
soccerdata

Release 1.8.8

Pieter Robberechts

Feb 01, 2026

CONTENTS

1 Getting Started	3
2 Overview of Data Sources	7
3 How-to Guides	99
4 Data Analysis Examples	103
5 API Reference	111
6 FAQ	133
7 Contributor Guide	135
8 Apache License	139
Index	141

Release v1.8.8. (pip install soccerdata)

SoccerData is a collection of scrapers to gather soccer data from popular websites, including Club Elo, ESPN, FBref, FiveThirtyEight, Football-Data.co.uk, Sofascore, SoFIFA, Understat and WhoScored.

```
import soccerdata as sd

# Create a scraper class instance for the 2020/21 Premier League
fbref = sd.FBref('ENG-Premier League', '2021')

# Fetch data
games = fbref.read_schedule()
team_season_stats = fbref.read_team_season_stats(stat_type="passing")
player_season_stats = fbref.read_player_season_stats(stat_type="standard")
```

Main features

- Access current and historical soccer fixtures, forecasts, detailed match stats, event stream data and more.
- All data is provided in the form of Pandas DataFrames with sensible, matching column names and identifiers across datasets to make working with the data and combining data from multiple sources easy.
- Data is only downloaded when needed and cached locally to speed up your analysis scripts.
- Integrates with the `socceraction` package to allow analysis of event stream data.

Do you like it? *Let's dive in!*

CHAPTER
ONE

GETTING STARTED

New to *soccerdata*? Well, you came to the right place: this tutorial will walk you through installing, configuring, and using the library. By the end of this tutorial, you will be able to scrape data from the top-5 European leagues and use it to create your own data-driven analyses.

1.1 Installation

SoccerData can be easily installed via pip:

```
python3 -m pip install soccerdata
```

1.2 Scraping data

Each of the *supported data sources* has its corresponding class for fetching data with a uniform API. For example, the *FBref* class is used to fetch data from [fbref.com](#).

```
import soccerdata as sd

# Create scraper class instance
fbref = sd.FBref()
```

Once you have a scraper class instance, you can use it to fetch data. See the the *examples* and *API reference* for the full list of options available for each scraper. For example, to fetch aggregated shooting stats for all teams:

```
# Create dataframes
season_stats = fbref.read_team_season_stats(stat_type='shooting')
```

The data is always returned as a convenient Pandas DataFrame.

league	season	team	#Pl	90s	Gls	Sh	SoT	Sh/SoT	Sh/G/S	G/S	Dist	FK	PK	PKs	xG	npxG	npxG	G-xG	npG-xG	
ENG	202	Ar-Pre Lea	29	38.0	53	455	141	31.0	11.9	3.71	0.1	0.32	16.9	23	6	6	53.4	49.0	0.11	- -
		se- nal																0.5	2.0	
		As-ton Vill	24	38.0	52	518	179	34.6	13.6	4.71	0.09	0.26	16.5	15	5	6	52.9	48.5	0.1	- -
																		0.9	1.5	
		Bri-	27	38.0	39	476	129	27.1	12.5	3.39	0.07	0.26	16.6	14	6	9	51.6	44.8	0.1	- -
																		12.6	11.8	
		Bur-ley	25	38.0	32	383	125	32.6	10.0	3.29	0.08	0.23	16.6	15	3	3	39.5	37.6	0.1	- -
																		7.9	8.6	
		Che	27	38.0	56	553	194	35.1	14.5	5.11	0.09	0.25	16.3	16	8	10	64.0	56.4	0.1	- -
																		8.0	8.4	

By default, the data for all available leagues and the five most recent seasons will be retrieved. However, in most cases, you would want to limit the data to specific leagues and / or seasons. This can be done by passing a list of leagues and seasons to the constructor of the scraper class. For example:

```
# Create scraper class instance filtering on specific leagues and seasons
fbref = sd.FBref(leagues=['ENG-Premier League'], seasons=['1718', '1819'])
# Retrieve data for the specified leagues and seasons
season_stats = fbref.read_team_season_stats(stat_type='shooting')
```

Note that only a limited number of leagues are supported out-of-the-box. The leagues available for each source can be listed with the `available_leagues()` class method.

```
sd.FBref.available_leagues()
>>> ['Big 5 European Leagues Combined', 'ENG-Premier League', 'ESP-La Liga', 'FRA-Ligue 1 ↴', 'GER-Bundesliga', 'ITA-Serie A', 'INT-World Cup', "INT-Women's World Cup"]
```

You can [add more leagues](#) but there are no guarantees that they will be scraped correctly.

1.3 Data caching

Data caching is used to speed up the runtime and to prevent exceeding the rate limit of web servers. By default, all downloaded data is cached to `~/soccerdata` on Linux and Mac OS, and to `C:\Users\yourusername\soccerdata` on Windows. A custom location can be set if desired. You can configure this using environment variables (see below) or on the level of an individual scraper by setting the `data_dir` parameter when creating the scraper class instance:

```
# Create scraper class instance with custom caching directory
fbref = sd.FBref(data_dir="/tmp/FBref")
```

This directory can be deleted at any time to reclaim disk space. However, this also means you will have to redownload the same data again if you need it, which will lead to reduced performance.

SoccerData has no knowledge of when the data on the server changes, so it is up to the user to decide when to refresh the cache. This can be done by deleting the cache directory or by setting the `no_cache` option to `True` when creating the scraper class instance:

```
# Create scraper class instance which always re-downloads the latest data
fbref = sd.FBref(no_cache=True)
```

Some methods will assume the cache is always out-of-date (for example, when scraping the fixture of the current season). Typically, these methods will have a `force_cache` option that can be set to True to force the cached data to be used. For example:

```
fbref = sd.FBref(leagues=['ENG-Premier League'], seasons=['2324'])
fbref.read_schedule(force_cache=True)
```

Caching can also be disabled entirely by setting the `no_store` option to True when creating the scraper class instance. However, it should almost always be left enabled.

```
# Create scraper class instance with caching disabled
fbref = sd.FBref(no_store=True)
```

1.4 Global configuration

Several settings can be configured globally using the following environment variables:

SOCCKERDATA_DIR

The directory where the downloaded data is cached and where logs are stored. By default, all data is stored to `~/soccerdata` on Linux / Mac OS and `C:\Users\yourusername\soccerdata` on Windows.

SOCCKERDATA_NOCACHE

If set to “true”, no cached data is returned. Note that no-cache does not mean “don’t cache”. All downloaded data is still cached and overwrites existing caches. If the sense of “don’t cache” that you want is actually “don’t store”, then `SOCCKERDATA_NOSTORE` is the option to use. By default, data is retrieved from the cache.

SOCCKERDATA_NOSTORE

If set to “true”, no data is stored. By default, data is cached.

SOCCKERDATA_MAXAGE

The maximum age of cached data in seconds. If the cached data is older than this, it will be re-downloaded. By default, this is set to infinity.

SOCCKERDATA_LOGLEVEL

The level of logging to use. By default, this is set to “INFO”.

Example:

```
# bash
export SOCCKERDATA_DIR = "~/soccerdata"
export SOCCKERDATA_NOCACHE = "False"
export SOCCKERDATA_NOSTORE = "False"
export SOCCKERDATA_LOGLEVEL = "INFO"
```

1.5 Uniform team names

Each data source uses a different set of team names, which makes it difficult to combine data from multiple sources. To mitigate this, SoccerData allows translating the team names to uniform names. This is done by providing a `SOCCKERDATA_DIR/config/teamname_replacements.json` file. This file should contain a mapping between a generic name for each team and the team name used by each data source that you want to support. The example below will map “Tottenham Hotspur”, “Tottenham Hotspur FC” and “Spurs” to “Tottenham” in all scraped data.

```
{  
    "Tottenham": ["Tottenham Hotspur", "Tottenham Hotspur FC", "Spurs"],  
}
```

1.6 Additional setup for scraping WhoScored data

WhoScored implements strong protection against scraping using Incapsula. To circumvent this, this scraper uses Selenium with the ChromeDriver extension to emulate a real user. Before using this scraper, you will have to [install Chrome](#). A Selenium driver matching your Chrome version will be downloaded automatically when you run the scraper.

1.7 Next steps

Look at you! You’re now basically an expert at SoccerData!

From this point you can:

- Look at the example notebooks for each *Data source*.
- Take a deep dive into the [API](#).
- Give us feedback or contribute, see [Contributing](#).

Have fun!

OVERVIEW OF DATA SOURCES

Currently, the following data sources are supported.

ClubElo

[URL](#) | [Example usage](#) | [API reference](#)

```
from soccerdata import ClubElo
```

Team's relative strengths as Elo ratings, for most European leagues. Recalculated after every round, includes history.

ESPN

[URL](#) | [Example usage](#) | [API reference](#)

```
from soccerdata import ESPN
```

Historical results, statistics and lineups.

FBref

[URL](#) | [Example usage](#) | [API reference](#)

```
from soccerdata import FBref
```

Historical results, lineups, and detailed aggregated statistics for teams and individual players based on Opta data.

Football-Data.co.uk

[URL](#) | [Example usage](#) | [API reference](#)

```
from soccerdata import MatchHistory
```

Historical results, betting odds and match statistics. Level of detail depends on league.

Sofascore

[URL](#) | [Example usage](#) | [API reference](#)

```
from soccerdata import Sofascore
```

Results, schedules, lineups, and detailed statistics for teams and individual players.

SoFIFA

[URL](#) | [Example usage](#) | [API reference](#)

```
from soccerdata import SoFIFA
```

Detailed scores on all player's abilities from EA Sports FC.

Understat

[URL](#) | [Example usage](#) | [API reference](#)

```
from soccerdata import Understat
```

Advanced statistics such as xG, xGBuildup and xGChain, and shot events with associated xG values for the top European leagues.

WhoScored

[URL](#) | [Example usage](#) | [API reference](#)

```
from soccerdata import WhoScored
```

Historical results, match preview data and detailed Opta event stream data for major leagues.

The following section was generated from doc/datasources/ClubElo.ipynb

```
[2]: import soccerdata as sd
```

2.1 ClubElo

```
[3]: elo = sd.ClubElo()
print(elo.__doc__)
```

Provides pd.DataFrames from CSV API at <http://api.clubelo.com>.

Data will be downloaded as necessary and cached locally in
`~/.soccerdata/data/ClubElo`.

Since the source does not provide league names, this class will not filter by league. League names will be inserted from the other sources where available. Leagues that are only covered by clubelo.com will have NaN values.

Parameters

(continues on next page)

(continued from previous page)

```

proxy : 'tor' or or dict or list(dict) or callable, optional
    Use a proxy to hide your IP address. Valid options are:
        - "tor": Uses the Tor network. Tor should be running in
            the background on port 9050.
        - dict: A dictionary with the proxy to use. The dict should be
            a mapping of supported protocols to proxy addresses. For example::

{
    'http': 'http://10.10.1.10:3128',
    'https': 'http://10.10.1.10:1080',
}

        - list(dict): A list of proxies to choose from. A different proxy will
            be selected from this list after failed requests, allowing rotating
            proxies.
        - callable: A function that returns a valid proxy. This function will
            be called after failed requests, allowing rotating proxies.

no_cache : bool
    If True, will not use cached data.
no_store : bool
    If True, will not store downloaded data.
data_dir : Path
    Path to directory where data will be cached.

```

2.1.1 ELO scores for all teams at specified date

[4]: current_elo = elo.read_by_date()
current_elo.head()

	rank	country	level	elo	from	to	\
team							
Liverpool	1.0	ENG	1	2047.083862	2022-04-20	2022-04-24	
Man City	2.0	ENG	1	2037.059937	2022-04-21	2022-04-23	
Bayern	3.0	GER	1	1984.775391	2022-04-18	2022-04-23	
Real Madrid	4.0	ESP	1	1969.584351	2022-04-21	2022-04-26	
Chelsea	5.0	ENG	1	1921.101440	2022-04-21	2022-04-24	
				league			
team							
Liverpool		ENG-Premier League					
Man City		ENG-Premier League					
Bayern		GER-Bundesliga					
Real Madrid		ESP-La Liga					
Chelsea		ENG-Premier League					

2.1.2 Full ELO history for one club

```
[5]: barca_elo = elo.read_team_history("Barcelona")
barca_elo.head()
```

```
[5]:      rank      team country level      elo      to
from
1939-10-22   NaN  Barcelona    ESP      1  1636.704590 1939-12-03
1939-12-04   NaN  Barcelona    ESP      1  1626.102173 1939-12-10
1939-12-11   NaN  Barcelona    ESP      1  1636.728271 1939-12-17
1939-12-18   NaN  Barcelona    ESP      1  1646.951660 1939-12-24
1939-12-25   NaN  Barcelona    ESP      1  1637.424316 1939-12-31
```

..... doc/datasources/ClubElo.ipynb ends here.

The following section was generated from doc/datasources/ESPN.ipynb

```
[2]: import soccerdata as sd
```

2.2 ESPN

```
[3]: espn = sd.ESPN(leagues="ENG-Premier League", seasons=2021)
print(espn.__doc__)

/cw/dtaijupiter/NoCsBack/dtai/pieterr/Projects/soccerdata/soccerdata/_common.py:466: UserWarning: Season id "2021" is ambiguous: interpreting as "20-21"
  warnings.warn(msg)
```

Provides pd.DataFrames from JSON api available at <http://site.api.espn.com>.

Data will be downloaded as necessary and cached locally in
`~/soccerdata/data/ESPN`.

Parameters

`leagues` : string or iterable, optional
IDs of leagues to include.

`seasons` : string, int or list, optional

Seasons to include. Supports multiple formats.

Examples: '16-17'; 2016; '2016-17'; [14, 15, 16]

`proxy` : 'tor' or or dict or list(dict) or callable, optional

Use a proxy to hide your IP address. Valid options are:

- "tor": Uses the Tor network. Tor should be running in the background on port 9050.

- dict: A dictionary with the proxy to use. The dict should be a mapping of supported protocols to proxy addresses. For example::

```
{
    'http': 'http://10.10.1.10:3128',
    'https': 'http://10.10.1.10:1080',
}
```

(continues on next page)

(continued from previous page)

- `list(dict)`: A list of proxies to choose from. A different proxy will be selected from this list after failed requests, allowing rotating proxies.
- `callable`: A function that returns a valid proxy. This function will be called after failed requests, allowing rotating proxies.

`no_cache` : bool
 If True, will not use cached data.

`no_store` : bool
 If True, will not store downloaded data.

`data_dir` : Path
 Path to directory where data will be cached.

2.2.1 Game schedule

```
[4]: epl_schedule = espn.read_schedule()  
epl_schedule.head()
```

(continues on next page)

(continued from previous page)

league	season	game	game_id	\
ENG-Premier League	2021	2019-08-09 Liverpool-Norwich City	541844	
		2019-08-10 AFC Bournemouth-Sheffield United	541840	
		2019-08-10 Burnley-Southampton	541841	
		2019-08-10 Crystal Palace-Everton	541839	
		2019-08-10 Tottenham Hotspur-Aston Villa	541837	
league_id				
league	season	game		
ENG-Premier League	2021	2019-08-09 Liverpool-Norwich City	eng.1	
		2019-08-10 AFC Bournemouth-Sheffield United	eng.1	
		2019-08-10 Burnley-Southampton	eng.1	
		2019-08-10 Crystal Palace-Everton	eng.1	
		2019-08-10 Tottenham Hotspur-Aston Villa	eng.1	

2.2.2 Match sheet data

[5]:	matchsheet = espn.read_matchsheet(match_id=541465)			
	matchsheet.head()			
[5]:			is_	
	↳ home \			
	league	season	game	team
	ENG-Premier League	2021	2020-07-26 West Ham United-Aston Villa	Aston Villa False
				West Ham United True
				venue \
	league	season	game	team
	ENG-Premier League	2021	2020-07-26 West Ham United-Aston Villa	Aston Villa None
				West Ham United None
	↳ attendance \			
	league	season	game	team
	ENG-Premier League	2021	2020-07-26 West Ham United-Aston Villa	Aston Villa
	↳ 0			West Ham United
	↳ 0			
	↳ capacity \			
	league	season	game	team
	ENG-Premier League	2021	2020-07-26 West Ham United-Aston Villa	Aston Villa None
				West Ham United None
	↳ roster \			
	league	season	game	team
	ENG-Premier League	2021	2020-07-26 West Ham United-Aston Villa	Aston Villa [{
	↳ 'active': True, 'starter': True, 'jersey': '...'			

(continues on next page)

(continued from previous page)

```

West Ham United [{

  ↵'active': True, 'starter': True, 'jersey': '...

    fouls_


  ↵committed \
league          season game              team
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa
  ↵      16
                                         West Ham United
  ↵      13

yellow_


  ↵cards \
league          season game              team
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa
  ↵      2
                                         West Ham United
  ↵      1

red_


  ↵cards \
league          season game              team
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa
  ↵      0
                                         West Ham United
  ↵      0

offsides_


  ↵offsides \
league          season game              team
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa
  ↵      0
                                         West Ham United
  ↵      1

won_


  ↵corners \
league          season game              team
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa
  ↵      0
                                         West Ham United
  ↵      7

saves_ \


league          season game              team
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa
  ↵      3
                                         West Ham United
  ↵      0

possession_pct_


  ↵possession_pct \
league          season game              team
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa
  ↵      63
                                         West Ham United

```

(continues on next page)

(continued from previous page)

```

↪ 37

total_
shots \
league      season game              team
ENG-Premier League 2021 2020-07-26 West Ham United-Aston Villa Aston Villa
↪ 10
                           West Ham United
↪ 13

shots_
on_target
league      season game              team
ENG-Premier League 2021 2020-07-26 West Ham United-Aston Villa Aston Villa
↪ 1
                           West Ham United
↪ 4

```

2.2.3 Line ups

```

[6]: lineups = espn.read_lineup(match_id=541465)
lineups.head()

[6]:
    is_home \
league      season game              team      player
ENG-Premier League 2021 2020-07-26 West Ham United-Aston Villa Aston Villa Anwar El_
↪ Ghazi     False
                           Conor_
↪ Hourihane False
                           Douglas_
↪ Luiz      False
                           Ezri Konsa_
↪ False
                           Frédéric_
↪ Guilbert False

position \
league      season game              team      player
ENG-Premier League 2021 2020-07-26 West Ham United-Aston Villa Aston Villa Anwar El_
↪ Ghazi     Substitute
                           Conor_
↪ Hourihane Center Left Midfielder
                           Douglas_
↪ Luiz      Defensive Midfielder
                           Ezri Konsa_
↪ Center Right Defender
                           Frédéric_
↪ Guilbert Right Back

```

(continues on next page)

(continued from previous page)

```

    ↵      formation_place \
league          season game                      team      player
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa Anwar Elu
    ↵Ghazi           0
                                         Conor
    ↵Hourihane       10
                                         Douglas
    ↵Luiz            4
                                         Ezri Konsa 
    ↵                  5
                                         Frédéric
    ↵Guilbert        2
                                         
```



```

    ↵      sub_in \
league          season game                      team      player
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa Anwar Elu
    ↵Ghazi          90
                                         Conor
    ↵Hourihane      start
                                         Douglas
    ↵Luiz           start
                                         Ezri Konsa 
    ↵                  start
                                         Frédéric
    ↵Guilbert       start
                                         
```



```

    ↵      sub_out \
league          season game                      team      player
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa Anwar Elu
    ↵Ghazi          end
                                         Conor
    ↵Hourihane      76
                                         Douglas
    ↵Luiz           end
                                         Ezri Konsa 
    ↵                  end
                                         Frédéric
    ↵Guilbert       76
                                         
```



```

    ↵      appearances \
league          season game                      team      player
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa Anwar Elu
    ↵Ghazi          1.0
                                         Conor
    ↵Hourihane      1.0
                                         Douglas
    ↵Luiz           1.0
                                         Ezri Konsa 
                                         
```

(continues on next page)

(continued from previous page)

↪	1.0		Frédéric ↵
↪	Guilbert	1.0	
 ↪ fouls_committed \			
league	season game	team	player
ENG-Premier League 2021	2020-07-26 West Ham United-Aston Villa	Aston Villa	Anwar El ↵
↪	Ghazi	0.0	
↪	Hourihane	0.0	Conor ↵
↪	Luiz	2.0	Douglas ↵
↪		0.0	Ezri Konsa ↵
↪	Guilbert	1.0	Frédéric ↵
 ↪ fouls_suffered \			
league	season game	team	player
ENG-Premier League 2021	2020-07-26 West Ham United-Aston Villa	Aston Villa	Anwar El ↵
↪	Ghazi	0.0	
↪	Hourihane	1.0	Conor ↵
↪	Luiz	2.0	Douglas ↵
↪		0.0	Ezri Konsa ↵
↪	Guilbert	1.0	Frédéric ↵
 ↪ own_goals \			
league	season game	team	player
ENG-Premier League 2021	2020-07-26 West Ham United-Aston Villa	Aston Villa	Anwar El ↵
↪	Ghazi	0.0	
↪	Hourihane	0.0	Conor ↵
↪	Luiz	0.0	Douglas ↵
↪		0.0	Ezri Konsa ↵
↪	Guilbert	0.0	Frédéric ↵
 ↪ red_cards \			
league	season game	team	player
ENG-Premier League 2021	2020-07-26 West Ham United-Aston Villa	Aston Villa	Anwar El ↵
↪	Ghazi	0.0	
↪		0.0	Conor ↵

(continues on next page)

(continued from previous page)

↳ Hourihane	0.0	Douglas
↳ Luiz	0.0	Ezri Konsa
↳	0.0	Frédéric
↳ Guilbert	0.0	
 ↳ sub_ins \		
league	season game	team player
ENG-Premier League	2021 2020-07-26	West Ham United-Aston Villa
↳ Ghazi	1.0	Anwar El-
↳ Hourihane	0.0	Conor
↳ Luiz	0.0	Douglas
↳	0.0	Ezri Konsa
↳ Guilbert	0.0	Frédéric
 ↳ yellow_cards \		
league	season game	team player
ENG-Premier League	2021 2020-07-26	West Ham United-Aston Villa
↳ Ghazi	0.0	Anwar El-
↳ Hourihane	0.0	Conor
↳ Luiz	0.0	Douglas
↳	0.0	Ezri Konsa
↳ Guilbert	0.0	Frédéric
 ↳ goal_assists \		
league	season game	team player
ENG-Premier League	2021 2020-07-26	West Ham United-Aston Villa
↳ Ghazi	0.0	Anwar El-
↳ Hourihane	0.0	Conor
↳ Luiz	0.0	Douglas
↳	0.0	Ezri Konsa
↳ Guilbert	0.0	Frédéric
 ↳ shots_on_target \		

(continues on next page)

(continued from previous page)

league	season	game	team	player
ENG-Premier League	2021	2020-07-26	West Ham United-Aston Villa	Anwar El-Ghazi
↳	0.0			Conor Hourihane
↳	0.0			Douglas Luiz
↳	0.0			Ezri Konsa
↳	0.0			Frédéric Guilbert
 ↳ total_goals \				
league	season	game	team	player
ENG-Premier League	2021	2020-07-26	West Ham United-Aston Villa	Anwar El-Ghazi
↳	0.0			Conor Hourihane
↳	0.0			Douglas Luiz
↳	0.0			Ezri Konsa
↳	0.0			Frédéric Guilbert
 ↳ total_shots \				
league	season	game	team	player
ENG-Premier League	2021	2020-07-26	West Ham United-Aston Villa	Anwar El-Ghazi
↳	0.0			Conor Hourihane
↳	0.0			Douglas Luiz
↳	0.0			Ezri Konsa
↳	0.0			Frédéric Guilbert
 ↳ goals_conceded \				
league	season	game	team	player
ENG-Premier League	2021	2020-07-26	West Ham United-Aston Villa	Anwar El-Ghazi
↳	0.0			Conor Hourihane
↳	0.0			Douglas Luiz
↳	1.0			Ezri Konsa
↳	1.0			Frédéric Guilbert

(continues on next page)

(continued from previous page)

↪ Guilbert	0.0	
saves \		
league	season game	team player
ENG-Premier League	2021 2020-07-26	West Ham United-Aston Villa Aston Villa Anwar El <u>...</u>
↪ Ghazi	NaN	Conor <u>...</u>
↪ Hourihane	NaN	Douglas <u>...</u>
↪ Luiz	NaN	Ezri Konsa <u>...</u>
↪	NaN	Frédéric <u>...</u>
↪ Guilbert	NaN	
shots_faced \		
league	season game	team player
ENG-Premier League	2021 2020-07-26	West Ham United-Aston Villa Aston Villa Anwar El <u>...</u>
↪ Ghazi	0.0	Conor <u>...</u>
↪ Hourihane	0.0	Douglas <u>...</u>
↪ Luiz	0.0	Ezri Konsa <u>...</u>
↪	0.0	Frédéric <u>...</u>
↪ Guilbert	0.0	
offsides		
league	season game	team player
ENG-Premier League	2021 2020-07-26	West Ham United-Aston Villa Aston Villa Anwar El <u>...</u>
↪ Ghazi	0.0	Conor <u>...</u>
↪ Hourihane	0.0	Douglas <u>...</u>
↪ Luiz	0.0	Ezri Konsa <u>...</u>
↪	0.0	Frédéric <u>...</u>
↪ Guilbert	0.0	

..... doc/datasources/ESPN.ipynb ends here.

The following section was generated from doc/datasources/FBref.ipynb

[3]: import soccerdata as sd

2.3 FBref

[4]: fbref = sd.FBref(leagues="ENG-Premier League", seasons=2021)
print(fbref.__doc__)

Provides pd.DataFrames from data at http://fbref.com.

Data will be downloaded as necessary and cached locally in
``~/soccerdata/data/FBref``.

Parameters

leagues : string or iterable, optional
IDs of leagues to include. For efficiently reading data from the Top-5 European leagues, use "Big 5 European Leagues Combined".

seasons : string, int or list, optional

Seasons to include. Supports multiple formats.

Examples: '16-17'; 2016; '2016-17'; [14, 15, 16]

proxy : 'tor' or dict or list(dict) or callable, optional

Use a proxy to hide your IP address. Valid options are:

- "tor": Uses the Tor network. Tor should be running in the background on port 9050.

- dict: A dictionary with the proxy to use. The dict should be a mapping of supported protocols to proxy addresses. For example::

```
{  
    'http': 'http://10.10.1.10:3128',  
    'https': 'http://10.10.1.10:1080',  
}
```

- list(dict): A list of proxies to choose from. A different proxy will be selected from this list after failed requests, allowing rotating proxies.

