Package: spiderorchid (via r-universe)

May 7, 2026

Title Download and wrangle publication data for Monash EBS academic staff

Version 0.1.0

Description Provide journal rankings from Monash Business School, ABDC, CORE, Scimago and ERA 2010. Fetch publication details from ORCID, Google Scholar, PURE and from DOIs. Find CRAN packages and download statistics for specified authors.

License MIT + file LICENSE

Encoding UTF-8

Imports dplyr, httr, jsonlite, pkgmeta, purrr, rorcid, scholar, stringr, tibble

Suggests fuzzyjoin, ggplot2, knitr, rmarkdown, testthat (>= 3.0.0)

Depends R (>= 4.1.0)

Remotes robjhyndman/pkgmeta

LazyData true

URL https://github.com/numbats/spiderorchid/, https://numbats.github.io/spiderorchid/

Roxygen list(markdown = TRUE)

VignetteBuilder knitr

Config/testthat/edition 3

Config/roxygen2/version 8.0.0

Config/pak/sysreqs libfontconfig1-dev libfreetype6-dev git libglpk-dev libicu-dev libxml2-dev libssl-dev

Repository https://robjhyndman.r-universe.dev

Date/Publication 2026-05-07 22:30:23 UTC

RemoteUrl https://github.com/numbats/spiderorchid

RemoteRef HEAD

RemoteSha 86d647f5e866c090ef0d5b86d3a1fdc5413f1617

Contents

---

abdc *ABDC Journal Quality List*

Description

This is a dataset that contains the quality list of rankings of the Australian Business Deans Council (ABDC) from 2025. You can read more about this list here.

Usage

```
data(abdc)
```

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 2651 rows and 7 columns.

Value

A data frame with 2651 observations on the following 7 variables:

`title`: Title of the journal

`publisher`: Publishing house

`issn`: International Standard Serial Number

`issn_online`: ISSN Online - as ISSN, but for the online, rather than print version

`year_inception`: Year the journal started

`field_of_research`: Field of Research Code as provided by the Australian Bureau of Statistics

`rank`: In order of best to lowest rank: A*, A, B, or C

Source

https://abdc.edu.au/abdc-journal-quality-list/

Examples

```
library(dplyr)
abdc |>
  filter(field_of_research == "4905") |>
  arrange(rank) |>
  select(title, rank)
```

`core` *CORE (Computing Research and Education) lists of conference and journal rankings*

Description

Two datasets are provided: `core` and `core_journals`, which contains lists of conference and journal rankings respectively, according to the CORE executive committee. The details of the CORE organisation, and its procedure for ranking are provided below.

Usage

```
core

core_journals
```

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 982 rows and 2 columns.

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 639 rows and 4 columns.

Details

CORE is an association of university departments of computer science in Australia and New Zealand. Prior to 2004 it was known as the Computer Science Association, CSA.

The CORE Conference Ranking provides assessments of major conferences in the computing disciplines. The rankings are managed by the CORE Executive Committee, with periodic rounds for submission of requests for addition or reranking of conferences. Decisions are made by academic committees based on objective data requested as part of the submission process. Conference rankings are determined by a mix of indicators, including citation rates, paper submission and acceptance rates, and the visibility and research track record of the key people hosting the conference and managing its technical program. A more detailed statement categorizing the ranks A*, A, B, and C can be found here.

Value

`core` is a data frame with 982 observations and two variables:

`title`: Title of the conference

`rank`: Conferences are assigned to one of the following categories:

- A*: flagship conference, a leading venue in a discipline area
- A: excellent conference, and highly respected in a discipline area
- B: good conference, and well regarded in a discipline area
- C: other ranked conference venues that meet minimum standards

`core_journals` is a data frame with 639 observations and 4 variables:

`title`: Title of the journal

`field_of_research`: Field of Research Code as provided by the Australian Bureau of Statistics

`issn`: International Standard Serial Number

`rank`: In order of best to lowest rank: A*, A, B, or C

Source

https://www.core.edu.au/conference-portal

Examples

```
core
core_journals
```

---

ebs_pure *Monash EBS PURE publications data*

Description

This dataset contains publications since 2018, downloaded from PURE on 16 January 2025. Additional data can be updated using the `fetch_pure()` function.

Usage

```
ebs_pure
```

Format

A data frame with 7 variables:

pure_id `character`. The unique identifier for the publication in PURE.

year `integer`. The year of publication.

authors `character`. The authors of the publication.

title `character`. The title of the publication.

journal character. The journal where the publication appeared.

subtype character. The subtype of the publication.

bib character. A bibliographic citation in the Harvard format.

doi character. The DOI of the publication where available.

Source

https://research.monash.edu/en/organisations/econometrics-business-statistics/publications/

---

era2010 *ERA2010 Journal List*

Description

This is a dataset that contains the list of journal rankings from the ARC Excellence in Research for Australia 2010 round.

