**FPL Project**

**Issues**

Save top 100 players for each metric 2023/24

**Immediate deliverables:**

1. Code to read in whole season fixtures.
2. Code to determine the fixture difficulty , given a list of dummy fixtures. NOT NEEDED- THE FDR IS PUBLISHED ON FANTASY WEBSITE- HOWEVER, THIS IS ONLY BASED OFF THE PREVIOUS 6 GAMES. MAYBE DO YOUR OWN FDR AND COMPARE THE MOST IMMEDIATE RESULTS TO THE WEBSITE ONE
3. Code to include difficulty of upcoming 10 fixtures in player reports? Maybe just a string of 10 numbers. (Could come back later and combine into a single calculated value. Would be good for automation).
4. Produce basic reports (excel or word?)

Note: conduct data analysis on stuff like FDR to produce report giving the acceptable range for each stat when choosing team.

Note: if you sort the above issues in time for start of season, you’ll be reasonably well set to have a good season. This should be the short-term goal.

Medium term goal: To achieve the medium term deliverables in time for next season.

[https://fantasy.premierleague.com/api/element-summary/{player-id}/](https://fantasy.premierleague.com/api/element-summary/%7Bplayer-id%7D/) - get upcoming fixtures etc for one player- way to do it if you want to manually write fixture difficulty algorithm

**Medium term deliverables:**

1. Code to write the output to an excel file/ Word doc (based on y/n prompt?)
2. Update code to include managers
3. Optimisation of the winning strategies (if not, just use blend of successful strategies as per internet with stats provided by my basic code) (Monte Carlo?). Stuff like optimal number of defenders (dynamic or static across season? Captain could be a defender? Prob not) Base the optimisation strategies off last year and only compare them in relative terms i.e. only compare optimisation strategies to each other, don’t use the output from 2023 strategy to predict the 2024 strategy. **RESEARCH THIS AREA.** Could apply strategies and models to training sets (i.e. 2016/2017) and take more recent years as test sets (i.e. 2021/2022). Would need to consider the average points of 75th percentile etc as per PPGPE here, to account for the fact that you won’t know exactly which player will get most points over whole season while still in season. For example, if you just use the top 11 PPGPE players (no average), your thinking will be slightly biased towards low-cost players (as they’ll show more regression to the mean?). During the season the cost will be a reasonably accurate predictor of future points (back this up with analysis?) and can be combined with past points/form for rough prediction of predict future points (NO- YOU’RE TRYING TO FIND ADVANTAGES OVER REST OF FPL MANAGERS BY CHOOSING HIGH PPGPE PLAYERS). Will need to make classes for the players to make this process efficient enough to be possible.
4. Do descriptive stats on all the main metrics and draw conclusions from these? Need a) average nonzero b) benchmark figures? **Histograms** could be key here- show spread of player costs, points, correlation between these etc. Finish online Python course before this.

**Long term deliverables**

1. Code to analyse your team, produce graphics quickly showing what is substandard. Maybe in future have a) delta in brackets in player report b) alerts based off deltas.
2. Code to return the player’s team in the player reports using the 2024/25 teams list. (Replace the numbers in this with team names in text). Team names of past players is not necessary.
3. Set up 1-3 overall metrics which combine form, PPGPE, fixtures etc so that players can be analysed in one go
4. Live fixtures read in from PL website or similar
5. Alert in report/email about players w/ less than 100/100 likelihood of playing?. Also include recent information and the date of the information.
6. Code to write to Word
7. Make a function to add up a team’s scores?
8. Load in XG and its variants (from other sources?)
9. Store all dataframes before they become outdated, so you can simulate strategies across a season
10. Delta % counter (If average PPGPE is above 0.0577, will achieve 2800 points (Check calc). Captain price halving will likely skew data upwards- average PPGPE of 0.052 would probably be sufficient. How to reflect captain cost in that? Separate evaluation system?
11. Past scores- weighting system based on what year the scores were earned. Read in the data from past seasons and condense down to ppg (divide points by no. games played this time). Maybe variance/consistency are better metrics for this. That being said, past performances will be very costly, so you’d be looking to exploit form as much as possible. MAYBE DON’T USE THIS SYSTEM- ARBITRARY
12. Search by player, team
13. Search by best in chosen metric by position- GK, def, mid, fwd, all
14. Include games and minutes played
15. Upload to Git
16. Code to automatically suggest replacements
17. Store all your predictions in a weekly document and note the delta- script to do this.
18. Regressions- Verbalise the main factors yourself, read books, articles, papers quantifying these, conduct tests yourself. Regressions on important factors as brainstormed above- only in very advanced stages (make sure they don’t add up to more than 1- should be statistically significant).
19. Take principles from investing and apply to this. Principles, algorithms etc. ML?
20. Form checker- function to allow you to determine when to drop a player. Include stuff like last 5 opponents, next 5 opponents, goals, assists in *nice graphic (key to this)*. (Maybe include whoscored.com data? Maybe too advanced for now). Include the XP in this- check past scores to what was predicted and if it was near the XP but future scores are predicted to be low, drop the player as they likely just had an easy run of games. Maybe the whole thing should just hinge on this. Have graph of points vs XP vs 6.68 ppg. Value.event\_points- important metric. Maybe a replacement for this would be the graph of points earned over past few games. Include team logos underneath each of these.
21. Replace the user input with a GUI
22. Completely automate the selection process, then just keep adding algorithms to account for heuristics/ common fpl strategies
23. V advanced- could import a pretrained DL model and use that as a team