- callable: A function that returns a valid proxy. This function will be called after failed requests, allowing rotating proxies.

no_cache : bool

If True, will not use cached data.

no_store : bool

If True, will not store downloaded data.

data_dir : Path

Path to directory where data will be cached.

```
/cw/dtaijupyter/NoCsBack/dtai/pieterr/Projects/soccerdata/soccerdata/_common.py:471:  
  ↵UserWarning: Season id "2021" is ambiguous: interpreting as "20-21"  
  warnings.warn(msg, stacklevel=1)
```

2.3.1 Team season stats

```
[5]: team_season_stats = fbref.read_team_season_stats(stat_type="passing")
team_season_stats.head()
```

league	season	team	players_used		90s	Total		\		
			Cmp	Att		Cmp%				
ENG-Premier League	2021	Arsenal	29	38	18041	21807	82.7			
		Aston Villa	24	38	12938	16996	76.1			
		Brighton	27	38	15772	19871	79.4			
		Burnley	25	38	10825	15562	69.6			
		Chelsea	27	38	21969	25799	85.2			
						Short	\			
			TotDist	PrgDist		Cmp	Att	Cmp%		
league	season	team								
ENG-Premier League	2021	Arsenal	306094	93540	8319	9236	90.1			
		Aston Villa	235927	93009	5729	6654	86.1			
		Brighton	283213	95248	6694	7654	87.5			
		Burnley	196816	82192	5117	6058	84.5			
		Chelsea	360450	112392	10454	11454	91.3			
						Medium	\			
			Cmp	Att	Cmp%	Long				
league	season	team								
ENG-Premier League	2021	Arsenal	7770	8814	88.2	1453	2621	55.4		
		Aston Villa	5222	6209	84.1	1571	3074	51.1		
		Brighton	6893	7960	86.6	1689	3090	54.7		
		Burnley	3688	4824	76.5	1553	3535	43.9		
		Chelsea	8965	9976	89.9	1667	2703	61.7		
			Ast	xAG	xA	A-xAG	KP	1/3	PPA	\
league	season	team								
ENG-Premier League	2021	Arsenal	38	35.1	35.3	2.9	332	1377	325	
		Aston Villa	38	37.4	31.4	0.6	403	1086	323	
		Brighton	24	33.2	34.9	-9.2	358	1132	358	
		Burnley	20	27.7	26.9	-7.7	267	923	292	
		Chelsea	38	41.5	37.1	-3.5	448	1441	377	
			CrsPA	PrgP						
league	season	team								
ENG-Premier League	2021	Arsenal	67	1760						
		Aston Villa	76	1420						
		Brighton	80	1516						
		Burnley	80	1012						
		Chelsea	70	1798						
										url
league	season	team								

(continues on next page)

(continued from previous page)

ENG-Premier League 2021	Arsenal	/en/squads/18bb7c10/2020-2021/Arsenal-Stats
	Aston Villa	/en/squads/8602292d/2020-2021/Aston-Villa-Stats
	Brighton	/en/squads/d07537b9/2020-2021/Brighton-and-Hov...
	Burnley	/en/squads/943e8050/2020-2021/Burnley-Stats
	Chelsea	/en/squads/cff3d9bb/2020-2021/Chelsea-Stats

2.3.2 Team match stats

[6]:	team_match_stats = fbref.read_team_match_stats(stat_type="schedule", team="Manchester ↳ City") team_match_stats.head()	
[6]:	date \ league season team game ENG-Premier League 2021 Manchester City 2020-09-21 Wolves-Manchester City 2020- ↳ 09-21 2020-09-27 Manchester City-Leicester City 2020- ↳ 09-27 2020-10-03 Leeds United-Manchester City 2020- ↳ 10-03 2020-10-17 Manchester City-Arsenal 2020- ↳ 10-17 2020-10-24 West Ham-Manchester City 2020- ↳ 10-24	
[6]:	time \ league season team game ENG-Premier League 2021 Manchester City 2020-09-21 Wolves-Manchester City 20: ↳ 15:00 2020-09-27 Manchester City-Leicester City 16: ↳ 30:00 2020-10-03 Leeds United-Manchester City 17: ↳ 30:00 2020-10-17 Manchester City-Arsenal 17: ↳ 30:00 2020-10-24 West Ham-Manchester City 12: ↳ 30:00	
[6]:	round \ league season team game ENG-Premier League 2021 Manchester City 2020-09-21 Wolves-Manchester City ↳ Matchweek 2 2020-09-27 Manchester City-Leicester City ↳ Matchweek 3 2020-10-03 Leeds United-Manchester City ↳ Matchweek 4 2020-10-17 Manchester City-Arsenal ↳ Matchweek 5	

(continues on next page)

(continued from previous page)

2020-10-24 West Ham-Manchester City

↳ Matchweek 6

day ↳

↳ \

league	season	team	game	day
ENG-Premier League	2021	Manchester City	2020-09-21 Wolves-Manchester City	Mon
			2020-09-27 Manchester City-Leicester City	Sun
			2020-10-03 Leeds United-Manchester City	Sat
			2020-10-17 Manchester City-Arsenal	Sat
			2020-10-24 West Ham-Manchester City	Sat

↳ venue \

league	season	team	game	venue
ENG-Premier League	2021	Manchester City	2020-09-21 Wolves-Manchester City	Away
			2020-09-27 Manchester City-Leicester City	Home
			2020-10-03 Leeds United-Manchester City	Away
			2020-10-17 Manchester City-Arsenal	Home
			2020-10-24 West Ham-Manchester City	Away

↳ result \

league	season	team	game	
ENG-Premier League	2021	Manchester City	2020-09-21 Wolves-Manchester City	
↳ W			2020-09-27 Manchester City-Leicester City	
↳ L			2020-10-03 Leeds United-Manchester City	
↳ D			2020-10-17 Manchester City-Arsenal	
↳ W			2020-10-24 West Ham-Manchester City	
↳ D				

GF ↳

↳ \

league	season	team	game	GF
ENG-Premier League	2021	Manchester City	2020-09-21 Wolves-Manchester City	3
			2020-09-27 Manchester City-Leicester City	2
			2020-10-03 Leeds United-Manchester City	1
			2020-10-17 Manchester City-Arsenal	1
			2020-10-24 West Ham-Manchester City	1

GA ↳

↳ \

league	season	team	game	GA
ENG-Premier League	2021	Manchester City	2020-09-21 Wolves-Manchester City	1
			2020-09-27 Manchester City-Leicester City	5
			2020-10-03 Leeds United-Manchester City	1
			2020-10-17 Manchester City-Arsenal	0
			2020-10-24 West Ham-Manchester City	1

(continues on next page)

(continued from previous page)

↳ opponent \					
league	season	team	game		
ENG-Premier League	2021	Manchester City	2020-09-21	Wolves-Manchester City	↳
↳ Wolves			2020-09-27	Manchester City-Leicester City	↳
↳ Leicester City			2020-10-03	Leeds United-Manchester City	↳
↳ Leeds United			2020-10-17	Manchester City-Arsenal	↳
↳ Arsenal			2020-10-24	West Ham-Manchester City	↳
↳ West Ham					
				xG	↳
↳ \					
league	season	team	game		
ENG-Premier League	2021	Manchester City	2020-09-21	Wolves-Manchester City	1.9
			2020-09-27	Manchester City-Leicester City	0.9
			2020-10-03	Leeds United-Manchester City	1.2
			2020-10-17	Manchester City-Arsenal	1.3
			2020-10-24	West Ham-Manchester City	1.0
				xGA	↳
↳ \					
league	season	team	game		
ENG-Premier League	2021	Manchester City	2020-09-21	Wolves-Manchester City	0.6
			2020-09-27	Manchester City-Leicester City	2.9
			2020-10-03	Leeds United-Manchester City	2.4
			2020-10-17	Manchester City-Arsenal	0.9
			2020-10-24	West Ham-Manchester City	0.3
				Poss	↳
↳ Poss \					
league	season	team	game		
ENG-Premier League	2021	Manchester City	2020-09-21	Wolves-Manchester City	65
			2020-09-27	Manchester City-Leicester City	72
			2020-10-03	Leeds United-Manchester City	49
			2020-10-17	Manchester City-Arsenal	58
			2020-10-24	West Ham-Manchester City	69
				Attendance	↳
↳ Attendance \					
league	season	team	game		
ENG-Premier League	2021	Manchester City	2020-09-21	Wolves-Manchester City	↳
↳ <NA>			2020-09-27	Manchester City-Leicester City	↳
↳ <NA>			2020-10-03	Leeds United-Manchester City	↳
↳ <NA>			2020-10-17	Manchester City-Arsenal	↳

(continues on next page)

(continued from previous page)

↪ <NA>			
	2020-10-24 West Ham-Manchester City		↳
↪ <NA>			↳
↪ Captain \			↳
league	season team	game	
ENG-Premier League 2021	Manchester City	2020-09-21 Wolves-Manchester City	↳
↪ Fernandinho		2020-09-27 Manchester City-Leicester City	↳
↪ Fernandinho		2020-10-03 Leeds United-Manchester City	↳
↪ Kevin De Bruyne		2020-10-17 Manchester City-Arsenal	↳
↪ Raheem Sterling		2020-10-24 West Ham-Manchester City	↳
↪ Raheem Sterling			↳
↪ Formation \			↳
league	season team	game	
ENG-Premier League 2021	Manchester City	2020-09-21 Wolves-Manchester City	4-
↪ 2-3-1		2020-09-27 Manchester City-Leicester City	4-
↪ 2-3-1		2020-10-03 Leeds United-Manchester City	↳
↪ 4-3-3		2020-10-17 Manchester City-Arsenal	3-
↪ 1-4-2		2020-10-24 West Ham-Manchester City	↳
↪ 4-3-3			↳
↪ Referee \			↳
league	season team	game	
ENG-Premier League 2021	Manchester City	2020-09-21 Wolves-Manchester City	↳
↪ Andre Marriner		2020-09-27 Manchester City-Leicester City	↳
↪ Michael Oliver		2020-10-03 Leeds United-Manchester City	↳
↪ Mike Dean		2020-10-17 Manchester City-Arsenal	↳
↪ Chris Kavanagh		2020-10-24 West Ham-Manchester City	↳
↪ Anthony Taylor			↳
↪ match_report \			↳
league	season team	game	
ENG-Premier League 2021	Manchester City	2020-09-21 Wolves-Manchester City	/en/
↪ matches/1c17eca3/Wolverhampton-Wanderers-M...		2020-09-27 Manchester City-Leicester City	/en/

(continues on next page)

(continued from previous page)

↳matches/31c2a061/Manchester-City-Leicester...	2020-10-03	Leeds United-Manchester City	/en/
↳matches/5ce15b58/Leeds-United-Manchester-C...	2020-10-17	Manchester City-Arsenal	/en/
↳matches/e95b8546/Manchester-City-Arsenal-0...	2020-10-24	West Ham-Manchester City	/en/
↳matches/2b0c0eca/West-Ham-United-Mancheste...			
			Notes
league	season	team	game
ENG-Premier League	2021	Manchester City	2020-09-21 Wolves-Manchester City <NA>
			2020-09-27 Manchester City-Leicester City <NA>
			2020-10-03 Leeds United-Manchester City <NA>
			2020-10-17 Manchester City-Arsenal <NA>
			2020-10-24 West Ham-Manchester City <NA>

2.3.3 Player season stats

```
[7]: player_season_stats = fbref.read_player_season_stats(stat_type="standard")
player_season_stats.head()
```

league	season	team	player	nation	pos	age	\
ENG-Premier League	2021	Arsenal	Ainsley Maitland-Niles	ENG	MF,DF	22	
			Alexandre Lacazette	FRA	FW	29	
			Bernd Leno	GER	GK	28	
			Bukayo Saka	ENG	FW, MF	18	
			Calum Chambers	ENG	DF	25	
				born	Playing Time	\	
					MP		
league	season	team	player				\
ENG-Premier League	2021	Arsenal	Ainsley Maitland-Niles	1997		11	
			Alexandre Lacazette	1991		31	
			Bernd Leno	1992		35	
			Bukayo Saka	2001		32	
			Calum Chambers	1995		10	
				Starts	Min	90s	\
league	season	team	player				
ENG-Premier League	2021	Arsenal	Ainsley Maitland-Niles	5	490	5.4	
			Alexandre Lacazette	22	1923	21.4	
			Bernd Leno	35	3131	34.8	
			Bukayo Saka	30	2553	28.4	
			Calum Chambers	8	753	8.4	
				Performance			\
				Gls	Ast	G+A	
league	season	team	player				

(continues on next page)

(continued from previous page)

league	season	team	player	G-PK	PK	PKatt	CrdY
ENG-Premier League 2021	Arsenal	Ainsley Maitland-Niles	0	0	0	0	0
		Alexandre Lacazette	13	2	15		
		Bernd Leno	0	0	0		
		Bukayo Saka	5	3	8		
		Calum Chambers	0	2	2		
 \\							
league	season	team	player	Expected	xG	npxG	
ENG-Premier League 2021	Arsenal	Ainsley Maitland-Niles	0	0	0	0	
		Alexandre Lacazette	10	3	3	3	
		Bernd Leno	0	0	0	0	
		Bukayo Saka	5	0	0	1	
		Calum Chambers	0	0	0	0	
 \\							
league	season	team	player	xAG	npxG+xAG		
ENG-Premier League 2021	Arsenal	Ainsley Maitland-Niles	0.1	0.1	0.1		
		Alexandre Lacazette	0	11.7	9.3		
		Bernd Leno	1	0.0	0.0		
		Bukayo Saka	0	6.9	6.9		
		Calum Chambers	0	0.4	0.4		
 \\							
league	season	team	player	Progression	PrgC	PrgP	
ENG-Premier League 2021	Arsenal	Ainsley Maitland-Niles	12	24			
		Alexandre Lacazette	15	55			
		Bernd Leno	0	2			
		Bukayo Saka	85	101			
		Calum Chambers	16	40			
 \\							
league	season	team	player	Per 90 Minutes	Gls		
ENG-Premier League 2021	Arsenal	Ainsley Maitland-Niles	21	0.0			
		Alexandre Lacazette	112	0.61			
		Bernd Leno	0	0.0			
		Bukayo Saka	269	0.18			
		Calum Chambers	49	0.0			
 \\							

(continues on next page)

(continued from previous page)

league	season	team	player	Ast	G+A	G-PK
ENG-Premier League	2021	Arsenal	Ainsley Maitland-Niles	0.0	0.0	0.0
			Alexandre Lacazette	0.09	0.7	0.47
			Bernd Leno	0.0	0.0	0.0
			Bukayo Saka	0.11	0.28	0.18
			Calum Chambers	0.24	0.24	0.0
league	season	team	player	G+A-PK	xG	xAG
ENG-Premier League	2021	Arsenal	Ainsley Maitland-Niles	0.0	0.01	0.02
			Alexandre Lacazette	0.56	0.55	0.1
			Bernd Leno	0.0	0.0	0.0
			Bukayo Saka	0.28	0.24	0.16
			Calum Chambers	0.24	0.05	0.13
league	season	team	player	xG+xAG	npxG	npxG+xAG
ENG-Premier League	2021	Arsenal	Ainsley Maitland-Niles	0.04	0.01	0.04
			Alexandre Lacazette	0.65	0.43	0.54
			Bernd Leno	0.0	0.0	0.0
			Bukayo Saka	0.4	0.24	0.4
			Calum Chambers	0.18	0.05	0.18

2.3.4 Player match stats

[8]:	player_match_stats = fbref.read_player_match_stats(stat_type="passing", match_id=
	↪ "db261cb0")
	player_match_stats.head()
[8]:	
	↪ jersey_number \
	league season game team player
	ENG-Premier League 2021 2020-09-12 Crystal Palace-Southampton Crystal Palace Andros_
	↪ Townsend 10 Cheikhoun_
	↪ Kouyaté 8 Eberechi_
	↪ Eze 25 James_
	↪ McArthur 18 James_
	↪ McCarthy 22 James_
	↪ nation \
	league season game team player

(continues on next page)

(continued from previous page)

ENG-Premier League 2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros <u>✓</u>
↳ Townsend	ENG			Cheikhou <u>✓</u>
↳ Kouyaté	SEN			Eberechi <u>✓</u>
↳ Eze	ENG			James <u>✓</u>
↳ McArthur	SCO			James <u>✓</u>
↳ McCarthy	IRL			
 ↳ pos \				
league	season	game	team	player
ENG-Premier League 2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros <u>✓</u>
↳ Townsend	RM			Cheikhou <u>✓</u>
↳ Kouyaté	CB			Eberechi <u>✓</u>
↳ Eze	LM			James <u>✓</u>
↳ McArthur	CM			James <u>✓</u>
↳ McCarthy	CM			
 ↳ age \				
league	season	game	team	player
ENG-Premier League 2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros <u>✓</u>
↳ Townsend	29-058			Cheikhou <u>✓</u>
↳ Kouyaté	30-266			Eberechi <u>✓</u>
↳ Eze	22-075			James <u>✓</u>
↳ McArthur	32-341			James <u>✓</u>
↳ McCarthy	29-305			
 ↳ min \				
league	season	game	team	player
ENG-Premier League 2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros <u>✓</u>
↳ Townsend	90			Cheikhou <u>✓</u>
↳ Kouyaté	90			Eberechi <u>✓</u>
↳ Eze	10			James <u>✓</u>

(continues on next page)

(continued from previous page)

↳ McArthur	90		James
↳ McCarthy	73		
↳ Total	\		
↳ Cmp			
league	season	game	team
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton
↳ Townsend	13		Crystal Palace
↳ Kouyaté	11		Andros
↳ Eze	8		Cheikhoun
↳ McArthur	15		Eberechi
↳ McCarthy	14		James
↳ \			James
↳ Att			
league	season	game	team
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton
↳ Townsend	30		Crystal Palace
↳ Kouyaté	19		Andros
↳ Eze	10		Cheikhoun
↳ McArthur	32		Eberechi
↳ McCarthy	23		James
↳ \			James
↳ Cmp%			
league	season	game	team
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton
↳ Townsend	43.3		Crystal Palace
↳ Kouyaté	57.9		Andros
↳ Eze	80.0		Cheikhoun
↳ McArthur	46.9		Eberechi
↳ McCarthy	60.9		James

(continues on next page)

(continued from previous page)

↔		\	↔	↔
TotDist				
league season game team player				
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros
↔ Townsend	184			Cheikhou
↔ Kouyaté	207			Eberechi
↔ Eze	177			James
↔ McArthur	322			James
↔ McCarthy	281			
↔		\	↔	↔
PrgDist				
league season game team player				
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros
↔ Townsend	68			Cheikhou
↔ Kouyaté	109			Eberechi
↔ Eze	46			James
↔ McArthur	159			James
↔ McCarthy	75			
↔		\	↔	↔
↔ Short	\			
Cmp				
league season game team player				
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros
↔ Townsend	8			Cheikhou
↔ Kouyaté	3			Eberechi
↔ Eze	2			James
↔ McArthur	2			James
↔ McCarthy	4			
↔		\	↔	↔
↔ Att				

(continues on next page)

(continued from previous page)

league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros <u>u</u>
→ Townsend	11			Cheikhou <u>u</u>
→ Kouyaté	4			Eberechi <u>u</u>
→ Eze	2			James <u>u</u>
→ McArthur	6			James <u>u</u>
→ McCarthy	6			
 → \				
 → Cmp%				
league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros <u>u</u>
→ Townsend	72.7			Cheikhou <u>u</u>
→ Kouyaté	75.0			Eberechi <u>u</u>
→ Eze	100.0			James <u>u</u>
→ McArthur	33.3			James <u>u</u>
→ McCarthy	66.7			
 → Medium \				
 → Cmp				
league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros <u>u</u>
→ Townsend	4			Cheikhou <u>u</u>
→ Kouyaté	3			Eberechi <u>u</u>
→ Eze	3			James <u>u</u>
→ McArthur	11			James <u>u</u>
→ McCarthy	6			
 → \				
 → Att				
league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros <u>u</u>
→ Townsend	8			Cheikhou <u>u</u>

(continues on next page)

(continued from previous page)

↳Kouyaté	6		Eberechi
↳Eze	3		James
↳McArthur	15		James
↳McCarthy	10		James
		\	
↳	Cmp%		
league	season	game	team
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton
↳Townsend	50.0		Crystal Palace
↳Kouyaté	50.0		Andros
↳Eze	100.0		Cheikhoun
↳McArthur	73.3		Eberechi
↳McCarthy	60.0		James
	Long	\	
↳	Cmp		
league	season	game	team
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton
↳Townsend	1		Crystal Palace
↳Kouyaté	3		Andros
↳Eze	2		Cheikhoun
↳McArthur	1		Eberechi
↳McCarthy	3		James
		\	
↳	Att		
league	season	game	team
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton
↳Townsend	6		Crystal Palace
↳Kouyaté	6		Andros
↳Eze	3		Cheikhoun
			Eberechi
			James

(continues on next page)

(continued from previous page)

↳ McArthur	7		James
↳ McCarthy	5		
	\		
↳ Cmp%			
league	season	game	team
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton
↳ Townsend	16.7		Crystal Palace
↳ Kouyaté	50.0		Andros
↳ Eze	66.7		Cheikhou
↳ McArthur	14.3		Eberechi
↳ McCarthy	60.0		James
	\		
↳ Ast	\		
league	season	game	team
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton
↳ Townsend	1		Crystal Palace
↳ Kouyaté	0		Andros
↳ Eze	0		Cheikhou
↳ McArthur	0		Eberechi
↳ McCarthy	0		James
	\		
↳ xAG	\		
league	season	game	team
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton
↳ Townsend	0.6		Crystal Palace
↳ Kouyaté	0.0		Andros
↳ Eze	0.0		Cheikhou
↳ McArthur	0.4		Eberechi
↳ McCarthy	0.0		James
	\		
↳ xA	\		

(continues on next page)

(continued from previous page)

league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros <u>u</u>
↳ Townsend	0.2			Cheikhou <u>u</u>
↳ Kouyaté	0.0			Eberechi <u>u</u>
↳ Eze	0.0			James <u>u</u>
↳ McArthur	0.0			James <u>u</u>
↳ McCarthy	0.1			

↳ KP \

league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros <u>u</u>
↳ Townsend	2			Cheikhou <u>u</u>
↳ Kouyaté	0			Eberechi <u>u</u>
↳ Eze	0			James <u>u</u>
↳ McArthur	1			James <u>u</u>
↳ McCarthy	0			

↳ 1/3 \

league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros <u>u</u>
↳ Townsend	0			Cheikhou <u>u</u>
↳ Kouyaté	1			Eberechi <u>u</u>
↳ Eze	1			James <u>u</u>
↳ McArthur	2			James <u>u</u>
↳ McCarthy	0			

↳ PPA \

league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros <u>u</u>
↳ Townsend	1			Cheikhou <u>u</u>
↳ Kouyaté	0			Eberechi <u>u</u>

(continues on next page)

(continued from previous page)

↳ Eze	0		James	
↳ McArthur	0		James	
↳ McCarthy	0			
 ↳ CrsPA \				
league season game team player				
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace
↳ Townsend	0		Andros	
↳ Kouyaté	0		Cheikhoun	
↳ Eze	0		Eberechi	
↳ McArthur	0		James	
↳ McCarthy	0		James	
 ↳ PrgP \				
league season game team player				
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace
↳ Townsend	1		Andros	
↳ Kouyaté	2		Cheikhoun	
↳ Eze	2		Eberechi	
↳ McArthur	3		James	
↳ McCarthy	2		James	
 ↳ game_id				
league season game team player				
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace
↳ Townsend	db261cb0		Andros	
↳ Kouyaté	db261cb0		Cheikhoun	
↳ Eze	db261cb0		Eberechi	
↳ McArthur	db261cb0		James	
↳ McCarthy	db261cb0		James	

2.3.5 Game schedule

```
[9]: epl_schedule = fbref.read_schedule()
epl_schedule.head()
```

league	season	game	week	day	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	1	Sat	
		2020-09-12 Fulham-Arsenal	1	Sat	
		2020-09-12 Liverpool-Leeds United	1	Sat	
		2020-09-12 West Ham-Newcastle Utd	1	Sat	
		2020-09-13 Tottenham-Everton	1	Sun	
					date \
league	season	game			
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	2020-09-12		
		2020-09-12 Fulham-Arsenal	2020-09-12		
		2020-09-12 Liverpool-Leeds United	2020-09-12		
		2020-09-12 West Ham-Newcastle Utd	2020-09-12		
		2020-09-13 Tottenham-Everton	2020-09-13		
					time \
league	season	game			
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	15:00		
		2020-09-12 Fulham-Arsenal	12:30		
		2020-09-12 Liverpool-Leeds United	17:30		
		2020-09-12 West Ham-Newcastle Utd	20:00		
		2020-09-13 Tottenham-Everton	16:30		
					home_team \
league	season	game			
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Crystal Palace		
		2020-09-12 Fulham-Arsenal	Fulham		
		2020-09-12 Liverpool-Leeds United	Liverpool		
		2020-09-12 West Ham-Newcastle Utd	West Ham		
		2020-09-13 Tottenham-Everton	Tottenham		
					home_xg \
league	season	game			
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	1.1		
		2020-09-12 Fulham-Arsenal	0.1		
		2020-09-12 Liverpool-Leeds United	2.7		
		2020-09-12 West Ham-Newcastle Utd	1.0		
		2020-09-13 Tottenham-Everton	1.1		
					score \
league	season	game			
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	1-0		
		2020-09-12 Fulham-Arsenal	0-3		
		2020-09-12 Liverpool-Leeds United	4-3		
		2020-09-12 West Ham-Newcastle Utd	0-2		
		2020-09-13 Tottenham-Everton	0-1		

(continues on next page)

(continued from previous page)

league	season	game	away_xg \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	0.9
		2020-09-12 Fulham-Arsenal	1.9
		2020-09-12 Liverpool-Leeds United	0.3
		2020-09-12 West Ham-Newcastle Utd	1.6
		2020-09-13 Tottenham-Everton	1.2
league	season	game	away_team \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Southampton
		2020-09-12 Fulham-Arsenal	Arsenal
		2020-09-12 Liverpool-Leeds United	Leeds United
		2020-09-12 West Ham-Newcastle Utd	Newcastle Utd
		2020-09-13 Tottenham-Everton	Everton
league	season	game	attendance \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	<NA>
		2020-09-12 Fulham-Arsenal	<NA>
		2020-09-12 Liverpool-Leeds United	<NA>
		2020-09-12 West Ham-Newcastle Utd	<NA>
		2020-09-13 Tottenham-Everton	<NA>
league	season	game	venue \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Selhurst <u>u</u>
↳ Park		2020-09-12 Fulham-Arsenal	Craven <u>u</u>
↳ Cottage		2020-09-12 Liverpool-Leeds United	<u>u</u>
↳ Anfield		2020-09-12 West Ham-Newcastle Utd	London <u>u</u>
↳ Stadium		2020-09-13 Tottenham-Everton	Tottenham Hotspur <u>u</u>
↳ Stadium			
league	season	game	referee \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Jonathan Moss
		2020-09-12 Fulham-Arsenal	Chris Kavanagh
		2020-09-12 Liverpool-Leeds United	Michael Oliver
		2020-09-12 West Ham-Newcastle Utd	Stuart Attwell
		2020-09-13 Tottenham-Everton	Martin Atkinson
league	season	game	match_report \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	/en/matches/db261cb0/
↳ Crystal-Palace-Southampto...		2020-09-12 Fulham-Arsenal	/en/matches/bf52349b/

(continues on next page)

(continued from previous page)

↳ Fulham-Arsenal-September...	2020-09-12 Liverpool-Leeds United	/en/matches/21b58926/
↳ Liverpool-Leeds-United-Se...	2020-09-12 West Ham-Newcastle Utd	/en/matches/78495ced/
↳ West-Ham-United-Newcastle...	2020-09-13 Tottenham-Everton	/en/matches/fc7f9aa1/
↳ Tottenham-Hotspur-Everton...		
		notes \
league	season game	
ENG-Premier League	2021 2020-09-12 Crystal Palace-Southampton	<NA>
	2020-09-12 Fulham-Arsenal	<NA>
	2020-09-12 Liverpool-Leeds United	<NA>
	2020-09-12 West Ham-Newcastle Utd	<NA>
	2020-09-13 Tottenham-Everton	<NA>
		game_id
league	season game	
ENG-Premier League	2021 2020-09-12 Crystal Palace-Southampton	db261cb0
	2020-09-12 Fulham-Arsenal	bf52349b
	2020-09-12 Liverpool-Leeds United	21b58926
	2020-09-12 West Ham-Newcastle Utd	78495ced
	2020-09-13 Tottenham-Everton	fc7f9aa1

