

Package ‘asciiSetupReader’

February 5, 2019

Type Package

Title Reads Fixed-Width ASCII Data Files (.txt or .dat) that Have
Accompanying Setup Files (.sps or .sas)

Version 2.0.1

Description Lets you open a fixed-width ASCII file (.txt or .dat) that has an accompanying setup file (.sps or .sas). These file combinations are sometimes referred to as .txt+.sps, .txt+.sas, .dat+.sps, or .dat+.sas. This will only run in a txt-sps or txt-sas pair in which the setup file contains instructions to open that text file. It will NOT open other text files, .sav, .sas, or .por data files. Fixed-width ASCII files with setup files are common in older (pre-2000) government data.

License MIT + file LICENSE

URL <https://github.com/jacobkap/asciiSetupReader>

BugReports <https://github.com/jacobkap/asciiSetupReader/issues>

Depends R (>= 3.1.0)

Imports data.table, haven, readr, stringr, zoo

Suggests covr, knitr, rmarkdown, spelling, testthat

VignetteBuilder knitr

Encoding UTF-8

Language en-US

LazyData true

RoxygenNote 6.1.1

NeedsCompilation no

Author Jacob Kaplan [aut, cre]

Maintainer Jacob Kaplan <jkkaplan6@gmail.com>

Repository CRAN

Date/Publication 2019-02-05 21:13:40 UTC

R topics documented:

make_sps_setup	2
parse_setup	3
read_ascii_setup	4
sas_ascii_reader	5
spss_ascii_reader	7

Index	9
--------------	----------

make_sps_setup	<i>Create an SPSS setup file (.sps) to use for reading in fixed-width text files</i>
----------------	--

Description

make_sps_setup() is used to create the setup file used in reading in fixed-width text files. Often the setup file comes with the data file but in some cases (usually with government data) you will need to create the setup file yourself.

Usage

```
make_sps_setup(file_name, col_positions, col_names = NULL,
               col_labels = NULL, value_labels = NULL, missing_values = NULL)
```

Arguments

file_name	Name of the file to be saved (e.g. "setup_file1"). There is no need to put the .sps extension in the file name.
col_positions	Either a vector of strings indicating the start and end position of each column (e.g. "1-3", "4-5") or a vector of the widths of the columns (e.g. 3, 2).
col_names	A vector of names for the columns. If none are provided, will automatically create names based on column number (e.g. V1, V2, V3).
col_labels	A vector of labels for the columns. These are often longer and more descriptive than the col_names. These are the values used as column names if real_names = TRUE in reading in the data.
value_labels	A vector with the value first, then an '=' then the label. Each new column should have the column named followed by '='.
missing_values	A vector of strings with the column name followed by the values to be replaced by NA.

Value

Does not return any object. Saves the .sps file that is created.

Examples

```
## Not run:
value_labels <- c("var1 = ",
                 "1 = label 1",
                 "2 = label 2",
                 "3 = label 3",
                 "4 = label 4",
                 "5 = label 5",
                 "var3 = ",
                 "1A = alpha",
                 "1B = bravo",
                 "1C = cat")
missing_values <- c("state name", "9", "-8", "county", "-8")
make_sps_setup(file_name = "example_name",
               col_positions = c(1, 3, 4, 2),
               col_names = c("var1", "var2", "var3", "var4"),
               col_labels = c("state name", "county",
                              "population", "census region code"),
               value_labels = value_labels,
               missing_values = missing_values)

## End(Not run)
```

parse_setup	<i>Parse the setup file (.sps or .sas).</i>
-------------	---

Description

Parse the setup file (.sps or .sas).

Usage

```
parse_setup(setup_file)
```

Arguments

setup_file	Name of the SPSS or SAS setup file - should be a .sps or .sas (.txt also accepted as are these files in zipped format)
------------	--

Value

A list of length 3. The first object ("setup") is a data frame containing 4 columns: first the non-descriptive name of each column, The second column is the descriptive name of the column. Columns three and four and the beginning and ending number of the column (used to determine the columns location in the fixed-width data file).

The second object ("value_labels") in the list is list of named vectors for the value labels. The list has a length equal to the number of columns with value labels. If there are no value labels, this will be NULL.

