# Package: nesstarR (via r-universe)

October 28, 2026

**Type** Package

**Title** Read and Export 'Nesstar' Binary Files

**Version** 0.1.0

**Description** Reads '.Nesstar' binary files into data frames. '.Nesstar' is the file format of 'Nesstar Publisher', survey documentation software from the International Household Survey Network <https://www.ihsn.org/node/557>. Statistical agencies, for example the Ministry of Statistics and Programme Implementation ('MoSPI', India), publish survey microdata in this format.

**License** MIT + file LICENSE

**URL** https://nesstarr.saketlab.org, https://github.com/saketkc/nesstarR

**BugReports** https://github.com/saketkc/nesstarR/issues

**Encoding** UTF-8

**Language** en-US

**Depends** R (>= 4.1.0)

**Imports** tools, utils

**Suggests** haven, knitr, rmarkdown, testthat (>= 3.0.0)

**Config/testthat/edition** 3

**Config/Needs/website** pkgdown, archive, saketkc/mospiR

**VignetteBuilder** knitr

**NeedsCompilation** yes

**Config/roxygen2/version** 8.1.0

**Author** Saket Choudhary [aut, cre, cph] (ORCID: <https://orcid.org/0000-0001-5202-7633>)

**Maintainer** Saket Choudhary <saketc@iitb.ac.in>

**Repository** https://cran.r-universe.dev

**Date/Publication** 2026-10-27 18:40:02 UTC

**RemoteUrl** https://github.com/cran/nesstarR

**RemoteRef** HEAD

**RemoteSha** e046e131c731eaa386f29acdda5c93110e744184

# Contents

nesstar_compare . . . . . 2
nesstar_datasets . . . . . 3
nesstar_export . . . . . 3
nesstar_metadata . . . . . 4
nesstar_parse . . . . . 5
nesstar_read_dataset . . . . . 6
nesstar_variables . . . . . 6

**Index** **8**

---

`nesstar_compare` *Compare a dataset with a reference data frame*

---

## Description

Compares one dataset of a Nesstar file with the same data from a different source, for example a Stata or SPSS export. Use this function to make sure that nesstarR decodes a file correctly.

## Usage

```
nesstar_compare(x, dataset_number, reference)
```

## Arguments

| | |
|---|---|
| `x` | A `nesstar_binary` object from `nesstar_parse`. |
| `dataset_number` | Integer dataset number. |
| `reference` | A data frame with the same rows, in the same order. |

## Details

The function matches the columns by name. Before the match, it changes each group of characters that are not letters or digits to _, removes _ from the start and the end, and changes the names to upper case. Thus, `Srl._No.` and `Srl_No` are the same column.

If all values in the two columns are numbers, the function compares the numbers. Thus, `"010"` and `10` are equal. Otherwise, it compares the values as text, without leading and trailing spaces. Empty strings and `NA` are equal.

Remove value labels from `reference` before the comparison, for example with `haven::zap_labels()`.

## Value

A data frame with one row for each column name and these columns:

`column` The column name after the change.

`in_nesstar`, `in_reference` `TRUE` if the column is in the source.

`different_values` The number of rows with different values. `NA` if the column is in only one source.

### Examples

```
path <- system.file("extdata", "example.Nesstar", package = "nesstarR")
x <- nesstar_parse(path)
reference <- nesstar_read_dataset(x, dataset_number = 1)
reference$AGE[1] <- 26
nesstar_compare(x, dataset_number = 1, reference = reference)
```

---

## nesstar_datasets *List the datasets in a Nesstar file*

---

### Description

List the datasets in a Nesstar file

### Usage

```
nesstar_datasets(x)
```

### Arguments

| | |
|---|---|
| x | A `nesstar_binary` object from `nesstar_parse`. |

### Value

A data frame with columns `dataset_number`, `row_count`, `variable_count`.

### Examples

```
path <- system.file("extdata", "example.Nesstar", package = "nesstarR")
x <- nesstar_parse(path)
nesstar_datasets(x)
```

---

## nesstar_export *Export datasets to CSV files*

---

### Description

Writes each dataset to a different CSV file. The file name contains the name of the Nesstar file and the dataset number.