Usage

```
data(era2010)
```

Format

An object of class tbl_df (inherits from tbl, data.frame) with 20712 rows and 5 columns.

Value

A data frame with 20712 rows and the following 5 variables:

eraid: ERA ID of the journal

title: Title of the journal

issn: International Standard Serial Number

field_of_research: Field of Research Code as provided by the Australian Bureau of Statistics at the time. Note that the codes have since changed.

rank: In order of best to lowest rank: A*, A, B, or C

Source

https://www.righttoknow.org.au/request/journal_list_relating_to_the_201

Examples

```
library(dplyr)
era2010 |>
  filter(field_of_research == "0104") |>
  arrange(rank)
```

`fetch_cran` *Fetch packages from CRAN*

Description

This function searches for CRAN packages by author names. Note that some authors may use different name variations on CRAN (e.g., "Di Cook" and "Dianne Cook"), so it may be necessary to call the function with several variations.

Usage

```
fetch_cran(
  author_names = NULL,
  package_names = NULL,
  downloads_from = "2000-01-01",
  downloads_to = Sys.Date()
)
```

Arguments

`author_names`	A character vector containing the authors' names in the form used on CRAN.
`package_names`	A character vector of package names. Ignored if `author_names` is provided.
`downloads_from`	A date or character string in the format "YYYY-MM-DD" specifying the date from which to start counting downloads. Default is "2000-01-01".
`downloads_to`	A date or character string in the format "YYYY-MM-DD" specifying the last date for counting downloads. Default is current date.

Value

A data frame returning meta data about a package including total downloads between `downloads_from` and `downloads_to`.

Examples

```
## Not run:
cran2024 <- fetch_cran(
   author_names = c("Michael Lydeamore", "Di Cook", "Dianne Cook", "Hyndman"),
   downloads_from = "2024-01-01",
   downloads_to = "2024-12-31"
)

## End(Not run)
```

fetch_doi *Fetch article information given a DOI*

Description

Retrieves publications for a given list of DOIs using the DOI API and formats them into a structured tibble.

Usage

```
fetch_doi(doi)
```

Arguments

doi	A character vector of DOIs.

Value

A tibble containing the article information.

Examples

```
fetch_doi("10.1016/j.ijforecast.2023.10.010")
```

fetch_orcid *Fetch publications from ORCID*

Description

Retrieves publications for given ORCID IDs, and returns them as a tibble. Only publications with DOIs are returned. The function uses the ORCID API to fetch the DOIs, and then uses the DOI API to fetch the publication details for each DOI.

Usage

```
fetch_orcid(orcid_ids)
```

Arguments

orcid_ids	A character vector of ORCID IDs.

Details

This function requires authentication on ORCID. If you have not previously authenticated, it will prompt you to do so when first run. If you just follow the prompts, you will be authenticated, but only for downloading your own papers. If you want to download papers from other ORCID IDs, you will need to authenticate with a 2-legged OAuth. Follow the instructions at https://info.orcid.org/register-a-client-application-production-member-api/. To avoid having to do this in each session, store the token obtained from `orcid_auth()` in your `.Renviron` file by running `usethis::edit_r_environ()`. It should be of the form `ORCID_TOKEN=<your token>`.

Value

A tibble containing all publications for the specified ORCID IDs.

Examples

```
## Not run:
fetch_orcid(c("0000-0003-2531-9408", "0000-0001-5738-1471"))

## End(Not run)
```

---

fetch_pure *Fetch publications from PURE*

Description

To download data from PURE, it is necessary to have access to the API via Simon Angus and the Astro team https://astro.monash.edu/. The API key is stored in the environment variable `PURE_API_KEY`. You can add it to your environment using `edit_r_environ()`. This is end-point restricted to Monash IP addresses. So either use it on campus or invoke the VPN before using it off campus.

Usage

```
fetch_pure(years)
```

Arguments

`years`	A numeric vector of publication years. All publications between the minimum year and the maximum year are returned.

Value

A data frame containing the data fetched from the PURE API covering the specified publication years.

See Also

`ebs_pure`

fetch_scholar *Fetch publications from Google Scholar*

Description

Retrieves publications for given Google Scholar IDs, and returns them as a tibble. This function retrieves publications for a given Google Scholar ID and formats them into a structured tibble.

Usage

```
fetch_scholar(scholar_id)
```

Arguments

scholar_id	A character vector of Google Scholar IDs.

Value

A tibble containing all publications for the specified Google Scholar IDs.

Examples

```
## Not run:
fetch_scholar("vamErfkAAAAJ")

## End(Not run)
```

journal_ranking *Find rankings of journals from the Monash Business School, ABDC, CORE, SCImago or ERA2010.*

Description

Given a list of journal titles, this function will return their ranking from various lists. Data sets used are:

***monash*:** Monash Business School

***abdc*:** Australian Business Deans' Council

***era2010*:** ERA 2010

***core*:** CORE

***scimago*:** SCImago

This function is used in the Journal rankings shiny app.