**End goals:**

1. Fully automated team management assistant
2. Quantitative model which makes predictions and gives uncertainty

**Thoughts**

Get over 2000 points should be the short-term aim, not 2800

Note: Good way to check fuzzy matching threshold matching would be to print out the desired and matched names beside each other.

**Make a function to return expected points from a player over the course of a season. Doesn’t have to be massively complicated- maybe just base off past seasons and upcoming fixtures etc. Use this to perform optimisation of winning strategies? Could introduce expected variations into this, have rules on when to drop player.**

Will have to watch for double/ zero gameweeks yourself- fixtures import script won’t do that for you.

**Assumptions of model/ strategy:**

* Sustained prior performance is a small indicator of future success- nothing else is significant (as per investing, past performance is not…)
* The only other significant indicator of future success is upcoming fixture difficulties- equivalent of knowledge about future market conditions
* Optimal strategy is to buy a number of big hitters and populate the rest of team with high PPGPE players.
* Deviation of player form from expected points is only statistically significant when routinely outside expected variation (standard deviation at certain confidence interval? Work out the maths of this).

**Flights of fancy**

* API to take odds from betting websites? Only in very advanced stages, when you’re looking at optimizing the delta. Maybe just use it sparingly to compare specific things, or as a tester for your regressions in future. Would need a function to remove margin and convert odds to probability. COULD BE VERY GOOD WAY TO RATE FIXTURE DIFFICULTIES TO HIGH DEGREE OF ACCURACY
* Past injury frequency- weight XP according to this (check papers to see likelihood of reinjury for weighting. Probably too precise to make a difference here.)
* Factor- teams (ELO), injuries, form, position etc

**Code cleanliness issues/ error handling:**

* Make integers of all columns and change the print statements in player\_report from this.
* Merge first and surnames using pandas (fuzzy merge and try/catch)- it’s picking up the wrong web\_name variables at present.
* Sweep through project, inserting error handling techniques throughout
* Code to make sure that you’re reading in data as of the current week.
* Script to store the current week’s data and name it accordingly
* Try/catch to account for players not being in prem in past and/or current teams not being in prem in past

**Bugs/issues solved:**

* Gobbledygook appearing in front of the first stout statement- just make the output window narrower
* Player names disappearing (especially the last player’s name)- problem disappeared when I rebooted the computer.
* Get\_ranking() giving incorrect values even after sorting- dataframes keep the original indices, even after sorting. Have to use the following code to reset it and thus reflect the new order: sorted\_df.reset\_index(drop=True, inplace=True)

**Heuristics based on traditional fpl strategy**

* Track XG and its variants for each player- positive (XG-G) means keep, negative means sell as they're a player for bandwagoners. Don't pay attention to points.
* Trust a core group of players long term- maybe this is an unsubstantiated heuristic
* Planfpl.com and fpl.team (captaincy, subs, transfers). Fantasy football fix and FPL review as validation
* Noise/information theory
* In the second half of the season, use wildcard in preparation for double and blank gameweeks.
* Only consider effective ownership in real 50/50 decisions. Should you be cautious or risk? F1 pitstops probably have answer
* Hold transfers whenever possible
* Delay moves till end of week- set up a 2 day protocol
* Captaincy- form (want them to do well that week)
* Transfers- form/fixture/home (want them to do well longer term)
* Triple captain- good player for a great team has a double
* Bench boost for double gameweek- especially with goalkeeper
* First wildcard- used whenever you can see a big fixture swing. Allows you to go relatively short term with initial picks
* Free hit- navigate a blank week
* Second wildcard- doubles, blanks, and fixture swings
* Fixtures are king
* Fixtures create form
* Make sure it's actually a good fixture though i.e. is it a good defensive fixture or attacking
* Better to get 2-3 points from a bench player than take a hit for a starting player who has a small injury/suspension

