Package 'lucas'

August 21, 2020

Type Package

Title Package to Download and Create the DB of LUCAS Data Harmonized

Version 1.0

Date 2020-08-21

Description Reproduces the harmonized DB of the ESTAT survey of the same name. The survey data is served as separate spreadsheats with noticeable differences in the collected attributes. The tool here presented carries out a series of instructions that harmonize the attributes in terms of name, meaning, and occurance, while also introducing a series of new variables, instrumental to adding value to the product. Outputs include one harmonized table with all the years, and three separate geometries, corresponding to the theoretical point, the gps location where the measurement was made and the 250m east-facing transect.

Depends R (>= 3.4)

Imports RPostgreSQL,
 rpostgis,
 plyr,
 utils,
 DBI

License GPL (>= 3)

LazyData TRUE

RoxygenNote 7.0.2

Encoding UTF-8

R topics documented:

Add_geom   *Add geometries and calculated distance*

Description

Add geometries to lucas harmonized table: - location of theoretical point(th_geom) from fields th_long, the_lat - location of lucas survey (gps_geom) from fields gps_long, gps_lat - lucas transect geometr (trans_geom) from fields gps_long, gps_lat - distance between theoretical and survey point (th_gps_dist)

Usage

```
Add_geom(con, save_dir)
```

Arguments

con	Connection to db
save_dir	Dirrectory where to save geometries

See Also

To create the conection please see Connect_to_db

Examples

```
## Not run:
Add_geom(con)
## End(Not run)
```

Add_new_cols *Add new columns to tables*

Description

Adds new columns to all table that will be necessary for when tables are merged. Includes letter group - first level of LUCAS land cover/land use classification system year - year of survey file_path_gisco_n/s/e/w/p - file path to full HD images on ESTAT GISCO cloud service for North, South, East, West, and Point images

Usage

```
Add_new_cols(con, years)
```

Arguments

con	Connection to db
years	Numeric vector of years to be harmonised

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Add_new_cols(con, c(2006, 2009, 2012, 2015, 2018))
## End(Not run)
```

Add_photo_fields_2006 *Add photo fields 2006*

Description

Add photo fields 2006

Add missing columns

Usage

```
Add_photo_fields_2006(con)

Add_missing_cols(con, years)
```

Arguments

con	Connection to database
years	Numeric vector of years to be harmonised

Functions

- `Add_photo_fields_2006`: missing columns photo_n/e/s/w in 2006 data from the information of the exif DB
- `Add_missing_cols`: missing columns to all tables before merge

See Also

To create the conection please see [Connect_to_db](#)

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Add_photo_fields_2006n(con)
## End(Not run)
## Not run:
Add_missing_cols(con, c(2006, 2009, 2012, 2015, 2018))
## End(Not run)
```

Add_revisit   *Add revisit column*

Description

Adds revisit column to lucas harmonized table to show the number of times between the years when the point was revisited.

Usage

```
Add_revisit(con)
```

Arguments

con	Connection to db

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Add_revisit(con)
## End(Not run)
```

`Align_Map_CSVs` *Align mapping CSVs*

Description

Corrects any typo, spelling mistake, or spelling difference in the user-created mapping CSVs, used to generate labels in subsequent User_friendly() function by alligning them to the mapping CSV of the latest survey.

Usage

```
Align_Map_CSVs(mapp_csv_folder, years)
```

Arguments

`mapp_csv_folder`

Directory where mapping files are stored

`years`   Numeric vector of years to be harmonised

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Align_Map_CSVs('/data/LUCAS_harmo/data/mappings', c(2006, 2009, 2012, 2015, 2018))
## End(Not run)
```

`Assert_files` *Assert files*

Description

Check that the user has downoad all the files needed

Usage

```
Assert_files(data_dir)
```

Arguments

`data_dir`   Character. Folder where you saved all the micro data downloaded from EURO-STAT

Value

Nothing if OK error if failed

Check_Map_CSVs   *Check mapping CSVs*

Description

Consistency check for Allign_map_CSVs function and creation of a new document with the explicit legends documented in document C3 (Classification) from LUCAS surveys.

Usage

```
Check_Map_CSVs(mapp_csv_folder, years, save_dir)
```

Arguments

mapp_csv_folder

                Directory where mapping files are stored

years   Numeric vector of years to be harmonised

save_dir   Directory where new C3 legends will be saved. Idealy (for consistency's sake) this directory should be the support_dir from main.R script.

