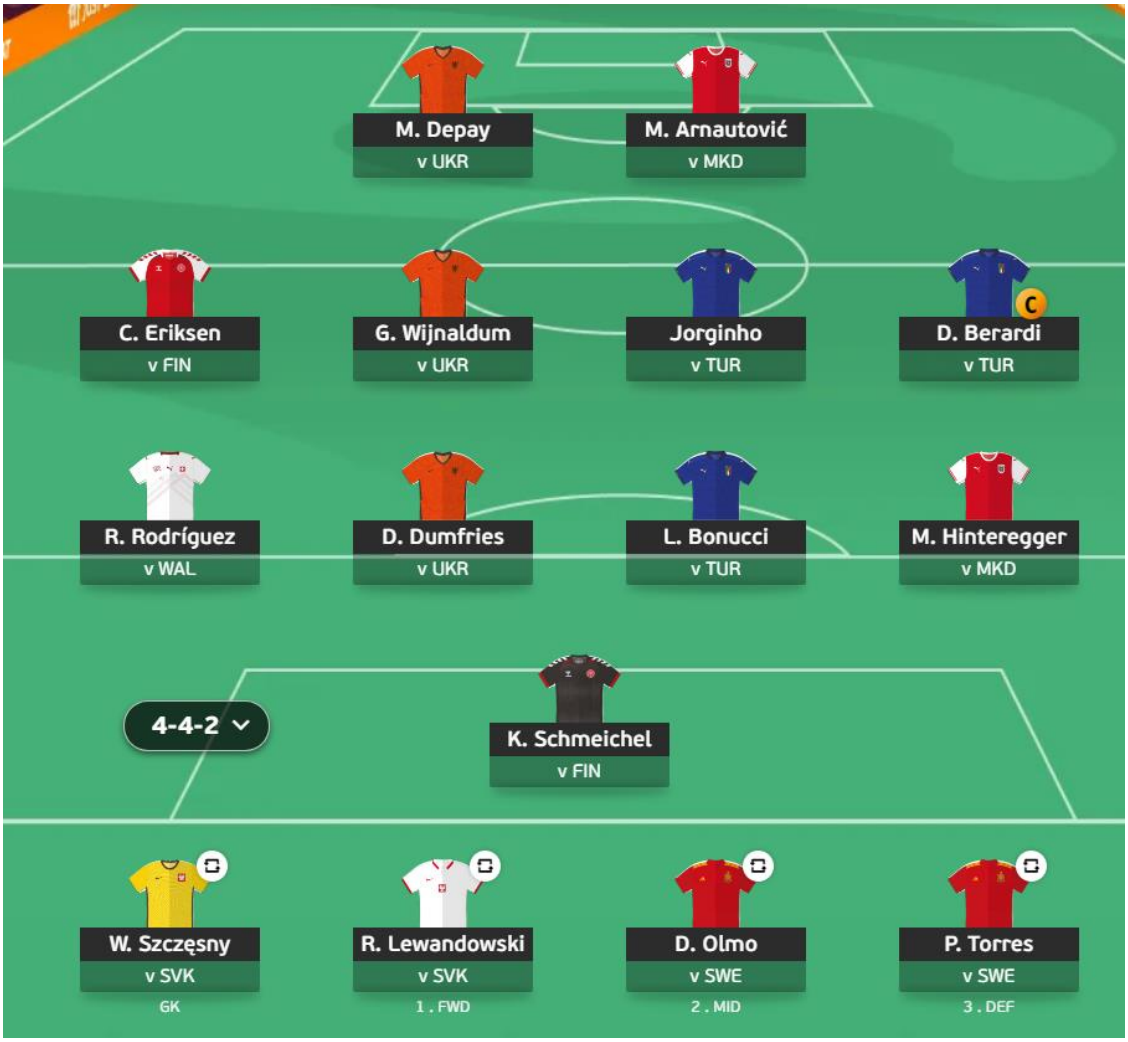
# EURO 2020 FANTASY DRAFT USING DATA MODEL

Euro 2020 begins soon and due to the shorter time duration of the tournament compared to FPL, it is even more important to be involved throughout the process. Manual substitutions and captaincy changes make for excrutiating tales of changing the captain to a player who gets dismissed and the ecstasy of captaining a player on the last day of an MD (Match Day) for a megahaul in equal measure!

Due to lack of reliable resources such as FPL Review, and a lack of expected data for international events such as World Cup 22 Qualifiers, Euro 20 Qualifiers, Nations League, etc. from the normally reliable FBREF, I decided to build my ow data model for the Euro Fantasy game. There are tremendous individuals in the FPL community who know their stuff when it comes to statistical analysis, so if someone wants to build their own data model, I would highly recommend Corridor of Uncertainty FPL podcast as the starting point. My current draft is the result of running optimisation algorithm on points predicted by my data model.

Expected value (EV) = 71.02 with a budget of 100m, assuming that Lewandowski will be eventually captained with a bench of Szczesny, Dumfries, Pau Torres and Jorginho.



