FinTech – Project 1

Soccer "Money" Ball

Meet the project team

- Adam Freeman

- Marcus Whitelock

- Mohamed Dallol

- Saroj Buragadda

- Tracey Martin

The WHY

Defining why are we doing this is a key step in any project. The why forms the reason and primary motivation for doing this. It is the origin and objective all other tasks stem from.

"Create The Best Fantasy Football Teams on the Globe"

The WHAT

Building on the why, defining the what gives us clarity of scope for this project. What are the deliverables.

"Deliver Notebooks, Reports, Dashboards and Images using programmatic statistical analysis to -

"Create The Best Fantasy Football Teams on the Globe"

The HOW

Now we understand the why and the what, how will we do this.

"Our Project team will utilise Markdown, Python, Pandas, Plotting libraries, Published APIs, Postgresql and CSV files to -

"Deliver Notebooks, Reports, Dashboards and Images using programmatic statistical analysis to -

"Create The Best Fantasy Football Teams on the Globe"

We Need Data – But where is it ?

In this project we found many sources of data and data types. We chose the following data sources.

Kaggle European Soccer Database - SQLDBlite Database

RapidAPI - FOOTBALL API

FutDB - Database API for FIFA 22

ISO.org - ISO.org CSV File

We Need Data – But where is it ?

The selection of these data sources was a deliberate decision by the project team. This intention is to demonstrate our ability using not just one, but three very prevalent data source types that you would expect to come across in the industry.
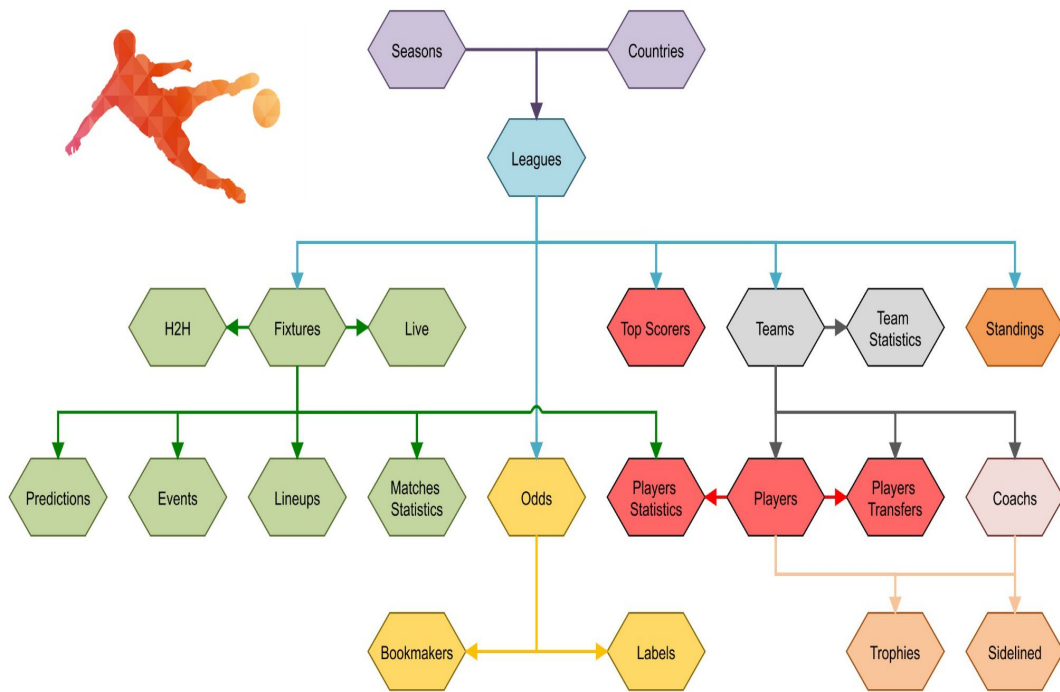
- SQL Databases
- API
- CSV Files