2.3.6 Line ups

[10]:	lineups = fbref.read_lineup(match_id="db261cb0")		
	lineups.head()		
[10]:		jersey_number \	
	league	season game	
	ENG-Premier League	2021 2020-09-12 Crystal Palace-Southampton	31
		2020-09-12 Crystal Palace-Southampton	2
		2020-09-12 Crystal Palace-Southampton	6
		2020-09-12 Crystal Palace-Southampton	8
		2020-09-12 Crystal Palace-Southampton	9
			player \
	league	season game	
	ENG-Premier League	2021 2020-09-12 Crystal Palace-Southampton	Vicente Guaita
		2020-09-12 Crystal Palace-Southampton	Joel Ward
		2020-09-12 Crystal Palace-Southampton	Scott Dann
		2020-09-12 Crystal Palace-Southampton	Cheikhou Kouyaté
		2020-09-12 Crystal Palace-Southampton	Jordan Ayew
			team \
	league	season game	
	ENG-Premier League	2021 2020-09-12 Crystal Palace-Southampton	Crystal Palace
		2020-09-12 Crystal Palace-Southampton	Crystal Palace
		2020-09-12 Crystal Palace-Southampton	Crystal Palace
		2020-09-12 Crystal Palace-Southampton	Crystal Palace

(continues on next page)

(continued from previous page)

2020-09-12 Crystal Palace-Southampton Crystal Palace

is_starter \

league	season	game	is_starter
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	True
		2020-09-12 Crystal Palace-Southampton	True
		2020-09-12 Crystal Palace-Southampton	True
		2020-09-12 Crystal Palace-Southampton	True
		2020-09-12 Crystal Palace-Southampton	True

position \

league	season	game	position
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	GK
		2020-09-12 Crystal Palace-Southampton	RB
		2020-09-12 Crystal Palace-Southampton	CB
		2020-09-12 Crystal Palace-Southampton	CB
		2020-09-12 Crystal Palace-Southampton	FW

minutes_played

league	season	game	minutes_played
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	90
		2020-09-12 Crystal Palace-Southampton	90
		2020-09-12 Crystal Palace-Southampton	90
		2020-09-12 Crystal Palace-Southampton	90
		2020-09-12 Crystal Palace-Southampton	90

2.3.7 Events

[11]: events = fbref.read_events(match_id="db261cb0")
events.head()

league	season	game	team
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Crystal Palace
		2020-09-12 Crystal Palace-Southampton	Crystal Palace
		2020-09-12 Crystal Palace-Southampton	Southampton
		2020-09-12 Crystal Palace-Southampton	Southampton
		2020-09-12 Crystal Palace-Southampton	Crystal Palace

minute score \

league	season	game	minute	score
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	13	1:0
		2020-09-12 Crystal Palace-Southampton	41	1:0
		2020-09-12 Crystal Palace-Southampton	46	1:0
		2020-09-12 Crystal Palace-Southampton	51	1:0
		2020-09-12 Crystal Palace-Southampton	58	1:0

player1 \

league	season	game	player1
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Wilfried Zaha
		2020-09-12 Crystal Palace-Southampton	Jeffrey Schlupp

(continues on next page)

(continued from previous page)

			2020-09-12 Crystal Palace-Southampton	Jannik Vestergaard
			2020-09-12 Crystal Palace-Southampton	Kyle Walker-Peters
			2020-09-12 Crystal Palace-Southampton	James McCarthy
				player2 \
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Andros Townsend	
		2020-09-12 Crystal Palace-Southampton	None	
		2020-09-12 Crystal Palace-Southampton	Jan Bednarek	
		2020-09-12 Crystal Palace-Southampton	None	
		2020-09-12 Crystal Palace-Southampton	None	
				event_type
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	goal	
		2020-09-12 Crystal Palace-Southampton	yellow_card	
		2020-09-12 Crystal Palace-Southampton	substitute_in	
		2020-09-12 Crystal Palace-Southampton	yellow_card	
		2020-09-12 Crystal Palace-Southampton	yellow_card	

2.3.8 Shot events

[12]: shots = fbref.read_shot_events(match_id="db261cb0")
shots.head()

[12]:				minute \
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	4	
		2020-09-12 Crystal Palace-Southampton	13	
		2020-09-12 Crystal Palace-Southampton	32	
		2020-09-12 Crystal Palace-Southampton	36	
		2020-09-12 Crystal Palace-Southampton	46	
				player \
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Jack Stephens	
		2020-09-12 Crystal Palace-Southampton	Wilfried Zaha	
		2020-09-12 Crystal Palace-Southampton	Danny Ings	
		2020-09-12 Crystal Palace-Southampton	Cheikhou Kouyaté	
		2020-09-12 Crystal Palace-Southampton	Nathan Redmond	
				team \
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Southampton	
		2020-09-12 Crystal Palace-Southampton	Crystal Palace	
		2020-09-12 Crystal Palace-Southampton	Southampton	
		2020-09-12 Crystal Palace-Southampton	Crystal Palace	
		2020-09-12 Crystal Palace-Southampton	Southampton	

(continues on next page)

(continued from previous page)

league	season	game	xG	PSxG	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	0.07	0.09	
		2020-09-12 Crystal Palace-Southampton	0.44	0.73	
		2020-09-12 Crystal Palace-Southampton	0.04	0.12	
		2020-09-12 Crystal Palace-Southampton	0.15	0.09	
		2020-09-12 Crystal Palace-Southampton	0.07	0.03	
league	season	game	outcome	\	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Saved		
		2020-09-12 Crystal Palace-Southampton	Goal		
		2020-09-12 Crystal Palace-Southampton	Saved		
		2020-09-12 Crystal Palace-Southampton	Saved		
		2020-09-12 Crystal Palace-Southampton	Saved		
league	season	game	distance	\	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	10		
		2020-09-12 Crystal Palace-Southampton	9		
		2020-09-12 Crystal Palace-Southampton	10		
		2020-09-12 Crystal Palace-Southampton	11		
		2020-09-12 Crystal Palace-Southampton	14		
league	season	game	body_part	\	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Head		
		2020-09-12 Crystal Palace-Southampton	Right Foot		
		2020-09-12 Crystal Palace-Southampton	Head		
		2020-09-12 Crystal Palace-Southampton	Right Foot		
		2020-09-12 Crystal Palace-Southampton	Right Foot		
league	season	game	notes	\	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	<NA>		
		2020-09-12 Crystal Palace-Southampton	Volley		
		2020-09-12 Crystal Palace-Southampton	<NA>		
		2020-09-12 Crystal Palace-Southampton	<NA>		
		2020-09-12 Crystal Palace-Southampton	<NA>		
league	season	game	SCA 1	\	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	James Ward-Prowse		
		2020-09-12 Crystal Palace-Southampton	Andros Townsend		
		2020-09-12 Crystal Palace-Southampton	Kyle Walker-Peters		

(continues on next page)

(continued from previous page)

			2020-09-12 Crystal Palace-Southampton	Andros Townsend
			2020-09-12 Crystal Palace-Southampton	Che Adams
				\
				event
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Pass (Dead)	
		2020-09-12 Crystal Palace-Southampton	Pass (Live)	
		2020-09-12 Crystal Palace-Southampton	Pass (Live)	
		2020-09-12 Crystal Palace-Southampton	Pass (Dead)	
		2020-09-12 Crystal Palace-Southampton	Pass (Live)	
				SCA 2 \
				player
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Kyle Walker-Peters	
		2020-09-12 Crystal Palace-Southampton	James McCarthy	
		2020-09-12 Crystal Palace-Southampton	William Smallbone	
		2020-09-12 Crystal Palace-Southampton	Wilfried Zaha	
		2020-09-12 Crystal Palace-Southampton	Danny Ings	
				event
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Pass (Dead)	
		2020-09-12 Crystal Palace-Southampton	Interception	
		2020-09-12 Crystal Palace-Southampton	Pass (Live)	
		2020-09-12 Crystal Palace-Southampton	Fouled	
		2020-09-12 Crystal Palace-Southampton	Pass (Live)	

doc/datasources/FBref.ipynb ends here.

The following section was generated from doc/datasources/MatchHistory.ipynb

[3]: import soccerdata as sd

2.4 Match History

```
[4]: mh = sd.MatchHistory(leagues="ENG-Premier League", seasons=2021)
print(mh.__doc__)

/cw/dtaijupiter/NoCsBack/dtai/pieterr/Projects/soccerdata/soccerdata/_common.py:466: UserWarning: Season id "2021" is ambiguous: interpreting as "20-21"
  warnings.warn(msg)
```

Provides pd.DataFrames from CSV files available at <http://www.football-data.co.uk/data.php>.

Data will be downloaded as necessary and cached locally in
`~/.soccerdata/data/MatchHistory`.

(continues on next page)

(continued from previous page)

```

Parameters
-----
leagues : string or iterable
    IDs of leagues to include.
seasons : string, int or list
    Seasons to include. Supports multiple formats.
    Examples: '16-17'; 2016; '2016-17'; [14, 15, 16]
proxy : 'tor' or dict or list(dict) or callable, optional
    Use a proxy to hide your IP address. Valid options are:
        - "tor": Uses the Tor network. Tor should be running in
            the background on port 9050.
        - dict: A dictionary with the proxy to use. The dict should be
            a mapping of supported protocols to proxy addresses. For example::
                {
                    'http': 'http://10.10.1.10:3128',
                    'https': 'http://10.10.1.10:1080',
                }
        - list(dict): A list of proxies to choose from. A different proxy will
            be selected from this list after failed requests, allowing rotating
            proxies.
        - callable: A function that returns a valid proxy. This function will
            be called after failed requests, allowing rotating proxies.
no_cache : bool
    If True, will not use cached data.
no_store : bool
    If True, will not store downloaded data.
data_dir : Path, optional
    Path to directory where data will be cached.

```

2.4.1 Historic match results and betting odds

```

[5]: hist = mh.read_games()
hist.head()

[5]:                                                 date \
league      season game
ENG-Premier League 2021  2020-01-11 Aston Villa-Southampton 2020-01-11 12:00:00
                                         2020-01-11 Man United-Arsenal      2020-01-11 16:30:00
                                         2020-01-11 Newcastle-Everton      2020-01-11 14:00:00
                                         2020-01-11 Tottenham-Brighton     2020-01-11 19:15:00
                                         2020-02-11 Fulham-West Brom      2020-02-11 17:30:00

                                                 home_team \
league      season game
ENG-Premier League 2021  2020-01-11 Aston Villa-Southampton  Aston Villa
                                         2020-01-11 Man United-Arsenal      Man United
                                         2020-01-11 Newcastle-Everton      Newcastle
                                         2020-01-11 Tottenham-Brighton     Tottenham

```

(continues on next page)

(continued from previous page)

2020-02-11 Fulham-West Brom				Fulham
league	season	game	away_team	\
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	Southampton	
		2020-01-11 Man United-Arsenal	Arsenal	
		2020-01-11 Newcastle-Everton	Everton	
		2020-01-11 Tottenham-Brighton	Brighton	
		2020-02-11 Fulham-West Brom	West Brom	
league	season	game	FTHG	FTAG FTR \
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	3	4 A
		2020-01-11 Man United-Arsenal	0	1 A
		2020-01-11 Newcastle-Everton	2	1 H
		2020-01-11 Tottenham-Brighton	2	1 H
		2020-02-11 Fulham-West Brom	2	0 H
league	season	game	HTHG	HTAG HTR \
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	0	3 A
		2020-01-11 Man United-Arsenal	0	0 D
		2020-01-11 Newcastle-Everton	0	0 D
		2020-01-11 Tottenham-Brighton	1	0 H
		2020-02-11 Fulham-West Brom	2	0 H
league	season	game	referee	HS \
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	D England	19
		2020-01-11 Man United-Arsenal	M Dean	8
		2020-01-11 Newcastle-Everton	S Attwell	11
		2020-01-11 Tottenham-Brighton	G Scott	9
		2020-02-11 Fulham-West Brom	S Hooper	13
league	season	game	AS	HST AST \
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	9	10 4
		2020-01-11 Man United-Arsenal	7	2 2
		2020-01-11 Newcastle-Everton	15	4 4
		2020-01-11 Tottenham-Brighton	6	3 2
		2020-02-11 Fulham-West Brom	10	6 1
league	season	game	HF	AF HC AC \
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	12	17 11 1
		2020-01-11 Man United-Arsenal	12	12 6 3
		2020-01-11 Newcastle-Everton	9	10 5 4
		2020-01-11 Tottenham-Brighton	14	13 4 5
		2020-02-11 Fulham-West Brom	13	11 5 2
league	season	game	HY	AY HR AR \

(continues on next page)

(continued from previous page)

ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	1	2	0	0
		2020-01-11	Man United-Arsenal	3	3	0	0
		2020-01-11	Newcastle-Everton	2	4	0	0
		2020-01-11	Tottenham-Brighton	2	1	0	0
		2020-02-11	Fulham-West Brom	3	2	0	0
league				B365H	B365D	\	
league				B365A	BWH	\	
league				BWD	BWA	\	
league				IWH	IWD	\	
league				IWA	PSH	\	
league				PSD	PSA	\	

(continues on next page)

(continued from previous page)

			2020-02-11 Fulham-West Brom	3.40	3.13	
				WHH	WHD	WHA \
league	season	game				
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.50	3.5	2.75	
		2020-01-11 Man United-Arsenal	2.05	3.6	3.60	
		2020-01-11 Newcastle-Everton	3.70	3.4	2.05	
		2020-01-11 Tottenham-Brighton	1.63	4.0	5.25	
		2020-02-11 Fulham-West Brom	2.40	3.3	3.00	
				VCH	VCD	\
league	season	game				
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.40	3.60		
		2020-01-11 Man United-Arsenal	2.05	3.60		
		2020-01-11 Newcastle-Everton	3.50	3.50		
		2020-01-11 Tottenham-Brighton	1.62	4.20		
		2020-02-11 Fulham-West Brom	2.45	3.25		
				VCA	MaxH	\
league	season	game				
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.80	2.60		
		2020-01-11 Man United-Arsenal	3.50	2.10		
		2020-01-11 Newcastle-Everton	2.05	4.00		
		2020-01-11 Tottenham-Brighton	5.25	1.67		
		2020-02-11 Fulham-West Brom	3.00	2.56		
				MaxD	MaxA	\
league	season	game				
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	3.84	2.85		
		2020-01-11 Man United-Arsenal	3.85	3.76		
		2020-01-11 Newcastle-Everton	3.59	2.15		
		2020-01-11 Tottenham-Brighton	4.30	5.75		
		2020-02-11 Fulham-West Brom	3.50	3.17		
				AvgH	AvgD	\
league	season	game				
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.46	3.60		
		2020-01-11 Man United-Arsenal	2.04	3.66		
		2020-01-11 Newcastle-Everton	3.71	3.47		
		2020-01-11 Tottenham-Brighton	1.63	4.17		
		2020-02-11 Fulham-West Brom	2.44	3.32		
				AvgA	B365>2.5	\
league	season	game				
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.78		1.66	
		2020-01-11 Man United-Arsenal	3.58		1.72	
		2020-01-11 Newcastle-Everton	2.06		1.80	
		2020-01-11 Tottenham-Brighton	5.29		1.66	
		2020-02-11 Fulham-West Brom	3.01		2.00	
				B365<2.5	P>2.5	\
league	season	game				

(continues on next page)

(continued from previous page)

ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	2.2	1.66	
	2020-01-11	Man United-Arsenal	2.1	1.76	
	2020-01-11	Newcastle-Everton	2.0	1.86	
	2020-01-11	Tottenham-Brighton	2.2	1.68	
	2020-02-11	Fulham-West Brom	1.8	2.13	
P<2.5 Max>2.5 \					
league	season	game			
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	2.37	1.69	
	2020-01-11	Man United-Arsenal	2.19	1.81	
	2020-01-11	Newcastle-Everton	2.07	1.88	
	2020-01-11	Tottenham-Brighton	2.32	1.71	
	2020-02-11	Fulham-West Brom	1.80	2.14	
Max<2.5 \					
league	season	game			
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	2.38		
	2020-01-11	Man United-Arsenal	2.22		
	2020-01-11	Newcastle-Everton	2.08		
	2020-01-11	Tottenham-Brighton	2.35		
	2020-02-11	Fulham-West Brom	1.85		
Avg>2.5 \					
league	season	game			
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	1.65		
	2020-01-11	Man United-Arsenal	1.74		
	2020-01-11	Newcastle-Everton	1.82		
	2020-01-11	Tottenham-Brighton	1.65		
	2020-02-11	Fulham-West Brom	2.06		
Avg<2.5 AHh \					
league	season	game			
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	2.27	0.00	
	2020-01-11	Man United-Arsenal	2.12	-0.50	
	2020-01-11	Newcastle-Everton	2.01	0.25	
	2020-01-11	Tottenham-Brighton	2.26	-1.00	
	2020-02-11	Fulham-West Brom	1.78	-0.25	
B365AHH \					
league	season	game			
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	1.83		
	2020-01-11	Man United-Arsenal	2.05		
	2020-01-11	Newcastle-Everton	2.05		
	2020-01-11	Tottenham-Brighton	2.02		
	2020-02-11	Fulham-West Brom	2.10		
B365AHA PAHH \					
league	season	game			
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	2.07	1.87	
	2020-01-11	Man United-Arsenal	1.85	2.05	
	2020-01-11	Newcastle-Everton	1.75	2.15	
	2020-01-11	Tottenham-Brighton	1.77	2.15	

(continues on next page)

(continued from previous page)

			2020-02-11 Fulham-West Brom	1.80	2.11	
league	season	game		PAHA	MaxAHH	\
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.07	1.87		
		2020-01-11 Man United-Arsenal	1.88	2.09		
		2020-01-11 Newcastle-Everton	1.80	2.17		
		2020-01-11 Tottenham-Brighton	1.80	2.16		
		2020-02-11 Fulham-West Brom	1.84	2.15		
				MaxAHA	AvgAHH	\
league	season	game		MaxAHA	AvgAHH	\
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.09	1.83		
		2020-01-11 Man United-Arsenal	1.90	2.04		
		2020-01-11 Newcastle-Everton	1.83	2.12		
		2020-01-11 Tottenham-Brighton	1.84	2.09		
		2020-02-11 Fulham-West Brom	1.86	2.10		
				AvgAHA	B365CH	\
league	season	game		AvgAHA	B365CH	\
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.05	2.60		
		2020-01-11 Man United-Arsenal	1.84	1.95		
		2020-01-11 Newcastle-Everton	1.77	3.30		
		2020-01-11 Tottenham-Brighton	1.80	1.55		
		2020-02-11 Fulham-West Brom	1.80	2.40		
				B365CD	B365CA	\
league	season	game		B365CD	B365CA	\
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	3.50	2.62		
		2020-01-11 Man United-Arsenal	3.60	3.80		
		2020-01-11 Newcastle-Everton	3.30	2.25		
		2020-01-11 Tottenham-Brighton	4.33	6.00		
		2020-02-11 Fulham-West Brom	3.30	3.00		
				BWCH	BWCD	\
league	season	game		BWCH	BWCD	\
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.50	3.60		
		2020-01-11 Man United-Arsenal	1.95	3.60		
		2020-01-11 Newcastle-Everton	3.30	3.30		
		2020-01-11 Tottenham-Brighton	1.57	4.25		
		2020-02-11 Fulham-West Brom	2.45	3.30		
				BWCA	IWCH	\
league	season	game		BWCA	IWCH	\
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.70	2.55		
		2020-01-11 Man United-Arsenal	3.90	2.00		
		2020-01-11 Newcastle-Everton	2.25	3.15		
		2020-01-11 Tottenham-Brighton	5.50	1.60		
		2020-02-11 Fulham-West Brom	2.95	2.40		
				IWCD	IWCA	\
league	season	game		IWCD	IWCA	\

(continues on next page)

(continued from previous page)

ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	3.50	2.70		
	2020-01-11	Man United-Arsenal	3.50	3.80		
	2020-01-11	Newcastle-Everton	3.35	2.30		
	2020-01-11	Tottenham-Brighton	4.20	5.75		
	2020-02-11	Fulham-West Brom	3.05	3.10		
			PSCH	PSCD	\	
league	season	game				
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	2.69	3.56		
	2020-01-11	Man United-Arsenal	2.02	3.65		
	2020-01-11	Newcastle-Everton	3.34	3.43		
	2020-01-11	Tottenham-Brighton	1.61	4.36		
	2020-02-11	Fulham-West Brom	2.48	3.32		
			PSCA	WHCH	\	
league	season	game				
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	2.73	2.60		
	2020-01-11	Man United-Arsenal	3.97	2.00		
	2020-01-11	Newcastle-Everton	2.33	3.40		
	2020-01-11	Tottenham-Brighton	5.77	1.57		
	2020-02-11	Fulham-West Brom	3.19	2.40		
			WHCD	WHCA	\	
league	season	game				
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	3.50	2.62		
	2020-01-11	Man United-Arsenal	3.50	3.80		
	2020-01-11	Newcastle-Everton	3.25	2.25		
	2020-01-11	Tottenham-Brighton	4.00	6.00		
	2020-02-11	Fulham-West Brom	3.20	3.10		
			VCCH	VCCD	\	
league	season	game				
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	2.50	3.60		
	2020-01-11	Man United-Arsenal	1.95	3.50		
	2020-01-11	Newcastle-Everton	3.30	3.30		
	2020-01-11	Tottenham-Brighton	1.57	4.20		
	2020-02-11	Fulham-West Brom	2.40	3.25		
			VCCA	MaxCH	\	
league	season	game				
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	2.70	2.70		
	2020-01-11	Man United-Arsenal	4.00	2.06		
	2020-01-11	Newcastle-Everton	2.25	3.62		
	2020-01-11	Tottenham-Brighton	5.75	1.63		
	2020-02-11	Fulham-West Brom	3.10	2.48		
			MaxCD	MaxCA	\	
league	season	game				
ENG-Premier League 2021	2020-01-11	Aston Villa-Southampton	3.70	2.80		
	2020-01-11	Man United-Arsenal	3.75	4.00		
	2020-01-11	Newcastle-Everton	3.50	2.34		
	2020-01-11	Tottenham-Brighton	4.45	6.39		

(continues on next page)

(continued from previous page)

			2020-02-11	Fulham-West Brom	3.45	3.22	
league	season	game			AvgCH	AvgCD	\
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	2.57	3.55		
		2020-01-11	Man United-Arsenal	1.99	3.59		
		2020-01-11	Newcastle-Everton	3.34	3.33		
		2020-01-11	Tottenham-Brighton	1.59	4.29		
		2020-02-11	Fulham-West Brom	2.41	3.27		
league	season	game			AvgCA	\	
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	2.69			
		2020-01-11	Man United-Arsenal	3.87			
		2020-01-11	Newcastle-Everton	2.27			
		2020-01-11	Tottenham-Brighton	5.62			
		2020-02-11	Fulham-West Brom	3.09			
league	season	game			B365C>2.5	\	
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	1.66			
		2020-01-11	Man United-Arsenal	1.90			
		2020-01-11	Newcastle-Everton	2.00			
		2020-01-11	Tottenham-Brighton	1.66			
		2020-02-11	Fulham-West Brom	2.05			
league	season	game			B365C<2.5	\	
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	2.20			
		2020-01-11	Man United-Arsenal	1.90			
		2020-01-11	Newcastle-Everton	1.80			
		2020-01-11	Tottenham-Brighton	2.20			
		2020-02-11	Fulham-West Brom	1.85			
league	season	game			PC>2.5	PC<2.5	\
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	1.70	2.29		
		2020-01-11	Man United-Arsenal	1.96	1.96		
		2020-01-11	Newcastle-Everton	2.12	1.81		
		2020-01-11	Tottenham-Brighton	1.68	2.32		
		2020-02-11	Fulham-West Brom	2.05	1.88		
league	season	game			MaxC>2.5	\	
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	1.71			
		2020-01-11	Man United-Arsenal	1.99			
		2020-01-11	Newcastle-Everton	2.14			
		2020-01-11	Tottenham-Brighton	1.72			
		2020-02-11	Fulham-West Brom	2.11			
league	season	game			MaxC<2.5	\	

(continues on next page)

(continued from previous page)

league	season	game	AvgC>2.5	\
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.37	
		2020-01-11 Man United-Arsenal	2.10	
		2020-01-11 Newcastle-Everton	1.92	
		2020-01-11 Tottenham-Brighton	2.37	
		2020-02-11 Fulham-West Brom	1.92	
league	season	game	AvgC<2.5	AHCh \
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	1.66	
		2020-01-11 Man United-Arsenal	1.88	
		2020-01-11 Newcastle-Everton	2.04	
		2020-01-11 Tottenham-Brighton	1.65	
		2020-02-11 Fulham-West Brom	2.02	
league	season	game	B365CAHH	\
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	2.24	0.00
		2020-01-11 Man United-Arsenal	1.95	-0.50
		2020-01-11 Newcastle-Everton	1.80	0.25
		2020-01-11 Tottenham-Brighton	2.27	-1.00
		2020-02-11 Fulham-West Brom	1.81	-0.25
league	season	game	B365CAHA	PCAH \
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	1.93	
		2020-01-11 Man United-Arsenal	2.00	
		2020-01-11 Newcastle-Everton	1.95	
		2020-01-11 Tottenham-Brighton	2.06	
		2020-02-11 Fulham-West Brom	2.06	
league	season	game	PCAH \	MaxCAHH \
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	1.97	1.95
		2020-01-11 Man United-Arsenal	1.90	2.03
		2020-01-11 Newcastle-Everton	1.95	1.93
		2020-01-11 Tottenham-Brighton	1.84	2.11
		2020-02-11 Fulham-West Brom	1.84	2.11
league	season	game	MaxCAHA	\
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	1.98	1.97
		2020-01-11 Man United-Arsenal	1.90	2.05
		2020-01-11 Newcastle-Everton	2.00	1.99
		2020-01-11 Tottenham-Brighton	1.84	2.13
		2020-02-11 Fulham-West Brom	1.84	2.12

(continues on next page)

(continued from previous page)

2020-02-11 Fulham-West Brom			1.87
league	season	game	AvgCAHH AvgCAHA
ENG-Premier League	2021	2020-01-11 Aston Villa-Southampton	1.89 1.98
		2020-01-11 Man United-Arsenal	2.00 1.88
		2020-01-11 Newcastle-Everton	1.93 1.95
		2020-01-11 Tottenham-Brighton	2.02 1.86
		2020-02-11 Fulham-West Brom	2.08 1.82

..... doc/datasources/MatchHistory.ipynb ends here.

The following section was generated from doc/datasources/Sofascore.ipynb

[1]: import pandas as pd

pd.set_option("display.max_columns", None)

```
%env SOCCERDATA_LOGLEVEL=ERROR
%env SOCCERDATA_NOCACHE=True
%env SOCCERDATA_NOSTORE=True

env: SOCCERDATA_LOGLEVEL=ERROR
env: SOCCERDATA_NOCACHE=True
env: SOCCERDATA_NOSTORE=True
```

[3]: import soccerdata as sd

2.5 Sofascore

[4]: sofascore = sd.Sofascore(leagues="ESP-La Liga", seasons="2022/2023")
print(sofascore.__doc__)

Provides pd.DataFrames from data available at <http://www.sofascore.com>.

Data will be downloaded as necessary and cached locally in
`~/soccerdata/data/Sofascore`.

