Package 'rcxl'

November 6, 2026

Title Read xlsx Files with a Native Parser

Version 0.1.2

Description Reads tabular data from xlsx files with a specialized C parser. Worksheet XML is scanned in a single pass into compact per-column buffers, with no intermediate document model, and the R vectors are built from those. Bundles the 'miniz' and 'libdeflate' decompressors to read the underlying archive.

License MIT + file LICENSE

Copyright file inst/COPYRIGHTS

URL https://github.com/vlshields/rcxl

BugReports https://github.com/vlshields/rcxl/issues

Encoding UTF-8

Depends R (>= 4.0)

Suggests readxl, cellranger

Config/roxygen2/version 8.1.0

NeedsCompilation yes

Author Vincent Shields [aut, cre, cph],
Rich Geldreich [ctb, cph] (author of the bundled miniz),
Tenacious Software LLC [cph] (copyright holder of the bundled miniz),
RAD Game Tools and Valve Software [cph] (copyright holder of the bundled miniz),
Martin Raiber [ctb, cph] (contributor to and copyright holder of the miniz ZIP reader),
Alex Evans [ctb] (author of the PNG writer in the bundled miniz),
Alistair Moffat [ctb] (author of the minimum-redundancy code in the bundled miniz),
Jyrki Katajainen [ctb] (author of the minimum-redundancy code in the bundled miniz),
Eric Biggers [ctb, cph] (author of the bundled libdeflate),
Google LLC [cph] (copyright holder of the bundled libdeflate)

Maintainer Vincent Shields <vince.shields913@gmail.com>

Repository CRAN

Date/Publication 2026-11-06 07:20:02 UTC

Contents

read_xlsx 2
read_xlsx_all 4
xlsx_sheets 5

Index **6**

`read_xlsx` *Read an xlsx worksheet*

Description

Reads one worksheet from an xlsx file into a data frame. Column types are guessed from the cells unless overridden with `col_types`.

Usage

```
read_xlsx(
  path,
  sheet = 1L,
  col_names = TRUE,
  trim_ws = TRUE,
  range = NULL,
  skip = 0L,
  n_max = Inf,
  col_types = NULL,
  na = "",
  name_repair = c("unique", "minimal", "check_unique")
)
```

Arguments

`path`	Path to an xlsx file. Tilde expansion is applied.
`sheet`	Worksheet to read, as a 1-based index or a sheet name. Overridden when `range` names a sheet.
`col_names`	`TRUE` to use the first row of the read area as column names, `FALSE` to number them `V1`, `V2`, ..., or a character vector of names, one per sheet column including any skipped via `col_types` (the skipped names are dropped from the result).
`trim_ws`	Trim leading and trailing whitespace from string cells and column names?

range	A1-style cell range to read: "B3:D87", a single cell "B3", whole columns "B:D", or whole rows "3:87", optionally sheet-qualified as "Sheet1!B3:D87" or "'My Sheet'!B3:D87". When given, skip and n_max are ignored. An axis the range bounds is read exactly, blank cells included; an open axis is trimmed to the sheet extent.
skip	Number of rows to skip before reading anything. Ignored when range is given.
n_max	Maximum number of data rows to read. The column name row does not count. Ignored when range is given.
col_types	NULL to guess every column, or a character vector of "guess", "skip", "logical", "numeric", "date", "text" or "list". A single value applies to every column; otherwise supply one entry per sheet column, counting "skip" columns. A "list" column keeps each cell's own type: length-one numeric, POSIXct, logical or character elements, with blank cells as logical NA.
na	Character vector of strings to read as NA. Matched against whitespace-trimmed cell text; entries that parse as numbers also blank numeric and date cells holding that value.
name_repair	How to resolve column names in the result. "unique" suffixes repeated names via make.unique() with separator "_", "minimal" keeps names as read, "check_unique" errors on any repeat.

