Capstone Project

Analysing Fantasy Premier League Football

// FLATIRON SCHOOL

Executive summary

- An introduction to FPL
- CB91's findings:
  - Customer quality by region
  - Effects of discounting
  - Increasing order frequency
  - Cost of shipping firms
- Possible next steps

Appendix contains additional analysis

An introduction to Fantasy Premier League

- Given a virtual budget of £100m, each 'manager' has to assemble a team of players, who score FPL points based on their performance in real life games
- Famous players who scored highly in previous seasons are more expensive
- Therefore, good FPL managers are those able to spot cheap, unknown players with the potential to score big

What can data science do for FPL?

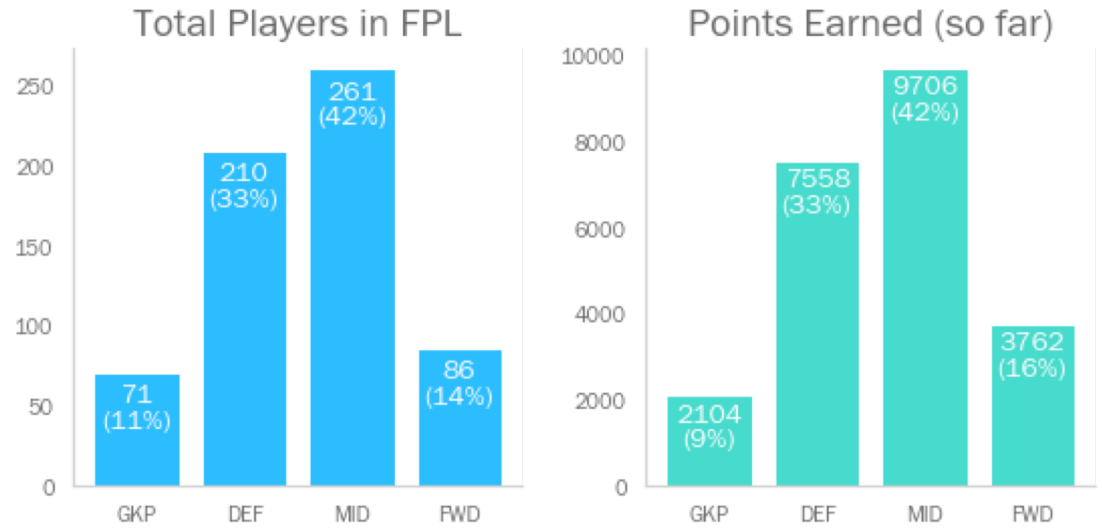
- Every football match generates a lot of data – both in terms of player performance and FPL manager decisions
- Managers can swap one player from their team every week to bring in a replacement (who may be in better form, or have an easier run of fixtures)
- The aim is to create a suite of tools that help managers make better picks throughout the season, based on the data

