

Package ‘nhanesdiva’

May 19, 2026

Type Package

Title NHANES Data Search, Preview, and Download Tools

Description Search, preview, and download datasets from the National Health and Nutrition Examination Survey (NHANES) across survey cycles. The package provides functions to identify relevant datasets by keyword, inspect available .XPT files before downloading, and organize retrieved data locally. Data are retrieved from the NHANES web services available at <https://www.cdc.gov/nchs/nhanes/>.

Version 1.0.0

License MIT + file LICENSE

Encoding UTF-8

Suggests curl, testthat (>= 3.0.0), withr

Config/testthat/edition 3

RoxygenNote 7.3.3

URL <https://github.com/Snowepi/nhanesdiva>

BugReports <https://github.com/Snowepi/nhanesdiva/issues>

NeedsCompilation no

Author Sushma Dahal [aut, cre]

Maintainer Sushma Dahal <sushdahal@gmail.com>

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2026-05-19 07:00:02 UTC

Contents

get_nhanes_data	2
preview_nhanes_downloads	3
search_nhanes	5

Index	7
--------------	----------

get_nhanes_data	<i>Download NHANES Data</i>
-----------------	-----------------------------

Description

Downloads NHANES datasets matching specified years, datasets, and components.

Usage

```
get_nhanes_data(years, datasets, components, base_dir = NULL)
```

Arguments

years	Vector of survey years (e.g., 2011, c(2011, 2013)).
datasets	Character vector of NHANES dataset categories ("demographics", "examination", "laboratory", "questionnaire").
components	Character vector of dataset components or dataset IDs (e.g., "DEMO", "PBCD", or specific component names like "demographic_variables_sample_weights"). The components can include the name of the component of the dataset such as "demographic_variables_sample_weights", "dietary_interview_individual_foods", or it can also include the name of the dataset_id "DEMO" or "DRXIFF" respectively. However for some datasets using dataset_id may also result the list of dataset that contains that id as a part of its dataset_id. For example in the dataset examination, component spirometry_pre_and_post_bronchodilator has the dataset_id "SPX". Likewise in the same dataset, component spirometry_raw_curve_data has dataset_id "SPXRAW". So if users write "SPX" as component, data for "SPXRAW" will also be downloaded. In cases like this it is recommended to write the exact component name instead of dataset_id. Users are encouraged to use <code>preview_nhanes_downloads()</code> to review the list of data that they are about to download.
base_dir	Directory where downloaded NHANES .XPT files will be saved. By default, <code>base_dir = NULL</code> and the data files are downloaded to a temporary directory in a subfolder named 'NHANES_data' (within <code>tempdir()</code>). Files stored in the temporary directory may be removed when the R session ends. Users may optionally specify <code>base_dir</code> to control where files are saved.

Details

It is recommended to use `preview_nhanes()` first to inspect available files before downloading.

This function retrieves NHANES data from official NHANES sources based on user-specified filters. Internet access is required.

Users can explore available datasets and components using: `nhanes_search("keyword")`.

If files are downloaded multiple times with the same parameters, existing files in the `base_dir` will be overwritten.

Users may also assign the output of `get_nhanes_data()` to an object, for example: `data_list <- get_nhanes_data(years = 1999, datasets = "demographics", components = "DEMO")`. This allows the downloaded file paths to be combined together to inspect the list of downloaded data.

Downloaded datasets can be combined or merged based on the needs of the analysis. Users should ensure appropriate alignment of identifiers and variables before combining datasets.

Value

Downloads NHANES .XPT files to a local folder. If no matches are found, no files are downloaded and an empty result is returned.

Examples

```
# Example 1. Download data of one component from one dataset of one cycle
#in a default temporary directory in a subfolder named "NHANES_data"
get_nhanes_data(
  years = 2011,
  datasets = "demographics",
  components = "DEMO",
  base_dir = NULL
)
# Example 2. Download data of one component from one dataset of one cycle
#in a user defined directory
get_nhanes_data(
  years = 2011,
  datasets = "demographics",
  components = "DEMO",
  base_dir = "userdefined directory"
)
# Example 2. Download data from multiple components from multiple
# datasets and cycles in the default NHANES_data folder in temporary
# directory
get_nhanes_data(
  years = c(1999, 2001, 2017),
  datasets = c("examination", "laboratory"),
  components = c("cardiovascular_fitness", "cadmium_lead_total_mercury_blood")
)
# Example 3. The above input can alternatively written using the dataset_id
# in the component section as follows
get_nhanes_data(
  years = c(1999:2002, 2017),
  datasets = c("examination", "laboratory"),
  components = c("CVX", "PBCD")
)
```

Description

Generates a preview of NHANES datasets that match the specified years, datasets, and components of the dataset without downloading any files. This helps users inspect what data will be retrieved before running a download operation, reducing unnecessary storage use.

