
soccerdata

Release 1.9.0

Pieter Robberechts

Jun 12, 2026

CONTENTS

1	Getting Started	3
2	Overview of Data Sources	7
3	How-to Guides	93
4	Data Analysis Examples	97
5	API Reference	105
6	FAQ	125
7	Contributor Guide	127
8	Apache License	131
	Index	133

Release v1.9.0. (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!*

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 *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	SoT	Sh/	SoT	G/9	G/9	Dis	FK	PK	PK%	xG	np	np	G-xG	np:G-xG
EN0	202	Ar-	29	38.0	53	455	141	31.0	11.9	3.71	0.1	0.33	16.9	23	6	6	53.4	49.0	0.11	-	-
Prei		se-																		0.5	2.0
Lea		nal																			
		As-	24	38.0	52	518	179	34.6	13.6	4.71	0.09	0.26	16.4	15	5	6	52.9	48.4	0.1	-	-
		ton																		0.9	1.5
		Vill																			
		Bri	27	38.0	39	476	129	27.1	12.4	3.39	0.07	0.26	16.6	14	6	9	51.6	44.8	0.1	-	-
		g																		12.6	11.8
		Bur	25	38.0	32	383	125	32.6	10.0	3.29	0.08	0.23	16.6	15	3	3	39.9	37.6	0.1	-	-
		ley																		7.9	8.6
		Che	27	38.0	56	553	194	35.1	14.4	5.11	0.09	0.24	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 re-download 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:

SOCCERDATA_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.

SOCCERDATA_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 `SOCCERDATA_NOSTORE` is the option to use. By default, data is retrieved from the cache.

SOCCERDATA_NOSTORE

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

SOCCERDATA_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.

SOCCERDATA_LOGLEVEL

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

Example:

```
# bash
export SOCCERDATA_DIR = "~/soccerdata"
export SOCCERDATA_NOCACHE = "False"
export SOCCERDATA_NOSTORE = "False"
export SOCCERDATA_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 `SOCCERDATA_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

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

(continues on next page)

(continued from previous page)

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()
```

```
[4]:
```

team	rank	country	level	elo	from	to	\
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]:
```

from	rank	team	country	level	elo	to
------	------	------	---------	-------	-----	----

(continues on next page)

(continued from previous page)

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 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

(continues on next page)

(continued from previous page)

```
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()
```

```
[4]:
↳ date \
league          season game
ENG-Premier League 2021  2019-08-09 Liverpool-Norwich City      2019-08-09 19:00:
↳00+00:00
                    2019-08-10 AFC Bournemouth-Sheffield United 2019-08-10 14:00:
↳00+00:00
                    2019-08-10 Burnley-Southampton              2019-08-10 14:00:
↳00+00:00
                    2019-08-10 Crystal Palace-Everton            2019-08-10 14:00:
↳00+00:00
                    2019-08-10 Tottenham Hotspur-Aston Villa    2019-08-10 16:30:
↳00+00:00

                    home_team_
↳ \
league          season game
ENG-Premier League 2021  2019-08-09 Liverpool-Norwich City      Liverpool
                    2019-08-10 AFC Bournemouth-Sheffield United  AFC Bournemouth
                    2019-08-10 Burnley-Southampton              Burnley
                    2019-08-10 Crystal Palace-Everton            Crystal Palace
                    2019-08-10 Tottenham Hotspur-Aston Villa    Tottenham Hotspur

                    away_team_
↳ \
league          season game
ENG-Premier League 2021  2019-08-09 Liverpool-Norwich City      Norwich City
                    2019-08-10 AFC Bournemouth-Sheffield United  Sheffield United
                    2019-08-10 Burnley-Southampton              Southampton
                    2019-08-10 Crystal Palace-Everton            Everton
                    2019-08-10 Tottenham Hotspur-Aston Villa    Aston Villa

                    game_id \
league          season game
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
```

(continues on next page)

(continued from previous page)

league	season	game	league_id
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]:
↳home \
league      season game
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa
                                  Aston Villa      False
                                  West Ham United  True

league      season game
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa
                                  Aston Villa      None
                                  West Ham United  None

↳attendance \
league      season game
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa
                                  Aston Villa
↳ 0
                                  West Ham United
↳ 0

↳capacity \
league      season game
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa
                                  Aston Villa      None
                                  West Ham United  None

↳ roster \
league      season game
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa
                                  Aston Villa      [{
↳ 'active': True, 'starter': True, 'jersey': '...
                                  West Ham United  [{
↳ 'active': True, 'starter': True, 'jersey': '...

↳committed \
league      season game
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa
                                  Aston Villa
↳ 16
                                  West Ham United
```

(continues on next page)

(continued from previous page)

							yellow_
↔	13						
↔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						
							off_
↔offsides \	league	season	game	team			
	ENG-Premier League	2021	2020-07-26	West Ham United-Aston Villa	Aston Villa	↔	
					West Ham United	↔	
						0	
						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		↔	
				West Ham United		↔	
						3	
						0	
							pos_
↔possession_pct \	league	season	game	team			
	ENG-Premier League	2021	2020-07-26	West Ham United-Aston Villa	Aston Villa	↔	
↔	63						
					West Ham United	↔	
↔	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						

(continues on next page)

(continued from previous page)

```

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

```

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
↪Hourihane      False
↪Luiz           False
↪              False
↪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
↪Hourihane      Center Left Midfielder
↪Luiz           Defensive Midfielder
↪              Center Right Defender
↪Guilbert      Right Back
↪      formation_place \
league          season game          team      player
ENG-Premier League 2021  2020-07-26 West Ham United-Aston Villa Aston Villa Anwar El_
↪Ghazi          0
↪Hourihane      10
↪Luiz           4
Ezri Konsa

```

(continues on next page)

(continued from previous page)

↪		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 El		
↪	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 El		
↪	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 El		
↪	Ghazi		1.0					Conor
↪	Hourihane		1.0					Douglas
↪	Luiz		1.0					Ezri Konsa
↪			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					Conor

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

(continues on next page)

(continued from previous page)

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

(continues on next page)

(continued from previous page)

```

↳Luiz      NaN
                                                    Ezri Kansa
↳          NaN
                                                    Frédéric
↳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
↳Ghazi      0.0
                                                    Conor
↳Hourihane      0.0
                                                    Douglas
↳Luiz          0.0
                                                    Ezri Kansa
↳          0.0
                                                    Frédéric
↳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
↳Ghazi      0.0
                                                    Conor
↳Hourihane      0.0
                                                    Douglas
↳Luiz          0.0
                                                    Ezri Kansa
↳          0.0
                                                    Frédéric
↳Guilbert      0.0

```

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

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

```
[3]: import soccerdata as sd
No entry for terminal type "xterm-color";
using dumb terminal settings.
```

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`.
```

(continues on next page)

(continued from previous page)

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.
- 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.