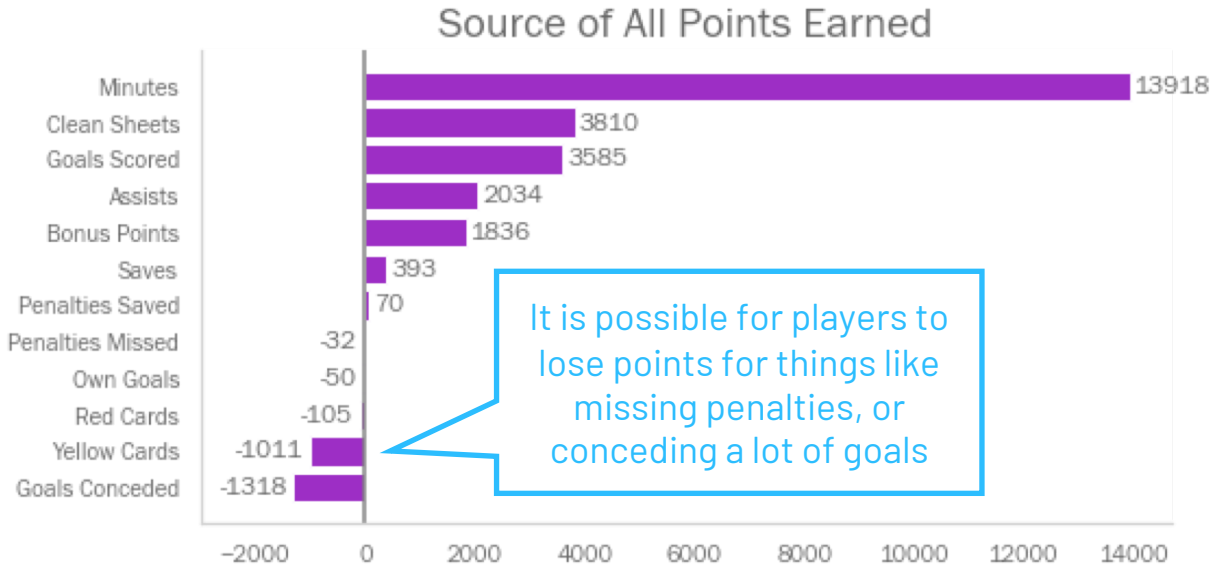
Data Sources

- Text commentary for every match scraped from the Premier League website
- 'Fantasy' data (e.g. FPL points, player transfers, etc.) from official API
- Combined, this gives a detailed view of every shot taken in every match, and an aggregate view of the decisions of each FPL manager

Where do FPL points come from?

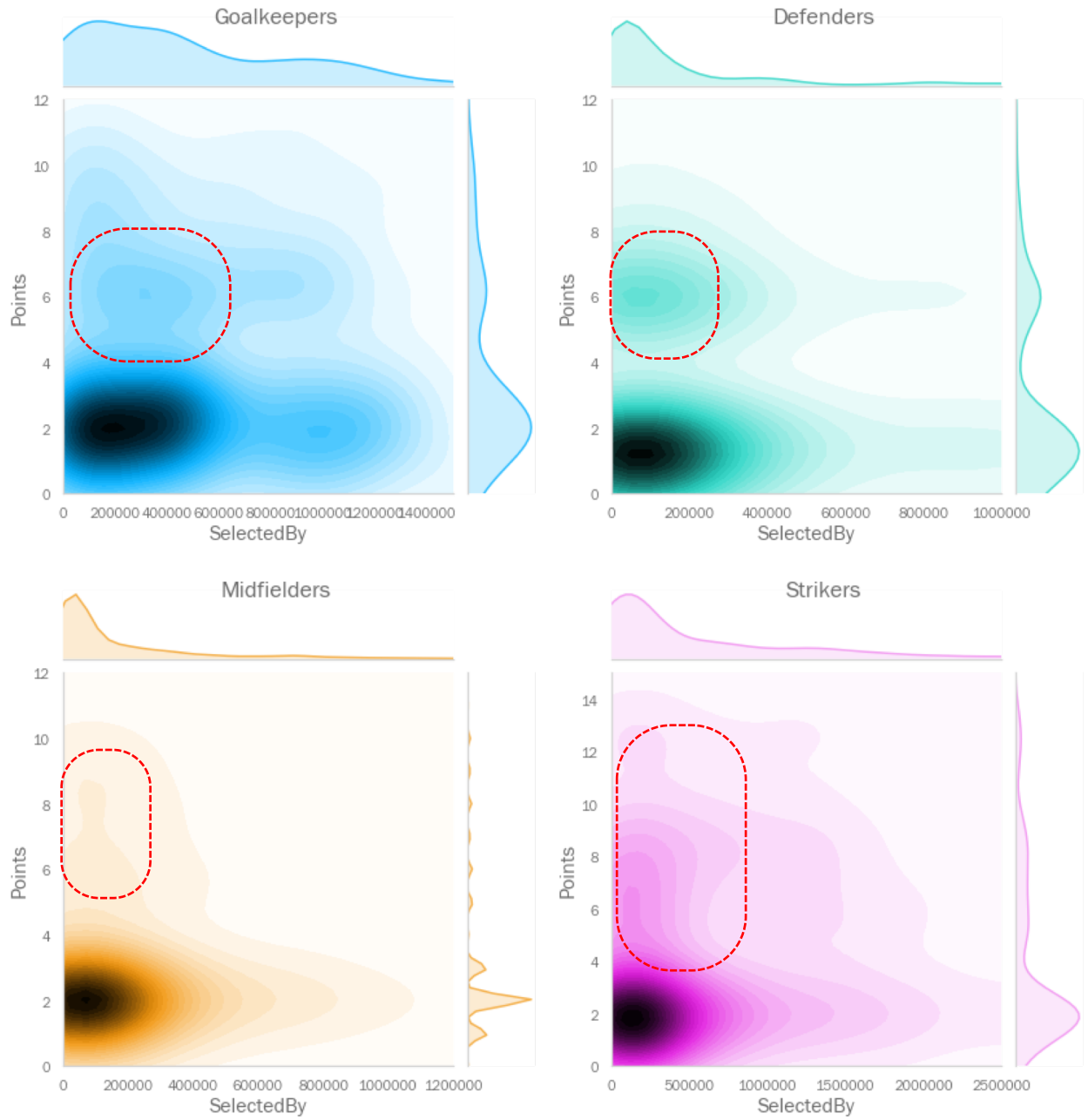
- The share of points earned by position reflects the number of players of each position
- More than half of FPL points are generated by 'Minutes Played', i.e. players simply being selected to play
- Keeping clean sheets and scoring goals are the most common ways to generate extra points - thus FPL managers need to be mindful of both attack and defence