Parameters

leagues : string or iterable, optional
IDs of Leagues to include.
seasons : string, int or list, optional
Seasons to include. Supports multiple formats.
Examples: '16-17'; 2016; '2016-17'; [14, 15, 16]
proxy : 'tor' or dict or list(dict) or callable, optional
Use a proxy to hide your IP address. Valid options are:
- 'tor': Uses the Tor network. Tor should be running in
the background on port 9050.
- dict: A dictionary with the proxy to use. The dict should be
a mapping of supported protocols to proxy addresses. For example::

(continues on next page)

(continued from previous page)

```

{
    'http': 'http://10.10.1.10:3128',
    'https': 'http://10.10.1.10:1080',
}

- list(dict): A list of proxies to choose from. A different proxy will
  be selected from this list after failed requests, allowing rotating
  proxies.
- callable: A function that returns a valid proxy. This function will
  be called after failed requests, allowing rotating proxies.
no_cache : bool
  If True, will not use cached data.
no_store : bool
  If True, will not store downloaded data.
data_dir : Path
  Path to directory where data will be cached.

```

2.5.1 Read league table

[5]: `league_table = sofascore.read_league_table()
league_table.head()`

league	season	team	MP	W	D	L	GF	GA	GD	Pts
ESP-La Liga	2223	Barcelona	38	28	4	6	70	20	50	88
	2223	Real Madrid	38	24	6	8	75	36	39	78
	2223	Atlético Madrid	38	23	8	7	70	33	37	77
	2223	Real Sociedad	38	21	8	9	51	35	16	71
	2223	Villarreal	38	19	7	12	59	40	19	64

2.5.2 Read schedule

[8]: `schedule = sofascore.read_schedule()
schedule.head()`

league	season	game	round	week	\
ESP-La Liga	2223	2022-08-12 Osasuna-Sevilla	None	1	
		2022-08-13 Barcelona-Rayo Vallecano	None	1	
		2022-08-13 Celta Vigo-Espanyol	None	1	
		2022-08-13 Real Valladolid-Villarreal	None	1	
		2022-08-14 Almería-Real Madrid	None	1	

league	season	game	date	\
ESP-La Liga	2223	2022-08-12 Osasuna-Sevilla	2022-08-12 15:00:00	
		2022-08-13 Barcelona-Rayo Vallecano	2022-08-13 15:00:00	
		2022-08-13 Celta Vigo-Espanyol	2022-08-13 11:00:00	

(continues on next page)

(continued from previous page)

2022-08-13	Real Valladolid-Villarreal	2022-08-13	13:00:00
2022-08-14	Almería-Real Madrid	2022-08-14	16:00:00

league	season	game	home_team	\
ESP-La Liga	2223	2022-08-12 Osasuna-Sevilla	Osasuna	
		2022-08-13 Barcelona-Rayo Vallecano	Barcelona	
		2022-08-13 Celta Vigo-Espanyol	Celta Vigo	
		2022-08-13 Real Valladolid-Villarreal	Real Valladolid	
		2022-08-14 Almería-Real Madrid	Almería	

league	season	game	away_team	\
ESP-La Liga	2223	2022-08-12 Osasuna-Sevilla	Sevilla	
		2022-08-13 Barcelona-Rayo Vallecano	Rayo Vallecano	
		2022-08-13 Celta Vigo-Espanyol	Espanyol	
		2022-08-13 Real Valladolid-Villarreal	Villarreal	
		2022-08-14 Almería-Real Madrid	Real Madrid	

league	season	game	home_score	\
ESP-La Liga	2223	2022-08-12 Osasuna-Sevilla	2	
		2022-08-13 Barcelona-Rayo Vallecano	0	
		2022-08-13 Celta Vigo-Espanyol	2	
		2022-08-13 Real Valladolid-Villarreal	0	
		2022-08-14 Almería-Real Madrid	1	

league	season	game	away_score	game_id
ESP-La Liga	2223	2022-08-12 Osasuna-Sevilla	1	10408559
		2022-08-13 Barcelona-Rayo Vallecano	0	10408557
		2022-08-13 Celta Vigo-Espanyol	2	10408645
		2022-08-13 Real Valladolid-Villarreal	3	10408563
		2022-08-14 Almería-Real Madrid	2	10408712

[]:

..... doc/datasources/Sofascore.ipynb ends here.

The following section was generated from doc/datasources/SoFIFA.ipynb

[3]: import soccerdata as sd

2.6 SoFIFA

```
[4]: sofifa = sd.SoFIFA(leagues="ENG-Premier League", versions="latest")
print(sofifa.__doc__)
```

Provides pd.DataFrames from data at <http://sofifa.com>.

Data will be downloaded as necessary and cached locally in
``~/soccerdata/data/SoFIFA``.

Parameters

leagues : string or iterable, optional
IDs of leagues to include.

versions : string, int or list of int, optional
FIFA releases to include. Should be specified by their ID used in the URL
(e.g., 230034). Alternatively, the string "all" can be used to include all
versions and "latest" to include the latest version only. Defaults to
"latest".

proxy : 'tor' or dict or list(dict) or callable, optional
Use a proxy to hide your IP address. Valid options are:
- "tor": Uses the Tor network. Tor should be running in
the background on port 9050.
- dict: A dictionary with the proxy to use. The dict should be
a mapping of supported protocols to proxy addresses. For example::

```
{  
    'http': 'http://10.10.1.10:3128',  
    'https': 'http://10.10.1.10:1080',  
}
```

- list(dict): A list of proxies to choose from. A different proxy will
be selected from this list after failed requests, allowing rotating
proxies.

- callable: A function that returns a valid proxy. This function will
be called after failed requests, allowing rotating proxies.

no_cache : bool
If True, will not use cached data.

no_store : bool
If True, will not store downloaded data.

data_dir : Path
Path to directory where data will be cached.

2.6.1 EA Sports FIFA team ratings

```
[5]: team_ratings = sofiya.read_team_ratings()
team_ratings.head()
```

league	team	overall	attack	midfield	defence	\
ENG-Premier League	AFC Bournemouth	74	75	73	73	
	Arsenal	82	82	84	81	
	Aston Villa	79	79	77	79	
	Brentford	75	74	75	76	
	Brighton & Hove Albion	77	75	77	78	
league	team	transfer_budget	players	\		
ENG-Premier League	AFC Bournemouth	€0	29			
	Arsenal	€0	32			
	Aston Villa	€0	32			
	Brentford	€0	31			
	Brighton & Hove Albion	€0	28			
league	team	fifa_edition	update			
ENG-Premier League	AFC Bournemouth	FIFA 23	May 26, 2023			
	Arsenal	FIFA 23	May 26, 2023			
	Aston Villa	FIFA 23	May 26, 2023			
	Brentford	FIFA 23	May 26, 2023			
	Brighton & Hove Albion	FIFA 23	May 26, 2023			

2.6.2 EA Sports FIFA player ratings

```
[6]: player_ratings = sofiya.read_player_ratings(team="Arsenal")
player_ratings.head()
```

player	fifa_edition	update	overallrating	potential	\	
Aaron Ramsdale	FIFA 23	May 26, 2023	83	86		
Ainsley Maitland-Niles	FIFA 23	May 26, 2023	73	75		
Albert-Mbogo Sambi Lokonga	FIFA 23	May 26, 2023	75	82		
Amario Cozier-Duberry	FIFA 23	May 26, 2023	59	77		
Arthur Okonkwo	FIFA 23	May 26, 2023	66	75		
player	crossing	finishing	headingaccuracy	shortpassing	\	
Aaron Ramsdale	15	14	15	56		
Ainsley Maitland-Niles	67	58	54	74		
Albert-Mbogo Sambi Lokonga	68	53	61	79		
Amario Cozier-Duberry	58	57	38	55		
Arthur Okonkwo	10	6	12	33		
player	volleys	dribbling	curve	fk_accuracy	longpassing	\

(continues on next page)

(continued from previous page)

Aaron Ramsdale	16	18	17	16	57
Ainsley Maitland-Niles	44	71	64	50	66
Albert-Mbogo Sambi Lokonga	58	78	71	68	77
Amario Cozier-Duberry	60	60	53	41	44
Arthur Okonkwo	7	9	12	20	20
 player					
Aaron Ramsdale	36		47	50	53
Ainsley Maitland-Niles	73		79	77	79
Albert-Mbogo Sambi Lokonga	81		63	69	78
Amario Cozier-Duberry	58		72	69	73
Arthur Okonkwo	17		44	44	40
 player					
Aaron Ramsdale	82	49	65	66	35
Ainsley Maitland-Niles	70	69	67	66	69
Albert-Mbogo Sambi Lokonga	72	68	68	71	77
Amario Cozier-Duberry	50	68	56	44	55
Arthur Okonkwo	63	31	47	50	29
 player					
Aaron Ramsdale	59	15	35	24	
Ainsley Maitland-Niles	67	62	65	70	
Albert-Mbogo Sambi Lokonga	63	63	72	77	
Amario Cozier-Duberry	45	45	55	27	
Arthur Okonkwo	50	7	20	11	
 player					
Aaron Ramsdale	18	65	25	65	
Ainsley Maitland-Niles	69	67	69	70	
Albert-Mbogo Sambi Lokonga	65	75	41	80	
Amario Cozier-Duberry	59	56	55	55	
Arthur Okonkwo	6	35	19	30	
 player					
Aaron Ramsdale	20		16	15	
Ainsley Maitland-Niles	66		74	72	
Albert-Mbogo Sambi Lokonga	64		73	68	
Amario Cozier-Duberry	28		28	24	
Arthur Okonkwo	13		13	12	
 player					
Aaron Ramsdale	84	79	87	79	
Ainsley Maitland-Niles	15	9	7	10	
Albert-Mbogo Sambi Lokonga	14	7	6	13	
Amario Cozier-Duberry	11	12	9	13	

(continues on next page)

(continued from previous page)

Arthur Okonkwo	64	65	63	67
gk_reflexes				
player				
Aaron Ramsdale	86			
Ainsley Maitland-Niles	13			
Albert-Mbogo Sambi Lokonga	9			
Amario Cozier-Duberry	7			
Arthur Okonkwo	66			

[7]: `sofifa.read_teams()`

team_id	team	league	fifa_edition	\
10	Manchester City	ENG-Premier League	FIFA 23	
9	Liverpool	ENG-Premier League	FIFA 23	
1	Arsenal	ENG-Premier League	FIFA 23	
5	Chelsea	ENG-Premier League	FIFA 23	
11	Manchester United	ENG-Premier League	FIFA 23	
18	Tottenham Hotspur	ENG-Premier League	FIFA 23	
13	Newcastle United	ENG-Premier League	FIFA 23	
2	Aston Villa	ENG-Premier League	FIFA 23	
19	West Ham United	ENG-Premier League	FIFA 23	
95	Leicester City	ENG-Premier League	FIFA 23	
110	Wolverhampton Wanderers	ENG-Premier League	FIFA 23	
14	Nottingham Forest	ENG-Premier League	FIFA 23	
1808	Brighton & Hove Albion	ENG-Premier League	FIFA 23	
7	Everton	ENG-Premier League	FIFA 23	
1799	Crystal Palace	ENG-Premier League	FIFA 23	
144	Fulham	ENG-Premier League	FIFA 23	
8	Leeds United	ENG-Premier League	FIFA 23	
17	Southampton	ENG-Premier League	FIFA 23	
1925	Brentford	ENG-Premier League	FIFA 23	
1943	AFC Bournemouth	ENG-Premier League	FIFA 23	
update				
team_id				
10	May 26, 2023			
9	May 26, 2023			
1	May 26, 2023			
5	May 26, 2023			
11	May 26, 2023			
18	May 26, 2023			
13	May 26, 2023			
2	May 26, 2023			
19	May 26, 2023			
95	May 26, 2023			
110	May 26, 2023			
14	May 26, 2023			
1808	May 26, 2023			
7	May 26, 2023			
1799	May 26, 2023			

(continues on next page)

(continued from previous page)

144	May 26, 2023
8	May 26, 2023
17	May 26, 2023
1925	May 26, 2023
1943	May 26, 2023

[8]: `sofifa.read_players(team="Arsenal")`

player_id	player	team	league	\
233934	Aaron Ramsdale	Arsenal	ENG-Premier League	
231936	Benjamin White	Arsenal	ENG-Premier League	
243715	William Saliba	Arsenal	ENG-Premier League	
232580	Gabriel dos S. Magalhães	Arsenal	ENG-Premier League	
227813	Oleksandr Zinchenko	Arsenal	ENG-Premier League	
205498	Luiz Frello Filho Jorge	Arsenal	ENG-Premier League	
222665	Martin Ødegaard	Arsenal	ENG-Premier League	
199503	Granit Xhaka	Arsenal	ENG-Premier League	
246669	Bukayo Saka	Arsenal	ENG-Premier League	
230666	Gabriel Fernando de Jesus	Arsenal	ENG-Premier League	
251566	Gabriel Teodoro Martinelli Silva	Arsenal	ENG-Premier League	
207421	Leandro Trossard	Arsenal	ENG-Premier League	
236988	Eddie Nketiah	Arsenal	ENG-Premier League	
209989	Thomas Partey	Arsenal	ENG-Premier League	
256958	Fábio Daniel Ferreira Vieira	Arsenal	ENG-Premier League	
226491	Kieran Tierney	Arsenal	ENG-Premier League	
265693	Jakub Kiwior	Arsenal	ENG-Premier League	
233267	Matt Turner	Arsenal	ENG-Premier League	
228295	Rob Holding	Arsenal	ENG-Premier League	
240273	Emile Smith Rowe	Arsenal	ENG-Premier League	
231448	Reiss Nelson	Arsenal	ENG-Premier League	
211454	Mohamed Elneny	Arsenal	ENG-Premier League	
232938	Takehiro Tomiyasu	Arsenal	ENG-Premier League	
247731	Matthew Smith	Arsenal	ENG-Premier League	
248732	Karl Hein	Arsenal	ENG-Premier League	
259403	Joel Ideho	Arsenal	ENG-Premier League	
260141	Cătălin Cîrjan	Arsenal	ENG-Premier League	
259095	George Lewis	Arsenal	ENG-Premier League	
271743	Amario Cozier-Duberry	Arsenal	ENG-Premier League	
271808	Lino Sousa	Arsenal	ENG-Premier League	
259097	Salah-Eddine Oulad M'Hand	Arsenal	ENG-Premier League	
274387	Reuell Walters	Arsenal	ENG-Premier League	
270390	Marcus Vincius Oliveira Alencar	Arsenal	ENG-Premier League	
260653	Brooke Norton-Cuffy	Arsenal	ENG-Premier League	
260144	Charlie Patino	Arsenal	ENG-Premier League	
251530	Nuno Albertino Varela Tavares	Arsenal	ENG-Premier League	
247463	Folarin Balogun	Arsenal	ENG-Premier League	
246671	Arthur Okonkwo	Arsenal	ENG-Premier League	
241928	Albert Sambi Lokonga	Arsenal	ENG-Premier League	
235755	Auston Trusty	Arsenal	ENG-Premier League	
226110	Nicolas Pépé	Arsenal	ENG-Premier League	
225782	Ainsley Maitland-Niles	Arsenal	ENG-Premier League	

(continues on next page)

(continued from previous page)

222562	Rúnar Alex Rúnarsson	Arsenal	ENG-Premier League
206654	Pablo Mari Villar	Arsenal	ENG-Premier League
201118	Cédric Ricardo Alves Soares	Arsenal	ENG-Premier League

player_id	fifa_edition	update
233934	FIFA 23	May 26, 2023
231936	FIFA 23	May 26, 2023
243715	FIFA 23	May 26, 2023
232580	FIFA 23	May 26, 2023
227813	FIFA 23	May 26, 2023
205498	FIFA 23	May 26, 2023
222665	FIFA 23	May 26, 2023
199503	FIFA 23	May 26, 2023
246669	FIFA 23	May 26, 2023
230666	FIFA 23	May 26, 2023
251566	FIFA 23	May 26, 2023
207421	FIFA 23	May 26, 2023
236988	FIFA 23	May 26, 2023
209989	FIFA 23	May 26, 2023
256958	FIFA 23	May 26, 2023
226491	FIFA 23	May 26, 2023
265693	FIFA 23	May 26, 2023
233267	FIFA 23	May 26, 2023
228295	FIFA 23	May 26, 2023
240273	FIFA 23	May 26, 2023
231448	FIFA 23	May 26, 2023
211454	FIFA 23	May 26, 2023
232938	FIFA 23	May 26, 2023
247731	FIFA 23	May 26, 2023
248732	FIFA 23	May 26, 2023
259403	FIFA 23	May 26, 2023
260141	FIFA 23	May 26, 2023
259095	FIFA 23	May 26, 2023
271743	FIFA 23	May 26, 2023
271808	FIFA 23	May 26, 2023
259097	FIFA 23	May 26, 2023
274387	FIFA 23	May 26, 2023
270390	FIFA 23	May 26, 2023
260653	FIFA 23	May 26, 2023
260144	FIFA 23	May 26, 2023
251530	FIFA 23	May 26, 2023
247463	FIFA 23	May 26, 2023
246671	FIFA 23	May 26, 2023
241928	FIFA 23	May 26, 2023
235755	FIFA 23	May 26, 2023
226110	FIFA 23	May 26, 2023
225782	FIFA 23	May 26, 2023
222562	FIFA 23	May 26, 2023
206654	FIFA 23	May 26, 2023
201118	FIFA 23	May 26, 2023

..... doc/datasources/SoFIFA.ipynb ends here.

The following section was generated from doc/datasources/Understat.ipynb

[2]: import soccerdata as sd

2.7 Understat

[3]: understat = sd.Understat(leagues="ENG-Premier League", seasons="2015/2016")
print(understat.__doc__)Provides pd.DataFrames from data at <https://understat.com>.

Data will be downloaded as necessary and cached locally in
`~/.soccerdata/data/Understat`.

Parameters

proxy : 'tor' or dict or list(dict) or callable, optional
Use a proxy to hide your IP address. Valid options are:
- "tor": Uses the Tor network. Tor should be running in
the background on port 9050.
- dict: A dictionary with the proxy to use. The dict should be
a mapping of supported protocols to proxy addresses. For example::

```
{  
    'http': 'http://10.10.1.10:3128',  
    'https': 'http://10.10.1.10:1080',  
}
```

- list(dict): A list of proxies to choose from. A different proxy will
be selected from this list after failed requests, allowing rotating
proxies.
- callable: A function that returns a valid proxy. This function will
be called after failed requests, allowing rotating proxies.

no_cache : bool
If True, will not use cached data.

no_store : bool
If True, will not store downloaded data.

data_dir : Path
Path to directory where data will be cached.

2.7.1 Leagues

[4]: leagues = understat.read_leagues()
leagues.head()[4]:

league	league_id	url
ENG-Premier League	1	https://understat.com/league/EPL

2.7.2 Seasons

```
[5]: seasons = understat.read_seasons()
seasons.head()

[5]:
          league_id  season_id \
league      season
ENG-Premier League 1516           1       2015

                                url
league      season
ENG-Premier League 1516 https://understat.com/league/EPL/2015
```

2.7.3 Schedule

```
[6]: schedule = understat.read_schedule()
schedule.head()

[6]:
          league_id \
league      season game
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa           1
                           2015-08-08 Chelsea-Swansea             1
                           2015-08-08 Everton-Watford            1
                           2015-08-08 Leicester-Sunderland        1
                           2015-08-08 Manchester United-Tottenham   1

                                season_id \
league      season game
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa       2015
                           2015-08-08 Chelsea-Swansea            2015
                           2015-08-08 Everton-Watford           2015
                           2015-08-08 Leicester-Sunderland        2015
                           2015-08-08 Manchester United-Tottenham   2015

                                game_id \
league      season game
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa         82
                           2015-08-08 Chelsea-Swansea           86
                           2015-08-08 Everton-Watford          83
                           2015-08-08 Leicester-Sunderland       84
                           2015-08-08 Manchester United-Tottenham   81

                                date \
league      season game
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa 2015-08-08 18:00:00
                           2015-08-08 Chelsea-Swansea 2015-08-08 20:30:00
                           2015-08-08 Everton-Watford 2015-08-08 18:00:00
                           2015-08-08 Leicester-Sunderland 2015-08-08 18:00:00
                           2015-08-08 Manchester United-Tottenham 2015-08-08 15:45:00

                                home_team_id \
league      season game

```

(continues on next page)

(continued from previous page)

league	season	game	away_team_id	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	73	
		2015-08-08 Chelsea-Swansea	80	
		2015-08-08 Everton-Watford	72	
		2015-08-08 Leicester-Sunderland	75	
		2015-08-08 Manchester United-Tottenham	89	
league	season	game	home_team	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	Bournemouth	
		2015-08-08 Chelsea-Swansea	Chelsea	
		2015-08-08 Everton-Watford	Everton	
		2015-08-08 Leicester-Sunderland	Leicester	
		2015-08-08 Manchester United-Tottenham	Manchester United	
league	season	game	away_team	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	Aston Villa	
		2015-08-08 Chelsea-Swansea	Swansea	
		2015-08-08 Everton-Watford	Watford	
		2015-08-08 Leicester-Sunderland	Sunderland	
		2015-08-08 Manchester United-Tottenham	Tottenham	
league	season	game	away_team_code	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	AVL	
		2015-08-08 Chelsea-Swansea	SWA	
		2015-08-08 Everton-Watford	WAT	
		2015-08-08 Leicester-Sunderland	SUN	
		2015-08-08 Manchester United-Tottenham	TOT	
league	season	game	home_team_code	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	BOU	
		2015-08-08 Chelsea-Swansea	CHE	
		2015-08-08 Everton-Watford	EVE	
		2015-08-08 Leicester-Sunderland	LEI	
		2015-08-08 Manchester United-Tottenham	MUN	
league	season	game	home_goals	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	0	
		2015-08-08 Chelsea-Swansea	2	
		2015-08-08 Everton-Watford	2	
		2015-08-08 Leicester-Sunderland	4	

(continues on next page)

(continued from previous page)

			2015-08-08 Manchester United-Tottenham	1
				away_goals \
league	season	game		
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa		1
		2015-08-08 Chelsea-Swansea		2
		2015-08-08 Everton-Watford		2
		2015-08-08 Leicester-Sunderland		2
		2015-08-08 Manchester United-Tottenham		0
				home_xg \
league	season	game		
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	0.876106	
		2015-08-08 Chelsea-Swansea	0.64396	
		2015-08-08 Everton-Watford	0.604226	
		2015-08-08 Leicester-Sunderland	2.56803	
		2015-08-08 Manchester United-Tottenham	0.627539	
				away_xg \
league	season	game		
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	0.782253	
		2015-08-08 Chelsea-Swansea	2.59203	
		2015-08-08 Everton-Watford	0.557892	
		2015-08-08 Leicester-Sunderland	1.45946	
		2015-08-08 Manchester United-Tottenham	0.6746	
				is_result \
league	season	game		
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	True	
		2015-08-08 Chelsea-Swansea	True	
		2015-08-08 Everton-Watford	True	
		2015-08-08 Leicester-Sunderland	True	
		2015-08-08 Manchester United-Tottenham	True	
				has_data \
league	season	game		
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	True	
		2015-08-08 Chelsea-Swansea	True	
		2015-08-08 Everton-Watford	True	
		2015-08-08 Leicester-Sunderland	True	
		2015-08-08 Manchester United-Tottenham	True	
				url
league	season	game		
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	https://understat.com/	
↳match/82		2015-08-08 Chelsea-Swansea	https://understat.com/	
↳match/86		2015-08-08 Everton-Watford	https://understat.com/	
↳match/83		2015-08-08 Leicester-Sunderland	https://understat.com/	

(continues on next page)

(continued from previous page)

↳match/84

2015-08-08 Manchester United-Tottenham <https://understat.com/>

↳match/81

2.7.4 Team match stats

[7]: team_match_stats = understat.read_team_match_stats()
team_match_stats.head()

league	season	game	league_id	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	1	
		2015-08-08 Chelsea-Swansea	1	
		2015-08-08 Everton-Watford	1	
		2015-08-08 Leicester-Sunderland	1	
		2015-08-08 Manchester United-Tottenham	1	

league	season	game	season_id	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	2015	
		2015-08-08 Chelsea-Swansea	2015	
		2015-08-08 Everton-Watford	2015	
		2015-08-08 Leicester-Sunderland	2015	
		2015-08-08 Manchester United-Tottenham	2015	

league	season	game	game_id	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	82	
		2015-08-08 Chelsea-Swansea	86	
		2015-08-08 Everton-Watford	83	
		2015-08-08 Leicester-Sunderland	84	
		2015-08-08 Manchester United-Tottenham	81	

league	season	game	date	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	2015-08-08 18:00:00	
		2015-08-08 Chelsea-Swansea	2015-08-08 20:30:00	
		2015-08-08 Everton-Watford	2015-08-08 18:00:00	
		2015-08-08 Leicester-Sunderland	2015-08-08 18:00:00	
		2015-08-08 Manchester United-Tottenham	2015-08-08 15:45:00	

league	season	game	home_team_id	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	73	
		2015-08-08 Chelsea-Swansea	80	
		2015-08-08 Everton-Watford	72	
		2015-08-08 Leicester-Sunderland	75	
		2015-08-08 Manchester United-Tottenham	89	

league	season	game	away_team_id	\

(continues on next page)

(continued from previous page)

ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	71
		2015-08-08	Chelsea-Swansea	84
		2015-08-08	Everton-Watford	90
		2015-08-08	Leicester-Sunderland	77
		2015-08-08	Manchester United-Tottenham	82
				home_team \
league	season	game		
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Bournemouth
		2015-08-08	Chelsea-Swansea	Chelsea
		2015-08-08	Everton-Watford	Everton
		2015-08-08	Leicester-Sunderland	Leicester
		2015-08-08	Manchester United-Tottenham	Manchester United
				away_team \
league	season	game		
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa
		2015-08-08	Chelsea-Swansea	Swansea
		2015-08-08	Everton-Watford	Watford
		2015-08-08	Leicester-Sunderland	Sunderland
		2015-08-08	Manchester United-Tottenham	Tottenham
				away_team_code \
league	season	game		
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	AVL
		2015-08-08	Chelsea-Swansea	SWA
		2015-08-08	Everton-Watford	WAT
		2015-08-08	Leicester-Sunderland	SUN
		2015-08-08	Manchester United-Tottenham	TOT
				home_team_code \
league	season	game		
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	BOU
		2015-08-08	Chelsea-Swansea	CHE
		2015-08-08	Everton-Watford	EVE
		2015-08-08	Leicester-Sunderland	LEI
		2015-08-08	Manchester United-Tottenham	MUN
				... \
league	season	game		...
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	...
		2015-08-08	Chelsea-Swansea	...
		2015-08-08	Everton-Watford	...
		2015-08-08	Leicester-Sunderland	...
		2015-08-08	Manchester United-Tottenham	...
				away_ppda \
league	season	game		
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	11.846154
		2015-08-08	Chelsea-Swansea	8.833333
		2015-08-08	Everton-Watford	17.157895
		2015-08-08	Leicester-Sunderland	9.555556

(continues on next page)

(continued from previous page)