2.3.1 Team season stats

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

```
[5]:
```

			players_used	90s	Standard	\		
					Gls	Sh	SoT	
league	season	team						
ENG-Premier League	2021	Arsenal	29	38	53	453	140	
		Aston Villa	24	38	52	514	178	
		Brighton	27	38	39	478	134	
		Burnley	25	38	32	381	124	
		Chelsea	27	38	56	546	193	

			SoT%	Sh/90	SoT/90	G/Sh	G/SoT	PK
league	season	team						
ENG-Premier League	2021	Arsenal	30.9	11.92	3.68	0.1	0.34	6
		Aston Villa	34.6	13.53	4.68	0.09	0.26	5
		Brighton	28.0	12.58	3.53	0.07	0.25	6

(continues on next page)

(continued from previous page)

league	season	team	PKatt	url
		Burnley	32.5	
		Chelsea	35.3	
		Arsenal	6	/en/squads/18bb7c10/2020-2021/Arsenal-Stats
		Aston Villa	6	/en/squads/8602292d/2020-2021/Aston-Villa-Stats
		Brighton	9	/en/squads/d07537b9/2020-2021/Brighton-and-Hov...
		Burnley	3	/en/squads/943e8050/2020-2021/Burnley-Stats
		Chelsea	10	/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 United-Manchester City
↳ 2020-10-24

↳ time \
league      season team      game      time
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
```

(continues on next page)

(continued from previous page)

				2020-10-24	West Ham United-Manchester City	12:	
↪	30:00						
↪	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			2020-10-24	West Ham United-Manchester City		
↪	Matchweek 6						
↪	day \						
league	season	team	game				
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 United-Manchester City		Sat	
↪	venue \						
league	season	team	game				
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 United-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 United-Manchester City		

(continues on next page)

(continued from previous page)

↪ D					GF
↪ \	league	season	team	game	
	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 United-Manchester City	1
↪ \					GA
↪ \	league	season	team	game	
	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 United-Manchester City	1
↪ 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 United-Manchester City	
			West Ham United		
↪ Poss \					
↪	league	season	team	game	
	ENG-Premier League	2021	Manchester City	2020-09-21 Wolves-Manchester City	
				2020-09-27 Manchester City-Leicester City	
				2020-10-03 Leeds United-Manchester City	
				2020-10-17 Manchester City-Arsenal	
				2020-10-24 West Ham United-Manchester City	
↪ Attendance \					
↪	league	season	team	game	
	ENG-Premier League	2021	Manchester City	2020-09-21 Wolves-Manchester City	

(continues on next page)

(continued from previous page)

↪	<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 United-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 United-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 United-Manchester City	└
↪	4-3-3					└
						Opp└
↪	Formation \					└
	league	season	team	game		
	ENG-Premier League	2021	Manchester City	2020-09-21	Wolves-Manchester City	└
↪	5-3-2			2020-09-27	Manchester City-Leicester City	└
↪	5-4-1			2020-10-03	Leeds United-Manchester City	└
↪	4-1-4-1			2020-10-17	Manchester City-Arsenal	└
↪	4-3-3			2020-10-24	West Ham United-Manchester City	└
↪	5-4-1					└

(continues on next page)

(continued from previous page)

```

↪ 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 United-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/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 United-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 United-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()
```

```
[7]:
           nation  pos age \
league          season team  player
ENG-Premier League 2021 Arsenal Ainsley Maitland-Niles  ENG  MF,DF  22
                               Alexandre Lacazette      FRA   FW  29

```

(continues on next page)

(continued from previous page)

league	season	team	player	born	Playing Time	MP	
			Bernd Leno	GER	GK	28	
			Bukayo Saka	ENG	MF	18	
			Calum Chambers	ENG	DF	25	
league	season	team	player	born	Playing Time	MP	
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	
league	season	team	player	Starts	Min	90s	
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	
league	season	team	player	Performance			
				Gls	Ast	G+A	
ENG-Premier League	2021	Arsenal	Ainsley Maitland-Niles	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	G-PK	PK	PKAtt	CrdY
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	Per 90 Minutes			
				CrdR	Gls		
ENG-Premier League	2021	Arsenal	Ainsley Maitland-Niles	0	0.0		
			Alexandre Lacazette	0	0.61		
			Bernd Leno	1	0.0		
			Bukayo Saka	0	0.18		
			Calum Chambers	0	0.0		
league	season	team	player	Ast	G+A	G-PK	

(continues on next page)

(continued from previous page)

league	season	team	player	G+A-PK
ENG-Premier League	2021	Arsenal	Ainsley Maitland-Niles	0.0
			Alexandre Lacazette	0.09
			Bernd Leno	0.0
			Bukayo Saka	0.11
			Calum Chambers	0.24

2.3.4 Player match stats

```
[9]: player_match_stats = fbref.read_player_match_stats(stat_type="summary", match_id="db261cb0")
player_match_stats.head()
```

```
[9]:
↳ jersey_number \
league          season game          team          player
ENG-Premier League 2021  2020-09-12 Crystal Palace-Southampton Crystal Palace Andros
↳Townsend          10
↳KouyatÃ©          8
↳Eze                25
↳McArthur           18
↳McCarthy           22
↳ nation \
league          season game          team          player
ENG-Premier League 2021  2020-09-12 Crystal Palace-Southampton Crystal Palace Andros
↳Townsend          ENG
↳KouyatÃ©          SEN
↳Eze                ENG
↳McArthur           SCO
↳McCarthy           IRL
```

(continues on next page)

(continued from previous page)

pos \				
league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros
↳Townsend	RM			
↳KouyatÃ©	CB			Cheikhou
↳Eze	FW,MF			Eberechi
↳McArthur	CM			James
↳McCarthy	CM			James
age \				
league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros
↳Townsend	29-058			
↳KouyatÃ©	30-266			Cheikhou
↳Eze	22-075			Eberechi
↳McArthur	32-341			James
↳McCarthy	29-305			James
min \				
league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros
↳Townsend	90			
↳KouyatÃ©	90			Cheikhou
↳Eze	10			Eberechi
↳McArthur	90			James
↳McCarthy	73			James
Performance \				
Gls				
league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros
↳Townsend	0			
				Cheikhou

(continues on next page)

(continued from previous page)

↪KouyatÃ©	0					Eberechi
↪Eze	0					James
↪McArthur	0					James
↪McCarthy	0					
↪		\				
↪		Ast				
league		season	game		team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros	
↪Townsend	1					Cheikhou
↪KouyatÃ©	0					Eberechi
↪Eze	0					James
↪McArthur	0					James
↪McCarthy	0					
↪		\				
↪		PK				
league		season	game		team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros	
↪Townsend	0					Cheikhou
↪KouyatÃ©	0					Eberechi
↪Eze	0					James
↪McArthur	0					James
↪McCarthy	0					
↪		\				
↪		PKatt				
league		season	game		team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros	
↪Townsend	0					Cheikhou
↪KouyatÃ©	0					Eberechi
↪Eze	0					James

(continues on next page)

(continued from previous page)

↪McArthur	0					James
↪McCarthy	0					
↪		\				
↪		Sh				
league		season	game		team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros	
↪Townsend	0					Cheikhou
↪KouyatÃ©	1					Eberechi
↪Eze	0					James
↪McArthur	1					James
↪McCarthy	0					
↪		\				
↪		SoT				
league		season	game		team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros	
↪Townsend	0					Cheikhou
↪KouyatÃ©	1					Eberechi
↪Eze	0					James
↪McArthur	0					James
↪McCarthy	0					
↪		\				
↪		CrdY				
league		season	game		team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros	
↪Townsend	0					Cheikhou
↪KouyatÃ©	0					Eberechi
↪Eze	0					James
↪McArthur	0					James
↪McCarthy	1					

(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
↳Townsend	1			
				Cheikhou
↳KouyatÃ©	0			Eberechi
↳Eze	0			James
↳McArthur	0			James
↳McCarthy	0			
↳	\			
↳	Crs			
league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros
↳Townsend	3			
				Cheikhou
↳KouyatÃ©	0			Eberechi
↳Eze	3			James
↳McArthur	0			James
↳McCarthy	4			
↳	\			
↳	TklW			
league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros
↳Townsend	0			
				Cheikhou
↳KouyatÃ©	0			Eberechi
↳Eze	0			James
↳McArthur	1			James
↳McCarthy	1			
↳	\			
↳	Int			
league	season	game	team	player
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace Andros
↳Townsend	1			
				Cheikhou

(continues on next page)

(continued from previous page)

↪KouyatÃ©	0					Eberechi
↪Eze	0					James
↪McArthur	3					James
↪McCarthy	2					
↪	\					
↪	OG					
league		season	game		team	player
ENG-Premier	League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros
↪Townsend	0					Cheikhou
↪KouyatÃ©	0					Eberechi
↪Eze	0					James
↪McArthur	0					James
↪McCarthy	0					
↪	\					
↪	PKwon					
league		season	game		team	player
ENG-Premier	League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros
↪Townsend	<NA>					Cheikhou
↪KouyatÃ©	<NA>					Eberechi
↪Eze	<NA>					James
↪McArthur	<NA>					James
↪McCarthy	<NA>					
↪	\					
↪	PKcon					
league		season	game		team	player
ENG-Premier	League	2021	2020-09-12	Crystal Palace-Southampton	Crystal Palace	Andros
↪Townsend	<NA>					Cheikhou
↪KouyatÃ©	<NA>					Eberechi
↪Eze	<NA>					James

(continues on next page)

(continued from previous page)