Details

A guessed column becomes character if it holds any string cell, logical if it holds only booleans, POSIXct if it holds date cells and no plain numbers, and numeric otherwise. Every cell in the read area informs the guess. Columns with only blank cells become logical NA. Date cells are returned as POSIXct in UTC; both the 1900 and 1904 date systems are handled, including the nonexistent 29 Feb 1900 that the 1900 system counts.

Cells that cannot be coerced to a type forced through col_types become NA, with a warning giving the count and the first offending cell.

Value

A data.frame with one column per worksheet column read.

See Also

read_xlsx_all() to read several sheets in one workbook pass; xlsx_sheets() for the sheet names.

Examples

```
types <- system.file("extdata", "types.xlsx", package = "rcxl")
str(read_xlsx(types))

# read a rectangle rather than the whole sheet
read_xlsx(types, range = "A1:C3")

# force every column to character instead of guessing
```

```
str(read_xlsx(types, col_types = "text"))

# a sheet other than the first, by name or by position
multi <- system.file("extdata", "multisheet.xlsx", package = "rcxl")
read_xlsx(multi, sheet = "Beta")
```

`read_xlsx_all` *Read several worksheets in one pass*

Description

Reads multiple sheets from an xlsx file through one shared workbook pass, so the ZIP directory, shared strings and styles are parsed once. All arguments other than `sheets` apply to every sheet read.

Usage

```
read_xlsx_all(
  path,
  sheets = NULL,
  col_names = TRUE,
  trim_ws = TRUE,
  range = NULL,
  skip = 0L,
  n_max = Inf,
  col_types = NULL,
  na = "",
  name_repair = c("unique", "minimal", "check_unique")
)
```

Arguments

`path`	Path to an xlsx file. Tilde expansion is applied.
`sheets`	Sheets to read, as a vector of 1-based indices or sheet names. `NULL` reads every sheet in workbook order.
`col_names`	`TRUE` to use the first row of the read area as column names, `FALSE` to number them `V1`, `V2`, ..., or a character vector of names, one per sheet column including any skipped via `col_types` (the skipped names are dropped from the result).
`trim_ws`	Trim leading and trailing whitespace from string cells and column names?
`range`	As in `read_xlsx()`, except a sheet-qualified range is an error here; select sheets with `sheets`.
`skip`	Number of rows to skip before reading anything. Ignored when `range` is given.
`n_max`	Maximum number of data rows to read. The column name row does not count. Ignored when `range` is given.

col_types	NULL to guess every column, or a character vector of "guess", "skip", "logical", "numeric", "date", "text" or "list". A single value applies to every column; otherwise supply one entry per sheet column, counting "skip" columns. A "list" column keeps each cell's own type: length-one numeric, POSIXct, logical or character elements, with blank cells as logical NA.
na	Character vector of strings to read as NA. Matched against whitespace-trimmed cell text; entries that parse as numbers also blank numeric and date cells holding that value.
name_repair	As in read_xlsx(). A "check_unique" error names the offending sheet.

Value

A named list of data.frames, one per sheet read, named by sheet name.

Examples

```
multi <- system.file("extdata", "multisheet.xlsx", package = "rcxl")
sheets <- read_xlsx_all(multi)
names(sheets)
sheets[["Alpha"]]

# a subset of the sheets, still in one workbook pass
names(read_xlsx_all(multi, sheets = c("Alpha", "Gamma")))
```

---

xlsx_sheets *List worksheet names*

Description

List worksheet names

Usage

```
xlsx_sheets(path)
```

Arguments

path	Path to an xlsx file. Tilde expansion is applied.

Value

A character vector of sheet names in workbook order.

Examples

```
xlsx_sheets(system.file("extdata", "multisheet.xlsx", package = "rcxl"))
```

Index

make.unique(), *3*

read_xlsx, 2
read_xlsx(), *4, 5*
read_xlsx_all, 4
read_xlsx_all(), *3*

xlsx_sheets, 5
xlsx_sheets(), *3*