

Package ‘nhlscrape’

October 13, 2022

Type Package

Title Scrapes the 'NHL' API for Statistical Analysis

Version 0.1.3

Author Adam Azoulay [aut, cre]

Maintainer Adam Azoulay <adam1azoulay@gmail.com>

URL <https://github.com/adamazoulay/nhlscrape>

BugReports <https://github.com/adamazoulay/nhlscrape/issues>

Description Add game events to a database file to use for statistical analysis of hockey games. This means we only call the 'NHL' API once for each game we want to add. We will have very fast retrieval of data once games have been added since the data is stored locally. We use the API located at <<https://statsapi.web.nhl.com/api/v1/teams>> with supplemental data from <<https://www.nhl.com/scores/>>. Other endpoints can be found at <<https://gitlab.com/dword4/nhlapi>>.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Imports httr, jsonlite, RSQLite, rvest, DBI, xml2

RoxygenNote 7.0.1

NeedsCompilation no

Suggests testthat

Repository CRAN

Date/Publication 2020-02-25 17:30:02 UTC

R topics documented:

AddAllTeamsDb	2
AddGameEvents	2
AddTeamRoster	3
GetApiJson	3

GetDbPath	4
GetGameIdRange	4
GetHeatmapCoords	5
GetPlayerId	6
GetPlayerStats	6
GetTeamId	7
QueryDb	8
SetDbPath	8

Index 10

AddAllTeamsDb	<i>Add team metadata to the database</i>
---------------	--

Description

Add all teams to the database. The teams will get stored in the database in the 'teams' table. Will not write if teams already exist in the database!

Usage

```
AddAllTeamsDb()
```

Examples

```
SetDbPath(example = TRUE)
AddAllTeamsDb()
```

AddGameEvents	<i>Adds all game events for game_id to the database for analysis</i>
---------------	--

Description

Add all events from a game_id to the 'events' table. Also adds all players in the game to the 'players' table to allow for searching by name to retrieve player_id. Finally, adds the total time on ice for each player to the 'player_toi' table. This allows for calculation of certain statistics based on time usage. Will not write if game already exists in the database!

Usage

```
AddGameEvents(game_ids)
```

Arguments

game_ids List, list of game_ids. Use GetGameIdRange to find

Examples

```
SetDbPath(example = TRUE)
AddGameEvents(2019020001)
```

AddTeamRoster	<i>Adds rosters for team id and a specific season</i>
---------------	---

Description

Add a teams roster to the 'roster' table for season. Can be accessed with a primary key of 'season_teamid_personid'

Usage

```
AddTeamRoster(team_id, season)
```

Arguments

team_id	Int, identity number for the team. Use GetTeamId to find
season	Int, A year range you want to add

Examples

```
SetDbPath(example = TRUE)
AddTeamRoster(10, 20192020)
```

GetApiJson	<i>Call the nhl api and parse the return</i>
------------	--

Description

Send a call to the 'NHL' api and get the result back in a list

Usage

```
GetApiJson(call)
```

Arguments

call	character, contains the api query
------	-----------------------------------

Value

list, contains all the data in the api call

Examples

```
GetApiJson("game/2019020001/feed/live")
```

GetDbPath	<i>Retrieve database location on system</i>
-----------	---

Description

Returns the current path to the local database file

Usage

```
GetDbPath()
```

Value

Character, system path to the database.

Examples

```
GetDbPath()
```

GetGameIdRange	<i>Find game ids for a specific team and date range</i>
----------------	---

Description

Gets a list of game ids for team_id in a specific date range, inclusive

Usage

```
GetGameIdRange(team_id, start_date, end_date)
```

Arguments

team_id	Int, team ID number
start_date	Starting date of the games, inclusive. Format: "yyyy-mm-dd"
end_date	Ending date of the games, inclusive. Format: "yyyy-mm-dd"

Value

List of ints, each element is a game ID in selected range

Examples

```
GetGameIdRange(10, "2019-09-30", "2019-12-16")
```

GetHeatmapCoords	<i>Get the heatmap coordinates for certain events in the database</i>
------------------	---