2015-08-08 Manchester United-Tottenham 8.21875

away_deep_completions ↴

league	season	game	
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	2
		2015-08-08 Chelsea-Swansea	5
		2015-08-08 Everton-Watford	4
		2015-08-08 Leicester-Sunderland	6
		2015-08-08 Manchester United-Tottenham	10

home_points \

league	season	game	
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	0
		2015-08-08 Chelsea-Swansea	1
		2015-08-08 Everton-Watford	1
		2015-08-08 Leicester-Sunderland	3
		2015-08-08 Manchester United-Tottenham	3

home_expected_points \

league	season	game	
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	1.3912
		2015-08-08 Chelsea-Swansea	0.1836
		2015-08-08 Everton-Watford	1.3301
		2015-08-08 Leicester-Sunderland	2.1323
		2015-08-08 Manchester United-Tottenham	1.2482

home_goals \

league	season	game	
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	0
		2015-08-08 Chelsea-Swansea	2
		2015-08-08 Everton-Watford	2
		2015-08-08 Leicester-Sunderland	4
		2015-08-08 Manchester United-Tottenham	1

home_xg \

league	season	game	
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	0.876106
		2015-08-08 Chelsea-Swansea	0.64396
		2015-08-08 Everton-Watford	0.604226
		2015-08-08 Leicester-Sunderland	2.56803
		2015-08-08 Manchester United-Tottenham	0.627539

home_np_xg \

league	season	game	
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	0.876106
		2015-08-08 Chelsea-Swansea	0.64396
		2015-08-08 Everton-Watford	0.604226
		2015-08-08 Leicester-Sunderland	1.80686
		2015-08-08 Manchester United-Tottenham	0.627539

home_np_xg_difference ↴

(continues on next page)

(continued from previous page)

league	season	game	home_ppda	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	0.093853	
		2015-08-08 Chelsea-Swansea	-1.1869	
		2015-08-08 Everton-Watford	0.046334	
		2015-08-08 Leicester-Sunderland	0.3474	
		2015-08-08 Manchester United-Tottenham	-0.047061	
league	season	game	home_ppda	\
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	6.9	
		2015-08-08 Chelsea-Swansea	10.363636	
		2015-08-08 Everton-Watford	6.65	
		2015-08-08 Leicester-Sunderland	10.88	
		2015-08-08 Manchester United-Tottenham	13.826087	
league	season	game	home_deep_completions	
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	11	
		2015-08-08 Chelsea-Swansea	10	
		2015-08-08 Everton-Watford	5	
		2015-08-08 Leicester-Sunderland	5	
		2015-08-08 Manchester United-Tottenham	4	

[5 rows x 26 columns]

2.7.5 Player season stats

```
[8]: player_season_stats = understat.read_player_season_stats()
player_season_stats.head()
```

league	season	player	league_id	season_id	team	\
ENG-Premier League	1516	Aaron Cresswell	1	2015	West Ham	
		Aaron Lennon	1	2015	Everton	
		Aaron Ramsey	1	2015	Arsenal	
		Abdul Rahman Baba	1	2015	Chelsea	
		Adam Bogdan	1	2015	Liverpool	
league	season	player	team_id	player_id	position	\
ENG-Premier League	1516	Aaron Cresswell	81	534	D M	
		Aaron Lennon	72	593	F M S	
		Aaron Ramsey	83	504	M S	
		Abdul Rahman Baba	80	684	D S	
		Adam Bogdan	87	698	GK	
league	season	player	matches	minutes	goals	\
ENG-Premier League	1516	Aaron Cresswell	37	3319	2	
		Aaron Lennon	25	1534	5	

(continues on next page)

(continued from previous page)

			xg	np_goals	np_xg	\
league	season	player				
ENG-Premier League	1516	Aaron Cresswell	1.092563	2	1.092563	
		Aaron Lennon	2.226095	5	2.226095	
		Aaron Ramsey	8.614323	5	8.614323	
		Abdul Rahman Baba	0.379115	0	0.379115	
		Adam Bogdan	0.0	0	0.0	
			assists	xa	shots	\
league	season	player				
ENG-Premier League	1516	Aaron Cresswell	4	3.294618	28	
		Aaron Lennon	0	1.697951	18	
		Aaron Ramsey	4	4.046242	69	
		Abdul Rahman Baba	1	0.179241	6	
		Adam Bogdan	0	0.0	0	
			key_passes	yellow_cards		\
league	season	player				
ENG-Premier League	1516	Aaron Cresswell	40	1		
		Aaron Lennon	23	1		
		Aaron Ramsey	37	4		
		Abdul Rahman Baba	3	1		
		Adam Bogdan	0	0		
			red_cards	xg_chain	xg_buildup	
league	season	player				
ENG-Premier League	1516	Aaron Cresswell	0	10.339367	6.831139	
		Aaron Lennon	0	5.535392	1.90374	
		Aaron Ramsey	0	23.282566	14.176444	
		Abdul Rahman Baba	0	3.869669	3.490554	
		Adam Bogdan	0	0.120216	0.120216	

2.7.6 Player match stats

[9]:	player_match_stats = understat.read_player_match_stats() player_match_stats.head()
[9]:	<pre> ↵ league_id \ league season game team player ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Ashley Westwood ↵ 1 Brad Guzan ↵ 1 Carlos Sánchez ↵ 1 Ciaran Clark </pre>

(continues on next page)

(continued from previous page)

↳	league	season_id	team	player
↳	ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	Aston Villa Ashley Westwood
↳		2015		
↳		2015		Brad Guzan
↳		2015		Carlos Sánchez
↳		2015		Ciaran Clark
↳		2015		Gabriel
↳	Agbonlahor	2015		
↳	game_id	team	player	
↳	league	season game		
↳	ENG-Premier League	1516 2015-08-08 Bournemouth-Aston Villa	Aston Villa Ashley Westwood	
↳		82		Brad Guzan
↳		82		Carlos Sánchez
↳		82		Ciaran Clark
↳		82		Gabriel
↳	Agbonlahor	82		
↳	team_id	team	player	
↳	league	season game		
↳	ENG-Premier League	1516 2015-08-08 Bournemouth-Aston Villa	Aston Villa Ashley Westwood	
↳		71		Brad Guzan
↳		71		Carlos Sánchez
↳		71		Ciaran Clark
↳		71		Gabriel
↳	Agbonlahor	71		
↳	player_id	team	player	
↳	league	season game		
↳	ENG-Premier League	1516 2015-08-08 Bournemouth-Aston Villa	Aston Villa Ashley Westwood	
↳		669		Brad Guzan

(continues on next page)

(continued from previous page)

↔	662		Carlos Sánchez	✉
↔	667		Ciaran Clark	✉
↔	875		Gabriel	✉
↔ Agbonlahor	890			
↔ position \				✉
league	season	game	team	player
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa
↔ MC			Ashley Westwood	✉
↔ GK			Brad Guzan	✉
↔ Sub			Carlos Sánchez	✉
↔ DC			Ciaran Clark	✉
↔ Agbonlahor	FW		Gabriel	✉
↔ position_id \				✉
league	season	game	team	player
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa
↔ 9			Ashley Westwood	✉
↔ 1			Brad Guzan	✉
↔ 17			Carlos Sánchez	✉
↔ 3			Ciaran Clark	✉
↔ Agbonlahor	15		Gabriel	✉
↔ minutes \				✉
league	season	game	team	player
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa
↔ 90			Ashley Westwood	✉
↔ 90			Brad Guzan	✉
↔ 18			Carlos Sánchez	✉
↔ 90			Ciaran Clark	✉
↔ Agbonlahor	90		Gabriel	✉
↔ goals \				✉

(continues on next page)

(continued from previous page)

league	season	game	team	player
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa
←	0		Ashley Westwood	...
←	0		Brad Guzan	...
←	0		Carlos Sánchez	...
←	0		Ciaran Clark	...
←	0		Gabriel	...
← Agbonlahor	0			
own_goals \				
league	season	game	team	player
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa
←	0		Ashley Westwood	...
←	0		Brad Guzan	...
←	0		Carlos Sánchez	...
←	0		Ciaran Clark	...
←	0		Gabriel	...
← Agbonlahor	0			
shots \				
league	season	game	team	player
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa
←	0		Ashley Westwood	...
←	0		Brad Guzan	...
←	0		Carlos Sánchez	...
←	0		Ciaran Clark	...
←	0		Gabriel	...
← Agbonlahor	2			
xg \				
league	season	game	team	player
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa
←	0.0		Ashley Westwood	...
←	0.0		Brad Guzan	...
←	0.0		Carlos Sánchez	...
←	0.0		Ciaran Clark	...
←	0.0		Gabriel	...

(continues on next page)

2.7.7 Shot events

```
[10]: shot_events = understat.read_shot_events()
shot_events.head()

[10]:
    ↵   league_id \
league          season game                      team      player
ENG-Premier League 1516  2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
    ↵Agbonlahor        1
                                         Gabriel
    ↵Agbonlahor        1
                                         Gabriel
    ↵                1
                                         Idrissa Gueye
    ↵                1
                                         Idrissa Gueye
    ↵                1
                                         Leandro Bacuna
    ↵                1

    ↵   season_id \
league          season game                      team      player
ENG-Premier League 1516  2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
    ↵Agbonlahor        2015
                                         Gabriel
    ↵Agbonlahor        2015
                                         Idrissa Gueye
    ↵                2015
                                         Idrissa Gueye
    ↵                2015
                                         Leandro Bacuna
    ↵                2015

    ↵   game_id \
league          season game                      team      player
ENG-Premier League 1516  2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
    ↵Agbonlahor        82
                                         Gabriel
    ↵Agbonlahor        82
                                         Idrissa Gueye
    ↵                82
                                         Idrissa Gueye
    ↵                82
                                         Leandro Bacuna
    ↵                82

    ↵   date \
league          season game                      team      player
ENG-Premier League 1516  2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
    ↵Agbonlahor 2015-08-08 18:00:00
                                         Gabriel
```

(continues on next page)

(continued from previous page)

```

↳ Agbonlahor 2015-08-08 18:00:00
    ↳ 2015-08-08 18:00:00 Idrissa Gueye ▾
    ↳ 2015-08-08 18:00:00 Idrissa Gueye ▾
    ↳ 2015-08-08 18:00:00 Leandro Bacuna ▾
    ↳ 2015-08-08 18:00:00 ▾

    ↳ shot_id \
league          season game      team      player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel ▾
↳ Agbonlahor 487146
    ↳ 487146 Gabriel ▾
↳ Agbonlahor 487150
    ↳ 487136 Idrissa Gueye ▾
    ↳ 487144 Idrissa Gueye ▾
    ↳ 487141 Leandro Bacuna ▾
    ↳ 487141 ▾

    ↳ team_id \
league          season game      team      player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel ▾
↳ Agbonlahor 71
    ↳ 71 Gabriel ▾
↳ Agbonlahor 71
    ↳ 71 Idrissa Gueye ▾
    ↳ 71 Idrissa Gueye ▾
    ↳ 71 Leandro Bacuna ▾
    ↳ 71 ▾

    ↳ player_id \
league          season game      team      player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel ▾
↳ Agbonlahor 890
    ↳ 890 Gabriel ▾
↳ Agbonlahor 890
    ↳ 668 Idrissa Gueye ▾
    ↳ 668 Idrissa Gueye ▾
    ↳ 674 Leandro Bacuna ▾
    ↳ 674 ▾

    ↳ assist_player_id \

```

(continues on next page)

(continued from previous page)

```
league      season game      team      player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
↪Agbonlahor          <NA>
↪Agbonlahor           548651
↪
↪      548653
↪      548657
↪      <NA>

↪      assist_player \
league      season game      team      player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
↪Agbonlahor          <NA>
↪Agbonlahor          Micah Richards
↪
↪      Jordan Veretout
↪      Gabriel Agbonlahor
↪      <NA>

↪      xg \
league      season game      team      player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
↪Agbonlahor  0.075484
↪Agbonlahor  0.054676
↪
↪      0.018269
↪      0.113668
↪      0.097371

↪      location_x \
league      season game      team      player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
↪Agbonlahor  0.86
↪Agbonlahor  0.795
↪
↪      0.742
↪      0.883
```

(continues on next page)

(continued from previous page)

↳	location_y	\
league	season	game
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
↳ Agbonlahor	0.503	Gabriel
↳ Agbonlahor	0.507	Gabriel
↳	0.598	Idrissa Gueye
↳	0.561	Idrissa Gueye
↳	0.437	Leandro Bacuna
↳	minute	\
league	season	game
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
↳ Agbonlahor	67	Gabriel
↳ Agbonlahor	81	Idrissa Gueye
↳	26	Idrissa Gueye
↳	47	Leandro Bacuna
↳	42	
↳	body_part	\
league	season	game
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
↳ Agbonlahor	Left Foot	Gabriel
↳ Agbonlahor	Left Foot	Idrissa Gueye
↳	Right Foot	Idrissa Gueye
↳	Left Foot	Leandro Bacuna
↳	Right Foot	
↳	situation	\
league	season	game
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
↳ Agbonlahor	Open Play	Gabriel
↳ Agbonlahor	Open Play	Idrissa Gueye

(continues on next page)

(continued from previous page)

↳	Open Play	Idrissa Gueye
↳	Open Play	Leandro Bacuna
↳	Direct Freekick	
 ↳ result		
league season game team player		
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel		
↳	Agbonlahor Blocked Shot	Gabriel
↳	Agbonlahor Saved Shot	Idrissa Gueye
↳	Missed Shot	Idrissa Gueye
↳	Saved Shot	Leandro Bacuna
↳	Blocked Shot	

..... doc/datasources/Understat.ipynb ends here.

The following section was generated from doc/datasources/WhoScored.ipynb

[3]: import soccerdata as sd

2.8 WhoScored

[4]: ws = sd.WhoScored(leagues="ENG-Premier League", seasons=2021)
print(ws.__doc__)

```
/cw/dtaijupiter/NoCsBack/dtai/pieterr/Projects/soccerdata/soccerdata/_common.py:493:  
↳UserWarning: Season id "2021" is ambiguous: interpreting as "20-21"  
    warnings.warn(msg, stacklevel=1)
```

Provides pd.DataFrames from data available at <http://whoscored.com>.

Data will be downloaded as necessary and cached locally in
`~/.soccerdata/data/WhoScored`.

Parameters

leagues : string or iterable, optional
IDs of Leagues to include.
seasons : string, int or list, optional
Seasons to include. Supports multiple formats.
Examples: '16-17'; 2016; '2016-17'; [14, 15, 16]
proxy : 'tor' or dict or list(dict) or callable, optional
Use a proxy to hide your IP address. Valid options are:
- "tor": Uses the Tor network. Tor should be running in
the background on port 9050.

(continues on next page)

(continued from previous page)

- dict: A dictionary with the proxy to use. The dict should be a mapping of supported protocols to proxy addresses. For example::

```
{
    'http': 'http://10.10.1.10:3128',
    'https': 'http://10.10.1.10:1080',
}

- list(dict): A list of proxies to choose from. A different proxy will be selected from this list after failed requests, allowing rotating proxies.
- callable: A function that returns a valid proxy. This function will be called after failed requests, allowing rotating proxies.

no_cache : bool
    If True, will not use cached data.
no_store : bool
    If True, will not store downloaded data.
data_dir : Path
    Path to directory where data will be cached.
path_to_browser : Path, optional
    Path to the Chrome executable.
headless : bool, default: True
    If True, will run Chrome in headless mode. Setting this to False might help to avoid getting blocked. Only supported for Selenium <4.13.
```

2.8.1 Game schedule

[5]: epl_schedule = ws.read_schedule()
epl_schedule.head()

league	season	game	stage_id	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	18685	
		2020-09-12 Fulham-Arsenal	18685	
		2020-09-12 Liverpool-Leeds United	18685	
		2020-09-12 West Ham United-Newcastle	18685	
		2020-09-13 Tottenham-Everton	18685	
				game_id \
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	1485186	
		2020-09-12 Fulham-Arsenal	1485187	
		2020-09-12 Liverpool-Leeds United	1485188	
		2020-09-12 West Ham United-Newcastle	1485191	
		2020-09-13 Tottenham-Everton	1485189	
				status \
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	6	
		2020-09-12 Fulham-Arsenal	6	

(continues on next page)

(continued from previous page)

2020-09-12	Liverpool-Leeds United	6
2020-09-12	West Ham United-Newcastle	6
2020-09-13	Tottenham-Everton	6

league	season	game	start_time	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	2020-09-12T15:00:00	
		2020-09-12 Fulham-Arsenal	2020-09-12T12:30:00	
		2020-09-12 Liverpool-Leeds United	2020-09-12T17:30:00	
		2020-09-12 West Ham United-Newcastle	2020-09-12T20:00:00	
		2020-09-13 Tottenham-Everton	2020-09-13T16:30:00	
league	season	game	home_team_id	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	162	
		2020-09-12 Fulham-Arsenal	170	
		2020-09-12 Liverpool-Leeds United	26	
		2020-09-12 West Ham United-Newcastle	29	
		2020-09-13 Tottenham-Everton	30	
league	season	game	home_team	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Crystal Palace	
		2020-09-12 Fulham-Arsenal	Fulham	
		2020-09-12 Liverpool-Leeds United	Liverpool	
		2020-09-12 West Ham United-Newcastle	West Ham United	
		2020-09-13 Tottenham-Everton	Tottenham	
league	season	game	home_yellow_cards	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	2	
		2020-09-12 Fulham-Arsenal	2	
		2020-09-12 Liverpool-Leeds United	1	
		2020-09-12 West Ham United-Newcastle	2	
		2020-09-13 Tottenham-Everton	1	
league	season	game	home_red_cards	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	0	
		2020-09-12 Fulham-Arsenal	0	
		2020-09-12 Liverpool-Leeds United	0	
		2020-09-12 West Ham United-Newcastle	0	
		2020-09-13 Tottenham-Everton	0	
league	season	game	away_team_id	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	18	
		2020-09-12 Fulham-Arsenal	13	
		2020-09-12 Liverpool-Leeds United	19	
		2020-09-12 West Ham United-Newcastle	23	
		2020-09-13 Tottenham-Everton	31	

(continues on next page)

(continued from previous page)

league	season	game	away_team	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	Southampton	
		2020-09-12 Fulham-Arsenal	Arsenal	
		2020-09-12 Liverpool-Leeds United	Leeds United	
		2020-09-12 West Ham United-Newcastle	Newcastle	
		2020-09-13 Tottenham-Everton	Everton	
league	season	game	away_yellow_cards	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	1	
		2020-09-12 Fulham-Arsenal	2	
		2020-09-12 Liverpool-Leeds United	0	
		2020-09-12 West Ham United-Newcastle	2	
		2020-09-13 Tottenham-Everton	0	
league	season	game	away_red_cards	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	0	
		2020-09-12 Fulham-Arsenal	0	
		2020-09-12 Liverpool-Leeds United	0	
		2020-09-12 West Ham United-Newcastle	0	
		2020-09-13 Tottenham-Everton	0	
league	season	game	has_incidents_summary	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	True	
		2020-09-12 Fulham-Arsenal	True	
		2020-09-12 Liverpool-Leeds United	True	
		2020-09-12 West Ham United-Newcastle	True	
		2020-09-13 Tottenham-Everton	True	
league	season	game	has_preview	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	True	
		2020-09-12 Fulham-Arsenal	True	
		2020-09-12 Liverpool-Leeds United	True	
		2020-09-12 West Ham United-Newcastle	True	
		2020-09-13 Tottenham-Everton	True	
league	season	game	score_changed_at	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	2020-09-12 15:14:31Z	
		2020-09-12 Fulham-Arsenal	2020-09-12 13:48:13Z	
		2020-09-12 Liverpool-Leeds United	2020-09-12 19:15:39Z	
		2020-09-12 West Ham United-Newcastle	2020-09-12 21:45:39Z	
		2020-09-13 Tottenham-Everton	2020-09-13 17:41:16Z	
league	season	game	elapsed	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	FT	
		2020-09-12 Fulham-Arsenal	FT	

(continues on next page)

(continued from previous page)

			FT
			FT
			FT
last_scorer \			
league	season	game	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	0.0
		2020-09-12 Fulham-Arsenal	1.0
		2020-09-12 Liverpool-Leeds United	0.0
		2020-09-12 West Ham United-Newcastle	1.0
		2020-09-13 Tottenham-Everton	1.0
is_top_match \			
league	season	game	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	False
		2020-09-12 Fulham-Arsenal	True
		2020-09-12 Liverpool-Leeds United	True
		2020-09-12 West Ham United-Newcastle	False
		2020-09-13 Tottenham-Everton	True
home_team_country_code \			
league	season	game	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	gb-eng
		2020-09-12 Fulham-Arsenal	gb-eng
		2020-09-12 Liverpool-Leeds United	gb-eng
		2020-09-12 West Ham United-Newcastle	gb-eng
		2020-09-13 Tottenham-Everton	gb-eng
away_team_country_code \			
league	season	game	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	gb-eng
		2020-09-12 Fulham-Arsenal	gb-eng
		2020-09-12 Liverpool-Leeds United	gb-eng
		2020-09-12 West Ham United-Newcastle	gb-eng
		2020-09-13 Tottenham-Everton	gb-eng
comment_count \			
league	season	game	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	6
		2020-09-12 Fulham-Arsenal	15
		2020-09-12 Liverpool-Leeds United	61
		2020-09-12 West Ham United-Newcastle	10
		2020-09-13 Tottenham-Everton	32
is_lineup_confirmed \			
league	season	game	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	True
		2020-09-12 Fulham-Arsenal	True
		2020-09-12 Liverpool-Leeds United	True
		2020-09-12 West Ham United-Newcastle	True
		2020-09-13 Tottenham-Everton	True

(continues on next page)

(continued from previous page)

league	season	game	is_stream_available \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	False
		2020-09-12 Fulham-Arsenal	False
		2020-09-12 Liverpool-Leeds United	False
		2020-09-12 West Ham United-Newcastle	False
		2020-09-13 Tottenham-Everton	False
league	season	game	match_is_opta \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	False
		2020-09-12 Fulham-Arsenal	False
		2020-09-12 Liverpool-Leeds United	False
		2020-09-12 West Ham United-Newcastle	False
		2020-09-13 Tottenham-Everton	False
league	season	game	home_team_country_name \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	England
		2020-09-12 Fulham-Arsenal	England
		2020-09-12 Liverpool-Leeds United	England
		2020-09-12 West Ham United-Newcastle	England
		2020-09-13 Tottenham-Everton	England
league	season	game	away_team_country_name \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	England
		2020-09-12 Fulham-Arsenal	England
		2020-09-12 Liverpool-Leeds United	England
		2020-09-12 West Ham United-Newcastle	England
		2020-09-13 Tottenham-Everton	England
date \			
league	season	game	
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	2020-09-12 14:00:00+00:00
		2020-09-12 Fulham-Arsenal	2020-09-12 11:30:00+00:00
		2020-09-12 Liverpool-Leeds United	2020-09-12 16:30:00+00:00
		2020-09-12 West Ham United-Newcastle	2020-09-12 19:00:00+00:00
		2020-09-13 Tottenham-Everton	2020-09-13 15:30:00+00:00
league	season	game	home_score \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	1
		2020-09-12 Fulham-Arsenal	0
		2020-09-12 Liverpool-Leeds United	4
		2020-09-12 West Ham United-Newcastle	0
		2020-09-13 Tottenham-Everton	0
league	season	game	away_score \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	0

(continues on next page)

(continued from previous page)

2020-09-12	Fulham-Arsenal	3
2020-09-12	Liverpool-Leeds United	3
2020-09-12	West Ham United-Newcastle	2
2020-09-13	Tottenham-Everton	1

league	season	game	incidents \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	[{'minute': '13', 'type': 'red', 'subType': 1, 'player': 'Fulham', 'team': 'Fulham'}, {'minute': '8', 'type': 'yellow', 'subType': 1, 'player': 'Arsenal', 'team': 'Arsenal'}]
		2020-09-12 Fulham-Arsenal	[{'minute': '13', 'type': 'red', 'subType': 1, 'player': 'Arsenal', 'team': 'Arsenal'}, {"minute": "8", "type": "yellow", "subType": 1, "player": "Fulham", "team": "Fulham"}]
		2020-09-12 Liverpool-Leeds United	[{'minute': '4', 'type': 'red', 'subType': 1, 'player': 'Liverpool', 'team': 'Liverpool'}, {"minute": "4", "type": "yellow", "subType": 1, "player": "Leeds United", "team": "Leeds United"}]
		2020-09-12 West Ham United-Newcastle	[{'minute': '56', 'type': 'red', 'subType': 1, 'player': 'West Ham United', 'team': 'West Ham United'}, {"minute": "56", "type": "yellow", "subType": 1, "player": "Newcastle", "team": "Newcastle"}]
		2020-09-13 Tottenham-Everton	[{'minute': '55', 'type': 'red', 'subType': 1, 'player': 'Tottenham', 'team': 'Tottenham'}, {"minute": "55", "type": "yellow", "subType": 1, "player": "Everton", "team": "Everton"}]
league	season	game	bets \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	None
		2020-09-12 Fulham-Arsenal	None
		2020-09-12 Liverpool-Leeds United	None
		2020-09-12 West Ham United-Newcastle	None
		2020-09-13 Tottenham-Everton	None
league	season	game	aggregate_winner_field \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	None
		2020-09-12 Fulham-Arsenal	None
		2020-09-12 Liverpool-Leeds United	None
		2020-09-12 West Ham United-Newcastle	None
		2020-09-13 Tottenham-Everton	None
league	season	game	winner_field \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	0.0
		2020-09-12 Fulham-Arsenal	1.0
		2020-09-12 Liverpool-Leeds United	0.0
		2020-09-12 West Ham United-Newcastle	1.0
		2020-09-13 Tottenham-Everton	1.0
league	season	game	period \
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	7
		2020-09-12 Fulham-Arsenal	7
		2020-09-12 Liverpool-Leeds United	7
		2020-09-12 West Ham United-Newcastle	7
		2020-09-13 Tottenham-Everton	7
league	season	game	extra_result_field \

(continues on next page)

(continued from previous page)

league	season	game	home_extratime_score	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	None	
		2020-09-12 Fulham-Arsenal	None	
		2020-09-12 Liverpool-Leeds United	None	
		2020-09-12 West Ham United-Newcastle	None	
		2020-09-13 Tottenham-Everton	None	
league	season	game	away_extratime_score	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	None	
		2020-09-12 Fulham-Arsenal	None	
		2020-09-12 Liverpool-Leeds United	None	
		2020-09-12 West Ham United-Newcastle	None	
		2020-09-13 Tottenham-Everton	None	
league	season	game	home_penalty_score	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	None	
		2020-09-12 Fulham-Arsenal	None	
		2020-09-12 Liverpool-Leeds United	None	
		2020-09-12 West Ham United-Newcastle	None	
		2020-09-13 Tottenham-Everton	None	
league	season	game	away_penalty_score	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	None	
		2020-09-12 Fulham-Arsenal	None	
		2020-09-12 Liverpool-Leeds United	None	
		2020-09-12 West Ham United-Newcastle	None	
		2020-09-13 Tottenham-Everton	None	
league	season	game	started_at_utc	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	2020-09-12T14:01:42Z	
		2020-09-12 Fulham-Arsenal	2020-09-12T11:32:23Z	
		2020-09-12 Liverpool-Leeds United	2020-09-12T16:30:21Z	
		2020-09-12 West Ham United-Newcastle	2020-09-12T19:00:32Z	
		2020-09-13 Tottenham-Everton	2020-09-13T15:30:20Z	
league	season	game	first_half_ended_at_utc	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	None	
		2020-09-12 Fulham-Arsenal	None	

(continues on next page)

(continued from previous page)

			None
			None
			None
second_half_started_at_			
utc \			
league	season game		
ENG-Premier League 2021	2020-09-12 Crystal Palace-Southampton	2020-09-12T15:04:	
01Z	2020-09-12 Fulham-Arsenal	2020-09-12T12:35:	
50Z	2020-09-12 Liverpool-Leeds United	2020-09-12T17:32:	
57Z	2020-09-12 West Ham United-Newcastle	2020-09-12T20:03:	
20Z	2020-09-13 Tottenham-Everton	2020-09-13T16:31:	
33Z			
stage			
league	season game		
ENG-Premier League 2021	2020-09-12 Crystal Palace-Southampton	None	
	2020-09-12 Fulham-Arsenal	None	
	2020-09-12 Liverpool-Leeds United	None	
	2020-09-12 West Ham United-Newcastle	None	
	2020-09-13 Tottenham-Everton	None	