**Getting around the "past performance is no..." issue**

1. Quantitative analysis- base on current and future conditions, rather than past performance

2. Fundamental analysis

3. Hedging strategies (i.e. don't go completely alternative, need to mitigate EO somewhat

4. Active management

5. Enhanced due diligence

**Other thoughts**

* Strategy- set target points per week and go off that
* 1 or 2 big players, choose the rest off efficiency?? Number of big players can be chosen w/ optimisation process. Definitely 1 v big player for captain.
* Calculate\_scores() may not be needed in my code. Probably not, given we already have an XP.
* *Possibly pursue a Southampton strategy in opening weeks to earn a greater budget? Would have to start dirt cheap with players who will be facing shite teams at start of season. Investigate the price changes from start to end of season in best case scenario, divide by length of time and adjust expectations for optimism bias then.* Answer: Only 1.6% of players change by more than 6 mill, so probably not worth it
* If you want to do a really unconventional team/series of moves, use the selected\_by\_percent.
* Not all EP\_next are given. Same for EP\_this. May need to do your own method for predicting this?
* If average PPGPE is above 0.0577, will achieve 2800 points

**Column names**

chance\_of\_playing\_next\_round

chance\_of\_playing\_this\_round

code

cost\_change\_event

cost\_change\_event\_fall

cost\_change\_start

cost\_change\_start\_fall

dreamteam\_count

element\_type

ep\_next

ep\_this

event\_points

first\_name

form

id

in\_dreamteam

news

news\_added

now\_cost

photo

points\_per\_game

second\_name

selected\_by\_percent

special

squad\_number

status

team

team\_code

total\_points

transfers\_in

transfers\_in\_event

transfers\_out

transfers\_out\_event

value\_form

value\_season

web\_name

minutes

goals\_scored

assists

clean\_sheets

goals\_conceded

own\_goals

penalties\_saved

penalties\_missed

yellow\_cards

red\_cards

saves

bonus

bps

influence

creativity

threat

ict\_index

starts

expected\_goals

expected\_assists

expected\_goal\_involvements

expected\_goals\_conceded

influence\_rank

influence\_rank\_type

creativity\_rank

creativity\_rank\_type

threat\_rank

threat\_rank\_type

ict\_index\_rank

ict\_index\_rank\_type

corners\_and\_indirect\_freekicks\_order

corners\_and\_indirect\_freekicks\_text

direct\_freekicks\_order

direct\_freekicks\_text

penalties\_order

penalties\_text

expected\_goals\_per\_90

saves\_per\_90

expected\_assists\_per\_90

expected\_goal\_involvements\_per\_90

expected\_goals\_conceded\_per\_90

goals\_conceded\_per\_90

now\_cost\_rank

now\_cost\_rank\_type

form\_rank

form\_rank\_type

points\_per\_game\_rank

points\_per\_game\_rank\_type

selected\_rank

selected\_rank\_type

starts\_per\_90

clean\_sheets\_per\_90

Note: 4.56 million players at the start of week 1

Note: Leave team as it is for the season?

Research

Player: Cole Palmer

Position: Mid

Points: 244 (1)

Cost: 10.5 (3)

PPG: 7.2 (1)

PPGPE: 0.686 (52)

EP this: None (None%) (185)

EP next: 2.4 (nan%) (122)

2023/24 Points: 244 PPG: 7.2

2022/23 Points: 19 PPG: 1.4

2021/22 Points: 9 PPG: 2.2

2020/21 Points: 0 PPG: 0.0

2019/20 Points: 0 PPG: 0.0

Player: Erling Haaland

Position: Fwd

Points: 217 (5)

Cost: 15.0 (1)

PPG: 7.0 (2)

PPGPE: 0.467 (232)

EP this: None (None%) (364)

EP next: 5.0 (nan%) (6)

2023/24 Points: 217 PPG: 7.0

2022/23 Points: 272 PPG: 7.8

Player: Phil Foden

Position: Mid

Points: 230 (2)

Cost: 9.5 (7)

PPG: 6.6 (3)

PPGPE: 0.695 (42)

EP this: None (None%) (361)

EP next: 4.2 (nan%) (14)