```

↪McArthur      <NA>
↪McCarthy      <NA>

↪
game_id

league          season game          team          player
ENG-Premier League 2021  2020-09-12 Crystal Palace-Southampton Crystal Palace Andros
↪Townsend      db261cb0
↪KouyatÃ©      db261cb0
↪Eze           db261cb0
↪McArthur      db261cb0
↪McCarthy      db261cb0
    
```

2.3.5 Game schedule

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

```
[10]:
league          season game          week \
ENG-Premier League 2021  2020-09-12 Crystal Palace-Southampton      1
                                2020-09-12 Fulham-Arsenal                1
                                2020-09-12 Liverpool-Leeds United          1
                                2020-09-12 West Ham United-Newcastle United  1
                                2020-09-13 Tottenham Hotspur-Everton        1

league          season game          day \
ENG-Premier League 2021  2020-09-12 Crystal Palace-Southampton      Sat
                                2020-09-12 Fulham-Arsenal                Sat
                                2020-09-12 Liverpool-Leeds United          Sat
                                2020-09-12 West Ham United-Newcastle United  Sat
                                2020-09-13 Tottenham Hotspur-Everton        Sun

league          season game          date \
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 United-Newcastle United  2020-09-12
                                2020-09-13 Tottenham Hotspur-Everton        2020-09-13

league          season game          time \
ENG-Premier League 2021  2020-09-12 Crystal Palace-Southampton      15:00 (16:00)
                                2020-09-12 Fulham-Arsenal                12:30 (13:30)
    
```

(continues on next page)

(continued from previous page)

		2020-09-12	Liverpool-Leeds United	17:30 (18:30)
		2020-09-12	West Ham United-Newcastle United	20:00 (21:00)
		2020-09-13	Tottenham Hotspur-Everton	16:30 (17:30)
				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 United-Newcastle United	West Ham United
			2020-09-13 Tottenham Hotspur-Everton	Tottenham Hotspur
				score \
	league	season	game	
	ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	1^a€"0
			2020-09-12 Fulham-Arsenal	0^a€"3
			2020-09-12 Liverpool-Leeds United	4^a€"3
			2020-09-12 West Ham United-Newcastle United	0^a€"2
			2020-09-13 Tottenham Hotspur-Everton	0^a€"1
				away_team_
→ \	league	season	game	
	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 United	Newcastle United
			2020-09-13 Tottenham Hotspur-Everton	Everton
				attendance \
	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 United-Newcastle United	<NA>
			2020-09-13 Tottenham Hotspur-Everton	<NA>
→ venue \	league	season	game	
	ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	→
→ Selhurst Park			2020-09-12 Fulham-Arsenal	Craven_
→ Cottage			2020-09-12 Liverpool-Leeds United	→
→ Anfield			2020-09-12 West Ham United-Newcastle United	London_
→ Stadium			2020-09-13 Tottenham Hotspur-Everton	Tottenham Hotspur_
→ Stadium				

(continues on next page)

(continued from previous page)

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 United-Newcastle United	Stuart Attwell	
		2020-09-13 Tottenham Hotspur-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/Fulham-Arsenal-September-...		
		2020-09-12 Liverpool-Leeds United	/en/matches/	
		↳ 21b58926/Liverpool-Leeds-United-Se...		
		2020-09-12 West Ham United-Newcastle United	/en/matches/	
		↳ 78495ced/West-Ham-United-Newcastle...		
		2020-09-13 Tottenham Hotspur-Everton	/en/matches/	
		↳ fc7f9aa1/Tottenham-Hotspur-Everton...		

league	season	game	notes	\
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 United-Newcastle United	<NA>	
		2020-09-13 Tottenham Hotspur-Everton	<NA>	

league	season	game	game_id
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 United-Newcastle United	78495ced
		2020-09-13 Tottenham Hotspur-Everton	fc7f9aa1

2.3.6 Line ups

```
[11]: lineups = fbref.read_lineup(match_id="db261cb0")
lineups.head()
```

```
[11]:
```

league	season	game	jersey_number	\	player	\
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			

(continues on next page)

(continued from previous page)

```

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
                               2020-09-12 Crystal Palace-Southampton  Crystal Palace

```

```

                               is_starter \
league          season game
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
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
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

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

```
[12]:
                               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  Southampton
                               2020-09-12 Crystal Palace-Southampton  Southampton
                               2020-09-12 Crystal Palace-Southampton  Crystal Palace

```

(continues on next page)

(continued from previous page)

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

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

league	season	game	player2	\
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	

league	season	game	event_type
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

..... 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]:
```

league	season	game	date	\
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	

league	season	game	home_team	\
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)

league	season	game	away_team	\			
		2020-02-11 Fulham-West Brom	Fulham				
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
				B365H	B365D	\	
league	season	game					
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	2.45	3.5		
		2020-01-11	Man United-Arsenal	2.00	3.6		
		2020-01-11	Newcastle-Everton	3.70	3.4		
		2020-01-11	Tottenham-Brighton	1.61	4.0		
		2020-02-11	Fulham-West Brom	2.40	3.3		
				B365A	BWH	\	
league	season	game					
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	2.80	2.35		
		2020-01-11	Man United-Arsenal	3.60	2.00		
		2020-01-11	Newcastle-Everton	2.05	3.75		
		2020-01-11	Tottenham-Brighton	5.50	1.62		
		2020-02-11	Fulham-West Brom	3.00	2.40		
				BWD	BWA	\	
league	season	game					
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	3.60	2.85		
		2020-01-11	Man United-Arsenal	3.70	3.60		
		2020-01-11	Newcastle-Everton	3.40	2.05		
		2020-01-11	Tottenham-Brighton	4.25	5.00		
		2020-02-11	Fulham-West Brom	3.30	3.00		
				IWH	IWD	\	
league	season	game					
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	2.50	3.35		
		2020-01-11	Man United-Arsenal	2.05	3.45		
		2020-01-11	Newcastle-Everton	3.55	3.30		
		2020-01-11	Tottenham-Brighton	1.65	3.90		
		2020-02-11	Fulham-West Brom	2.45	3.10		
				IWA	PSH	\	
league	season	game					
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	2.75	2.53		
		2020-01-11	Man United-Arsenal	3.55	2.05		
		2020-01-11	Newcastle-Everton	2.10	3.82		
		2020-01-11	Tottenham-Brighton	5.00	1.65		
		2020-02-11	Fulham-West Brom	3.00	2.47		
				PSD	PSA	\	
league	season	game					
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	3.72	2.81		
		2020-01-11	Man United-Arsenal	3.78	3.73		
		2020-01-11	Newcastle-Everton	3.56	2.09		
		2020-01-11	Tottenham-Brighton	4.26	5.54		

(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		
					PAHA	MaxAHH	\	
league	season	game						
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						
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						
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						
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						
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						
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						

(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	
				AvgCH	AvgCD	\
league	season	game				
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	
				AvgCA		\
league	season	game				
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		
				B365C>2.5		\
league	season	game				
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		
				B365C<2.5		\
league	season	game				
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		
				PC>2.5	PC<2.5	\
league	season	game				
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	
				MaxC>2.5		\
league	season	game				
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		
				MaxC<2.5		\
league	season	game				

(continues on next page)

(continued from previous page)

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		
					AvgC>2.5	\
league	season	game				
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		
					AvgC<2.5	AHCh \
league	season	game				
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	
						B365CAHH \
league	season	game				
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		
						B365CAHA PCAHH \
league	season	game				
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	
						PCAHA MaxCAHH \
league	season	game				
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	
						MaxCAHA \
league	season	game				
ENG-Premier League	2021	2020-01-11	Aston Villa-Southampton	2.04		
		2020-01-11	Man United-Arsenal	1.92		
		2020-01-11	Newcastle-Everton	2.03		
		2020-01-11	Tottenham-Brighton	1.99		

(continues on next page)

(continued from previous page)

league	season	game	AvgCAHH	AvgCAHA
	2020-02-11	Fulham-West Brom	1.87	
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)
```

```
[2]: %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()
```

```
[5]:
```