2.8.2 Injured and suspended players

[6]:	missing_players = ws.read_missing_players(match_id=1485184)	
missing_players.head()		
[6]:		
game_id \		
league	season game	team
ENG-Premier League 2021	2021-01-12 Burnley-Manchester United	Burnley
Taylor 1485184		Charlie
McNeil 1485184		Dwight
Rodriguez 1485184		Jay
Dunne 1485184		Jimmy
Bailly 1485184		Manchester United Eric
player_id \		
league	season game	team
ENG-Premier League 2021	2021-01-12 Burnley-Manchester United	Burnley
Taylor 107462		Charlie
		Dwight

(continues on next page)

(continued from previous page)

↳ McNeil	357427		Jay ↴
↳ Rodriguez	33891		Jimmy ↴
↳ Dunne	366743		Manchester United Eric ↴
↳ Bailly	243814		
 ↳ reason \			
league	season game	team	player
ENG-Premier League 2021	2021-01-12 Burnley-Manchester United	Burnley	Charlie ↴
↳ Taylor	injured doubtful		Dwight ↴
↳ McNeil	injured doubtful		Jay ↴
↳ Rodriguez	injured doubtful		Jimmy ↴
↳ Dunne	injured doubtful		Manchester United Eric ↴
↳ Bailly	injured doubtful		
 ↳ status			
league	season game	team	player
ENG-Premier League 2021	2021-01-12 Burnley-Manchester United	Burnley	Charlie ↴
↳ Taylor	Doubtful		Dwight ↴
↳ McNeil	Doubtful		Jay ↴
↳ Rodriguez	Doubtful		Jimmy ↴
↳ Dunne	Doubtful		Manchester United Eric ↴
↳ Bailly	Doubtful		

2.8.3 Match event stream data

```
[7]: events = ws.read_events(match_id=1485184)
events.head()
```

```
[7]:
```

league	season game	id	game_id \
ENG-Premier League 2021	2021-01-12 Burnley-Manchester United	2253458317	1485184
		2253458375	1485184
		2253487469	1485184
		2253487473	1485184
		2253487625	1485184

league	season game	id	period \

(continues on next page)

(continued from previous page)

league	season	game	id	minute \
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	PreMatch
			2253458375	PreMatch
			2253487469	FirstHalf
			2253487473	FirstHalf
			2253487625	FirstHalf
league	season	game	id	second \
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	0
			2253458375	0
			2253487469	0
			2253487473	0
			2253487625	0
league	season	game	id	expanded_
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	
↳ 0			2253458375	
↳ 0			2253487469	
↳ 0			2253487473	
↳ 0			2253487625	
league	season	game	id	type ↳
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	FormationSet
			2253458375	FormationSet
			2253487469	Start
			2253487473	Start
			2253487625	Pass
league	season	game	id	outcome_type \
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	Successful
			2253458375	Successful
			2253487469	Successful
			2253487473	Successful
			2253487625	Successful

(continues on next page)

(continued from previous page)

league	season	game	id	team_id
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	184
			2253458375	32
			2253487469	184
			2253487473	32
			2253487625	184
league	season	game	id	team
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	Burnley
			2253458375	Man Utd
			2253487469	Burnley
			2253487473	Man Utd
			2253487625	Burnley
league	season	game	id	player_id
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	NaN
			2253458375	NaN
			2253487469	NaN
			2253487473	NaN
			2253487625	79050.0
league	season	game	id	x
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	0.0
→NaN			2253458375	0.0
→NaN			2253487469	0.0
→NaN			2253487473	0.0
→NaN			2253487625	Ashley
→Westwood				Westwood
league	season	game	id	y
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	0.0
			2253458375	0.0
			2253487469	0.0
			2253487473	0.0
			2253487625	50.3

(continues on next page)

(continued from previous page)

2253487473	0.0
2253487625	50.3

league	season	game	id	end_x	\
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	NaN	
			2253458375	NaN	
			2253487469	NaN	
			2253487473	NaN	
			2253487625	30.5	
league	season	game	id	end_y	\
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	NaN	
			2253458375	NaN	
			2253487469	NaN	
			2253487473	NaN	
			2253487625	50.3	
league	season	game	id	goal_mouth_y	✉
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	NaN	
			2253458375	NaN	
			2253487469	NaN	
			2253487473	NaN	
			2253487625	NaN	
league	season	game	id	goal_mouth_z	✉
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	NaN	
			2253458375	NaN	
			2253487469	NaN	
			2253487473	NaN	
			2253487625	NaN	
league	season	game	id	blocked_x	\
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	NaN	
			2253458375	NaN	
			2253487469	NaN	
			2253487473	NaN	
			2253487625	NaN	
league	season	game	id	blocked_y	\
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	NaN	
			2253458375	NaN	
			2253487469	NaN	
			2253487473	NaN	
			2253487625	NaN	

(continues on next page)

(continued from previous page)

```

    ↵          qualifiers \
league      season game           id
ENG-Premier League 2021 2021-01-12 Burnley-Manchester United 2253458317 [{"type": {
    ↵ "displayName": "TeamPlayerFormation...
                                2253458375 [{"type": {
    ↵ "displayName": "CaptainPlayerId", ...
                                2253487469
    ↵ []
                                2253487473
    ↵ []
                                2253487625 [{"type": {
    ↵ "displayName": "Angle", "value": 21...
                                is_touch \
league      season game           id
ENG-Premier League 2021 2021-01-12 Burnley-Manchester United 2253458317 False
                                2253458375 False
                                2253487469 False
                                2253487473 False
                                2253487625 True
                                is_shot \
league      season game           id
ENG-Premier League 2021 2021-01-12 Burnley-Manchester United 2253458317 NaN
                                2253458375 NaN
                                2253487469 NaN
                                2253487473 NaN
                                2253487625 NaN
                                is_goal \
league      season game           id
ENG-Premier League 2021 2021-01-12 Burnley-Manchester United 2253458317 NaN
                                2253458375 NaN
                                2253487469 NaN
                                2253487473 NaN
                                2253487625 NaN
                                card_type \
league      season game           id
ENG-Premier League 2021 2021-01-12 Burnley-Manchester United 2253458317 NaN
                                2253458375 NaN
                                2253487469 NaN
                                2253487473 NaN
                                2253487625 NaN
                                related_event_
    ↵ id \
league      season game           id
ENG-Premier League 2021 2021-01-12 Burnley-Manchester United 2253458317
    ↵ NaN

```

(continues on next page)

(continued from previous page)

↪NaN		2253458375	↳
↪NaN		2253487469	↳
↪NaN		2253487473	↳
↪NaN		2253487625	↳
↪NaN			
↪player_id			related_
league	season game	id	
ENG-Premier League	2021 2021-01-12 Burnley-Manchester United	2253458317	↳
↪NaN		2253458375	↳
↪NaN		2253487469	↳
↪NaN		2253487473	↳
↪NaN		2253487625	↳
↪NaN			

Match event stream data can be returned in various formats, which can be selected with the “output_fmt” parameter.

- events (default): Returns a dataframe with all events.
- raw: Returns the original unformatted WhoScored JSON.
- spadl: Returns a dataframe with the [SPADL representation](#) of the original events.
- atomic-spadl: Returns a dataframe with the [Atomic-SPADL representation](#) of the original events.
- loader: Returns a `socceration.data.opta.OptaLoader` instance

```
[8]: events = ws.read_events(match_id=1485184, output_fmt="raw")

import json # noqa

print(json.dumps(events[1485184][0], indent=2))
{
    "eventId": 2,
    "expandedMinute": 0,
    "id": 2253487473,
    "isTouch": false,
    "minute": 0,
    "outcomeType": {
        "displayName": "Successful",
        "value": 1
    },
    "period": {
        "displayName": "FirstHalf",
        "value": 1
    },
}
```

(continues on next page)

(continued from previous page)

```
"qualifiers": [],
"satisfiedEventsTypes": [],
"second": 0,
"teamId": 32,
"type": {
    "displayName": "Start",
    "value": 32
},
"x": 0,
"y": 0
}
```

[9]: actions = ws.read_events(match_id=1485184, output_fmt="spadl")
actions.head()

```
/cw/dtaijupiter/NoCsBack/dtai/pieterr/Projects/soccerdata/.venv/lib/python3.11/site-
→ packages/socceraction/spadl/opta.py:219: FutureWarning: Downcasting object dtype_
→ arrays on .fillna, .ffill, .bfill is deprecated and will change in a future version.
→ Call result.infer_objects(copy=False) instead. To opt-in to the future behavior, set_
→ `pd.set_option('future.no_silent_downcasting', True)`
).bfill()
```

[9]:

	game_id	original_event_id	period_id	time_seconds	team_id	player_id	\
0	1485184	2253487625	1	0.0	184	79050.0	
1	1485184	2253487639	1	2.0	184	131464.0	
2	1485184		Nan	4.5	184	80067.0	
3	1485184	2253487685	1	7.0	184	80067.0	
4	1485184	2253487689	1	11.0	184	93473.0	

	start_x	end_x	start_y	end_y	type_id	result_id	bodypart_id	\
0	52.815	32.025	34.204	34.204	0	1	0	
1	31.080	38.220	36.312	15.844	0	1	0	
2	38.220	43.365	15.844	12.512	21	1	0	
3	43.365	90.300	12.512	49.708	0	1	0	
4	90.300	105.000	49.708	38.828	11	0	4	

	action_id	player	team
0	0	Ashley Westwood	Burnley
1	1	James Tarkowski	Burnley
2	2	Matthew Lowton	Burnley
3	3	Matthew Lowton	Burnley
4	4	Robbie Brady	Burnley

[10]: atomic_actions = ws.read_events(match_id=1485184, output_fmt="atomic-spadl")
atomic_actions.head()

```
/cw/dtaijupiter/NoCsBack/dtai/pieterr/Projects/soccerdata/.venv/lib/python3.11/site-
→ packages/socceraction/spadl/opta.py:219: FutureWarning: Downcasting object dtype_
→ arrays on .fillna, .ffill, .bfill is deprecated and will change in a future version.
→ Call result.infer_objects(copy=False) instead. To opt-in to the future behavior, set_
→ `pd.set_option('future.no_silent_downcasting', True)`
).bfill()
```

```
[10]:
```

	game_id	original_event_id	action_id	period_id	time_seconds	team_id	\
0	1485184	2253487625	0	1	0.00	184	
1	1485184	2253487625	1	1	1.00	184	
2	1485184	2253487639	2	1	2.00	184	
3	1485184	2253487639	3	1	3.25	184	
4	1485184	NaN	4	1	4.50	184	

	player_id	x	y	dx	dy	type_id	bodypart_id	\
0	79050.0	52.815	34.204	-20.790	0.000	0	0	
1	131464.0	32.025	34.204	0.000	0.000	23	0	
2	131464.0	31.080	36.312	7.140	-20.468	0	0	
3	80067.0	38.220	15.844	0.000	0.000	23	0	
4	80067.0	38.220	15.844	5.145	-3.332	21	0	

	player	team
0	Ashley Westwood	Burnley
1	James Tarkowski	Burnley
2	James Tarkowski	Burnley
3	Matthew Lowton	Burnley
4	Matthew Lowton	Burnley

```
[11]: # Scrape all games and return a socceraction.data.opta.OptaLoader
loader = ws.read_events(output_fmt="loader")

# Now use this loader to load the data
print("Games:")
df_games = loader.games(competition_id="ENG-Premier League", season_id="2021")
display(df_games.head())

print("Teams:")
df_teams = loader.teams(game_id=1485184)
display(df_teams.head())

print("Players:")
df_players = loader.players(game_id=1485184)
display(df_players.head())

print("Events:")
df_events = loader.events(game_id=1485184)
display(df_events.head())

# You can use the socceraction package to convert the events
# to SPADL and to compute xT or VAEP action values
```

Games:

	game_id	season_id	competition_id	game_day	game_date	\
0	1485494	2021	ENG-Premier League	None	2021-04-04 12:00:00	
1	1485300	2021	ENG-Premier League	None	2020-12-16 20:00:00	
2	1485264	2021	ENG-Premier League	None	2020-12-06 19:15:00	
3	1485519	2021	ENG-Premier League	None	2021-05-16 16:30:00	
4	1485436	2021	ENG-Premier League	None	2021-03-19 20:00:00	

(continues on next page)

(continued from previous page)

	home_team_id	away_team_id	home_score	away_score	duration	\		
0	18	184	3	2	98			
1	170	211	0	0	95			
2	26	161	4	0	97			
3	175	26	1	2	102			
4	170	19	1	2	100			
	referee	venue	attendance	home_manager	away_manager	\		
0	Andre Marriner	St. Mary's Stadium	0	Ralph Hasenhüttl	Sean Dyche			
1	Robert Jones	Craven Cottage	0	Scott Parker	Graham Potter			
2	Craig Pawson	Anfield	2000	Jürgen Klopp	Nuno Espírito Santo			
3	Mike Dean	The Hawthorns	0	Sam Allardyce	Jürgen Klopp			
4	David Coote	Craven Cottage	0	Scott Parker	Marcelo Bielsa			
	team_id	team_name						
0	184	Burnley						
1	32	Man Utd						
	game_id	team_id	player_id	player_name	is_starter	minutes_played	\	
0	1485184	184	105720	Nick Pope	True	102		
1	1485184	184	80067	Matthew Lowton	True	102		
2	1485184	184	94935	Ben Mee	True	102		
3	1485184	184	131464	James Tarkowski	True	102		
4	1485184	184	24148	Erik Pieters	True	102		
	jersey_number	starting_position						
0	1	GK						
1	2	DR						
2	6	DC						
3	5	DC						
4	23	DL						
	game_id	event_id	period_id	team_id	player_id	type_id	\	
0	1485184	2253487473	1	32	NaN	32		
1	1485184	2253487469	1	184	NaN	32		
2	1485184	2253487625	1	184	79050.0	1		
3	1485184	2253487639	1	184	131464.0	1		
4	1485184	2253487685	1	184	80067.0	1		
	timestamp	minute	second	outcome	start_x	start_y	end_x	\
0	2021-01-12 20:15:00	0	0	True	0.0	0.0	0.0	
1	2021-01-12 20:15:00	0	0	True	0.0	0.0	0.0	

(continues on next page)

(continued from previous page)

2	2021-01-12	20:15:00	0	0	True	50.3	50.3	30.5
3	2021-01-12	20:15:02	0	2	True	29.6	53.4	36.4
4	2021-01-12	20:15:07	0	7	True	41.3	18.4	86.0
			end_y		qualifiers	\		
0			0.0			{}		
1			0.0			{}		
2			50.3	{213: '3.1', 178: True, 141: '50.3', 212: '20...				
3			23.3	{178: True, 213: '5.0', 212: '21.7', 141: '23...				
4			73.1	{1: True, 213: '0.7', 56: 'Center', 178: True,...				
			related_player_id	touch	goal	shot	type_name	
0				NaN	False	False	False	start
1				NaN	False	False	False	start
2				NaN	True	False	False	pass
3				NaN	True	False	False	pass
4				NaN	True	False	False	pass

[]:

..... doc/datasources/WhoScored.ipynb ends here.

HOW-TO GUIDES

Here you'll find short answers to "How do I...?" types of questions. These how-to guides don't cover topics in depth – you'll find that material in the [API Reference](#). However, these guides will help you quickly accomplish common tasks.

3.1 How to add custom leagues

SoccerData has built-in support to scrape data from the top-5 European leagues and the major international tournaments. The leagues available for each source can be listed with the `available_leagues()` class method.

```
import soccerdata as sd
sd.FBref.available_leagues()
>>> ['ENG-Premier League', 'ESP-La Liga', 'FRA-Ligue 1', 'GER-Bundesliga', 'ITA-Serie A']
```

This documentation explains how to add custom leagues.

⚠ Warning

Note that you might encounter errors when trying to scrape data for the leagues you added yourself. This is because the data provided for these leagues might have a different structure. If you encounter such an error, please do not open an issue on GitHub, but try to fix it yourself.

3.1.1 Adding a new league

Additional leagues can be configured in `SOCCKERDATA_DIR/config/league_dict.json`. This file should contain a mapping between a generic name for the league and the identifier used internally by each data source (see below) that you want to support. For example, for the Dutch Eredivisie this would be:

```
{  
    "NED-Eredivisie": {  
        "ClubElo": "NED_1",  
        "MatchHistory": "N1",  
        "SoFIFA": "[Netherlands] Eredivisie",  
        "FBref": "Eredivisie",  
        "ESPN": "ned.1",  
        "FiveThirtyEight": "eredivisie",  
        "WhoScored": "Netherlands - Eredivisie",  
        "Sofascore": "Eredivisie",  
        "season_start": "Aug",  
    },  
}
```

(continues on next page)

(continued from previous page)

```
    "season_end": "May"  
}  
}
```

The `season_end` and `season_start` fields are optional. This should be the month in which the last game and first game of a season are played, respectively. If they are not provided, June is used as the last month of the season and July as the first one.

Now, restart your Python session and check whether it is added to available leagues by running the command below.

```
>>> import soccerdata as sd  
>>> sd.FBref.available_leagues()  
[..., 'NED-Eredivisie', ...]
```

3.1.2 Internal identifiers

Below are instructions on how to find the internal identifiers for each data source.

ClubElo

The internal identifier has the format `{country_code}_{level}`. To get the country code, go to <https://clubelo.com/>, click on the league you want to add and take the three-letter code in the URL. For example, the URL for the Dutch Eredivisie is <http://clubelo.com/NED> which means that the country identifier is NED. The level is the number of the league, starting with 1 for the top league. The internal identifier for the Dutch Eredivisie is therefore `NED_1`.

MatchHistory

The internal identifier has the format `{country_code}{level}`. Download the CSV file corresponding corresponding to the league you would like to add from <https://www.football-data.co.uk/data.php> and take the value in the Div column.

SoFIFA

The internal identifier has the format `[{region}] {league name}`. Go to <https://sofifa.com/api/league> to get the list of available leagues. The `{region}` corresponds to the `nationName` field in the JSON response. The `{league name}` corresponds to the value field.

FBref

Go to <https://fbref.com/en/comps/> and take the value in the Competition Name column.

ESPN

The internal identifier has the format `{country_code}.{level}`. Go to <https://www.espn.com/soccer/competitions>, click on the league you want to add and take the value in the URL after `/league/_/name/`.

FiveThirtyEight

Go to <https://projects.fivethirtyeight.com/soccer-predictions/>, select the relevant league and take the value in the URL after `/soccer-predictions/`.

WhoScored

Go to <https://www.whoscored.com> and use the JavaScript console to get the value of the `allRegions` variable. The internal identifier has the format `{region name} - {league name}`.

Sofascore

Go to <https://api.sofascore.com/api/v1/config/unique-tournaments/EN/football> to get the list of major leagues and tournaments. Access `uniqueTournaments` in the JSON response, and the `{league name}` corresponds to the `name` field.

3.1.3 Troubleshooting

If you add a new league and it doesn't show up in the list of available leagues, there are a few things you can do to debug the problem.

1. Make sure to reload the soccerdata module after you modify the `league_dict.json` file. The most straightforward way to do this is to restart your notebook or Python interpreter.
2. Check whether your `league_dict.json` file is at the correct location. If so, you should see this appear in the log messages when importing the soccerdata library.

```
>>> import soccerdata as sd
[11/25/22 11:49:12] INFO      Custom team name replacements loaded from <path>/
↳teamname_replacements.json.
↳                                         _config.py:83
[11/25/22 11:49:12] INFO      Custom league dict loaded from <path>/league_dict.json.
↳                                         _config.py:153
```

3. Check whether the content of your `league_dict.json` file is valid JSON. You can check the file's syntax using Python's built-in `json.tool` module.

```
$ cat config/league_dict.json | python -m json.tool
Expecting ',' delimiter: line 1 column 10 (char 9)
```

3.2 How to use a proxy server

You can setup a SOCKS5 proxy with Tor. Checkout the [installation guide](#) on the Tor website for installation instructions. After installing Tor, make sure to start it up before scraping. This can easily be done by running the `tor` command from your terminal (in a separate window), Tor will start up and run on “localhost:9050” by default. Once Tor is running, you can enable the extension by setting `proxy='tor'`.

```
ws = sd.WhoScored(proxy='tor')
```

The code snippet above assumes you have a Tor proxy running on “localhost:9050”. Many distributions indeed default to having a SOCKS proxy listening on port 9050, but some may not. In particular, the Tor Browser Bundle defaults to listening on port 9150. You can specify a custom host and port as

```
ws = sd.WhoScored(proxy={
    "http": "socks5://127.0.0.1:9150",
    "https": "socks5://127.0.0.1:9150",
})
```


DATA ANALYSIS EXAMPLES

Below is a gallery of examples of data analysis using the `soccerdata` package. If you have an example you would like to share, please submit a pull request to the SoccerData GitHub repository.

The following section was generated from doc/examples/ClubElo - Evolution of current top teams.ipynb

[2]: `import soccerdata as sd`

[3]: `%matplotlib inline`
`%config InlineBackend.figure_format = 'retina'`

`import matplotlib.pyplot as plt`
`import pandas as pd`
`import seaborn as sns`

`sns.set_context("notebook")`
`sns.set_style("whitegrid")`

4.1 Evolution of top team's Elo ratings

How did the current top 5 teams in the world develop over time?

[4]: `elo = sd.ClubElo()`
`current_elo = elo.read_by_date()`
`current_elo.head()`

/cw/dtaijupiter/NoCsBack/dtai/pieterr/Projects/soccerdata/soccerdata/clubelo.py:18:
↳ FutureWarning: The argument 'infer_datetime_format' is deprecated and will be removed.
↳ in a future version. A strict version of it is now the default, see https://pandas.
↳ pydata.org/pdeps/0004-consistent-to-datetime-parsing.html. You can safely remove this.
↳ argument.
 return pd.read_csv(

[4]:

team	rank	country	level	elo	from	to	\
Man City	1.0	ENG	1	2079.984619	2023-06-11	2023-08-11	
Liverpool	2.0	ENG	1	1949.932861	2023-06-11	2023-08-13	
Bayern	3.0	GER	1	1937.012451	2023-05-28	2023-08-18	
Arsenal	4.0	ENG	1	1928.216187	2023-06-11	2023-08-12	
Real Madrid	5.0	ESP	1	1907.589233	2023-06-08	2023-08-12	

(continues on next page)

(continued from previous page)

	league
team	
Man City	ENG-Premier League
Liverpool	ENG-Premier League
Bayern	GER-Bundesliga
Arsenal	ENG-Premier League
Real Madrid	ESP-La Liga

```
[5]: num_teams = 5
smoothing = 100
elo_top_development = pd.concat(
    [
        elo.read_team_history(team)["elo"].rolling(smoothing).mean()
        for team in current_elo.reset_index()["team"][:num_teams]
    ],
    axis=1,
)
elo_top_development.columns = current_elo.reset_index()["team"][:num_teams]
elo_top_development.fillna(method="ffill")

fig = plt.figure(figsize=(16, 10))
ax1 = fig.add_subplot(111, ylabel="ELO rolling avg.", xlabel="Date")
elo_top_development.plot(ax=ax1)
ax1.legend(loc="upper left", frameon=False, bbox_to_anchor=(0, 1.05), ncol=num_teams)
sns.despine();

/cw/dtaijupiter/NoCsBack/dtai/pieterr/Projects/soccerdata/soccerdata/clubelo.py:18: FutureWarning: The argument 'infer_datetime_format' is deprecated and will be removed in a future version. A strict version of it is now the default, see https://pandas.pydata.org/pdeps/0004-consistent-to-datetime-parsing.html. You can safely remove this argument.
    return pd.read_csv(
/cw/dtaijupiter/NoCsBack/dtai/pieterr/Projects/soccerdata/soccerdata/clubelo.py:18: FutureWarning: The argument 'infer_datetime_format' is deprecated and will be removed in a future version. A strict version of it is now the default, see https://pandas.pydata.org/pdeps/0004-consistent-to-datetime-parsing.html. You can safely remove this argument.
    return pd.read_csv(
/cw/dtaijupiter/NoCsBack/dtai/pieterr/Projects/soccerdata/soccerdata/clubelo.py:18: FutureWarning: The argument 'infer_datetime_format' is deprecated and will be removed in a future version. A strict version of it is now the default, see https://pandas.pydata.org/pdeps/0004-consistent-to-datetime-parsing.html. You can safely remove this argument.
    return pd.read_csv(
/cw/dtaijupiter/NoCsBack/dtai/pieterr/Projects/soccerdata/soccerdata/clubelo.py:18: FutureWarning: The argument 'infer_datetime_format' is deprecated and will be removed in a future version. A strict version of it is now the default, see https://pandas.pydata.org/pdeps/0004-consistent-to-datetime-parsing.html. You can safely remove this argument.
    return pd.read_csv(
/cw/dtaijupiter/NoCsBack/dtai/pieterr/Projects/soccerdata/soccerdata/clubelo.py:18: FutureWarning: The argument 'infer_datetime_format' is deprecated and will be removed in a future version. A strict version of it is now the default, see https://pandas.pydata.org/pdeps/0004-consistent-to-datetime-parsing.html. You can safely remove this argument.
```

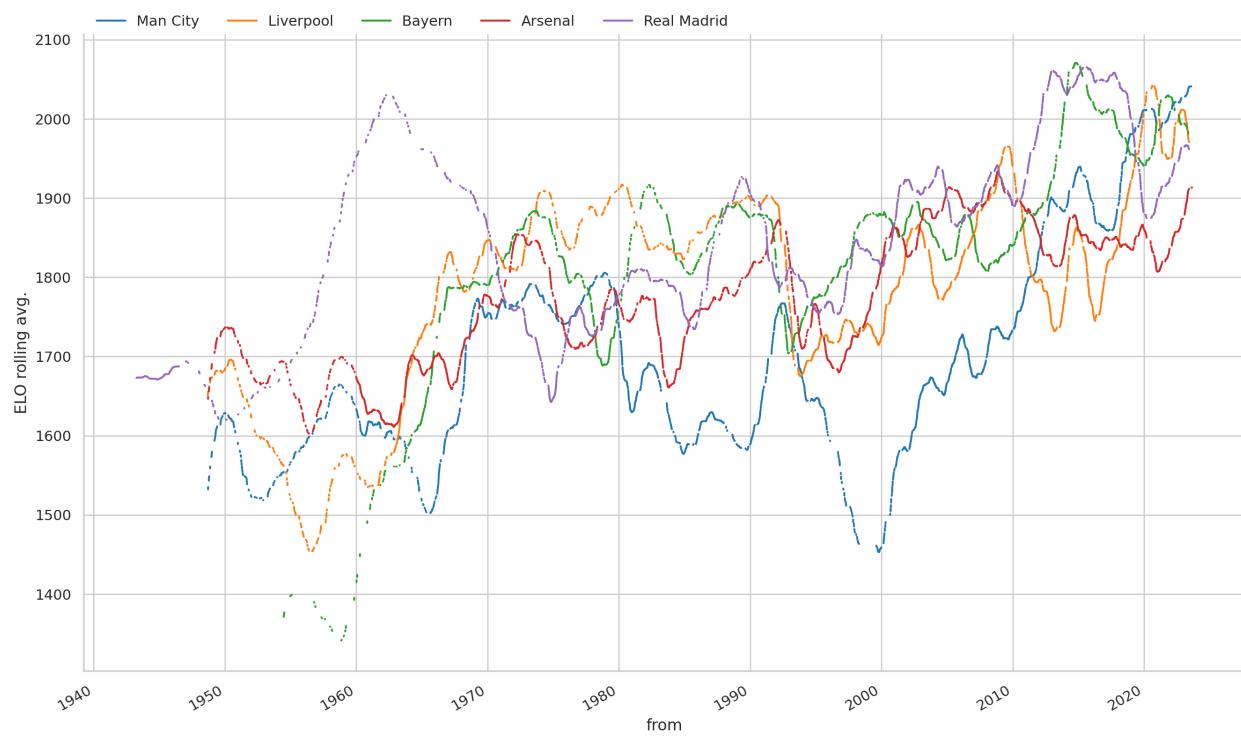
(continues on next page)

