Package 'unexcel'

October 8, 2025

Title Revert Excel Serial Dates Back to Intended Day.Month Numerics

Version 0.1.0

Description Detects values imported from spreadsheets that were auto-converted to Excel date serials and reconstructs the originally intended day.month decimals (for example, '30.3' that Excel displayed as '30/03/2025'). The functions work in a vectorized manner, preserve non-serial values, and support both the 1900 and 1904 date systems.

License MIT + file LICENSE

Depends R (>= 3.6)

Suggests testthat (>= 3.0.0), knitr, rmarkdown

VignetteBuilder knitr

Encoding UTF-8

RoxygenNote 7.3.2

URL https://github.com/drhrf/unexcel

BugReports https://github.com/drhrf/unexcel/issues

Language en-US

NeedsCompilation no

Author Hercules Freitas [aut, cre] (ORCID: <https://orcid.org/0000-0003-1584-9157>)

Maintainer Hercules Freitas <hercules.freitas@uerj.br>

Repository CRAN

Date/Publication 2025-10-08 19:50:03 UTC

Contents

`fix_serial_columns` *Fix likely-serial columns in a data frame*

Description

Applies `restore_day_month()` only to columns that look dominated by Excel serials, controlled by a minimum fraction threshold.

Usage

```
fix_serial_columns(
  df,
  pmin = 0.8,
  low_serial = 20000,
  high_serial = 65000,
  year_window = 1990:2035,
  origin_mode = "auto",
  ref_date = Sys.Date()
)
```

Arguments

`df`	A data frame.
`pmin`	Minimum fraction of in-range integers to flag a column.
`low_serial`	Lower bound for plausible serials (inclusive).
`high_serial`	Upper bound for plausible serials (inclusive).
`year_window`	Integer vector of years that, when resolved, will be considered valid to revert. This guards against accidental matches.
`origin_mode`	One of "auto", "1900", or "1904". In "1900" mode the origin is "1899-12-30" (Excel's 1900 system with the leap-year quirk compensated). In "1904" mode the origin is "1904-01-01". In "auto" mode, the origin yielding dates with median proximity to a reference date is chosen; the reference can be controlled via `ref_date` for tests.
`ref_date`	Reference date for origin selection when origin_mode="auto". Defaults to Sys.Date(); set to a fixed Date in tests for determinism.

Value

The data frame with corrected columns where applicable.

Examples

```
df <- data.frame(a = c(45812, 44730), b = c(1.2, 3.4))
fix_serial_columns(df)
```

`restore_day_month` *Revert Excel date serials to intended day.month numerics*

Description

Many spreadsheets auto-convert entries like '30.3' into dates. After import, those values arrive as Excel date serials (integers). This function detects such serials and reconstructs the intended 'day.month' decimals while leaving other entries intact. Both 1900 and 1904 systems are supported.

Usage

```
restore_day_month(
  x,
  low_serial = 20000,
  high_serial = 65000,
  year_window = 1990:2035,
  origin_mode = c("auto", "1900", "1904"),
  ref_date = Sys.Date()
)
```

Arguments

`x`	A vector (numeric, integer, character, or Date).
`low_serial`	Lower bound for plausible serials (inclusive).
`high_serial`	Upper bound for plausible serials (inclusive).
`year_window`	Integer vector of years that, when resolved, will be considered valid to revert. This guards against accidental matches.
`origin_mode`	One of "auto", "1900", or "1904". In "1900" mode the origin is "1899-12-30" (Excel's 1900 system with the leap-year quirk compensated). In "1904" mode the origin is "1904-01-01". In "auto" mode, the origin yielding dates with median proximity to a reference date is chosen; the reference can be controlled via `ref_date` for tests.
`ref_date`	Reference date for origin selection when origin_mode="auto". Defaults to Sys.Date(); set to a fixed Date in tests for determinism.

Value

Returns a numeric vector when restoration is unambiguous; otherwise character vector, preserving both restored and untouched values.

Examples

```
restore_day_month(c(45812, 12.5, 44730), origin_mode = "1900")
restore_day_month(c(45812, 44730), origin_mode = "auto")
```

Index