2023/24 Points: 230 PPG: 6.6

2022/23 Points: 142 PPG: 4.4

2021/22 Points: 137 PPG: 4.9

2020/21 Points: 135 PPG: 4.8

2019/20 Points: 69 PPG: 3.0

Player: Jarrod Bowen

Position: Mid

Points: 182 (11)

Cost: 7.5 (21)

PPG: 5.4 (11)

PPGPE: 0.720 (27)

EP this: None (None%) (534)

EP next: 2.6 (nan%) (94)

2023/24 Points: 182 PPG: 5.4

2022/23 Points: 145 PPG: 3.8

2021/22 Points: 206 PPG: 5.7

2020/21 Points: 141 PPG: 3.7

2019/20 Points: 44 PPG: 3.4

Player: Martin Ødegaard

Position: Mid

Points: 186 (8)

Cost: 8.5 (10)

PPG: 5.3 (12)

PPGPE: 0.624 (93)

EP this: None (None%) (13)

EP next: 3.9 (nan%) (20)

2023/24 Points: 186 PPG: 5.3

Player: Bryan Mbeumo

Position: Mid

Points: 127 (47)

Cost: 7.0 (30)

PPG: 5.1 (15)

PPGPE: 0.729 (25)

EP this: None (None%) (102)

EP next: 2.4 (nan%) (112)

2023/24 Points: 127 PPG: 5.1

2022/23 Points: 150 PPG: 3.9

2021/22 Points: 119 PPG: 3.4

Player: Benjamin White

Position: Def

Points: 182 (10)

Cost: 6.5 (60)

PPG: 4.9 (17)

PPGPE: 0.754 (18)

EP this: None (None%) (23)

EP next: 4.6 (nan%) (8)

2023/24 Points: 182 PPG: 4.9

2022/23 Points: 156 PPG: 4.1

Player: Bruno Borges Fernandes

Position: Mid

Points: 166 (15)

Cost: 8.5 (11)

PPG: 4.7 (21)

PPGPE: 0.553 (159)

EP this: None (None%) (380)

EP next: 2.9 (nan%) (78)

2023/24 Points: 166 PPG: 4.7

2022/23 Points: 176 PPG: 4.8

Player: Richarlison de Andrade

Position: Fwd

Points: 122 (56)

Cost: 7.0 (31)

PPG: 4.4 (26)

PPGPE: 0.629 (89)

EP this: None (None%) (517)

EP next: 2.7 (nan%) (92)

2023/24 Points: 122 PPG: 4.4

2022/23 Points: 54 PPG: 2.0

2021/22 Points: 125 PPG: 4.2

2020/21 Points: 123 PPG: 3.6

2019/20 Points: 165 PPG: 4.6

Player: Joško Gvardiol

Position: Def

Points: 123 (55)

Cost: 6.0 (74)

PPG: 4.4 (28)

PPGPE: 0.733 (20)

EP this: None (None%) (363)

EP next: 4.2 (nan%) (13)

2023/24 Points: 123 PPG: 4.4

Player: William Saliba

Position: Def

Points: 164 (18)

Cost: 6.0 (73)

PPG: 4.3 (32)

PPGPE: 0.717 (28)

EP this: None (None%) (18)

EP next: 4.2 (nan%) (15)

2023/24 Points: 164 PPG: 4.3

2022/23 Points: 117 PPG: 4.3

2021/22 Points: 0 PPG: 0.0

2020/21 Points: 0 PPG: 0.0

Player: Leandro Trossard

Position: Mid

Points: 138 (32)

Cost: 7.0 (36)

PPG: 4.1 (43)

PPGPE: 0.586 (126)

EP this: None (None%) (22)

EP next: 3.4 (nan%) (39)

2023/24 Points: 138 PPG: 4.1

2022/23 Points: 154 PPG: 4.3

2021/22 Points: 141 PPG: 4.1

2020/21 Points: 132 PPG: 3.8

2019/20 Points: 104 PPG: 3.4

Player: Gabriel dos Santos Magalhães

Position: Def

Points: 149 (24)

Cost: 6.0 (92)

PPG: 4.1 (44)

PPGPE: 0.683 (53)

EP this: None (None%) (3)

EP next: 4.2 (nan%) (16)

2023/24 Points: 149 PPG: 4.1

Player: Jordan Pickford

Position: GK

Points: 153 (21)

Cost: 5.0 (179)

PPG: 4.0 (48)

PPGPE: 0.800 (10)

EP this: None (None%) (240)

EP next: 3.1 (nan%) (70)