	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()
```

```
[8]:
```

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

(continues on next page)

(continued from previous page)

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

(continues on next page)

(continued from previous page)

```

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 = sofifa.read_team_ratings()
team_ratings.head()
```

```
[5]:
```

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	

(continues on next page)

(continued from previous page)

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 = sofifa.read_player_ratings(team="Arsenal")
player_ratings.head()
```

```
[6]:
```

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-Mboyo Sambu 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-Mboyo Sambu 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	\
Aaron Ramsdale	16	18	17	16	57	
Ainsley Maitland-Niles	44	71	64	50	66	
Albert-Mboyo Sambu Lokonga	58	78	71	68	77	
Amario Cozier-Duberry	60	60	53	41	44	
Arthur Okonkwo	7	9	12	20	20	

player	ballcontrol	acceleration	sprints	speed	agility	\
Aaron Ramsdale	36	47	50	53		
Ainsley Maitland-Niles	73	79	77	79		
Albert-Mboyo Sambu Lokonga	81	63	69	78		
Amario Cozier-Duberry	58	72	69	73		
Arthur Okonkwo	17	44	44	40		

player	reactions	balance	shotpower	jumping	stamina	\
Aaron Ramsdale	82	49	65	66	35	
Ainsley Maitland-Niles	70	69	67	66	69	
Albert-Mboyo Sambu Lokonga	72	68	68	71	77	

(continues on next page)

(continued from previous page)

Amario Cozier-Duberry	50	68	56	44	55
Arthur Okonkwo	63	31	47	50	29
	strength longshots aggression interceptions \				
player					
Aaron Ramsdale	59	15	35		24
Ainsley Maitland-Niles	67	62	65		70
Albert-Mboyo Sambu Lokonga	63	63	72		77
Amario Cozier-Duberry	45	45	55		27
Arthur Okonkwo	50	7	20		11
	positioning vision penalties composure \				
player					
Aaron Ramsdale	18	65	25		65
Ainsley Maitland-Niles	69	67	69		70
Albert-Mboyo Sambu Lokonga	65	75	41		80
Amario Cozier-Duberry	59	56	55		55
Arthur Okonkwo	6	35	19		30
	defensiveawareness standingtackle slidingtackle \				
player					
Aaron Ramsdale		20		16	15
Ainsley Maitland-Niles		66		74	72
Albert-Mboyo Sambu Lokonga		64		73	68
Amario Cozier-Duberry		28		28	24
Arthur Okonkwo		13		13	12
	gk_diving gk_handling gk_kicking gk_positioning \				
player					
Aaron Ramsdale	84		79		87
Ainsley Maitland-Niles	15		9		7
Albert-Mboyo Sambu Lokonga	14		7		6
Amario Cozier-Duberry	11		12		9
Arthur Okonkwo	64		65		63
	gk_reflexes				
player					
Aaron Ramsdale					86
Ainsley Maitland-Niles					13
Albert-Mboyo Sambu Lokonga					9
Amario Cozier-Duberry					7
Arthur Okonkwo					66

```
[7]: sofifa.read_teams()
```

```
[7]:
```

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	

(continues on next page)

(continued from previous page)

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		
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")
```

```
[8]:
```

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

(continues on next page)

(continued from previous page)

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 Cărișan	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
222562	Rúnar Alex Rúnarsson	Arsenal	ENG-Premier League
206654	Pablo Marí 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

(continues on next page)

(continued from previous page)

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

(continues on next page)

(continued from previous page)

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_id url
league
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

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

url
```

## 2.7.3 Schedule

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

[6]:

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	\
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	

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	

(continues on next page)

(continued from previous page)

		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
				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
				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 \

(continues on next page)

(continued from previous page)

```

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/
↪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()
```

```
[7]: 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

season_id \
league season game

```

(continues on next page)

(continued from previous page)

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		
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
				away_team_id \
league	season	game		
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

(continues on next page)

(continued from previous page)

		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	
		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	\

(continues on next page)

(continued from previous page)

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 ↵
↵ \			
league	season	game	
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
			home_ppda \
league	season	game	
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
			home_deep_completions
league	season	game	
ENG-Premier League	1516	2015-08-08 Bournemouth-Aston Villa	11
		2015-08-08 Chelsea-Swansea	10

(continues on next page)

(continued from previous page)

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()
```

```
[8]:
```

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	
		Aaron Ramsey	31	2624	5	
		Abdul Rahman Baba	15	1018	0	
		Adam Bogdan	2	180	0	

league	season	player	xg	np_goals	np_xg	\
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	

league	season	player	assists	xa	shots	\
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	

league	season	player	key_passes	yellow_cards	\
--------	--------	--------	------------	--------------	---

(continues on next page)

(continued from previous page)

league	season	player	red_cards	xg_chain	xg_buildup
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

league	season	player	red_cards	xg_chain	xg_buildup
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]:
↳ league_id \
league season game team player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Ashley Westwood
↳ 1
↳ 1
↳ 1
↳ 1
↳ 1
↳ Agbonlahor 1
↳ 1
↳ season_id \
league season game team player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Ashley Westwood
↳ 2015
↳ 2015
↳ 2015
↳ 2015
↳ 2015
↳ Agbonlahor 2015
↳ game_id \
league season game team player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Ashley Westwood
```

(continues on next page)

(continued from previous page)

↪	82						Brad Guzan	↪
↪	82						Carlos Sánchez	↪
↪	82						Ciaran Clark	↪
↪	82						Gabriel	↪
↪	Agbonlahor	82						↪
↪	team_id \							↪
league	season	game			team		player	
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa	Aston Villa		Ashley Westwood	↪
↪	71						Brad Guzan	↪
↪	71						Carlos Sánchez	↪
↪	71						Ciaran Clark	↪
↪	71						Gabriel	↪
↪	Agbonlahor	71						↪
↪	player_id \							↪
league	season	game			team		player	
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa	Aston Villa		Ashley Westwood	↪
↪	669						Brad Guzan	↪
↪	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	Aston Villa		Ashley Westwood	↪
↪	MC						Brad Guzan	↪
↪	GK						Carlos Sánchez	↪
↪	Sub						Ciaran Clark	↪
↪	DC						Gabriel	↪
↪	Agbonlahor	FW						↪

(continues on next page)

(continued from previous page)

position_id \	league	season	game	team	player
	ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Ashley Westwood
↪	9			Aston Villa	Brad Guzan
↪	1			Aston Villa	Carlos Sánchez
↪	17			Aston Villa	Ciaran Clark
↪	3			Aston Villa	Gabriel
↪	Agbonlahor	15			

minutes \	league	season	game	team	player
	ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Ashley Westwood
↪	90			Aston Villa	Brad Guzan
↪	90			Aston Villa	Carlos Sánchez
↪	18			Aston Villa	Ciaran Clark
↪	90			Aston Villa	Gabriel
↪	Agbonlahor	90			

goals \	league	season	game	team	player
	ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Ashley Westwood
↪	0			Aston Villa	Brad Guzan
↪	0			Aston Villa	Carlos Sánchez
↪	0			Aston Villa	Ciaran Clark
↪	0			Aston Villa	Gabriel
↪	Agbonlahor	0			

own_goals \	league	season	game	team	player
	ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Ashley Westwood
↪	0			Aston Villa	Brad Guzan
↪	0			Aston Villa	Carlos Sánchez
↪	0			Aston Villa	Ciaran Clark