Kaggle – Euro_soccer_db ERD
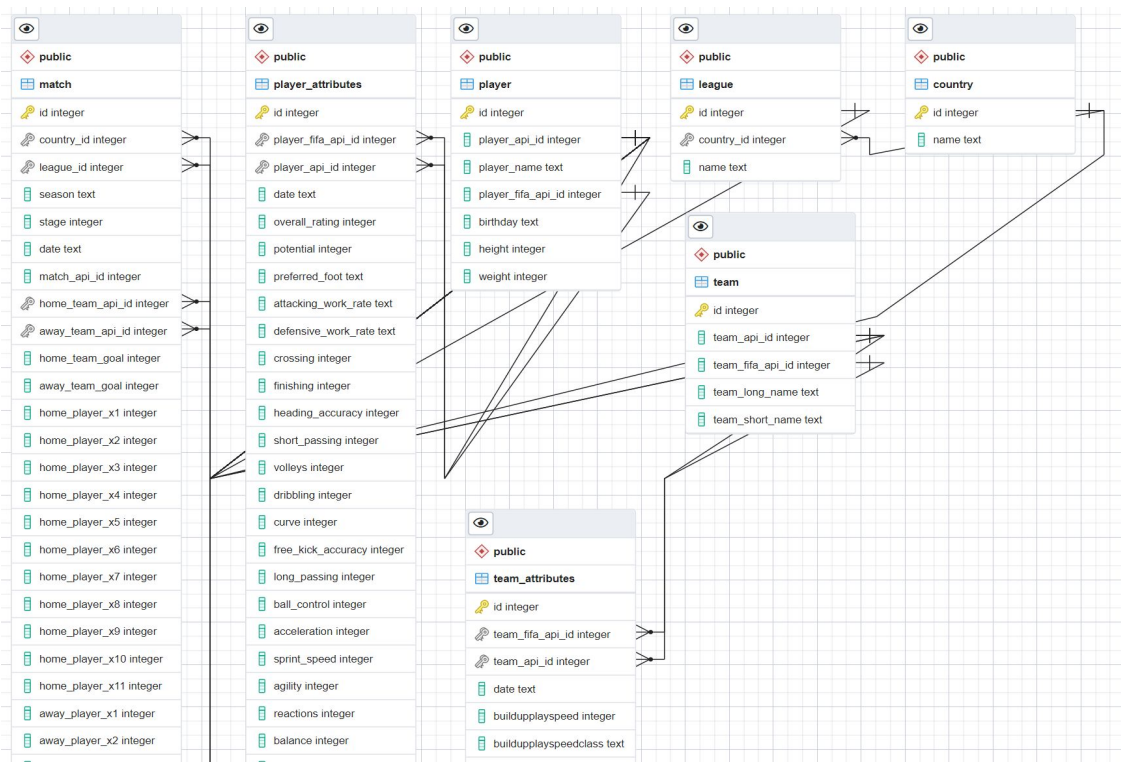
- 7 Tables
- 200,000 rows
- Inspected data
- Clean data
- Created Postgres DB
- Created schema
- Created constraints
- Generate ERD
- Seed data
- Run Queries

RapidApi – Football–API

- Created Pro access paid account
- 19 USD per month
- 7500 requests per day limit
- Overage charge 0.0025 cents per request
- 300 requests per minute rate limit

FutDB API – Database API for FIFA 22

Statistic

Real time statistics of our players that a stored at the moment in our database.

Players	Cards
18.251	40

Nations	Leagues	Clubs
159	49	668

- Created Premium access paid account
- 19.99 Euro for 30 day access
- 200,000 requests per day limit
- Rate Limit Unknown

ISO Org – Officially Assigned Country Codes

English short name	French short name	Alpha-2 code	Alpha-3 code
Afghanistan	Afghanistan (l')	AF	AFG
Albania	Albanie (l')	AL	ALB
Algeria	Algérie (l')	DZ	DZA
American Samoa	Samoa américaines (les)	AS	ASM
Andorra	Andorre (l')	AD	AND
Angola	Angola (l')	AO	AGO
Anguilla	Anguilla	AI	AIA
Antarctica	Antarctique (l')	AQ	ATA
Antigua and Barbuda	Antigua-et-Barbuda	AG	ATG

- Freely Available
- Copied to local csv
- Provided the Alpha-3 code for plotting

We Have Data ! Let's do some preliminary analysis

- What are the biggest Leagues in the World and How many Games do they play
- Are different leagues more attacking or defensive ?
- Are there any factors that unilaterally influence the number of goals scored in a game by a team
- Does the amount of home vs away advantage change per season ?
- Does the amount of home vs away advantage change per league ?

Let's look at the teams dashboard to answer these questions !

Looking good, but did we pick the right sport ?