Description

This function will take in a `team_id` and a list of `game_ids`, and a list of events to look for, and return a list of `x`, `y` coordinates, transformed to be attacking zone on the left of the rink, and defending zone on the right of the rink. Some anomalies due to how the NHL tracks positions

Usage

```
GetHeatmapCoords(team_id, gids, events_list)
```

Arguments

<code>team_id</code>	Int, id of the team to transform for
<code>gids</code>	Int, list of game ids to check for events
<code>events_list</code>	Character, string of events to select from the database Note that events must be quoted in single quotes inside the string

Value

List, containing all `x`, `y` pairs for events and additional metadata

Examples

```
SetDbPath(example = TRUE)  
GetHeatmapCoords(10, 201902001, "'Shot', 'Goal'")
```

GetPlayerId	<i>Search player id using name</i>
-------------	------------------------------------

Description

Gets a player id from their name. Will only work for players that were active in a game that has already been added to the database

Usage

```
GetPlayerId(player_name)
```

Arguments

player_name	character, players full name
-------------	------------------------------

Value

int, player id number

Examples

```
SetDbPath(example = TRUE)
AddGameEvents(2019020001)
GetPlayerId("John Tavares")
```

GetPlayerStats	<i>WIP - Get advanced statistics for player_id on team_id in a list of games.</i>
----------------	---

Description

This function will search all games in game_ids and return a list of stats for the player id selected. The current stats returned are: - Shots - Goals - Corsi - Fenwick

Usage

```
GetPlayerStats(player_id, game_ids, team_id)
```

Arguments

player_id	Int, player ID number
game_ids	List, game ids to check. Must already be in the database
team_id	Int, the ID of the team the player plays for

Value

List, contains a row of stats for even strength and for all situations

Examples

```
SetDbPath(example = TRUE)
AddGameEvents(2019020001)
GetPlayerStats(8475166, 2019020001, 10)
```

GetTeamId	<i>Search for team id from team name</i>
-----------	--

Description

Retrieve the team ID using the abbreviation, or full name of the team

Usage

```
GetTeamId(team_name)
```

Arguments

team_name Character, contains either abbreviation or full name

Value

Int, team ID number.

Examples

```
SetDbPath(example = TRUE)
AddAllTeamsDb()
GetTeamId("TOR")
GetTeamId("tor")
GetTeamId("Toronto Maple Leafs")
```

QueryDb	<i>Send query to database</i>
---------	-------------------------------

Description

Send a SQL query to the database. Returns the SQL result as a data.frame. Useful for seeing the structure of the database for construction of queries

Usage

```
QueryDb(query)
```

Arguments

query	A string containing the SQL query
-------	-----------------------------------

Value

List, contains the SQL query result.

Examples

```
SetDbPath(example = TRUE)
AddGameEvents(2019020001)
QueryDb("SELECT * FROM events")
QueryDb("SELECT result_description FROM events WHERE game_id=2019020001 AND player_id=8475166")
```

SetDbPath	<i>Set database location on system</i>
-----------	--

Description

Sets the path to the database file. This function must be run to allow operation of other functions which require saving to a database

Usage

```
SetDbPath(db_path = "nhl.sqlite", example = FALSE)
```

Arguments

db_path	Character, contains the system path to the db file
example	boolean, set to FALSE by default, if TRUE it will copy the example db file and set the path pointing to it

Value

Character, contains the path to the database.

Examples

```
SetDbPath(example = TRUE)
```

Index

AddAllTeamsDb, [2](#)
AddGameEvents, [2](#)
AddTeamRoster, [3](#)

GetApiJson, [3](#)
GetDbPath, [4](#)
GetGameIdRange, [4](#)
GetHeatmapCoords, [5](#)
GetPlayerId, [6](#)
GetPlayerStats, [6](#)
GetTeamId, [7](#)

QueryDb, [8](#)

SetDbPath, [8](#)