### Usage

```
nesstar_export(
  x,
  output_dir,
  datasets = NULL,
  compress = TRUE,
  chunk_size = 50000L
)
```

### Arguments

| | |
|---|---|
| `x` | A `nesstar_binary` object. |
| `output_dir` | Directory for the output files. If the directory does not exist, the function makes it. |
| `datasets` | Integer vector of the dataset numbers to export. If `NULL` (the default), the function exports all datasets. |
| `compress` | Logical. If `TRUE` (the default), the function writes gzip files (`.csv.gz`). |
| `chunk_size` | Number of rows to read at a time. A smaller value decreases the memory use. |

### Value

Invisibly, a character vector of output file paths.

### Examples

```
path <- system.file("extdata", "example.Nesstar", package = "nesstarR")
x <- nesstar_parse(path)
out <- nesstar_export(x, output_dir = tempdir())
read.csv(out[1])
```

---

## `nesstar_metadata` *Read the XML metadata of a Nesstar file*

---

### Description

Decodes the Huffman-coded XML blocks in the file. The XML blocks contain the dataset file names, the variable labels and the category labels.

### Usage

```
nesstar_metadata(x)
```

### Arguments

| | |
|---|---|
| `x` | A `nesstar_binary` object from `nesstar_parse`. |

### Details

Some older NSS files use different record IDs. For these files, the function cannot decode the metadata. It gives a warning and returns `NULL`.

### Value

A named list:

`datasets` A list with one element for each dataset. Each element contains `dataset_number`, `file_name` and `variables`. `variables` is a list with one list for each variable. Each variable list contains `name`, `label` and `categories`.

If the function cannot decode the metadata, it returns `NULL`.

## Examples

```
path <- system.file("extdata", "example.Nesstar", package = "nesstarR")
x <- nesstar_parse(path)
meta <- nesstar_metadata(x)
meta$datasets[[1]]$variables[[1]]$label
```

---

`nesstar_parse` *Parse a Nesstar file*

---

## Description

Decodes the resource index, the dataset descriptors and the variable directories. Does not decode row data.

## Usage

```
nesstar_parse(path)
```

## Arguments

| | |
|---|---|
| `path` | Path to a `.Nesstar` file. |

## Details

The function does not keep the file in memory. The other functions read only the bytes that they need from the file. Thus, the memory use does not increase with the file size. Do not move or change the file while you use the result.

## Value

A `nesstar_binary` list (S3 class).

## Examples

```
path <- system.file("extdata", "example.Nesstar", package = "nesstarR")
x <- nesstar_parse(path)
x
```

## nesstar_read_dataset *Read one dataset into an R data frame*

### Description

Numeric columns (mode 5) become `numeric`. String columns (mode 1) become `character`. Missing values become `NA`.

### Usage

```
nesstar_read_dataset(x, dataset_number, columns = NULL)
```

### Arguments

| | |
|---|---|
| `x` | A `nesstar_binary` object. |
| `dataset_number` | Integer dataset number. |
| `columns` | Character vector of the column names to read. If `NULL` (the default), the function reads all columns. |

### Value

A `data.frame`.

### Examples

```
path <- system.file("extdata", "example.Nesstar", package = "nesstarR")
x <- nesstar_parse(path)
nesstar_read_dataset(x, dataset_number = 1)
nesstar_read_dataset(x, dataset_number = 1, columns = c("AGE", "NAME"))
```

## nesstar_variables *List variables in one dataset*

### Description

List variables in one dataset

### Usage

```
nesstar_variables(x, dataset_number)
```

### Arguments

| | |
|---|---|
| `x` | A `nesstar_binary` object. |
| `dataset_number` | Integer dataset number (as shown by `nesstar_datasets`). |

## Value

A data frame with one row for each variable. The columns give the storage format of the variable.

## Examples

```
path <- system.file("extdata", "example.Nesstar", package = "nesstarR")
x <- nesstar_parse(path)
nesstar_variables(x, dataset_number = 1)
```

# Index

nesstar_compare, 2
nesstar_datasets, 3, *6*
nesstar_export, 3
nesstar_metadata, 4
nesstar_parse, *2–4*, 5
nesstar_read_dataset, 6
nesstar_variables, 6