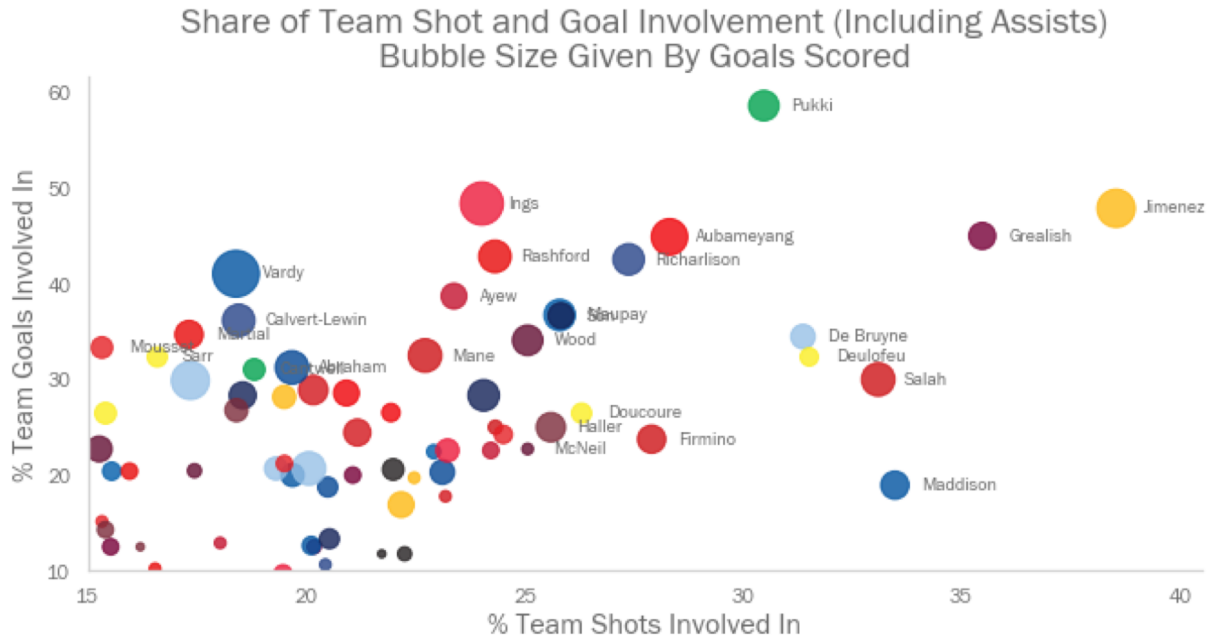
Picking 'differentials'