(continues on next page)

(continued from previous page)

↪	0							Gabriel	↪
↪	Agbonlahor	0							↪
↪	shots	\							↪
	league		season	game		team		player	
	ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa	Ashley Westwood			↪
↪	0								↪
↪	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	Ashley Westwood			↪
↪	0.0								↪
↪	0.0							Brad Guzan	↪
↪	0.0							Carlos Sánchez	↪
↪	0.0							Ciaran Clark	↪
↪	0.0							Gabriel	↪
↪	Agbonlahor	0.13016							↪
↪	xa	\							↪
	league		season	game		team		player	
	ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa	Ashley Westwood			↪
↪	0.374082								↪
↪	0.0							Brad Guzan	↪
↪	0.0							Carlos Sánchez	↪
↪	0.0							Ciaran Clark	↪
↪	0.0							Gabriel	↪
↪	Agbonlahor	0.113668							↪
↪	xg_chain	\							↪
	league		season	game		team		player	
	ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa	Ashley Westwood			↪
↪	0.131937								↪
↪								Brad Guzan	↪

(continues on next page)

(continued from previous page)

```

↪ 0.0
↪ 0.0
↪ 0.131937
↪Agbonlahor 0.243828

↪ xg_buildup
league season game team player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Ashley Westwood
↪ 0.131937
↪ 0.0
↪ 0.0
↪ 0.131937
↪Agbonlahor 0.0

```

### 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
↪Agbonlahor 1
↪ 1
↪ 1
↪ 1
↪ 1

↪ season_id \
league season game team player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
↪Agbonlahor 2015
↪Agbonlahor 2015
↪ 2015
↪ 2015

```

(continues on next page)

(continued from previous page)

							Leandro Bacuna	↳
↳	2015							↳
↳	game_id \							↳
league	season	game		team		player		
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	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	Aston Villa	Gabriel	↳	
↳Agbonlahor	2015-08-08	18:00:00				Gabriel	↳	
↳Agbonlahor	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	Aston Villa	Gabriel	↳	
↳Agbonlahor	487146					Gabriel	↳	
↳Agbonlahor	487150					Idrissa Gueye	↳	
↳	487136					Idrissa Gueye	↳	
↳	487144					Leandro Bacuna	↳	
↳	487141							↳
↳	team_id \							↳
league	season	game		team		player		
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa	Aston Villa	Gabriel	↳	
↳Agbonlahor	71					Gabriel	↳	
↳Agbonlahor	71							↳

(continues on next page)

(continued from previous page)

↔	71						Idrissa Gueye	↳
↔	71						Idrissa Gueye	↳
↔	71						Leandro Bacuna	↳
↔	player_id \							↳
	league	season	game		team		player	
	ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa		Gabriel	↳
↔	Agbonlahor	890					Gabriel	↳
↔	Agbonlahor	890					Idrissa Gueye	↳
↔	668						Idrissa Gueye	↳
↔	668						Leandro Bacuna	↳
↔	674							↳
↔	assist_player_id \							↳
	league	season	game		team		player	
	ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa		Gabriel	↳
↔	Agbonlahor		<NA>				Gabriel	↳
↔	Agbonlahor	548651					Idrissa Gueye	↳
↔	548653						Idrissa Gueye	↳
↔	548657						Leandro Bacuna	↳
↔	<NA>							↳
↔	assist_player \							↳
	league	season	game		team		player	
	ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa		Gabriel	↳
↔	Agbonlahor		<NA>				Gabriel	↳
↔	Agbonlahor	Micah Richards					Idrissa Gueye	↳
↔	Jordan Veretout						Idrissa Gueye	↳
↔	Gabriel Agbonlahor						Leandro Bacuna	↳
↔	<NA>							↳
↔	xg \							↳
	league	season	game		team		player	

(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	0.075484			Gabriel
↪ Agbonlahor	0.054676			Idrissa Gueye
↪	0.018269			Idrissa Gueye
↪	0.113668			Leandro Bacuna
↪	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			Gabriel
↪ Agbonlahor	0.795			Idrissa Gueye
↪	0.742			Idrissa Gueye
↪	0.883			Leandro Bacuna
↪	0.786			
↪ location_y \				
league	season	game	team	player
ENG-Premier League	1516	2015-08-08	Bournemouth-Aston Villa	Aston Villa Gabriel
↪ Agbonlahor	0.503			Gabriel
↪ Agbonlahor	0.507			Idrissa Gueye
↪	0.598			Idrissa Gueye
↪	0.561			Leandro Bacuna
↪	0.437			
↪ minute \				
league	season	game	team	player
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			

(continues on next page)

(continued from previous page)

```

↪ body_part \
league season game team player
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 team player
ENG-Premier League 1516 2015-08-08 Bournemouth-Aston Villa Aston Villa Gabriel
↪Agbonlahor Open Play
 Gabriel
↪Agbonlahor Open Play
 Idrissa Gueye
↪ 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.
  - 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()
```

```
[5]:
```

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	

league	season	game	game_id	\
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	

league	season	game	status	\
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	6	
		2020-09-12 Fulham-Arsenal	6	
		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	\
--------	--------	------	-------------------	---

(continues on next page)

(continued from previous page)

league	season	game	
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
			home_red_cards \
league	season	game	
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
			away_team_id \
league	season	game	
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
			away_team \
league	season	game	
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
			away_yellow_cards \
league	season	game	
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
			away_red_cards \
league	season	game	
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
			has_incidents_summary \
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

(continues on next page)

(continued from previous page)

		2020-09-12	West Ham United-Newcastle	True
		2020-09-13	Tottenham-Everton	True
				has_preview \
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
				score_changed_at \
league	season	game		
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
				elapsed \
league	season	game		
ENG-Premier League	2021	2020-09-12	Crystal Palace-Southampton	FT
		2020-09-12	Fulham-Arsenal	FT
		2020-09-12	Liverpool-Leeds United	FT
		2020-09-12	West Ham United-Newcastle	FT
		2020-09-13	Tottenham-Everton	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 \

(continues on next page)

(continued from previous page)

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 \
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 \
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
			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
			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
			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
			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

(continues on next page)

(continued from previous page)

		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	
home_score \				
league	season	game		
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
away_score \				
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton		0
		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
└─				
↪		incidents \		
league	season	game		
ENG-Premier League	2021	2020-09-12 Crystal Palace-Southampton	[{'minute': '13', 'type	
		↪': 1, 'subType': 1, 'pla...		
		2020-09-12 Fulham-Arsenal	[{'minute': '8', 'type':	
		↪ 1, 'subType': 1, 'pla...		
		2020-09-12 Liverpool-Leeds United	[{'minute': '4', 'type':	
		↪ 1, 'subType': 2, 'pla...		
		2020-09-12 West Ham United-Newcastle	[{'minute': '56', 'type	
		↪': 1, 'subType': 1, 'pl...		
		2020-09-13 Tottenham-Everton	[{'minute': '55', 'type	
		↪': 1, 'subType': 1, 'pl...		
bets \				
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	
aggregate_winner_field \				
league	season	game		

(continues on next page)

(continued from previous page)

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
				winner_field \
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
				period \
league	season	game		
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
				extra_result_field \
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
				home_extratime_score \
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
				away_extratime_score \
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
				home_penalty_score \
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

(continues on next page)

(continued from previous page)

league	season	game	away_penalty_score
		2020-09-13 Tottenham-Everton	None
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
		2020-09-12 Liverpool-Leeds United	None
		2020-09-12 West Ham United-Newcastle	None
		2020-09-13 Tottenham-Everton	None

league	season	game	second_half_started_at_utc
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

league	season	game	stage
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	player
	ENG-Premier League	2021	2021-01-12	Burnley-Manchester United	Charlie Taylor
1485184				Burnley	Dwight McNeil
					Jay Rodriguez
					Jimmy Dunne
				Manchester United	Eric Bailly

player_id \	league	season	game	team	player
	ENG-Premier League	2021	2021-01-12	Burnley-Manchester United	Charlie Taylor
107462				Burnley	Dwight McNeil
					Jay Rodriguez
					Jimmy Dunne
				Manchester United	Eric Bailly

reason \	league	season	game	team	player
	ENG-Premier League	2021	2021-01-12	Burnley-Manchester United	Charlie Taylor
injured doubtful				Burnley	Dwight McNeil
					Jay Rodriguez
					Jimmy Dunne
				Manchester United	Eric Bailly

status	league	season	game	team	player
	ENG-Premier League	2021	2021-01-12	Burnley-Manchester United	Charlie Taylor
Doubtful				Burnley	Dwight McNeil

(continues on next page)

(continued from previous page)

↔Rodriguez	Doubtful	Jay
↔Dunne	Doubtful	Jimmy
↔Bailly	Doubtful	Manchester United Eric

### 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 \
ENG-Premier League 2021 2021-01-12 Burnley-Manchester United 2253458317 PreMatch
 2253458375 PreMatch
 2253487469 FirstHalf
 2253487473 FirstHalf
 2253487625 FirstHalf

league season game id minute \
ENG-Premier League 2021 2021-01-12 Burnley-Manchester United 2253458317 0
 2253458375 0
 2253487469 0
 2253487473 0
 2253487625 0

league season game id second \
ENG-Premier League 2021 2021-01-12 Burnley-Manchester United 2253458317 0.0
 2253458375 0.0
 2253487469 0.0
 2253487473 0.0
 2253487625 0.0

↔minute \
league season game id expanded_
ENG-Premier League 2021 2021-01-12 Burnley-Manchester United 2253458317
↔0
 2253458375
↔0
 2253487469
```