(continued from previous page)

```

-> FutureWarning: The argument 'infer_datetime_format' is deprecated and will be removed.
-> in a future version. A strict version of it is now the default, see https://pandas.
-> pydata.org/pdeps/0004-consistent-to-datetime-parsing.html. You can safely remove this.
-> argument.
    return pd.read_csv(

```



The following section was generated from doc/examples/MatchHistory - Home advantage.ipynb

```
[2]: import soccerdata as sd
```

```
[3]: %matplotlib inline
%config InlineBackend.figure_format = 'retina'

import pandas as pd
import seaborn as sns

sns.set_context("notebook")
sns.set_style("whitegrid")
```

4.2 Home team advantage in the Italian Serie A

We all know sports teams have an advantage when playing at home. Here's a look at home team advantage for 5 years of the Serie A.

```
[4]: seriea_hist = sd.MatchHistory("ITA-Serie A", range(2018, 2023))
games = seriea_hist.read_games()
games.sample(5)
```

```
[4]:                                         date \
league      season game
ITA-Serie A 2223    2023-05-20 Cremonese-Bologna 2023-05-20 14:00:00
              2021    2020-12-06 Crotone-Napoli      2020-12-06 17:00:00
              1819    2019-05-26 Fiorentina-Genoa   2019-05-26 12:00:00
              2021    2020-11-01 Udinese-Milan     2020-11-01 11:30:00
              1819    2019-03-10 Frosinone-Torino   2019-03-10 12:00:00
```

league	season	game	home_team	away_team	FTHG	\
ITA-Serie A	2223	2023-05-20	Cremonese-Bologna	Cremonese	Bologna	1
	2021	2020-12-06	Crotone-Napoli	Crotone	Napoli	0
	1819	2019-05-26	Fiorentina-Genoa	Fiorentina	Genoa	0
	2021	2020-11-01	Udinese-Milan	Udinese	Milan	1
	1819	2019-03-10	Frosinone-Torino	Frosinone	Torino	1

league	season	game		FTAG	FTR	HTHG	HTAG	HTR	HS	\
ITA-Serie A	2223	2023-05-20	Cremonese-Bologna	5	A	0	3	A	13	
	2021	2020-12-06	Crotone-Napoli	4	A	0	1	A	4	
	1819	2019-05-26	Fiorentina-Genoa	0	D	0	0	D	5	
	2021	2020-11-01	Udinese-Milan	2	A	0	1	A	7	
	1819	2019-03-10	Frosinone-Torino	2	A	1	0	H	9	

league	season	game		...	AvgC<2.5	AHCh	\
ITA-Serie A	2223	2023-05-20	Cremonese-Bologna	...	1.8	0.50	
	2021	2020-12-06	Crotone-Napoli	...	2.3	1.25	
	1819	2019-05-26	Fiorentina-Genoa	...	NaN	NaN	
	2021	2020-11-01	Udinese-Milan	...	2.0	0.50	
	1819	2019-03-10	Frosinone-Torino	...	NaN	NaN	

league	season	game	B365CAHH	B365CAHA	PCAHH	\
ITA-Serie A	2223	2023-05-20 Cremonese-Bologna	1.84	2.06	1.85	
	2021	2020-12-06 Crotone-Napoli	1.99	1.94	1.98	
	1819	2019-05-26 Fiorentina-Genoa	NaN	NaN	NaN	
	2021	2020-11-01 Udinese-Milan	1.95	1.98	1.93	
	1819	2019-03-10 Frosinone-Torino	NaN	NaN	NaN	

league	season	game	PCAHA	MaxCAHH	MaxCAHA	\
ITA-Serie A	2223	2023-05-20 Cremonese-Bologna	2.08	1.90	2.15	
	2021	2020-12-06 Crotone-Napoli	1.94	2.03	1.96	

(continues on next page)

(continued from previous page)

1819	2019-05-26	Fiorentina-Genoa	NaN	NaN	NaN
2021	2020-11-01	Udinese-Milan	1.99	1.99	2.01
1819	2019-03-10	Frosinone-Torino	NaN	NaN	NaN
AvgCAHH AvgCAHA					
league	season	game			
ITA-Serie A	2223	2023-05-20 Cremonese-Bologna	1.83	2.03	
	2021	2020-12-06 Crotone-Napoli	1.96	1.90	
	1819	2019-05-26 Fiorentina-Genoa	NaN	NaN	
	2021	2020-11-01 Udinese-Milan	1.93	1.94	
	1819	2019-03-10 Frosinone-Torino	NaN	NaN	

[5 rows x 121 columns]

```
[5]: def home_away_results(games: pd.DataFrame):
    """Returns aggregated home/away results per team"""
    res = pd.melt(
        games.reset_index(),
        id_vars=["date", "FTR"],
        value_name="team",
        var_name="is_home",
        value_vars=["home_team", "away_team"],
    )

    res.is_home = res.is_home.replace(["home_team", "away_team"], ["Home", "Away"])
    res["win"] = res["lose"] = res["draw"] = 0
    res.loc[(res["is_home"] == "Home") & (res["FTR"] == "H"), "win"] = 1
    res.loc[(res["is_home"] == "Away") & (res["FTR"] == "A"), "win"] = 1
    res.loc[(res["is_home"] == "Home") & (res["FTR"] == "A"), "lose"] = 1
    res.loc[(res["is_home"] == "Away") & (res["FTR"] == "H"), "lose"] = 1
    res.loc[res["FTR"] == "D", "draw"] = 1

    groups = res.groupby(["team", "is_home"])
    win = groups.win.agg(["sum", "mean"]).rename(columns={"sum": "n_win", "mean": "win_pct"})
    loss = groups.lose.agg(["sum", "mean"]).rename(columns={"sum": "n_lose", "mean": "lose_pct"})
    draw = groups.draw.agg(["sum", "mean"]).rename(columns={"sum": "n_draw", "mean": "draw_pct"})

    res = pd.concat([win, loss, draw], axis=1)
    return res
```

```
[6]: results = home_away_results(games)
results.head(6)
```

team	is_home	n_win	win_pct	n_lose	lose_pct	n_draw	draw_pct
Atalanta	Away	52	0.547368	16	0.168421	27	0.284211
	Home	56	0.589474	23	0.242105	16	0.168421
Benevento	Away	10	0.263158	18	0.473684	10	0.263158
	Home	4	0.105263	20	0.526316	14	0.368421

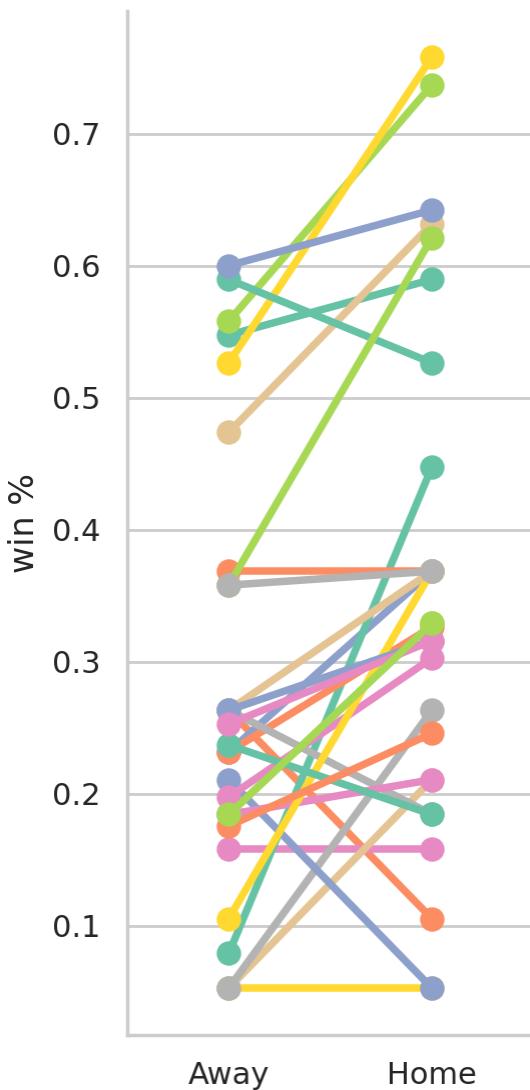
(continues on next page)

(continued from previous page)

Bologna	Away	22	0.231579	48	0.505263	25	0.263158
	Home	35	0.368421	29	0.305263	31	0.326316

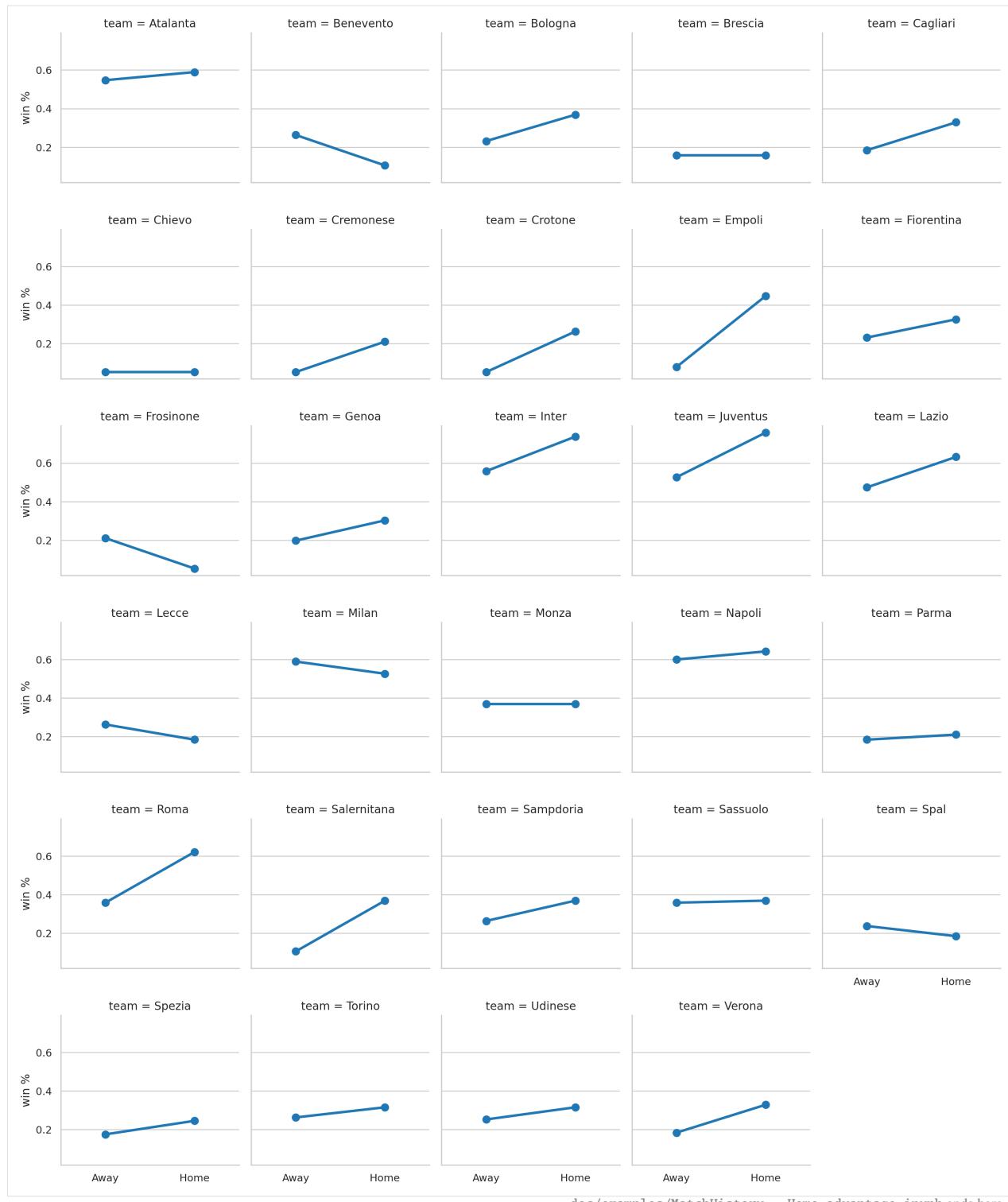
The overall picture shows most teams have a clear advantage at home:

```
[7]: g = sns.FacetGrid(results.reset_index(), hue="team", palette="Set2", height=6, aspect=0.5)
g.map(sns.pointplot, "is_home", "win_pct", order=["Away", "Home"])
g.set_axis_labels("", "win %");
```



But there are a few exceptions.

```
[8]: g = sns.FacetGrid(results.reset_index(), col="team", col_wrap=5)
g.map(sns.pointplot, "is_home", "win_pct", order=["Away", "Home"])
g.set_axis_labels("", "win %");
```



doc/examples/MatchHistory - Home advantage.ipynb ends here.

API REFERENCE

This part of the documentation covers all the interfaces of the implemented data scrapers.

5.1 Club Elo

```
class soccerdata.ClubElo(proxy=None, no_cache=False, no_store=False,  
                         data_dir=PosixPath('/home/docs/soccerdata/data/ClubElo'))
```

Provides pd.DataFrame from CSV API at <http://api.clubelo.com>.

Data will be downloaded as necessary and cached locally in `~/soccerdata/data/ClubElo`.

Since the source does not provide league names, this class will not filter by league. League names will be inserted from the other sources where available. Leagues that are only covered by clubelo.com will have NaN values.

Parameters

- **proxy** ('tor' or dict or list(dict) or callable, optional) –

Use a proxy to hide your IP address. Valid options are:

- "tor": Uses the Tor network. Tor should be running in the background on port 9050.
- str: The address of the proxy server to use.
- list(str): A list of proxies to choose from. A different proxy will be selected from this list after failed requests, allowing rotating proxies.
- callable: A function that returns a valid proxy. This function will be called after failed requests, allowing rotating proxies.

- **no_cache** (bool) – If True, will not use cached data.

- **no_store** (bool) – If True, will not store downloaded data.

- **data_dir** (Path) – Path to directory where data will be cached.

read_by_date(date=None)

Retrieve ELO scores for all teams at specified date.

ELO scores are available as early as 1939. Values before 1960 should be considered provisional.

Parameters

date (datetime object or string like 'YYYY-MM-DD') – Date for which to retrieve ELO scores. If no date is specified, get today's scores.

Raises

- **TypeError** – If date is not a date string or datetime object.

- **ValueError** – If data is an invalid date string.

Return type

pd.DataFrame

read_team_history(*team*, *max_age*=*I*)

Retrieve full ELO history for one club.

For the exact spelling of a club's name, check the result of `read_by_date()` or `clubelo.com`. You can also use alternative team names specified in `teamname_replacements.json`. Values before 1960 should be considered provisional.

Parameters

- **team (str)** – The club's name.
- **max_age (int for age in days, or timedelta object)** – The max. age of locally cached file before re-download. To disable re-downloading, set to None.

Raises

- **TypeError** – If max_age is not an integer or timedelta object.
- **ValueError** – If no ratings for the given team are available.

Return type

pd.DataFrame

classmethod available_leagues()

Return a list of league IDs available for this source.

Return type

list[str]

get(*url*, *filepath*=None, *max_age*=None, *no_cache*=False, *var*=None)

Load data from *url*.

By default, the source of *url* is downloaded and saved to *filepath*. If *filepath* exists, the *url* is not visited and the cached data is returned.

Parameters

- **url (str)** – URL to download.
- **filepath (Path, optional)** – Path to save downloaded file. If None, downloaded data is not cached.
- **max_age (int for age in days, or timedelta object)** – The max. age of locally cached file before re-download.
- **no_cache (bool)** – If True, will not use cached data. Overrides the class property.
- **var (str or list of str, optional)** – Return a JavaScript variable instead of the page source.

Raises

TypeError – If max_age is not an integer or timedelta object.

Returns

File-like object of downloaded data.

Return type

io.BufferedIOBase

property leagues: list[str]

Return a list of selected leagues.

property seasons: list[str]

Return a list of selected seasons.

5.2 ESPN

```
class soccerdata.ESPN(leagues=None, seasons=None, proxy=None, no_cache=False, no_store=False,  
                      data_dir=PosixPath('/home/docs/soccerdata/data/ESPN'))
```

Provides pd.DataFrames from JSON api available at <http://site.api.espn.com>.

Data will be downloaded as necessary and cached locally in `~/soccerdata/data/ESPN`.

Parameters

- **leagues** (*string or iterable, optional*) – IDs of leagues to include.
- **seasons** (*string, int or list, optional*) – Seasons to include. Supports multiple formats. Examples: ‘16-17’; 2016; ‘2016-17’; [14, 15, 16]
- **proxy** (*‘tor’ or str or dict or list(dict) or callable, optional*) –
Use a proxy to hide your IP address. Valid options are:
 - “tor”: Uses the Tor network. Tor should be running in the background on port 9050.
 - str: The address of the proxy server to use.
 - list(str): A list of proxies to choose from. A different proxy will be selected from this list after failed requests, allowing rotating proxies.
 - callable: A function that returns a valid proxy. This function will be called after failed requests, allowing rotating proxies.
- **no_cache** (*bool*) – If True, will not use cached data.
- **no_store** (*bool*) – If True, will not store downloaded data.
- **data_dir** (*Path*) – Path to directory where data will be cached.

property seasons: list[str]

Return a list of selected seasons.

read_schedule(force_cache=False)

Retrieve the game schedule for the selected leagues and seasons.

Parameters

force_cache (*bool*) – By default no cached data is used for the current season. If True, will force the use of cached data anyway.

Return type

`pd.DataFrame`

read_matchesheet(match_id=None)

Retrieve match sheets for the selected leagues and seasons.

Parameters

match_id (*int or list of int, optional*) – Retrieve the match sheet for a specific game.

Raises

ValueError – If no games with the given IDs were found for the selected seasons and leagues.

Return type

pd.DataFrame.

read_lineup(*match_id=None*)

Retrieve lineups for the selected leagues and seasons.

Parameters

match_id (*int or list of int, optional*) – Retrieve the lineup for a specific game.

Raises

ValueError – If no games with the given IDs were found for the selected seasons and leagues.

Return type

pd.DataFrame.

classmethod available_leagues()

Return a list of league IDs available for this source.

Return type

list[str]

get(*url, filepath=None, max_age=None, no_cache=False, var=None*)

Load data from *url*.

By default, the source of *url* is downloaded and saved to *filepath*. If *filepath* exists, the *url* is not visited and the cached data is returned.

Parameters

- **url** (*str*) – URL to download.
- **filepath** (*Path, optional*) – Path to save downloaded file. If None, downloaded data is not cached.
- **max_age** (*int for age in days, or timedelta object*) – The max. age of locally cached file before re-download.
- **no_cache** (*bool*) – If True, will not use cached data. Overrides the class property.
- **var** (*str or list of str, optional*) – Return a JavaScript variable instead of the page source.

Raises

TypeError – If max_age is not an integer or timedelta object.

Returns

File-like object of downloaded data.

Return type

io.BufferedIOBase

property leagues: list[str]

Return a list of selected leagues.

5.3 FBref

```
class soccerdata.FBref(leagues=None, seasons=None, proxy=None, no_cache=False, no_store=False,
                       data_dir=PosixPath('/home/docs/soccerdata/data/FBref'))
```

Provides pd.DataFrames from data at <http://fbref.com>.

Data will be downloaded as necessary and cached locally in `~/soccerdata/data/FBref`.

Parameters

- **leagues** (*string or iterable, optional*) – IDs of leagues to include. For efficiently reading data from the Top-5 European leagues, use “Big 5 European Leagues Combined”.
- **seasons** (*string, int or list, optional*) – Seasons to include. Supports multiple formats. Examples: ‘16-17’; 2016; ‘2016-17’; [14, 15, 16]
- **proxy** (*'tor' or str or dict or list(dict) or callable, optional*) – **Use a proxy to hide your IP address. Valid options are:**
 - “tor”: Uses the Tor network. Tor should be running in the background on port 9050.
 - str: The address of the proxy server to use.
 - list(str): A list of proxies to choose from. A different proxy will be selected from this list after failed requests, allowing rotating proxies.
 - callable: A function that returns a valid proxy. This function will be called after failed requests, allowing rotating proxies.
- **no_cache** (*bool*) – If True, will not use cached data.
- **no_store** (*bool*) – If True, will not store downloaded data.
- **data_dir** (*Path*) – Path to directory where data will be cached.

read_leagues (*split_up_big5=False*)

Retrieve selected leagues from the datasource.

Parameters

split_up_big5 (*bool*) – If True, it will load the “Big 5 European Leagues Combined” instead of each league individually.

Return type

pd.DataFrame

read_seasons (*split_up_big5=False*)

Retrieve the selected seasons for the selected leagues.

Parameters

split_up_big5 (*bool*) – If True, it will load the “Big 5 European Leagues Combined” instead of each league individually.

Return type

pd.DataFrame

read_team_season_stats (*stat_type='standard', opponent_stats=False*)

Retrieve aggregated season stats for all teams in the selected leagues and seasons.

The following stat types are available:

- ‘standard’
- ‘keeper’

- ‘keeper_adv’
- ‘shooting’
- ‘passing’
- ‘passing_types’
- ‘goal_shot_creation’
- ‘defense’
- ‘possession’
- ‘playing_time’
- ‘misc’

Parameters

- **stat_type** (*str*) – Type of stats to retrieve.
- **opponent_stats** (*bool*) – If True, will retrieve opponent stats.

Raises

ValueError – If `stat_type` is not valid.

Return type

`pd.DataFrame`

read_team_match_stats(stat_type='schedule', opponent_stats=False, team=None, force_cache=False)

Retrieve the match logs for all teams in the selected leagues and seasons.

The following stat types are available:

- ‘schedule’
- ‘keeper’
- ‘shooting’
- ‘passing’
- ‘passing_types’
- ‘goal_shot_creation’
- ‘defense’
- ‘possession’
- ‘misc’

Parameters

- **stat_type** (*str*) – Type of stats to retrieve.
- **opponent_stats** (*bool*) – If True, will retrieve opponent stats.
- **team** (*str or list of str, optional*) – Team(s) to retrieve. If None, will retrieve all teams.
- **force_cache** (*bool*) – By default no cached data is used for the current season. If True, will force the use of cached data anyway.

Raises

ValueError – If `stat_type` is not valid.

Return type

pd.DataFrame

read_player_season_stats(stat_type='standard')

Retrieve players from the datasource for the selected leagues and seasons.

The following stat types are available:

- ‘standard’
- ‘shooting’
- ‘passing’
- ‘passing_types’
- ‘goal_shot_creation’
- ‘defense’
- ‘possession’
- ‘playing_time’
- ‘misc’
- ‘keeper’
- ‘keeper_adv’

Parameters**stat_type (str)** – Type of stats to retrieve.**Raises****TypeError** – If `stat_type` is not valid.**Return type**

pd.DataFrame

read_schedule(force_cache=False)

Retrieve the game schedule for the selected leagues and seasons.

Parameters**force_cache (bool)** – By default no cached data is used for the current season. If True, will force the use of cached data anyway.**Return type**

pd.DataFrame

read_player_match_stats(stat_type='summary', match_id=None, force_cache=False)

Retrieve the match stats for the selected leagues and seasons.

The following stat types are available:

- ‘summary’
- ‘keepers’
- ‘passing’
- ‘passing_types’
- ‘defense’
- ‘possession’

- ‘misc’

Parameters

- **stat_type** (*str*) – Type of stats to retrieve.
- **match_id** (*int or list of int, optional*) – Retrieve the event stream for a specific game.
- **force_cache** (*bool*) – By default no cached data is used to scrape the list of available games for the current season. If True, will force the use of cached data anyway.

Raises

- **ValueError** – If no games with the given IDs were found for the selected seasons and leagues.
- **TypeError** – If **stat_type** is not valid.

Return type

pd.DataFrame

`read_lineup(match_id=None, force_cache=False)`

Retrieve lineups for the selected leagues and seasons.

Parameters

- **match_id** (*int or list of int, optional*) – Retrieve the lineup for a specific game.
- **force_cache** (*bool*) – By default no cached data is used to scrape the list of available games for the current season. If True, will force the use of cached data anyway.

Raises

- **ValueError** – If no games with the given IDs were found for the selected seasons and leagues.

Return type

pd.DataFrame

`read_events(match_id=None, force_cache=False)`

Retrieve match events for the selected seasons or selected matches.

The data returned includes the timing of goals, cards and substitutions. Also includes the players who are involved in the event.

Parameters

- **match_id** (*int or list of int, optional*) – Retrieve the events for a specific game.
- **force_cache** (*bool*) – By default no cached data is used to scrape the list of available games for the current season. If True, will force the use of cached data anyway.

Raises

- **ValueError** – If no games with the given IDs were found for the selected seasons and leagues.

Return type

pd.DataFrame

`read_shot_events(match_id=None, force_cache=False)`

Retrieve shooting data for the selected seasons or selected matches.

The data returned includes who took the shot, when, with which body part and from how far away. Additionally, the player creating the chance and also the creation before this are included in the data.

Parameters

- **match_id** (*int or list of int, optional*) – Retrieve the shots for a specific game.
- **force_cache** (*bool*) – By default no cached data is used to scrape the list of available games for the current season. If True, will force the use of cached data anyway.

Raises

ValueError – If no games with the given IDs were found for the selected seasons and leagues.

Return type

pd.DataFrame.

classmethod available_leagues()

Return a list of league IDs available for this source.

Return type

list[str]

5.4 MatchHistory

```
class soccerdata.MatchHistory(leagues=None, seasons=None, proxy=None, no_cache=False,  
                               no_store=False,  
                               data_dir=PosixPath('/home/docs/soccerdata/data/MatchHistory'))
```

Provides pd.DataFrames from CSV files available at <http://www.football-data.co.uk/data.php>.

Data will be downloaded as necessary and cached locally in `~/soccerdata/data/MatchHistory`.

Parameters

- **leagues** (*string or iterable*) – IDs of leagues to include.
- **seasons** (*string, int or list*) – Seasons to include. Supports multiple formats. Examples: ‘16-17’; 2016; ‘2016-17’; [14, 15, 16]
- **proxy** (*'tor' or str or dict or list(dict) or callable, optional*) –
Use a proxy to hide your IP address. Valid options are:
 - “tor”: Uses the Tor network. Tor should be running in the background on port 9050.
 - str: The address of the proxy server to use.
 - list(str): A list of proxies to choose from. A different proxy will be selected from this list after failed requests, allowing rotating proxies.
 - callable: A function that returns a valid proxy. This function will be called after failed requests, allowing rotating proxies.
- **no_cache** (*bool*) – If True, will not use cached data.
- **no_store** (*bool*) – If True, will not store downloaded data.
- **data_dir** (*Path, optional*) – Path to directory where data will be cached.

property seasons: list[str]

Return a list of selected seasons.

read_games()

Retrieve game history for the selected leagues and seasons.