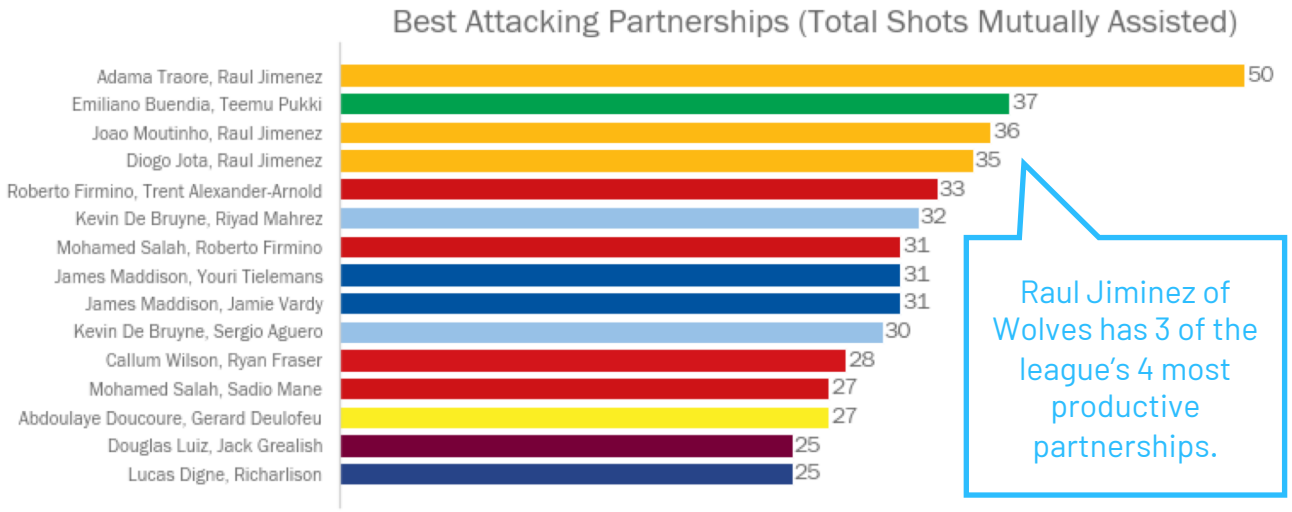
- Differential players are those owned by a small number of FPL managers
- Consistently picking 'differentials' is key to FPL success – if you own a high-scoring player, this isn't overly advantageous if the same player is owned by 50% of managers
- Data suggests that there are such 'differential' players in the game (though they are in the minority)



Player importance

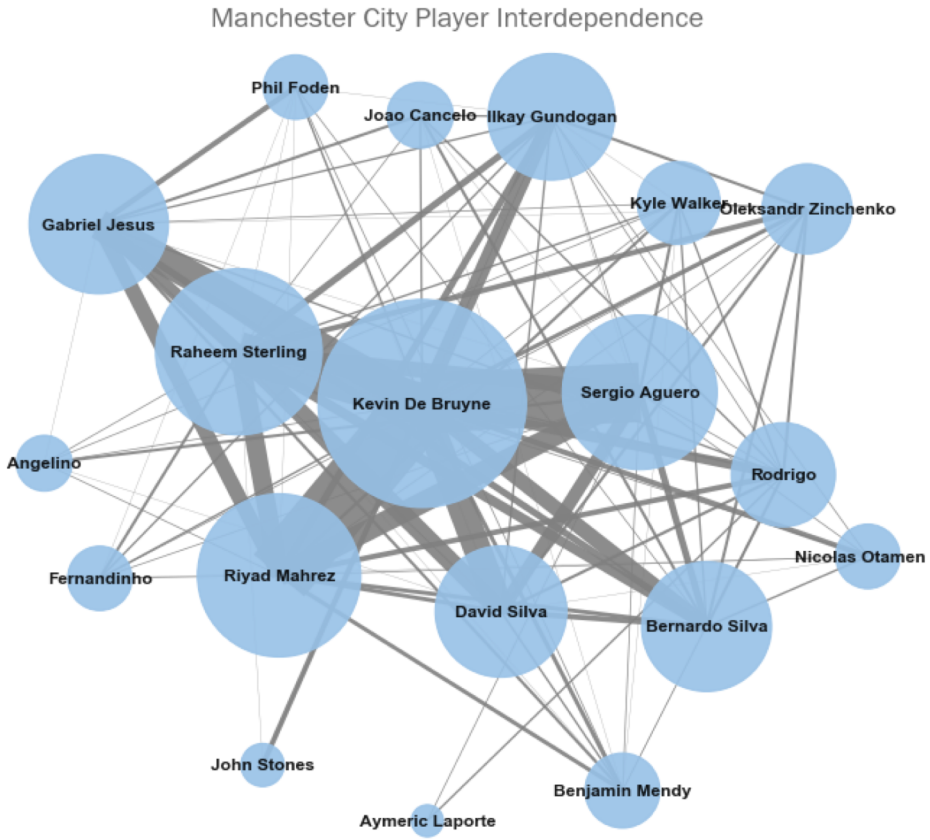
- Picking 'important' players of less well known teams could be a good way to spot differentials
- We can start to explore which players appear to be 'intrinsic' to their teams' performances
- How can we analyse this more rigorously?