(continues on next page)

(continued from previous page)

↔0					2253487473		
↔0					2253487625		
↔0							
↔\						type	
league	season	game			id		
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	
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	
↔player	\						
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		↔
↔Westwood				2253487625	Ashley	↔
					x \	
league	season	game		id		
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	
					y \	
league	season	game		id		
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	
					end_x \	
league	season	game		id		
ENG-Premier League	2021	2021-01-12	Burnley-Manchester United	2253458317	NaN	
				2253458375	NaN	
				2253487469	NaN	
				2253487473	NaN	
				2253487625	30.5	
					end_y \	
league	season	game		id		
ENG-Premier League	2021	2021-01-12	Burnley-Manchester United	2253458317	NaN	
				2253458375	NaN	
				2253487469	NaN	
				2253487473	NaN	
				2253487625	50.3	
					goal_mouth_y ↔	
↔ \						
league	season	game		id		
ENG-Premier League	2021	2021-01-12	Burnley-Manchester United	2253458317	NaN	
				2253458375	NaN	
				2253487469	NaN	
				2253487473	NaN	
				2253487625	NaN	
					goal_mouth_z ↔	
↔ \						

(continues on next page)

(continued from previous page)

league	season	game	id	
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

league	season	game	id	qualifiers
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	[{'type': {
			2253458375	[{'type': {
			2253487469	[
			2253487473	[
			2253487625	[{'type': {

league	season	game	id	is_touch
ENG-Premier League	2021	2021-01-12 Burnley-Manchester United	2253458317	False
			2253458375	False
			2253487469	False
			2253487473	False
			2253487625	True

league	season	game	id	is_shot
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)

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

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

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

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

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 `socceraction.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
 },
 "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	1	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	

(continues on next page)

(continued from previous page)

```

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())

```

(continues on next page)

(continued from previous page)

```
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	

	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	\
0	Andre Marriner	St. Mary's Stadium	0	Ralph Hasenhüttl	
1	Robert Jones	Craven Cottage	0	Scott Parker	
2	Craig Pawson	Anfield	2000	Jürgen Klopp	
3	Mike Dean	The Hawthorns	0	Sam Allardyce	
4	David Coote	Craven Cottage	0	Scott Parker	

	away_manager
0	Sean Dyche
1	Graham Potter
2	Nuno Espírito Santo
3	Jürgen Klopp
4	Marcelo Bielsa

Teams:

	team_id	team_name
0	184	Burnley
1	32	Man Utd

Players:

	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	

(continues on next page)

(continued from previous page)

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

Events:

	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	
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 `SOCCERDATA_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",
 "season_end": "May"
 }
}
```

(continues on next page)

(continued from previous page)

```
}
}
```

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
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]:
```

	rank	country	level	elo	from	to	\
team							
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)



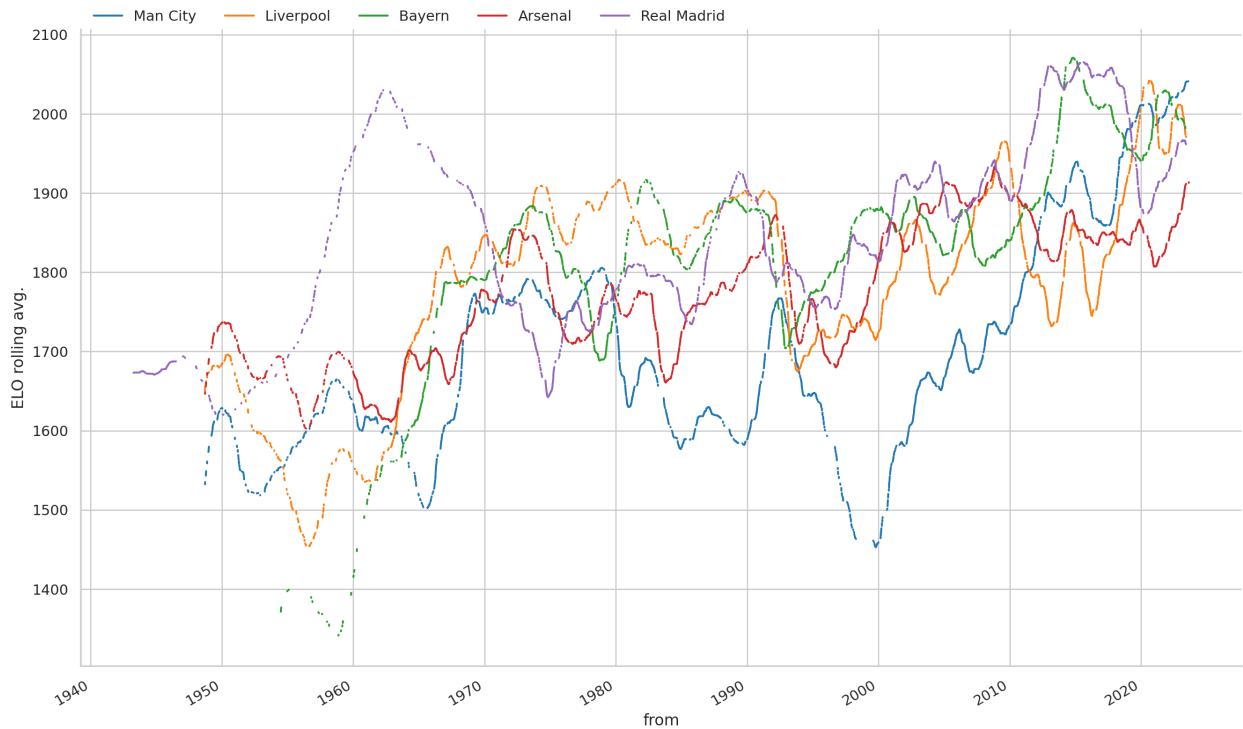
(continued from previous page)