- Where do these Soccer players come from is it truly a global sport ?
- Which Countries have professional Soccer Leagues and how many ?
- Are there continents where Soccer is more popular ?
- Is there a relation between developed countries and Soccer Leagues ?
- Does this sport transcend a countries economic status

Let's look at the teams dashboard to answer these questions !

Preliminary analysis complete – Let's play Soccer "Money" Ball

From the ground work we completed we now have a better understanding of the data. Through this process we narrowed our scope to the following leagues

- English Premier League
- Italy Serie A
- France Ligue 1
- Spanish LIGA

We know these are the most attacking leagues and play the most games. We also understand that there is a definite home team advantage, and there is a loose trend of more goals being scored each year on year. We have also confirmed that Soccer is truly a global sport played in many countries and represents a good choice of sport, as there will be many fantasy football competitions across the globe we could potentially enter.

Using all of this data along with programmatic statistical analysis by the team, we have assembled the following fantasy football teams for each of the 4 Top Leagues

Fantasy Team Plot Per League – England and France

Premier League Best Fantasy Team

GK	DEF	MID	ST

Virgil van Dijk
Liverpool
Rating: 89

Kevin De Bruyne
Manchester City
Rating: 91

Trent Alexander-Arnold
Liverpool
Rating: 87

Sadio Mane
Liverpool
Rating: 89

Harry Kane
Tottenham
Rating: 90

Alisson
Liverpool
Rating: 89

Ruben Dias
Manchester City
Rating: 87

Mohamed Salah
Liverpool
Rating: 89

Cristiano Ronaldo
Manchester United
Rating: 91

Andrew Robertson
Liverpool
Rating: 87

Heung Min Son
Tottenham
Rating: 89

Ligue 1 Best Fantasy Team

| GK | DEF | MID | ST |

Sergio Ramos Garcia
PSG
Rating: 88

Lionel Messi
PSG
Rating: 93

Marquinhos
PSG
Rating: 87

Marco Verratti
PSG
Rating: 87

Kylian Mbappe
PSG
Rating: 86

Gianluigi Donnarumma
PSG
Rating: 89

Achraf Hakimi
PSG
Rating: 85

Angel Di Maria
PSG
Rating: 87

Neymar Jr
PSG
Rating: 93

Presnel Kimpembe
PSG
Rating: 83

Georginio Wijnaldum
PSG
Rating: 84

Fantasy Team Plot Per League – Italy and Spain

English Premier League – Player Statistics

Player1

Player2

Player3

Player4

Player5

Player6

Player7

Player8

Player9

Player10

Player11

France Ligue 1 – Player Statistics

Player1

Player2

Player3

Player4

Player5

Player6

Player7

Player8

Player9

Player10

Player11

Italy Seria A – Player Statistics

Spain La Liga – Player Statistics

Player1

Player2

Player3

Player4

Player5

Player6

Player7

Player8

Player9

Player10

Player11

Challenges and Lessons Learnt

In any project its important to reflect and take some time to determine what important lessons can be taken away and learnt from. We also took time to reflect on what were some of the biggest challenges we faced. Here is our top 7.

- Persistent data (Databases) is king for large data analysis !
- API are great for targeted data
- You often need to pay for good data
- Be aware of API throttling and overage charges (we got banned for 24 hours)
- Test for loop conditions with simple code first
- JSON Normalisation is key with API data and Pandas
- Things don't ever go to plan !!

What did we do well

Also important is to reflect on what we did well, here is a list of our top 7 achievements

- Produced 5 notebooks - 5500+ lines of code written
- Used 1 DB , 2 API , 1 CSV, 2 new libs, new plotting library
- 30+ Git Pushes
- For loop to gather data from API into persistent format
- Discerning with what we used (only 75% of the code we wrote made it to final dashboard)
- Collaborated well and even across time zones (once we did 44hr straight with 2 hour break)
- Completed all project requirements and more, on time and in budget :)

What features would we add ?

If given the chance to create soccer "money" ball ver 2, we would focus on the following enhancements

- Use Dash or PowerBI instead of panel for Dashboard,
- We would add most improved players tab
- We would add player cost analytics, this would allow
  - Player cost per point of rating
  - Std deviation of rating over players career
  - Correlation of player cost, vs experience, vs trophies won, vs rating
- Betting Algorithm based on results
  - Results for $100 bet on all homes teams
  - Results for $100 bet on all favourites
  - Results for $100 bet on for best on ground for each of our fantasy team players

FinTech – Project 1

The End

Thank you from Project Team 4 for your participation

Link to Project Files here