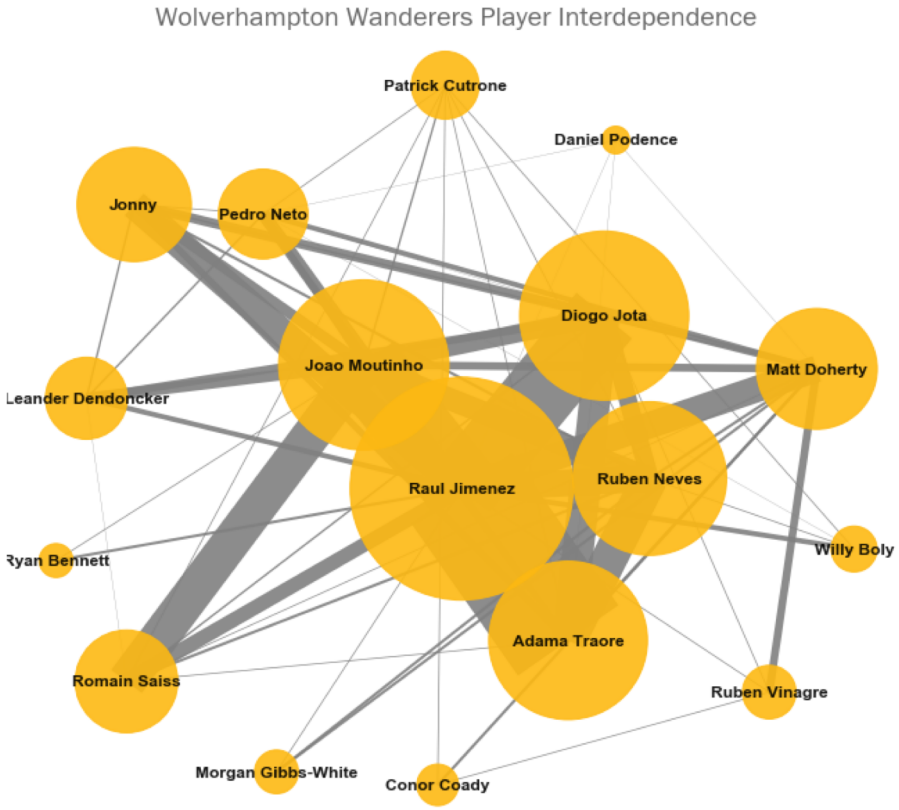




Team networks

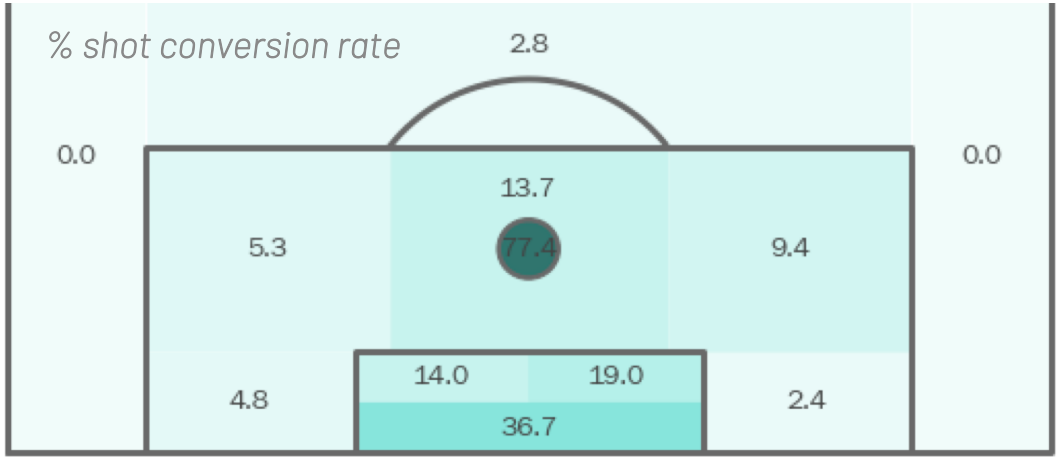
- We can think about teams like they were social networks. Then the strength of connection between two players can be given by number of times they've assisted each other's shots
- Then we can use 'betweenness' metrics to work out who the most 'enmeshed' (and thus, 'important') players are on a team by team basis