```

→ 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(
```



..... doc/examples/ClubElo - Evolution of current top teams.ipynb ends here.

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]:

league	season	game	date \						
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	1				
	2021	2020-12-06	Crotone	Napoli	0				
	1819	2019-05-26	Fiorentina	Genoa	0				
	2021	2020-11-01	Udinese	Milan	1				
	1819	2019-03-10	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			
	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			
league	season	game	AvgCAHH	AvgCAHA	\				
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				

(continues on next page)

(continued from previous page)

```

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)
```

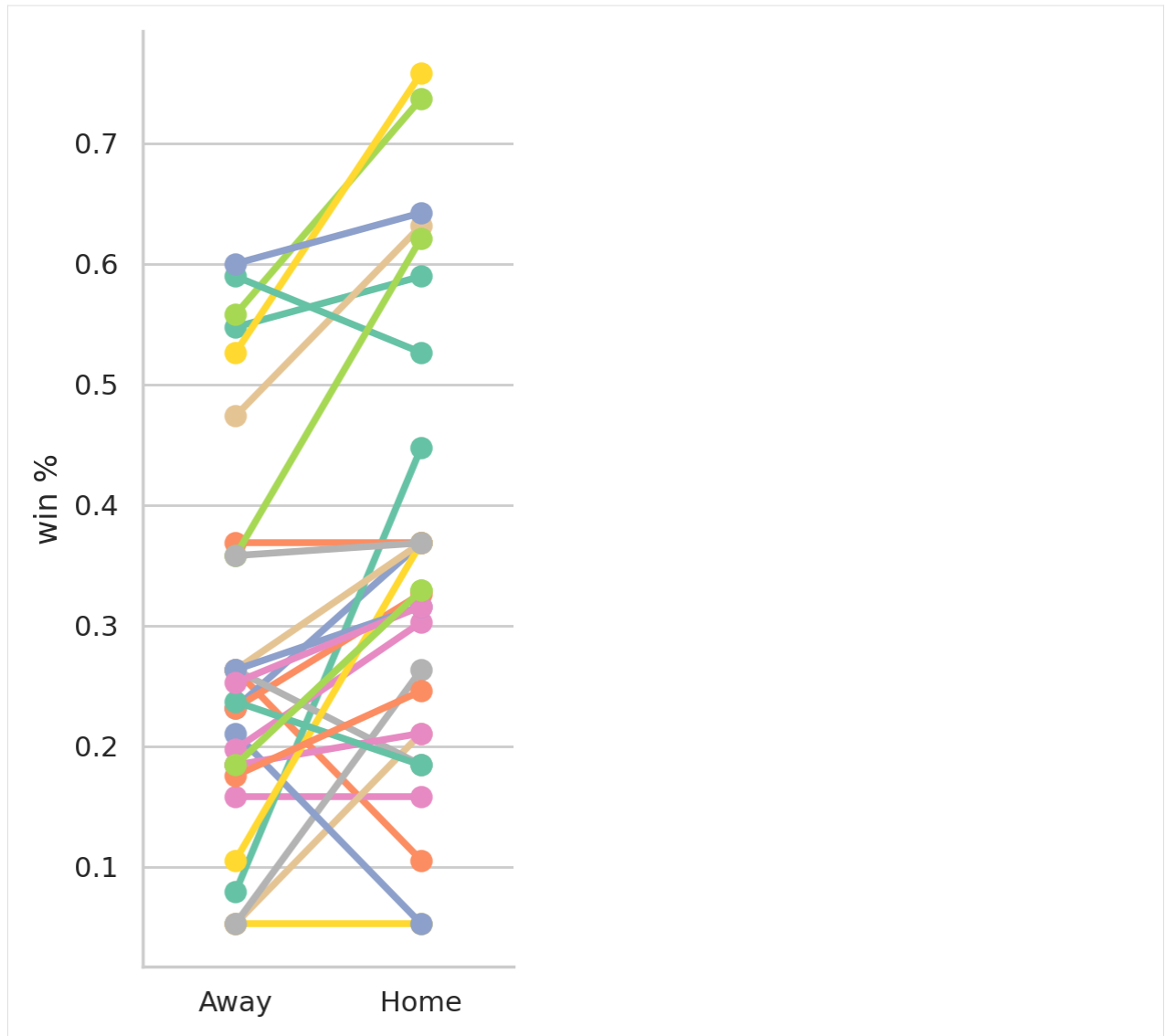
```
[6]:
```

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
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:

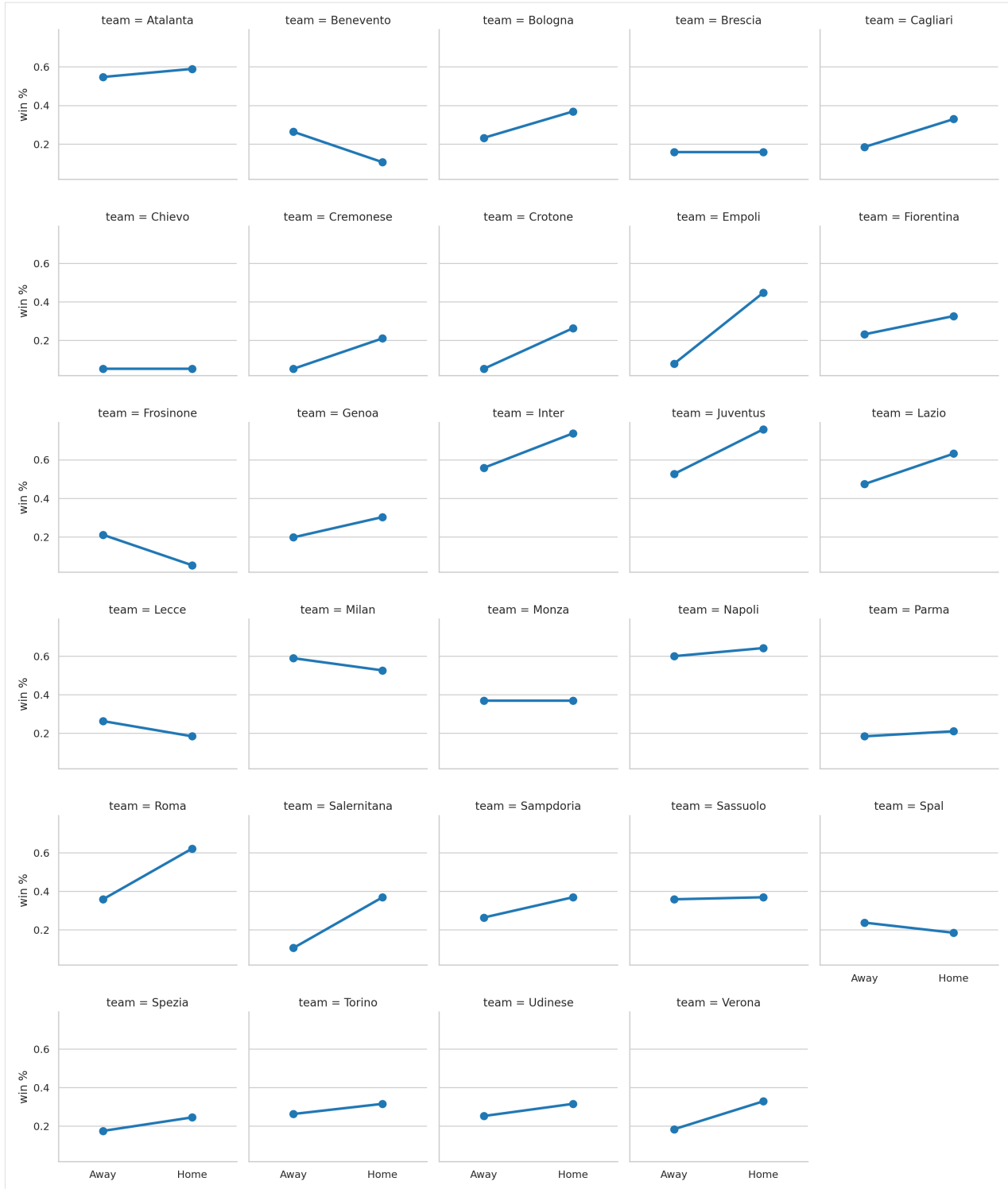
```
[7]: g = sns.FacetGrid(results.reset_index(), hue="team", palette="Set2", height=6, aspect=0.
 ↪5)
 g.map(sns.pointplot, "is_home", "win_pct", order=["Away", "Home"])
 g.set_axis_labels("", "win %");

```



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.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

- **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=1*)

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 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\_matchsheet**(*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'), path_to_browser=None,
 headless=True)
```

