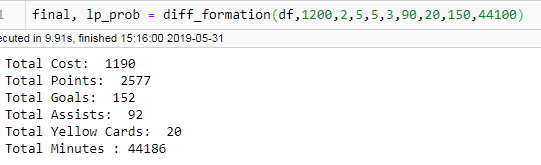
Above are main function to get the mathematical equation of each constraint and append in into the result which in here lp\_prob.



Then all the constaint mathematical equations is being solve by optimization function which solve the Linear Optimization equations as shown in above code.



Then lastly it will use above code decision function to itterate the mathematical equations and form dataframe for all the players/variable.



Above snapshot shows that the main code run successfully and print some of the constraints that we use.

1. **Justification**

After running the whole script, below are the output and basically the chose player.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No | Name | Team | Position | Cost | Assists | Goals\_score | Yellow\_cards | Minutes | Points |
| 1 | Azpilicueta | CHE | DEF | 65 | 6 | 2 | 1 | 3330 | 175 |
| 2 | Bertrand | SOU | DEF | 50 | 5 | 0 | 2 | 3135 | 103 |
| 3 | Christensen | CHE | DEF | 55 | 0 | 0 | 0 | 2067 | 79 |
| 4 | Daniels | BOU | DEF | 45 | 3 | 1 | 0 | 3001 | 94 |
| 5 | Eriksen | TOT | MID | 95 | 10 | 11 | 0 | 3218 | 199 |
| 6 | Fabianski | WHU | GKP | 45 | 0 | 0 | 0 | 3420 | 157 |
| 7 | Firmino | LIV | FWD | 95 | 8 | 15 | 1 | 2760 | 181 |
| 8 | Kane | TOT | FWD | 125 | 2 | 29 | 5 | 3074 | 217 |
| 9 | Mahrez | MCI | MID | 90 | 13 | 12 | 2 | 2948 | 195 |
| 10 | Ryan | BHA | GKP | 45 | 0 | 0 | 0 | 3420 | 146 |
| 11 | Salah | LIV | MID | 130 | 12 | 32 | 1 | 2905 | 303 |
| 12 | Son | TOT | MID | 85 | 8 | 12 | 0 | 2292 | 178 |
| 13 | Sterling | MCI | MID | 110 | 17 | 18 | 3 | 2584 | 229 |
| 14 | Vardy | LEI | FWD | 90 | 2 | 20 | 3 | 3248 | 183 |
| 15 | Walker | MCI | DEF | 65 | 6 | 0 | 2 | 2784 | 138 |

So we test our output based on the constrain that we define previously.

|  |  |  |  |
| --- | --- | --- | --- |
| *Constraint* | Plan | Actual | Status |
| *Cost* | <= 1200 | 1190 | OK |
| *Goals* | >= 150 | 152 | OK |
| *Assists* | >= 90 | 92 | OK |
| *Yellow Cards* | <= 20 | 20 | OK |
| *Minutes* | >= 44100 | 44186 | OK |

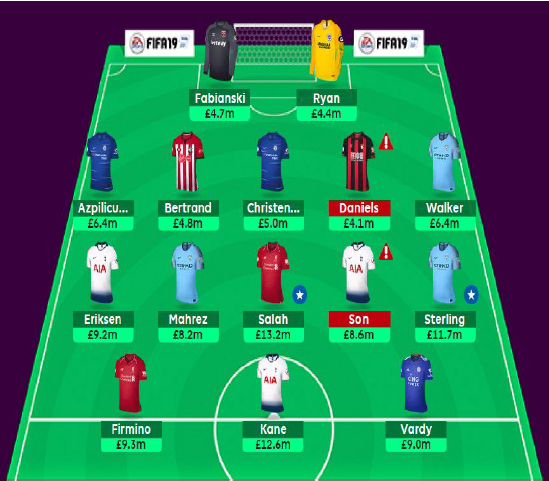
We also test for the position constraint that we define earlier

|  |  |  |  |
| --- | --- | --- | --- |
| Constraint | Plan | Actual | Status |
| Goal Keeper (GKP) | 2 | 2 | OK |
| Defender (DEF) | 5 | 5 | OK |
| Mid Field (MID) | 5 | 5 | OK |
| Forward (FWD) | 3 | 3 | OK |

In the end, we got **2577** points.

1. **Conclusion**

Linear Optimization can be used as a tool to solve for most of constraint problem. Below are the 15 players that being chose by the algorithm.



*Figure 3: Final selected player in Fantasy Football GUI*