Package: spanish (via r-universe)

August 31, 2024

Type Package

Title Translate Quantities from Strings Spelled in Spanish to Integer

Version 0.5.0

Date 2017-12-14

URL https://github.com/verajosemanuel

BugReports https://github.com/verajosemanuel/spanish/issues

Imports magrittr, xml2

Description Character vector to numerical translation in Euros from Spanish spelled monetary quantities. Text must be previously cleaned & removed extraneous words, symbols or cents. Quantities MUST be written in a correct Spanish cause this isn't a grammar tool. Upper limit is up to the millions range. Reverse translation using to_words. Geocoding via Cadastral site.

License GPL-3

Encoding UTF-8

LazyData true

Collate 'geocode_cadastral.R' 'cadastral_references-data.R' 'to_number.R' 'cantidades-data.R' 'to_words.R' 'spanish.R'

RoxygenNote 6.1.1

Suggests testthat, tidyr

Repository https://ropenspain.r-universe.dev

RemoteUrl https://github.com/rOpenSpain/spanish

RemoteRef HEAD

RemoteSha 27b4f64a59b0db13471bdb316133fcffd4e1f4b4

Contents

`cadastral_references` *Cadastral references test data*

Description

Randomly selected data from catastro to test geocode_cadastral function

Usage

```
data(cadastral_references)
```

Format

A data frame.

Source

Sede Electrónica del Catastro

References

Catastro. Ministerio de Hacienda y función pública. (Catastro)

Examples

```
## Not run:
## source is cadastral reference number ##

 geocode_cadastral("0636105UF3403N", parse_files = FALSE)

## Use lapply to geocode cadastral references from dataframe columns.

 cadastral_references$new <- lapply(cadastral_references$cadref1, geocode_cadastral)

## separate previously generated "new" data into columns usign tidyr

library(tidyr)
separate(cadastral_references, new, into = c('longitude','latitude'), sep = "," )
```

```
## source is folder. A loop is needed to process each kml file ##

files <- list.files("folder", full.names = T)

for (f in files) {
 coords <- geocode_cadastral(f, parse_files = TRUE)
 d <- as.data.frame(rbind(d , as.data.frame(coords, stringsAsFactors = F )))
}

# separate lat/lon into columns if you prefer using tidyr
d <- tidyr::separate(coords, into = c("longitude","latitude"), sep = "," )

## End(Not run)
```

`cantidades` *Cantidades test data*

Description

Randomly generated spanish spelled monetary integers to test to_number function

Usage

```
data(cantidades)
```

Format

A data frame.

Examples

```
to_number("mil trescientos noventa y dos")


## testing provided dataframe: cantidades

cantidades$var3 <- lapply(cantidades$var2, to_number)
```

`geocode_cadastral` *geocode by longitude and latitude from cadastral references.*

Description

Get longitude/latitude from valid cadastral ref. or kml files from catastro.

Arguments

x	A valid spanish cadastral reference.
parse_files	bool. Default to FALSE. Set TRUE if source are KML files.

Value

A string for longitude/latitude if found. NA if not found.

Warning

You may be banned if many requests in short time are made.

Examples

```
## source is cadastral reference number ##
## Not run:
## geocode_cadastral("0636105UF3403N", parse_files = FALSE)

##"36.5209422288168,-4.89298751473745"

## Use lapply to geocode cadastral references from dataframe columns.

cadastral_references$new <- lapply(cadastral_references$cadref1, geocode_cadastral)

## separate previously generated "new" data into columns usign tidyr

# library(tidyr)
# separate(cadastral_references, new, into = c('longitude','latitude'), sep = "," )

## source is folder. A loop is needed to process each kml file ##

# files <- list.files("folder", full.names = T)

# for (f in files) {
#  coords <- geocode_cadastral(f, parse_files = TRUE)
# d <- as.data.frame(rbind(d , as.data.frame(coords, stringsAsFactors = F )))
# }

# separate lat/lon into columns if you prefer using tidyr
# d <- tidyr::separate(coords, into = c("longitude","latitude"), sep = "," )

## End(Not run)
```

---

spanish *spanish: A package for spanish related data functions.*

Description

The spanish package provides some functions for spanish data: to_number() geocode_cadastral() to_words()

to_number()

Translate spanish spelled quantities into their integer counterparts. Allows you to translate to integer numerical words spelled in spanish. Text must be previously cleaned & removed extraneous words or symbols. Quantities MUST be written in a correct Spanish (this is not a grammar tool) The upper limit is up to the millions range. Cents must be removed. (in my TODO list to parse cents part)

to_words()

Translate to spanish spelled quantities from integers

geocode_cadastral()

geocode by longitude and latitude from cadastral references. Get longitude/latitude from valid cadastral ref. or kml files from catastro.

`to_number` *translate spanish spelled quantities into their integer counterparts.*

Description

Allows you to translate to integer numerical words spelled in spanish. Text must be previously cleaned & removed extraneous words or symbols Quantities MUST be written in a correct Spanish (this is not a grammar tool) The upper limit is up to the millions range. Cents must be removed.

Arguments

`x`	A spanish spelled number.

Examples

```
to_number("mil trescientos noventa y dos")

# Example table is provided: cantidades
cantidades$var3 <- lapply(cantidades$var2, to_number)
```

`to_words` *From integers to spanish spelled quantities.*

Description

Takes any integer from zero to millions range and returns spanish characters

Arguments

`x`	A valid integer amount.

Value

A string for the same integer number in spanish.

Index