Usage

```
journal_ranking(
  title,
  source = c("monash", "abdc", "era2010", "core", "scimago"),
  fuzzy = TRUE,
  only_best = length(title) > 1,
  ...
)
```

Arguments

title	A character vector containing (partial) journal names.
source	A character string indicating which ranking data base to use. Default "monash".
fuzzy	Should fuzzy matching be used. If FALSE, partial exact matching is used. Otherwise, full fuzzy matching is used.
only_best	If TRUE, only the best matching journal is returned.
...	Other arguments are passed to agrepl (if fuzzy is TRUE), or grepl otherwise.

Value

A data frame containing the journal title, rank and source for each matching journal.

Author(s)

Rob J Hyndman

Examples

```
# Return ranking for individual journals or conferences
journal_ranking("Annals of Statistics")
journal_ranking("Annals of Statistics", "abdc")
journal_ranking("International Conference on Machine Learning")
journal_ranking("International Conference on Machine Learning", "core")
journal_ranking("R Journal", "scimago", only_best = TRUE)
```

---

monash *Monash Business School Journal Quality List*

Description

This is a dataset that contains the list of quality journal rankings from the Monash Business School. In most cases, it follows ABDC with A* equal to Group 1 and A equal to Group 2. The "Group 1+" category contains a small set of the highest rank journals. The data set is updated from time to time when journals not on the ABDC list are classified. See https://www.intranet.monash/business/research-services/research-standards for the latest information.

Usage

```
data(monash)
```

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 4489 rows and 2 columns.

Value

A data frame with 4489 observations on the following 2 variables:

`title`: Title of the journal

`rank`: In order of best to lowest rank: Group 1+, Group 1, Group 2

Source

Monash Business School

Examples

```
library(dplyr)
library(stringr)
monash |>
  filter(str_detect(title, "Statist")) |>
  arrange(rank)
```

---

scimago *SCImago Journal Rank for all journals indexed by Scopus*

Description

This data was taken from https://www.scimagojr.com/journalrank.php

Usage

```
data(scimago)
```

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 32193 rows and 29 columns.

Value

A tibble with 32193 rows and 29 variables:

year Year of SCImago Journal Ranking calculation.

rank Rank of the journal among all journals.

sourceid Database ID of the journal.

title Jounal's title.

type Type: "journal", "book series", "trade journal", or "conference and proceedings"

issn ISSN journal identifier.

sjr SCImago Journal Rank indicator. It expresses the average number of weighted citations received in the selected year by the documents published in the selected journal in the three previous years, –i.e. weighted citations received in year X to documents published in the journal in years X-1, X-2 and X-3. See detailed description of SJR (PDF).

sjr_best_quartile Highest quartile of the journal among all categories it belongs to.

h_index Hirsch index of the journal. The h index expresses the journal's number of articles (h) that have received at least h citations. It quantifies both journal scientific productivity and scientific impact and it is also applicable to scientists, countries, etc. (see H-index wikipedia definition).

total_docs_year Total number of published documents within a specific year. All types of documents are considered, including citable and non citable documents.

total_docs_3years Published documents in the three previous years (selected year documents are excluded), i.e.when the year X is selected, then X-1, X-2 and X-3 published documents are retrieved. All types of documents are considered, including citable and non citable documents.

total_refs Total number of citations received by a journal to the documents published within a specific year.

total_cites_3years Number of citations received in the seleted year by a journal to the documents published in the three previous years, –i.e. citations received in year X to documents published in years X-1, X-2 and X-3. All types of documents are considered.

citable_docs_3years Number of citable documents published by a journal in the three previous years (selected year documents are excluded). Exclusively articles, reviews and conference papers are considered..

cites_doc_2years Average citations per document in a 2 year period. It is computed considering the number of citations received by a journal in the current year to the documents published in the two previous years, –i.e. citations received in year X to documents published in years X-1 and X-2. Comparable to Journal Impact Factor.

ref_doc Average number of references per document in the selected year..

country Country of the publisher.

publisher Publisher of the journal.

categories Categories the jounal belongs to.

highest_category Category in which the journal ranks highest by percentile.

highest_rank Rank of journal in `highest_category`.

highest_percentile Highest percentile of journal in any category.

Author(s)

Rob Hyndman

Source

SCImago Journal & Country Rank. Retrieved from https://www.scimagojr.com/journalrank.php

---

`staff_ids` *Monash EBS academic research staff IDs*

Description

This dataset contains the mappings between researcher names and their respective ORCID and Google Scholar IDs. It is useful for identifying and linking academic profiles across different platforms.

Usage

```
staff_ids
```

Format

A data frame with 4 variables:

first_name `character`. The first name of the individual.

last_name `character`. The last name of the individual.

orcid_id `character`. The ORCID identifier.

scholar_id `character`. The Google Scholar user ID.

Source

https://www.monash.edu/business/ebs/our-people/staff-directory

Index