The third object ("missing") in the list is a data.frame with two columns. The first column says the variable name and the second column says the value that is missing and will be replaced with NA.

Examples

```
## Not run:
sas_name <- system.file("extdata", "example_setup.sas",
                        package = "asciiSetupReader")
sas_example <- parse_setup(sas_name)

sps_name <- system.file("extdata", "example_setup.sps",
                        package = "asciiSetupReader")
sps_example <- parse_setup(sps_name)

## End(Not run)
```

read_ascii_setup	<i>Read fixed-width ASCII file using SPSS or SAS Setup file.</i>
------------------	--

Description

read_ascii_setup() is used when you need to read an fixed-width ASCII (text) file that comes with a setup file. The setup file provides instructions on how to create and name the columns, and fix the key-value pairs (sometimes called value labels). This is common in government data, particular data produced before 2010.

Usage

```
read_ascii_setup(data, setup_file, use_value_labels = TRUE,
                 use_clean_names = TRUE, select_columns = NULL,
                 coerce_numeric = TRUE)
```

Arguments

data	Name of the ASCII (.txt or .dat) file that contains the data. This file may be zipped with a file extension of .zip.
setup_file	Name of the SPSS or SAS setup file - should be a .sps or .sas (.txt also accepted as are these files in zipped format)
use_value_labels	If TRUE, fixes value labels of the data. e.g. If a column is "sex" and has values of 0 or 1, and the setup file says 0 = male and 1 = female, it will make that change. Using this parameter for enormous files may slow down the package considerably.
use_clean_names	If TRUE fixes column names from default column name in the setup file (e.g. V1, V2) to the descriptive label for the column provided in the file (e.g. age, sex, etc.).

- `select_columns` Specify which columns from the dataset you want. If NULL, will return all columns. Accepts the column number (e.g. 1:5), column name (e.g. V1, V2, etc.) or column label (e.g. VICTIM_NAME, CITY, etc.).
- `coerce_numeric` If TRUE (default) will make columns where all values can be made numeric into numeric columns. Useful as FALSE if variables have leading zeros - such as US Census FIPS codes.

Value

data.frame of the data from the ASCII file

Examples

```
# Text file is zipped to save space.
dataset_name <- system.file("extdata", "example_data.zip",
  package = "asciiSetupReader")
sps_name <- system.file("extdata", "example_setup.sps",
  package = "asciiSetupReader")

## Not run:
example <- read_ascii_setup(data = dataset_name,
  setup_file = sps_name)

# Does not fix value labels
example2 <- read_ascii_setup(data = dataset_name,
  setup_file = sps_name, use_value_labels = FALSE)

# Keeps original column names
example3 <- read_ascii_setup(data = dataset_name,
  setup_file = sps_name, use_clean_names = FALSE)

## End(Not run)

# Only returns the first 5 columns
example4 <- read_ascii_setup(data = dataset_name,
  setup_file = sps_name, select_columns = 1:5)
```

sas_ascii_reader

Read fixed-width ASCII file using SAS Setup file.

Description

`sas_ascii_reader()` and `spss_ascii_reader()` are used when you need to read an fixed-width ASCII (text) file that comes with a setup file. These file combinations are sometimes referred to as .txt+.sps, .txt+.sas, .dat+.sps, or .dat+.sas. The setup file provides instructions on how to create and name the columns, and fix the key-value pairs (sometimes called value labels). This is common in government data, particular data produced before 2010.

Usage

```
sas_ascii_reader(dataset_name, sas_name, value_label_fix = TRUE,
  real_names = TRUE, keep_columns = NULL, coerce_numeric = TRUE)
```

Arguments

dataset_name	Name of the ASCII (.txt) file that contains the data. This file may be zipped with a file extension of .zip.
sas_name	Name of the SAS Setup file - should be a .sas or .txt file.
value_label_fix	If TRUE, fixes value labels of the data. e.g. If a column is "sex" and has values of 0 or 1, and the setup file says 0 = male and 1 = female, it will make that change. The reader is much faster if this parameter is FALSE.
real_names	If TRUE fixes column names from default column name in the SPSS setup file (e.g. V1, V2) to the name it says the column is called (e.g. age, sex, etc.).
keep_columns	Specify which columns from the dataset you want. If NULL, will return all columns. Accepts the column number (e.g. 1:5), column name (e.g. V1, V2, etc.) or column label (e.g. VICTIM_NAME, CITY, etc.).
coerce_numeric	If TRUE (default) will make columns where all values can be made numeric into numeric columns. Useful as FALSE if variables have leading zeros - such as US Census FIPS codes.