2023/24 Points: 153 PPG: 4.0

2022/23 Points: 124 PPG: 3.4

2021/22 Points: 116 PPG: 3.3

2020/21 Points: 114 PPG: 3.7

2019/20 Points: 117 PPG: 3.1

Player: Kieran Trippier

Position: Def

Points: 111 (75)

Cost: 6.0 (76)

PPG: 4.0 (49)

PPGPE: 0.667 (61)

EP this: None (None%) (433)

EP next: 4.2 (nan%) (10)

2023/24 Points: 111 PPG: 4.0

2022/23 Points: 198 PPG: 5.2

2021/22 Points: 28 PPG: 4.7

2019/20 Points: 0 PPG: 0.0

Player: Alisson Ramses Becker

Position: GK

Points: 107 (85)

Cost: 5.5 (167)

PPG: 3.8 (57)

PPGPE: 0.691 (45)

EP this: None (None%) (323)

EP next: 5.0 (nan%) (3)

2023/24 Points: 107 PPG: 3.8

2022/23 Points: 162 PPG: 4.4

2021/22 Points: 176 PPG: 4.9

2020/21 Points: 140 PPG: 4.2

2019/20 Points: 122 PPG: 4.2

Player: Manuel Akanji

Position: Def

Points: 112 (72)

Cost: 5.5 (129)

PPG: 3.7 (69)

PPGPE: 0.673 (56)

EP this: None (None%) (353)

EP next: 3.8 (nan%) (22)

2023/24 Points: 112 PPG: 3.7

2022/23 Points: 82 PPG: 2.8

Player: Dwight McNeil

Position: Mid

Points: 124 (50)

Cost: 5.5 (97)

PPG: 3.5 (76)

PPGPE: 0.636 (86)

EP this: None (None%) (236)

EP next: 2.0 (nan%) (161)

2023/24 Points: 124 PPG: 3.5

2022/23 Points: 126 PPG: 3.5

2021/22 Points: 84 PPG: 2.2

2020/21 Points: 113 PPG: 3.1

2019/20 Points: 123 PPG: 3.2

Player: Conor Gallagher

Position: Mid

Points: 129 (46)

Cost: 6.0 (71)

PPG: 3.5 (80)

PPGPE: 0.583 (128)

EP this: None (None%) (172)

EP next: 1.2 (nan%) (377)

2023/24 Points: 129 PPG: 3.5

2022/23 Points: 66 PPG: 1.9

2021/22 Points: 140 PPG: 4.1

2020/21 Points: 71 PPG: 2.4

Player: David Raya Martin

Position: GK

Points: 135 (36)

Cost: 5.5 (150)

PPG: 4.2 (38)

PPGPE: 0.764 (12)

EP this: None (None%) (15)

EP next: 5.0 (nan%) (4)

2023/24 Points: 135 PPG: 4.2

2022/23 Points: 166 PPG: 4.4

2021/22 Points: 95 PPG: 4.0

Player: Rodrigo 'Rodri' Hernandez

Position: Mid

Points: 159 (19)

Cost: 6.5 (65)

PPG: 4.7 (22)

PPGPE: 0.723 (26)

EP this: None (None%) (373)

EP next: 3.3 (nan%) (46)

2023/24 Points: 159 PPG: 4.7

Player: William Saliba

Position: Def

Points: 164 (18)

Cost: 6.0 (73)

PPG: 4.3 (32)

PPGPE: 0.717 (28)

EP this: None (None%) (18)

EP next: 4.2 (nan%) (15)

2023/24 Points: 164 PPG: 4.3

2022/23 Points: 117 PPG: 4.3

2021/22 Points: 0 PPG: 0.0

2020/21 Points: 0 PPG: 0.0

Player: Phil Foden

Position: Mid

Points: 230 (2)

Cost: 9.5 (7)

PPG: 6.6 (3)

PPGPE: 0.695 (42)

EP this: None (None%) (361)

EP next: 4.2 (nan%) (14)

2023/24 Points: 230 PPG: 6.6

2022/23 Points: 142 PPG: 4.4

2021/22 Points: 137 PPG: 4.9

2020/21 Points: 135 PPG: 4.8

2019/20 Points: 69 PPG: 3.0

Player: Ollie Watkins

Position: Fwd

Points: 228 (3)

Cost: 9.0 (8)

PPG: 6.2 (6)

PPGPE: 0.689 (50)

EP this: None (None%) (58)

EP next: 2.7 (nan%) (91)

2023/24 Points: 228 PPG: 6.2

2022/23 Points: 175 PPG: 4.7

2021/22 Points: 131 PPG: 3.7

2020/21 Points: 168 PPG: 4.5