

# Package ‘cricketdata’

August 5, 2022

**Version** 0.2.0

**Title** International Cricket Data

**Description** Data on international and other major cricket matches from ESPNCricinfo <<https://www.espncricinfo.com>> and Cricsheet <<https://cricsheet.org>>. This package provides some functions to download the data into tibbles ready for analysis.

**Depends** R (>= 3.5.0)

**Imports** dplyr, janitor, lubridate, magrittr, progress, readr, rvest, stringr, tibble, tidyr, xml2

**Suggests** gghighlight, ggplot2, ggtext, glue, knitr, paletteer, patchwork, rmarkdown, showtext, tidyverse

**License** GPL-3

**Encoding** UTF-8

**ByteCompile** true

**URL** <http://pkg.robjhyndman.com/cricketdata/>,  
<https://github.com/robjhyndman/cricketdata>

**LazyData** true

**VignetteBuilder** knitr

**RoxygenNote** 7.2.1

**NeedsCompilation** no

**Author** Rob Hyndman [aut, cre],  
Timothy Hyndman [aut],  
Charles Gray [aut],  
Sayani Gupta [aut],  
Jacquie Tran [aut],  
Hassan Rafique [aut],  
Puwasala Gamakumara [ctb],  
Alex Whan [ctb]

**Maintainer** Rob Hyndman <[Rob.Hyndman@monash.edu](mailto:Rob.Hyndman@monash.edu)>

**Repository** CRAN

**Date/Publication** 2022-08-05 09:50:10 UTC

## R topics documented:

cricketdata . . . . .	2
fetch_cricinfo . . . . .	2
fetch_cricsheet . . . . .	3
fetch_player_data . . . . .	4
fetch_player_meta . . . . .	5
find_player_id . . . . .	6
player_meta . . . . .	7
update_player_meta . . . . .	8

<b>Index</b>	<b>9</b>
--------------	----------

---

cricketdata	<i>Cricket Data from international and other major cricket matches.</i>
-------------	---

---

### Description

Data is provided by ESPNCricinfo and Cricsheet. This package provides some functions to download the data into tibbles ready for analysis. Please respect the terms of use for each website: [ESPNCricinfo](#), [Cricsheet](#).

### Details

Package:	cricketdata
Type:	Package
License:	GPL3
LazyLoad:	yes

---

fetch_cricinfo	<i>Fetch Data from Cricinfo</i>
----------------	---------------------------------

---

### Description

Fetch data from ESPNCricinfo and return a tibble. All arguments are case-insensitive and partially matched.

### Usage

```
fetch_cricinfo(
  matchtype = c("test", "odi", "t20"),
  sex = c("men", "women"),
```

```

    activity = c("batting", "bowling", "fielding"),
    type = c("career", "innings"),
    country = NULL
  )

```

### Arguments

matchtype	Character indicating test (default), odi, or t20.
sex	Character indicating men (default) or women.
activity	Character indicating batting (default), bowling or fielding.
type	Character indicating innings-by-innings or career (default) data
country	Character indicating country. The default is to fetch data for all countries.

### Value

A tibble object, similar to a data.frame.

### Author(s)

Rob J Hyndman, Timothy Hyndman, Charles Gray

### Examples

```

## Not run:
auswt20 <- fetch_cricinfo("T20", "Women", country="Aust")
IndiaODIBowling <- fetch_cricinfo("ODI", "men", "bowling", country="india")

## End(Not run)

```

---

fetch_cricsheet	<i>Fetch ball-by-ball, match and player data from Cricsheet and return a tibble.</i>
-----------------	--

---

### Description

Download csv data from Cricsheet <https://cricsheet.org/downloads/>. Data must be specified by three factors: (a) type of data: bbb (ball-by-ball), match or player. (b) gender; (c) competition. See <https://cricsheet.org/downloads/> for what the competition character codes mean.

### Usage

```

fetch_cricsheet(
  type = c("bbb", "match", "player"),
  gender = c("female", "male"),
  competition = c("tests", "multi_day", "odis", "odms", "t20is", "t20is_unofficial",
    "apl", "bbl", "bpl", "county", "edwards_cup", "cpl", "the_hundred", "ipl", "lpl",

```

```
"msl", "t20_blast", "psl", "heyhoe_flint_trophy", "sheffield_shield", "super_smash",  
  "wbb1", "wt20c")  
)
```

### Arguments

type	Character string giving type of data: ball-by-ball, match info or player info.
gender	Character string giving player gender: female or male.
competition	Character string giving name of competition.

### Value

A tibble object, similar to a `data.frame`.

### Author(s)

Jacquie Tran, Hassan Rafique and Rob J Hyndman

### Examples

```
## Not run:  
wbb1_bbb <- fetch_cricsheet(competition = "wbb1", type = "bbb")  
wbb1_match <- fetch_cricsheet(competition = "wbb1", type = "match")  
wbb1_player <- fetch_cricsheet(competition = "wbb1", type = "player")  
  
## End(Not run)
```

---

fetch_player_data	<i>Fetch Player Data</i>
-------------------	--------------------------

---

### Description

Fetch individual player data from all matches played. The function will scrape the data from ESPN-Cricinfo and return a tibble with one line per innings for all games a player has played. To identify a player, use their Cricinfo player ID. The simplest way to find this is to look up their Cricinfo Profile page. The number at the end of the URL is the ID. For example, Meg Lanning's profile page is <http://www.espnricinfo.com/australia/content/player/329336.html>, so her ID is 329336.