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.
  - 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\_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’
- ‘shooting’
- ‘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'
- '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'
- 'playing\_time'
- 'keeper'
- 'misc'

**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'

**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.

**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 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 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'), path_to_browser=None,
 headless=True)
```

Provides `pd.DataFrame`s 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.
  - 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_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.BufferedReader

**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 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 scrape 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 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"*) – Wheter 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 type**

See 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, headers=None)
```

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*)
- **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.

## 5.10 Utilities

`_common.SeasonCode`(*names=None*, \* (*Keyword-only parameters separator (PEP 3102)*), *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

**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 3.10
$ 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 Data Version Control (DVC)

The project uses `DVC` to manage the test data used by the test suite. This data is stored in the `tests/appdata/data` directory and is tracked by the `tests/appdata/data.dvc` file.

Using `DVC` ensures that tests are fast and reproducible by providing a cached version of the data, thus avoiding the need to scrape external websites during every test run. This is particularly important for CI, but also helps during local development.

To pull the latest test data from the remote storage, run:

```
$ uv run dvc pull
```

If you've added a new scraper or modified an existing one, you might need to update the test data. To do this, run the test suite (or a specific test) with the `SOC CERDATA_DIR` environment variable set to `tests/appdata`. This will cause the scrapers to save the downloaded data to the test data directory.

```
$ SOC CERDATA_DIR=tests/appdata uv run pytest tests/test_MyNewScraper.py
```

After running the tests, you can add the new data to `DVC`:

```
$ uv run dvc add tests/appdata/data
```

This will update the `tests/appdata/data.dvc` file, which you should then include in your pull request. Note that only maintainers have write access to the `DVC` remote, so you won't be able to run `dvc push`. The maintainers will push the new data once your pull request is merged.

`DVC` is included in the `test` dependency group and should be available if you've followed the environment setup instructions.

#### 7.4.5 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.

#### 7.4.6 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.

## 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.



## A

available\_leagues() (*soccerdata.\_common.BaseRequestsReader* class method), 121

available\_leagues() (*soccerdata.\_common.BaseSeleniumReader* class method), 122

available\_leagues() (*soccerdata.ClubElo* class method), 106

available\_leagues() (*soccerdata.ESPN* class method), 108

available\_leagues() (*soccerdata.FBref* class method), 112

available\_leagues() (*soccerdata.MatchHistory* class method), 113

available\_leagues() (*soccerdata.SoFIFA* class method), 116

available\_leagues() (*soccerdata.Understat* class method), 118

available\_leagues() (*soccerdata.WhoScored* class method), 120

## B

BaseRequestsReader (class in *soccerdata.\_common*), 121

BaseSeleniumReader (class in *soccerdata.\_common*), 122

## C

check\_proxy() (*soccerdata.\_common* method), 123

ClubElo (class in *soccerdata*), 105

## E

ESPN (class in *soccerdata*), 107

## F

FBref (class in *soccerdata*), 108

## G

get() (*soccerdata.\_common.BaseRequestsReader* method), 121

get() (*soccerdata.\_common.BaseSeleniumReader* method), 122

get() (*soccerdata.ClubElo* method), 106

get() (*soccerdata.ESPN* method), 108

get() (*soccerdata.MatchHistory* method), 113

get() (*soccerdata.Understat* method), 118

get\_proxy() (*soccerdata.\_common* method), 123

## L

leagues (*soccerdata.\_common.BaseRequestsReader* property), 122

leagues (*soccerdata.\_common.BaseSeleniumReader* property), 123

leagues (*soccerdata.ClubElo* property), 106

leagues (*soccerdata.ESPN* property), 108

leagues (*soccerdata.MatchHistory* property), 113

leagues (*soccerdata.Understat* property), 118

## M

make\_game\_id() (*soccerdata.\_common* method), 123

MatchHistory (class in *soccerdata*), 112

MULTI\_YEAR (in module *soccerdata*), 123

## R

read\_by\_date() (*soccerdata.ClubElo* method), 105

read\_events() (*soccerdata.FBref* method), 111

read\_events() (*soccerdata.WhoScored* method), 120

read\_games() (*soccerdata.MatchHistory* method), 113

read\_league\_table() (*soccerdata.Sofascore* method), 114

read\_leagues() (*soccerdata.FBref* method), 109

read\_leagues() (*soccerdata.Sofascore* method), 114

read\_leagues() (*soccerdata.SoFIFA* method), 115

read\_leagues() (*soccerdata.Understat* method), 117

read\_lineup() (*soccerdata.ESPN* method), 108

read\_lineup() (*soccerdata.FBref* method), 111

read\_matchsheet() (*soccerdata.ESPN* method), 107

read\_missing\_players() (*soccerdata.WhoScored* method), 119

read\_player\_match\_stats() (*soccerdata.FBref* method), 111

`read_player_match_stats()` (*soccerdata.Understat method*), 117  
`read_player_ratings()` (*soccerdata.SoFIFA method*), 116  
`read_player_season_stats()` (*soccerdata.FBref method*), 110  
`read_player_season_stats()` (*soccerdata.Understat method*), 117  
`read_players()` (*soccerdata.SoFIFA method*), 115  
`read_schedule()` (*soccerdata.ESPN method*), 107  
`read_schedule()` (*soccerdata.FBref method*), 111  
`read_schedule()` (*soccerdata.Sofascore method*), 114  
`read_schedule()` (*soccerdata.Understat method*), 117  
`read_schedule()` (*soccerdata.WhoScored method*), 119  
`read_seasons()` (*soccerdata.FBref method*), 109  
`read_seasons()` (*soccerdata.Sofascore method*), 114  
`read_seasons()` (*soccerdata.Understat method*), 117  
`read_shot_events()` (*soccerdata.Understat method*), 118  
`read_team_history()` (*soccerdata.ClubElo method*), 106  
`read_team_match_stats()` (*soccerdata.FBref method*), 110  
`read_team_match_stats()` (*soccerdata.Understat method*), 117  
`read_team_ratings()` (*soccerdata.SoFIFA method*), 116  
`read_team_season_stats()` (*soccerdata.FBref method*), 109  
`read_teams()` (*soccerdata.SoFIFA method*), 115  
`read_versions()` (*soccerdata.SoFIFA method*), 115

## S

`SeasonCode()` (*soccerdata.\_common method*), 123  
`seasons` (*soccerdata.\_common.BaseRequestsReader property*), 122  
`seasons` (*soccerdata.\_common.BaseSeleniumReader property*), 123  
`seasons` (*soccerdata.ClubElo property*), 106  
`seasons` (*soccerdata.ESPN property*), 107  
`seasons` (*soccerdata.MatchHistory property*), 112  
`seasons` (*soccerdata.Understat property*), 117  
`SINGLE_YEAR` (*soccerdata.\_common attribute*), 123  
`Sofascore` (*class in soccerdata*), 113  
`SoFIFA` (*class in soccerdata*), 114  
`standardize_colnames()` (*soccerdata.\_common method*), 123

## U

`Understat` (*class in soccerdata*), 116

## W

`WhoScored` (*class in soccerdata*), 119