While it is impossible to look underneath the hood of a data model in the length of an article, it is possible to explain the tenets. The data model used for computing expected points (xPts) comprises primarily of the following:

**1. Expected Minutes (xMin.)/Squad information:** xMin is king when it comes to fantasy football. Based on predicted lineups by reputed media sources, with weightage assigned to each source, each player's expected minutes are determined. Apart from this, set-piece taker probability was also determined which increases the expected points significantly for penalty takers and also boosts the expected value for freekick and corner specialists.

**2. Attacking data:** Underlying stats for the past 3 seasons in the domestic league (with each season being wieghted according to their recency) formed a part of the metric. Along with that, data from the past 7 international events, i.e., Euro 2016, World Cup 2018, Nations League 2018-19, Euro 2020 Qualifiers, Nations League 2020-21, World Cup 2022 Qualifiers and 2021 International Friendlies have also been used with different weightage assigned according to their importance and recency. Out of all the international events, expected data was available only for World Cup 2018.

**3. Offensive data adjustment:** Robert Lewandowski is arguably the best player on the planet in a Bayern Munich shirt, but how does that translate to the national team ? Using FiveThirtyEight's offensive ratings and a number of attacking and defensive stats for all countries since Euro 2016, each player's underlying club stats were adjusted to fit his national squad.

**4. Odds/Probability**: In order to understand the likelihood of a cleansheet, defensive ratings from FiveThirtyEight were combined with the Cleansheet odds available on various reputable websites. This was used to determine the expected value of a cleansheet for each player. Similarly, anytime goalscorer and assist odds were used to supplement the data available for each player in order to predict attacking returns.

**5. Home/Away factor:** Looking at the expected goals data of the Big 5 leagues (Premier League, Serie A, Ligue 1, La Liga and Bundesliga) from 2017-19 (precovid seasons), there is a distinct home advantage which is quantified by the ratio of xG accumulated by the home team to the xG accumulated by the away side. This has been used to adjust the expected returns for players playing in front of a home crowd, or on the road with a gruelling travel schedule.

Due to the KO nature of the tournament, it is essential to use the Limitless chip in the group stages coupled with the fact that the budget increases to 105m for the RO16. Even though managers may adopt different chip strategies, there will be a core of players that will be common across them. In order to take advantage of MD1 fixtures, it is important to have captaincy options for each day within the match day, as can be seen in my current draft. The beauty with having 3 Italian players is that the lineups will be announced before the deadline, so in case Berardi is benched, I can move him to Chiesa (7.0m) or Carrasco (7.0m) and downgrade Szczesny to Hradecky (4.0m). I can also swap Bonucci (5.5) for Spinazzola (5.5m) if the latter starts. After an Italian captain on the first day,

the armband will move to Eriksen on the second day, Depay on the third and Lewandowski on the fourth. The most obvious thing in this draft is a lack of player that plays on the fifth day of MD1, although by that point one of your four captains is bound to return! In case you need the additional security of a Group F player, you can swap Szczesny (5m), Dumfries (5.5m) and Olmo (7m) for a slightly suboptimal trio of Bachmann (4.5m), Wijndal (4.5m) and Jota (8.5m) which would result in an EV loss of around 0.13 only.

Conspicuously absent from this draft are Romelu Lukaku (11.0m) and Cristiano Ronaldo (12.0m). While both are excellent options and might end up outscoring other forwards due to large variance, they are not the optimal choice as suggested by the model. To caveat this, the current optimisation algorithm asks me to tranfer Lukaku in for MD3, in place of Lewandowski.

According to the model projections, the top players for MD1 are*:

| Player | Day # of MD1 | Country | Price | Pos | MD1 xPts |
|---|---|---|---|---|---|
| **Robert Lewandowski** | 4 | Poland | € 11.5 | FWD | 8.4270 |
| **Memphis Depay** | 3 | Netherlands | € 10.0 | FWD | 7.4027 |
| **Cristiano Ronaldo** | 5 | Portugal | € 12.0 | FWD | 6.3911 |
| **Christian Eriksen** | 2 | Denmark | € 9.0 | MID | 6.2739 |
| **Romelu Lukaku** | 2 | Belgium | € 11.0 | FWD | 6.0582 |
| **Harry Kane** | 3 | England | € 11.5 | FWD | 5.7716 |
| **Georginio Wijnaldum** | 3 | Netherlands | € 8.0 | MID | 5.7056 |
| **Domenico Berardi** | 1 | Italy | € 6.0 | MID | 5.5918 |
| **Lorenzo Insigne** | 1 | Italy | € 8.5 | MID | 5.5041 |
| **Diogo Jota** | 5 | Portugal | € 8.5 | MID | 5.2045 |

*Assuming the player starts

These players cover all 5 days within MD1, which should suit any strategy chosen by a manager. Also, if your expectation of home team advantage is higher than the assumed value, there are 6 players out of those 10 that play in front of a home crowd (Netherlands, Denmark, England and Italy).


@Schadenfreud1st

(FPL Schadenfreudist)

Masters' student with a penchant for numbers