Usage

```
preview_nhanes_downloads(years, datasets, components)
```

Arguments

years	Vector of survey years to search (e.g., c(2011, 2013)).
datasets	Character vector of NHANES dataset names to include in the search. This will be any of the four datasets: "demographics", "examination", "laboratory", or "questionnaire".
components	Character vector of NHANES components. The components can include the name of the component of the dataset such as "demographic_variables_sample_weights", "dietary_interview_individual_foods", or it can also include the name of the dataset_id "DEMO" or "DRXIFF" respectively. However for some datasets using dataset_id may also result the list of dataset that contains that id as a part of its dataset_id. For example in the dataset examination, component spirometry_pre_and_post_bronchodilator has the dataset_id "SPX". Likewise in the same dataset, component spirometry_raw_curve_data had dataset_id "SPXRAW". So if users write "SPX" as component, list of data for "SPXRAW" will also be shown. In cases like this it is recommended to write the exact component name instead of dataset_id. Users can look at the names of the component or dataset_id by typing their desired search query in the nhanes_search to see the list of dataset, components, dataset_id. for their search term e.g nhanes_search("tuberculosis"), nhanes_search("dietary"), nhanes_search("spirometry").

Details

This function uses the internal NHANES mapping table to resolve dataset availability across survey cycles. It does not perform any downloads or external requests.

Value

A data frame listing matching NHANES files and associated metadata. If no matches are found, an empty data frame is returned.

Examples

```
# Example 1. Preview dataset that have dataset_id "DEMO" within the
# demographics dataset for year 2011 and 2013
preview_nhanes_downloads(
  years = c(2011, 2013),
  datasets = "demographics",
  components = "DEMO"
```

```

)

# Example 2. Preview dataset that have component name
# "cardiovascular_fitness", and "cadmium_lead_total_mercury_blood"
# within dataset examination and laboratory for the years 1999, 2001 and 2017
preview_nhanes_downloads(
  years = c(1999, 2001, 2017),
  datasets = c("examination", "laboratory"),
  components = c("cardiovascular_fitness", "cadmium_lead_total_mercury_blood")
)

# Example 3. The above input can alternatively written using the dataset_id
# in the component section as follows using CVX which is the dataset_id for
# cardiovascular_fitness and PBCD which is dataset_id for
# cadmium_lead_total_mercury_blood. The names of the dataset_id and the
# components can be found using \code{search_nhanes()}
preview_nhanes_downloads(
  years = c(1999:2002, 2017),
  datasets = c("examination", "laboratory"),
  components = c("CVX", "PBCD")
)

```

search_nhanes

Search NHANES Data Using Your Search Query

Description

Generates a list of all the NHANES datasets that match the your search term.

Usage

```
search_nhanes(query)
```

Arguments

query	Any search term in text that the user wants information. There should be one search query at a time e.g "demographic", "weight", "spirometry". To get a list of all data available in a given publicly available dataset, user can search for the dataset name e.g. "demographics", or "examination", or "laboratory" or "questionnaire".
-------	---

Details

This helps users to inspect what datasets are present in the NHANES data related to their search term. They can then use the result to identify the dataset that they want. Then using the search result users can pick the year, dataset, component, or dataset_id of that particular data they want. These inputs will go into the preview function and the data download function.

This function uses the internal NHANES mapping table to resolve dataset availability across survey cycles. It does not perform any downloads or external requests.

Value

A data frame listing matching NHANES files and associated metadata. If no matches are found, user will get a message that "No matches found for:".

Examples

```
# Example 1. Search for term demographic
search_nhanes("demographic")

# Example 2. Search for term spirometry
search_nhanes("spirometry")

# Example 3. Search for term tuberculosis
search_nhanes("tuberculosis")

# Example 4. Search for term blood_pressure
search_nhanes("blood_pressure")

# Example 5. Search for the whole list within the laboratory dataset
search_nhanes("laboratory")
```

Index

`get_nhanes_data`, [2](#)

`preview_nhanes_downloads`, [3](#)

`search_nhanes`, [5](#)