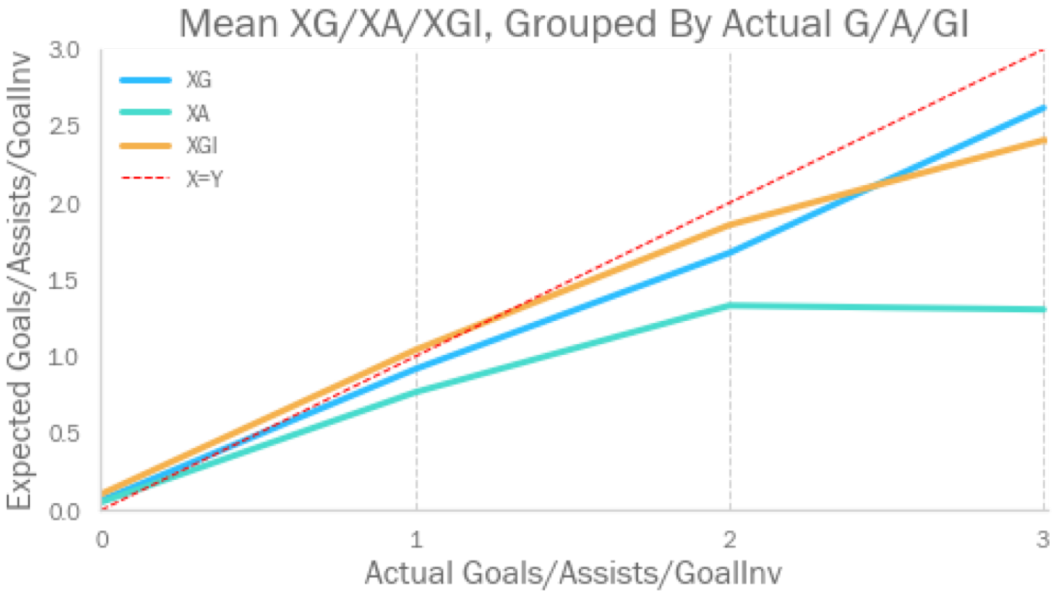




Shot quality and 'expected goals'

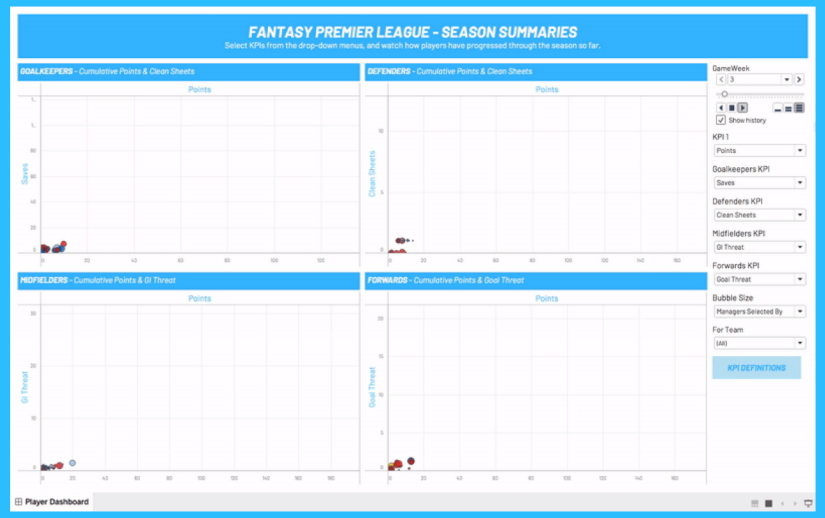
- Can we remove some of the sport's intrinsic 'randomness' to help make our analysis more reliable?
- Having defined the different types of possible shots we can see the goal conversion rate of each
- We can then sum the expected conversion rates for shots that a player took in a match to give a grand 'Expected Goals' (XG)