### Usage

```
fetch_player_data(  
  playerid,  
  matchtype = c("test", "odi", "t20"),  
  activity = c("batting", "bowling", "fielding")  
)
```

**Arguments**

playerid	The player ID as given in the Cricinfo profile. Integer or character.
matchtype	Which type of cricket matches do you want? Tests, ODIs or T20s? Not case-sensitive.
activity	Which type of activities do you want? Batting, Bowling or Fielding? Not case-sensitive.

**Value**

A tibble containing data on the selected player, with one row for every innings of every match in which they have played.

**Author(s)**

Rob J Hyndman and Sayani Gupta

**See Also**

[find\\_player\\_id\(\)](#) to find a player ID by searching on their name, and [fetch\\_player\\_meta\(\)](#) to download meta data for players.

**Examples**

```
## Not run:
# Download data on some players
EllysePerry <- fetch_player_data(275487, "T20", "batting")
RahulDravid <- fetch_player_data(28114, "ODI", "fielding")
LasithMalinga <- fetch_player_data(49758, "Test", "bowling")

# Create a plot for Ellyse Perry's T20 scores
library(dplyr)
library(ggplot2)
EllysePerry |> filter(!is.na(Runs)) |>
ggplot(aes(x=Start_Date, y=Runs, col = Dismissal, na.rm = TRUE)) + geom_point() +
  ggtitle("Ellyse Perry's T20 Scores")

## End(Not run)
```

---

fetch\_player\_meta      *Fetch Player Meta Data*

---

**Description**

Fetch player meta data from ESPNcricinfo and return a tibble with one line per player. To identify the players, use their Cricinfo player IDs. The simplest way to find this is to look up their Cricinfo Profile page. The number at the end of the URL is the ID. For example, Meg Lanning's profile page is <http://www.espnricinfo.com/australia/content/player/329336.html>, so her ID is 329336.

**Usage**

```
fetch_player_meta(playerid)
```

**Arguments**

playerid      A vector of player IDs as given in Cricinfo profiles. Integer or character.

**Value**

A tibble containing meta data on the selected players, with one row for each player.

**Author(s)**

Hassan Rafique and Rob J Hyndman

**See Also**

It is usually simpler to just use the saved data set [player\\_meta](#) which contains the meta data for all players on ESPNcricinfo as at 5 August 2022. To find a player ID, use [find\\_player\\_id\(\)](#). Use [fetch\\_player\\_data\(\)](#) to download playing statistics for a player.

**Examples**

```
## Not run:  
# Download meta data on Meg Lanning and Ellyse Perry  
aus_women <- fetch_player_meta(c(329336, 275487))  
  
## End(Not run)
```

---

find_player_id	<i>Find a player id from cricinfo.com</i>
----------------	---

---

**Description**

Find a player id from cricinfo.com

**Usage**

```
find_player_id(searchstring)
```

**Arguments**

searchstring    Part of a player name(s) to search for. Can be a character vector.

**Value**

A table of matching players, their ids, and teams they played for.

**Author(s)**

Rob J Hyndman

**See Also**

[fetch\\_player\\_data\(\)](#) to download playing statistics for a player, and [fetch\\_player\\_meta\(\)](#) to download meta data on players.

**Examples**

```
## Not run:  
(perry <- find_player_id("Perry"))  
EllysePerry <- fetch_player_data(perry[2,"ID"], "test")  
  
## End(Not run)
```

---

player\_meta

*Meta data on players listed at ESPNCricinfo*

---

**Description**

A dataset containing the names and other attributes of players who appear on both [cricsheet](#) and [ESPNCricinfo](#), as at 5 August 2022.

**Usage**

```
player_meta
```

**Format**

A data frame with 11355 rows and 10 variables.

**Source**

<https://www.espncriinfo.com>

---

update_player_meta	<i>Update player_meta</i>
--------------------	---------------------------

---

### Description

The [player\\_meta](#) data set contains the names and other attributes of players who appear on both [cricsheet](#) and [ESPNCricinfo](#) as at 5 August 2022. This function returns an updated version of the data set based on information currently available online.

### Usage

```
update_player_meta(start_again = FALSE)
```

### Arguments

`start_again` If TRUE, downloads all data from ESPNCricinfo without using `player_meta` as a starting point. This can take a long time.

### Value

A tibble containing meta data on cricket players.

### Author(s)

Hassan Rafique and Rob J Hyndman

### See Also

[player\\_meta](#), [fetch\\_player\\_meta\(\)](#).

### Examples

```
## Not run:  
# Update data to current  
new_player_meta <- update_player_meta()  
  
## End(Not run)
```



# Index

\* **datasets**

player\_meta, 7

\* **package**

cricketdata, 2

cricketdata, 2

fetch\_cricinfo, 2

fetch\_cricsheet, 3

fetch\_player\_data, 4

fetch\_player\_data(), 6, 7

fetch\_player\_meta, 5

fetch\_player\_meta(), 5, 7, 8

find\_player\_id, 6

find\_player\_id(), 5, 6

player\_meta, 6, 7, 8

update\_player\_meta, 8