Column names are explained here: <http://www.football-data.co.uk/notes.txt>

Return type
pd.DataFrame

classmethod available_leagues()

Return a list of league IDs available for this source.

Return type
list[str]

get(url, filepath=None, max_age=None, no_cache=False, var=None)

Load data from *url*.

By default, the source of *url* is downloaded and saved to *filepath*. If *filepath* exists, the *url* is not visited and the cached data is returned.

Parameters

- **url (str)** – URL to download.
- **filepath (Path, optional)** – Path to save downloaded file. If None, downloaded data is not cached.
- **max_age (int for age in days, or timedelta object)** – The max. age of locally cached file before re-download.
- **no_cache (bool)** – If True, will not use cached data. Overrides the class property.
- **var (str or list of str, optional)** – Return a JavaScript variable instead of the page source.

Raises

TypeError – If max_age is not an integer or timedelta object.

Returns

File-like object of downloaded data.

Return type
io.BufferedIOBase

property leagues: list[str]

Return a list of selected leagues.

5.5 Sofascore

```
class soccerdata.Sofascore(leagues=None, seasons=None, proxy=None, no_cache=False, no_store=False,  
                           data_dir=PosixPath('/home/docs/soccerdata/data/Sofascore'))
```

Provides pd.DataFrames from data available at <http://www.sofascore.com>.

Data will be downloaded as necessary and cached locally in `~/soccerdata/data/Sofascore`.

Parameters

- **leagues (string or iterable, optional)** – IDs of Leagues to include.
- **seasons (string, int or list, optional)** – Seasons to include. Supports multiple formats. Examples: ‘16-17’; 2016; ‘2016-17’; [14, 15, 16]
- **proxy ('tor' or or dict or list(dict) or callable, optional)** –

Use a proxy to hide your IP address. Valid options are:

- “tor”: Uses the Tor network. Tor should be running in the background on port 9050.

- str: The address of the proxy server to use.
 - list(str): A list of proxies to choose from. A different proxy will be selected from this list after failed requests, allowing rotating proxies.
 - callable: A function that returns a valid proxy. This function will be called after failed requests, allowing rotating proxies.
- **no_cache** (*bool*) – If True, will not use cached data.
 - **no_store** (*bool*) – If True, will not store downloaded data.
 - **data_dir** (*Path*) – Path to directory where data will be cached.

`read_leagues()`

Retrieve the selected leagues from the datasource.

Return type

pd.DataFrame

`read_seasons()`

Retrieve the selected seasons for the selected leagues.

Return type

pd.DataFrame

`read_league_table(force_cache=False)`

Retrieve the league table for the selected leagues.

Parameters

force_cache (*bool*) – By default no cached data is used for the current season. If True, will force the use of cached data anyway.

Return type

pd.DataFrame

`read_schedule(force_cache=False)`

Retrieve the game schedule for the selected leagues and seasons.

Parameters

force_cache (*bool*) – By default no cached data is used for the current season. If True, will force the use of cached data anyway.

Return type

pd.DataFrame

5.6 SoFIFA

```
class soccerdata.SoFIFA(leagues=None, versions='latest', proxy=None, no_cache=False, no_store=False,
                           data_dir=PosixPath('/home/docs/soccerdata/data/SoFIFA'))
```

Provides pd.DataFrames from data at <http://sofifa.com>.

Data will be downloaded as necessary and cached locally in `~/soccerdata/data/SoFIFA`.

Parameters

- **leagues** (*string or iterable, optional*) – IDs of leagues to include.

- **versions** (*string, int or list of int, optional*) – FIFA releases to include. Should be specified by their ID used in the URL (e.g., 230034). Alternatively, the string “all” can be used to include all versions and “latest” to include the latest version only. Defaults to “latest”.
- **proxy** ('tor' or or dict or list(dict) or callable, optional) –
Use a proxy to hide your IP address. Valid options are:
 - “tor”: Uses the Tor network. Tor should be running in the background on port 9050.
 - str: The address of the proxy server to use.
 - list(str): A list of proxies to choose from. A different proxy will be selected from this list after failed requests, allowing rotating proxies.
 - callable: A function that returns a valid proxy. This function will be called after failed requests, allowing rotating proxies.
- **no_cache** (*bool*) – If True, will not use cached data.
- **no_store** (*bool*) – If True, will not store downloaded data.
- **data_dir** (*Path*) – Path to directory where data will be cached.

read_leagues()

Retrieve selected leagues from the datasource.

Return type

pd.DataFrame

read_versions(max_age=1)

Retrieve available FIFA releases and rating updates.

Parameters

max_age (*int for age in days, or timedelta object*) – The max. age of the locally cached release history before a new version is downloaded.

Raises

TypeError – If max_age is not an integer or timedelta object.

Return type

pd.DataFrame

read_teams()

Retrieve all teams for the selected leagues.

Return type

pd.DataFrame

read_players(team=None)

Retrieve all players for the selected leagues.

Parameters

team (*str or list of str, optional*) – Team(s) to retrieve. If None, will retrieve all teams.

Raises

ValueError – If no data is found for the given team(s) in the selected leagues.

Return type

pd.DataFrame

classmethod available_leagues()

Return a list of league IDs available for this source.

Return type

list[str]

read_team_ratings()

Retrieve ratings for all teams in the selected leagues.

Return type

pd.DataFrame

read_player_ratings(team=None, player=None)

Retrieve ratings for players.

Parameters

- **team** (*str or list of str, optional*) – Team(s) to retrieve. If None, will retrieve all teams.
- **player** (*int or list of int, optional*) – Player(s) to retrieve. If None, will retrieve all players.

Return type

pd.DataFrame

5.7 Understat

```
class soccerdata.Understat(leagues=None, seasons=None, proxy=None, no_cache=False, no_store=False,  
                           data_dir=PosixPath('/home/docs/soccerdata/data/Understat'))
```

Provides pd.DataFrames from data at <https://understat.com>.

Data will be downloaded as necessary and cached locally in ~/soccerdata/data/Understat.

Parameters

- **proxy** ('tor' or dict or list(dict) or callable, optional) –
Use a proxy to hide your IP address. Valid options are:
 - "tor": Uses the Tor network. Tor should be running in the background on port 9050.
 - str: The address of the proxy server to use.
 - list(str): A list of proxies to choose from. A different proxy will be selected from this list after failed requests, allowing rotating proxies.
 - callable: A function that returns a valid proxy. This function will be called after failed requests, allowing rotating proxies.
- **no_cache** (bool) – If True, will not use cached data.
- **no_store** (bool) – If True, will not store downloaded data.
- **data_dir** (Path) – Path to directory where data will be cached.
- **leagues** (str / list[str] / None)
- **seasons** (str / int / Iterable[str / int] / None)

property seasons: list[str]

Return a list of selected seasons.

read_leagues()

Retrieve the selected leagues from the datasource.

Return type

pd.DataFrame

read_seasons()

Retrieve the selected seasons from the datasource.

Return type

pd.DataFrame

read_schedule(include_matches_without_data=True, force_cache=False)

Retrieve the matches for the selected leagues and seasons.

Parameters

- **include_matches_without_data (bool)** – By default matches with and without data are returned. If False, will only return matches with data.
- **force_cache (bool)** – By default no cached data is used for the current season. If True, will force the use of cached data anyway.

Return type

pd.DataFrame

read_team_match_stats(force_cache=False)

Retrieve the team match stats for the selected leagues and seasons.

Parameters

force_cache (bool) – By default no cached data is used for the current season. If True, will force the use of cached data anyway.

Return type

pd.DataFrame

read_player_season_stats(force_cache=False)

Retrieve the player season stats for the selected leagues and seasons.

Parameters

force_cache (bool) – By default no cached data is used for the current season. If True, will force the use of cached data anyway.

Return type

pd.DataFrame

read_player_match_stats(match_id=None)

Retrieve the player match stats for the selected leagues and seasons.

Parameters

match_id(int or list of int, optional) – Retrieve the player match stats for a specific match.

Raises

ValueError – If the given match_id could not be found in the selected seasons.

Return type

pd.DataFrame

read_shot_events(*match_id=None*)

Retrieve the shot events for the selected matches or the selected leagues and seasons.

Parameters

match_id (*int or list of int, optional*) – Retrieve the shot events for a specific match.

Raises

ValueError – If the given *match_id* could not be found in the selected seasons.

Return type

pd.DataFrame

classmethod available_leagues()

Return a list of league IDs available for this source.

Return type

list[str]

get(*url, filepath=None, max_age=None, no_cache=False, var=None*)

Load data from *url*.

By default, the source of *url* is downloaded and saved to *filepath*. If *filepath* exists, the *url* is not visited and the cached data is returned.

Parameters

- **url** (*str*) – URL to download.
- **filepath** (*Path, optional*) – Path to save downloaded file. If None, downloaded data is not cached.
- **max_age** (*int for age in days, or timedelta object*) – The max. age of locally cached file before re-download.
- **no_cache** (*bool*) – If True, will not use cached data. Overrides the class property.
- **var** (*str or list of str, optional*) – Return a JavaScript variable instead of the page source.

Raises

TypeError – If *max_age* is not an integer or timedelta object.

Returns

File-like object of downloaded data.

Return type

io.BufferedIOBase

property leagues: list[str]

Return a list of selected leagues.

5.8 WhoScored

```
class soccerdata.WhoScored(leagues=None, seasons=None, proxy=None, no_cache=False, no_store=False,
                             data_dir=PosixPath('/home/docs/soccerdata/data/WhoScored'),
                             path_to_browser=None, headless=False)
```

Provides pd.DataFrames from data available at <http://whoscored.com>.

Data will be downloaded as necessary and cached locally in `~/soccerdata/data/WhoScored`.

Parameters

- **leagues** (*string or iterable, optional*) – IDs of Leagues to include.
- **seasons** (*string, int or list, optional*) – Seasons to include. Supports multiple formats. Examples: ‘16-17’; 2016; ‘2016-17’; [14, 15, 16]
- **proxy** (*'tor' or str or dict or list(dict) or callable, optional*) –

Use a proxy to hide your IP address. Valid options are:

- “tor”: Uses the Tor network. Tor should be running in the background on port 9050.
 - str: The address of the proxy server to use.
 - list(str): A list of proxies to choose from. A different proxy will be selected from this list after failed requests, allowing rotating proxies.
 - callable: A function that returns a valid proxy. This function will be called after failed requests, allowing rotating proxies.
- **no_cache** (*bool*) – If True, will not use cached data.
 - **no_store** (*bool*) – If True, will not store downloaded data.
 - **data_dir** (*Path*) – Path to directory where data will be cached.
 - **path_to_browser** (*Path, optional*) – Path to the Chrome executable.
 - **headless** (*bool, default: True*) – If True, will run Chrome in headless mode. Setting this to False might help to avoid getting blocked. Only supported for Selenium <4.13.

```
read_schedule(force_cache=False)
```

Retrieve the game schedule for the selected leagues and seasons.

Parameters

force_cache (*bool*) – By default no cached data is used for the current season. If True, will force the use of cached data anyway.

Return type

`pd.DataFrame`

```
read_missing_players(match_id=None, force_cache=False)
```

Retrieve a list of injured and suspended players ahead of each game.

Parameters

- **match_id** (*int or list of int, optional*) – Retrieve the missing players for a specific game.
- **force_cache** (*bool*) – By default no cached data is used to scrapre the list of available games for the current season. If True, will force the use of cached data anyway.

Raises

ValueError – If the given match_id could not be found in the selected seasons.

Return type

pd.DataFrame

read_events(*match_id=None, force_cache=False, live=False, output_fmt='events', retry_missing=True, on_error='raise'*)

Retrieve the event data for each game in the selected leagues and seasons.

Parameters

- **match_id** (*int or list of int, optional*) – Retrieve the event stream for a specific game.
- **force_cache** (*bool*) – By default no cached data is used to scrape the list of available games for the current season. If True, will force the use of cached data anyway.
- **live** (*bool*) – If True, will not return a cached copy of the event data. This is useful to scrape live data.
- **output_fmt** (*str, default: 'events'*) –

The output format of the returned data. Possible values are:

- 'events' (default): Returns a dataframe with all events.
 - 'raw': Returns the original unformatted WhoScored JSON.
 - 'spadl': Returns a dataframe with the SPADL representation of the original events. See <https://socceraction.readthedocs.io/en/latest/documentation/SPADL.html#spadl>
 - 'atomic-spadl': Returns a dataframe with the Atomic-SPADL representation of the original events. See <https://socceraction.readthedocs.io/en/latest/documentation/SPADL.html#atomic-spadl>
 - 'loader': Returns a socceraction.data.opta.OptaLoader instance, which can be used to retrieve the actual data. See <https://socceraction.readthedocs.io/en/latest/modules/generated/socceraction.data.opta.OptaLoader.html#socceraction.data.opta.OptaLoader>
 - None: Doesn't return any data. This is useful to just cache the data without storing the events in memory.
- **retry_missing** (*bool*) – If no events were found for a game in a previous attempt, will retry to scrape the events
 - **on_error** ("raise" or "skip", *default: "raise"*) – Whether to raise an exception or to skip the game if an error occurs.

Raises

- **ValueError** – If the given match_id could not be found in the selected seasons.
- **ConnectionError** – If the match page could not be retrieved.
- **ImportError** – If the requested output format is 'spadl', 'atomic-spadl' or 'loader' but the socceraction package is not installed.

Return typeSee the description of the `output_fmt` parameter.**classmethod available_leagues()**

Return a list of league IDs available for this source.

Return type

list[str]

If you would like to extend the functionality of soccerdata, you might also be interested in the following modules:

5.9 Base Readers

The logic for downloading data from the web is implemented in the base classes that are documented here. The base classes are not intended to be used directly but rather to be subclassed by the specific readers which implement the logic to parse the data.

The `BaseRequestsReader` is a wrapper around the `requests` library and is used by scrapers that do not require JavaScript to be executed. The `BaseSeleniumReader` is a wrapper around the `selenium` library and is used by scrapers that require JavaScript to be executed.

```
class soccerdata._common.BaseRequestsReader(leagues=None, proxy=None, no_cache=False,
                                             no_store=False,
                                             data_dir=PosixPath('/home/docs/soccerdata/data'),
                                             headers=None)
```

Base class for readers that use the Python requests module.

Parameters

- `leagues` (`str` / `list[str]` / `None`)
- `proxy` (`str` / `list[str]` / `Callable[[], str]` / `None`)
- `no_cache` (`bool`)
- `no_store` (`bool`)
- `data_dir` (`Path`)
- `headers` (`dict[str, str]` / `None`)

```
classmethod available_leagues()
```

Return a list of league IDs available for this source.

Return type

`list[str]`

```
get(url, filepath=None, max_age=None, no_cache=False, var=None)
```

Load data from `url`.

By default, the source of `url` is downloaded and saved to `filepath`. If `filepath` exists, the `url` is not visited and the cached data is returned.

Parameters

- `url` (`str`) – URL to download.
- `filepath` (`Path`, `optional`) – Path to save downloaded file. If `None`, downloaded data is not cached.
- `max_age` (`int for age in days, or timedelta object`) – The max. age of locally cached file before re-download.
- `no_cache` (`bool`) – If `True`, will not use cached data. Overrides the class property.
- `var` (`str or list of str, optional`) – Return a JavaScript variable instead of the page source.

Raises

`TypeError` – If `max_age` is not an integer or `timedelta` object.

Returns

File-like object of downloaded data.

Return type

io.BufferedIOBase

property leagues: list[str]

Return a list of selected leagues.

property seasons: list[str]

Return a list of selected seasons.

```
class soccerdata._common.BaseSeleniumReader(leagues=None, proxy=None, no_cache=False,
                                              no_store=False,
                                              data_dir=PosixPath('/home/docs/soccerdata/data'),
                                              path_to_browser=None, headless=True)
```

Base class for readers that use Selenium.

Parameters

- **leagues** (str / list[str] / None)
- **proxy** (str / list[str] / Callable[], str] / None)
- **no_cache** (bool)
- **no_store** (bool)
- **data_dir** (Path)
- **path_to_browser** (Path / None)
- **headless** (bool)

classmethod available_leagues()

Return a list of league IDs available for this source.

Return type

list[str]

get(url, filepath=None, max_age=None, no_cache=False, var=None)

Load data from *url*.

By default, the source of *url* is downloaded and saved to *filepath*. If *filepath* exists, the *url* is not visited and the cached data is returned.

Parameters

- **url** (str) – URL to download.
- **filepath** (Path, optional) – Path to save downloaded file. If None, downloaded data is not cached.
- **max_age** (int for age in days, or timedelta object) – The max. age of locally cached file before re-download.
- **no_cache** (bool) – If True, will not use cached data. Overrides the class property.
- **var** (str or list of str, optional) – Return a JavaScript variable instead of the page source.

Raises

TypeError – If max_age is not an integer or timedelta object.

Returns

File-like object of downloaded data.

Return type

io.BufferedIOBase

property leagues: list[str]

Return a list of selected leagues.

property seasons: list[str]

Return a list of selected seasons.

5.10 Utilities

_common.SeasonCode(names=None, *, module=None, qualname=None, type=None, start=1, boundary=None)

How to interpret season codes.

SINGLE_YEAR

Type

The season code is a single year, e.g. ‘2021’.

soccerdata.MULTI_YEAR

Type

The season code is a range of years, e.g. ‘2122’.

_common.make_game_id()

Return a game id based on date, home and away team.

Parameters

row (Series)

Return type

str

_common.standardize_colnames(cols=None)

Convert DataFrame column names to snake case.

Parameters

- **df (DataFrame)**
- **cols (list[str] / None)**

Return type

DataFrame

_common.get_proxy()

Return a public proxy.

Return type

dict[str, str]

_common.check_proxy()

Check if proxy is working.

Parameters

proxy (dict)

Return type

bool

FAQ**Is web scraping legal?**

Even though web scraping is ubiquitous, its legal status remains unclear. That is because whether web scraping is legal will depend on many aspects. It is always best to consult with a lawyer or legal expert to ensure that your web scraping activities are legal and comply with all applicable laws and regulations.

Something doesn't work

1. Have you updated to the newest version of soccerdata?
2. Clear the cache or run your script without caching enabled.
3. Does the log produce any warnings that sound like they might be related? Maybe the data you are looking for is not available or can not be processed correctly.
4. Open an issue on GitHub.

CONTRIBUTOR GUIDE

This document lays out guidelines and advice for contributing to this project. If you're thinking of contributing, please start by reading this document and getting a feel for how contributing to this project works. If you have any questions, feel free to reach out to [Pieter Robberechts](#), the primary maintainer.

The guide is split into sections based on the type of contribution you're thinking of making.

7.1 Bug Reports

Bug reports are hugely important! Before you raise one, though, please check through the [GitHub issues](#), **both open and closed**, to confirm that the bug hasn't been reported before.

When filing an issue, make sure to answer these questions:

- Which Python version are you using?
- Which version of soccerdata are you using?
- What did you do?
- What did you expect to see?
- What did you see instead?

The best way to get your bug fixed is to provide a test case, and/or steps to reproduce the issue.

7.2 Feature Requests

If you believe there is a feature missing, feel free to raise a feature request on the [Issue Tracker](#).

7.3 Documentation Contributions

Documentation improvements are always welcome! The documentation files live in the `docs/` directory of the code-base. They're written in [reStructuredText](#), and use [Sphinx](#) to generate the full suite of documentation.

You do not have to set up a development environment to make small changes to the docs. Instead, you can edit files directly on [GitHub](#) and suggest changes.

When contributing documentation, please do your best to follow the style of the documentation files. This means a soft-limit of 79 characters wide in your text files and a semi-formal, yet friendly and approachable, prose style.

When presenting Python code, use double-quoted strings ("hello" instead of 'hello').

7.4 Code Contributions

If you intend to contribute code, do not feel the need to sit on your contribution until it is perfectly polished and complete. It helps everyone involved for you to seek feedback as early as you possibly can. Submitting an early, unfinished version of your contribution for feedback can save you from putting a lot of work into a contribution that is not suitable for the project.

7.4.1 Setting up your development environment

To test out code changes, you'll need to set up a Python environment with all required dependencies. It's recommended to use `uv` to set up your environment. On macOS and Linux, you can use `curl` to download the installation script and execute it:

```
$ curl -LsSf https://astral.sh/uv/install.sh | sh
```

Information about installation options and instructions for Windows can be found in `uv`'s [installation guide](#).

Next, create and activate a virtual environment using `uv` with a Python version that soccerdata supports:

```
$ uv venv --python 3.9
$ source .venv/bin/activate
```

Alternatively, you can use `pip`. You'll need to have at least the minimum Python version that soccerdata supports. Next, install the required dependencies in a virtual environment as follows:

```
$ python3 -m venv .venv
$ source .venv/bin/activate
$ python -m pip install -e .
$ python -m pip install -r requirements.txt
```

7.4.2 Steps for submitting Code

When contributing code, you'll want to follow this checklist:

1. Fork the repository on GitHub.
2. Run the tests to confirm they all pass on your system. If they don't, you'll need to investigate why they fail. If you're unable to diagnose this yourself, raise it as a bug report.
3. Write tests that demonstrate your bug or feature. Ensure that they fail.
4. Make your change.
5. Run the entire test suite again, confirming that all tests pass *including the ones you just added*.
6. Make sure your code follows the code style discussed below.
7. Send a GitHub Pull Request to the main repository's `master` branch. GitHub Pull Requests are the expected method of code collaboration on this project.

7.4.3 Testing the project

Run the full test suite:

```
$ make test
```

You can also run tests for a specific data source. For example, invoke the unit test suite like this to run tests for ClubElo only:

```
$ make test-class clubelo
```

Unit tests are located in the `tests` directory, and are written using the `pytest` testing framework.

7.4.4 Code style

The soccerdata codebase uses the [PEP 8](#) code style. In addition, we have a few guidelines:

- Line-length can exceed 79 characters, to 100, when convenient.
- Line-length can exceed 100 characters, when doing otherwise would be *terribly* inconvenient.
- Always use double-quoted strings (e.g. `"#soccer"`), unless a double-quote occurs within the string.

To ensure all code conforms to this format. You can format the code using the pre-commit hooks.

```
$ make pre-commit-test
```

Docstrings are to follow the [numpydoc](#) guidelines.

7.4.5 Submitting changes

Open a [pull request](#) to submit changes to this project.

Your pull request needs to meet the following guidelines for acceptance:

- The test suite must pass without errors and warnings.
- Include unit tests.
- If your changes add functionality, update the documentation accordingly.

Feel free to submit early, though. We can always iterate on this.

To run linting and code formatting checks before committing your change, you can install pre-commit as a Git hook by running the following command:

```
$ make pre-commit-test
```

To automatically run the pre-commit checks before committing, you can run

```
$ make pre-commit-update
```

It is recommended to open an issue before starting work on anything.

**CHAPTER
EIGHT**

APACHE LICENSE

Copyright (c) 2021 Pieter Robberechts

Licensed under the Apache License, Version 2.0 (the “License”); you may not use this file except in compliance with the License. You may obtain a copy of the License at

<http://www.apache.org/licenses/LICENSE-2.0>

Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an “AS IS” BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

This file incorporates code of the `footballdata` software package covered by the following copyright and permission notice:

Copyright (c) 2017 skagr

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the “Software”), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

INDEX

A

available_leagues() (soccerdata._common.BaseRequestsReader method), 128
available_leagues() (soccerdata._common.BaseSeleniumReader method), 129
available_leagues() (soccerdata.ClubElo method), 112
available_leagues() (soccerdata.ESPN method), 114
available_leagues() (soccerdata.FBref method), 119
available_leagues() (soccerdata.MatchHistory method), 120
available_leagues() (soccerdata.SoFIFA method), 122
available_leagues() (soccerdata.Understat method), 125
available_leagues() (soccerdata.WhoScored method), 127

get() (soccerdata._common.BaseSeleniumReader method), 129
get() (soccerdata.ClubElo method), 112
get() (soccerdata.ESPN method), 114
get() (soccerdata.MatchHistory method), 120
get() (soccerdata.Understat method), 125
get_proxy() (soccerdata._common method), 130

L

leagues (soccerdata._common.BaseRequestsReader property), 129
leagues (soccerdata._common.BaseSeleniumReader property), 130
leagues (soccerdata.ClubElo property), 112
leagues (soccerdata.ESPN property), 114
leagues (soccerdata.MatchHistory property), 120
leagues (soccerdata.Understat property), 125

M

make_game_id() (soccerdata._common method), 130
MatchHistory (class in soccerdata), 119
MULTI_YEAR (in module soccerdata), 130

R

read_by_date() (soccerdata.ClubElo method), 111
read_events() (soccerdata.FBref method), 118
read_events() (soccerdata.WhoScored method), 127
read_games() (soccerdata.MatchHistory method), 119
read_league_table() (soccerdata.Sofascore method), 121
read_leagues() (soccerdata.FBref method), 115
read_leagues() (soccerdata.Sofascore method), 121
read_leagues() (soccerdata.SoFIFA method), 122
read_leagues() (soccerdata.Understat method), 124
read_lineup() (soccerdata.ESPN method), 114
read_lineup() (soccerdata.FBref method), 118
read_matchesheet() (soccerdata.ESPN method), 113
read_missing_players() (soccerdata.WhoScored method), 126
read_player_match_stats() (soccerdata.FBref method), 117

B

BaseRequestsReader (class in soccerdata._common), 128
BaseSeleniumReader (class in soccerdata._common), 129

C

check_proxy() (soccerdata._common method), 130
ClubElo (class in soccerdata), 111

E

ESPN (class in soccerdata), 113

F

FBref (class in soccerdata), 115

G

get() (soccerdata._common.BaseRequestsReader method), 128

read_player_match_stats() (*soccerdata.Understat method*), 124
read_player_ratings() (*soccerdata.SoFIFA method*), 123
read_player_season_stats() (*soccerdata.FBref method*), 117
read_player_season_stats() (*soccerdata.Understat method*), 124
read_players() (*soccerdata.SoFIFA method*), 122
read_schedule() (*soccerdata.ESPN method*), 113
read_schedule() (*soccerdata.FBref method*), 117
read_schedule() (*soccerdata.Sofascore method*), 121
read_schedule() (*soccerdata.Understat method*), 124
read_schedule() (*soccerdata.WhoScored method*), 126
read_seasons() (*soccerdata.FBref method*), 115
read_seasons() (*soccerdata.Sofascore method*), 121
read_seasons() (*soccerdata.Understat method*), 124
read_shot_events() (*soccerdata.FBref method*), 118
read_shot_events() (*soccerdata.Understat method*), 124
read_team_history() (*soccerdata.ClubElo method*), 112
read_team_match_stats() (*soccerdata.FBref method*), 116
read_team_match_stats() (*soccerdata.Understat method*), 124
read_team_ratings() (*soccerdata.SoFIFA method*), 123
read_team_season_stats() (*soccerdata.FBref method*), 115
read_teams() (*soccerdata.SoFIFA method*), 122
read_versions() (*soccerdata.SoFIFA method*), 122

S

SeasonCode() (*soccerdata._common method*), 130
seasons (*soccerdata._common.BaseRequestsReader property*), 129
seasons (*soccerdata._common.BaseSeleniumReader property*), 130
seasons (*soccerdata.ClubElo property*), 113
seasons (*soccerdata.ESPN property*), 113
seasons (*soccerdata.MatchHistory property*), 119
seasons (*soccerdata.Understat property*), 123
SINGLE_YEAR (*soccerdata._common attribute*), 130
Sofascore (*class in soccerdata*), 120
SoFIFA (*class in soccerdata*), 121
standardize_colnames() (*soccerdata._common method*), 130

U

Understat (*class in soccerdata*), 123

W

WhoScored (*class in soccerdata*), 126