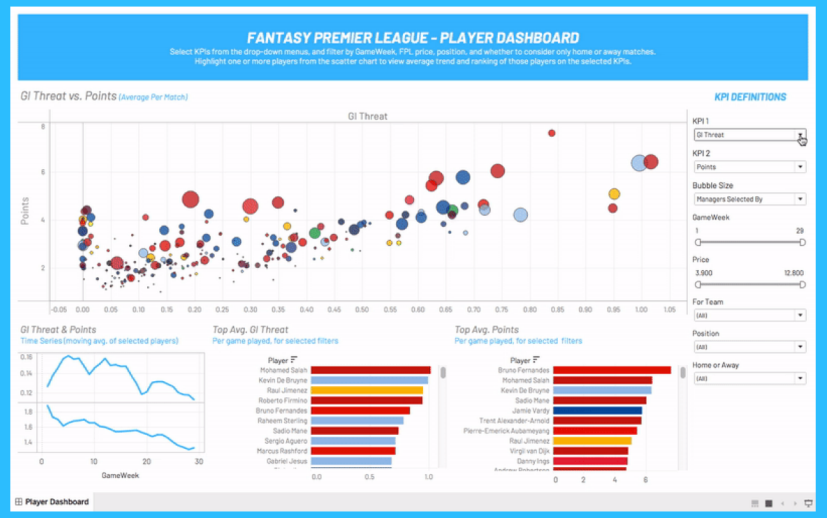
A suite of tools for FPL managers...

1) SEASON SUMMARY



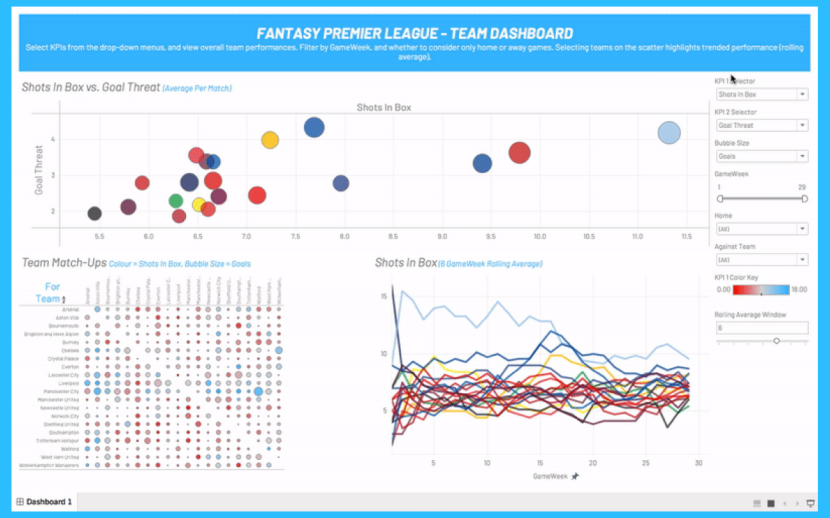
- Shows how players have progressed through the season so far (by position)
- Can be used to identify trends before they become obvious to every manager
- KPIs fully customisable

2) PLAYER DASHBOARD



- A fully customisable dashboard showing KPIs for each player
- Can filter performance based on player price, gameweek, and home/away matches
- View trends/rankings of players selected in the scatter plot

3) TEAM DASHBOARD



- A dashboard showing team-level performance across KPIs
- All views and KPIs are fully customisable and can be filtered as required
- Gives a view of how teams have played against each other

Possible next steps

There is further work that could be done using the gathered data:

- Create further dashboards (e.g. using the shot level data)
- Spend further time working on feature engineering to create predictive models
- Get more data from previous seasons to strengthen the dataset

Thank you for your time

Further detail on the model, and additional analysis of the dataset is included in further pages