See Also

[spss_ascii_reader](#) For using an SPSS setup file

Other ASCII Reader functions: [spss_ascii_reader](#)

Examples

```
# Text file is zipped to save space.
dataset_name <- system.file("extdata", "example_data.zip",
  package = "asciiSetupReader")
sas_name <- system.file("extdata", "example_setup.sas",
  package = "asciiSetupReader")

## Not run:
example <- sas_ascii_reader(dataset_name = dataset_name,
  sas_name = sas_name)

# Does not fix value labels
example2 <- sas_ascii_reader(dataset_name = dataset_name,
  sas_name = sas_name, value_label_fix = FALSE)

# Keeps original column names
example3 <- sas_ascii_reader(dataset_name = dataset_name,
  sas_name = sas_name, real_names = FALSE)
```

```
## End(Not run)
# Only returns the first 5 columns
example <- sas_ascii_reader(dataset_name = dataset_name,
  sas_name = sas_name, keep_columns = 1:5)
```

spss_ascii_reader *Read fixed-width ASCII file using SPSS Setup file.*

Description

spss_ascii_reader() and sas_ascii_reader() are used when you need to read an fixed-width ASCII (text) file that comes with a setup file. These file combinations are sometimes referred to as .txt+.sps, .txt+.sas, .dat+.sps, or .dat+.sas. The setup file provides instructions on how to create and name the columns, and fix the key-value pairs (sometimes called value labels). This is common in government data, particular data produced before 2010.

Usage

```
spss_ascii_reader(dataset_name, sps_name, value_label_fix = TRUE,
  real_names = TRUE, keep_columns = NULL, coerce_numeric = TRUE)
```

Arguments

dataset_name	Name of the ASCII (.txt) file that contains the data. This file may be zipped with a file extension of .zip.
sps_name	Name of the SPSS Setup file - should be a .sps or .txt (zipped text files also work) file.
value_label_fix	If TRUE, fixes value labels of the data. e.g. If a column is "sex" and has values of 0 or 1, and the setup file says 0 = male and 1 = female, it will make that change. The reader is much faster if this parameter is FALSE.
real_names	If TRUE fixes column names from default column name in the SPSS setup file (e.g. V1, V2) to the name it says the column is called (e.g. age, sex, etc.).
keep_columns	Specify which columns from the dataset you want. If NULL, will return all columns. Accepts the column number (e.g. 1:5), column name (e.g. V1, V2, etc.) or column label (e.g. VICTIM_NAME, CITY, etc.).
coerce_numeric	If TRUE (default) will make columns where all values can be made numeric into numeric columns. Useful as FALSE if variables have leading zeros - such as US Census FIPS codes.

Value

Data.frame of the data from the ASCII file

See Also

[sas_ascii_reader](#) For using an SAS setup file

Other ASCII Reader functions: [sas_ascii_reader](#)

Examples

```
# Text file is zipped to save space.
dataset_name <- system.file("extdata", "example_data.zip",
  package = "asciiSetupReader")
sps_name <- system.file("extdata", "example_setup.sps",
  package = "asciiSetupReader")

## Not run:
example <- spss_ascii_reader(dataset_name = dataset_name,
  sps_name = sps_name)

# Does not fix value labels
example2 <- spss_ascii_reader(dataset_name = dataset_name,
  sps_name = sps_name, value_label_fix = FALSE)

# Keeps original column names
example3 <- spss_ascii_reader(dataset_name = dataset_name,
  sps_name = sps_name, real_names = FALSE)

## End(Not run)
# Only returns the first 5 columns
example4 <- spss_ascii_reader(dataset_name = dataset_name,
  sps_name = sps_name, keep_columns = 1:5)
```


Index

`make_sps_setup`, [2](#)

`parse_setup`, [3](#)

`read_ascii_setup`, [4](#)

`sas_ascii_reader`, [5](#), [8](#)

`spss_ascii_reader`, [6](#), [7](#)