See Also

To create the conection please see [Connect_to_db](#)

Connect_to_db   *Conntect to DB*

Description

connect to the db where you want to upload all LUCAS points

Usage

```
Connect_to_db(user, host, port, password, dbname)
```

Arguments

user   Character. User of the database

host   Character. Host of the DB

port   Integer. Port to connect to usually 5432

password   Character. Password to access to the DB

dbname   Character

Value

conection to the db

Examples

```
## Not run:
con <- Connect_to_db("andrrap", "localhost", 5432,"andrrap","andrrap")
con <- Connect_to_db("martlur", "/var/run/postgresql", 5432,"martlur","postgres")
con <- Connect_to_db("postgres", "172.15.0.10", 5432,"test","postgres")
## End(Not run)
```

Consistency_check   *Consistency checks*

Description

Perform consistency checks on newly created tables to ensure conformity in terms of column order
and data types

Usage

```
Consistency_check(con, years, manChangedVars)
```

Arguments

con	Connection to db
years	Numeric vector of years to be harmonised
manChangedVars	File path to csv of attributes and relevant years to which manual manipulation has been done and thus cannot clear a consistency of values check

See Also

To create the conection please see Connect_to_db

Examples

```
## Not run:
Consistency_check(con, c(2006, 2009, 2012, 2015, 2018))
## End(Not run)
```

`Correct_long`	*Harmonize long values in all tables*

Description

Harmonize long values in all tables

Usage

```
Correct_long(lucas2009)
```

Arguments

`lucas2009`	Dataframe with the 2009 data

Functions

- `Correct_long`: the long values of 2009 data by applying th_ew to th_long and erase this column

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Correct_long(con)
## End(Not run)
```

`Correct_th_loc`	*Correct theoretical long lat*

Description

Applying a correction of the values of columns th_long and th_lat according to the latest LUCAS grid

Usage

```
Correct_th_loc(con, lucas_grid)
```

Arguments

`con`	Connection to db
`lucas_grid`	File path to the csv file of the latest LUCAS grid

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Correct_th_loc(con, lucas_grid)
## End(Not run)
```

Create_tags   *Create tags for harmonized table*

Description

Create database tags (primary key), index, and spatial index and a new id column for the harmonized table

Usage

```
Create_tags(con)
```

Arguments

con   Connection to db

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Create_tags(con)
## End(Not run)de
```

Final_order_cols   *Final column order*

Description

Re-order columns of final tables

Usage

```
Final_order_cols(con)
```

Arguments

con   Connection to db

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Final_order_cols(con)
## End(Not run)
```

Merge_harmo   *Merge all tables*

Description

Merge all tables into a single harmonized version containing all years and change to relevant data type, as mapped in the record descriptor

Usage

```
Merge_harmo(con, rd)
```

Arguments

con   Connection to db
rd   Record descriptor in CSV format

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Merge_harmo(con, '/data/LUCAS_harmo/data/supportDocs/LUCAS_harmo_RD.csv')
## End(Not run)
```

Order_cols   *Change column order*

Description

Changes order of columns to fit the last survey (2018) and set all column data type to character varying in order to prepare for merge

Usage

```
Order_cols(con, years)
```

Arguments

con	Connection to db
years	Numeric vector of years to be harmonised

See Also

To create the conection please see Connect_to_db

Examples

```
## Not run:
Order_cols(con, c(2006, 2009, 2012, 2015))
## End(Not run)
```

Recode_vars   *Update values to fit 2018*

Description

Updates values in all tables to fit the last survey (2018) in terms of the coding of different variables; update is based on pre-made mappings

Usage

```
Recode_vars(con, csv, years, nonHarmonizeableVars)
```

Arguments

con	Connection to db
csv	CSV file that holds the pre-prepared variable mappings
years	Numeric vector of years to be harmonised
nonHarmonizeableVars	
	Characer vector of variables that cannot be harmonized between the years. These pertain to the variables which have been collected at the earlier stages (before 2018) of survey as ordered categorical variables, and at later stages (at and later than (?) 2018) - as discrete numbers. Such attributes are lc1/2_perc, lu1/2_perc and soil_stones_perc.

See Also

To create the conection please see Connect_to_db

Examples

```
## Not run:
Recode_vars(con, '/data/LUCAS_harmo/data/mappings/RecodeVars.csv', c(2006, 2009, 2012, 2015, 2018))
## End(Not run)
```

Remove_vars   *Remove unwanted columns*

Description

Removes unwanted columns as specified by user

Usage

```
Remove_vars(con, vars)
```

Arguments

con	Connection to db
vars	Character vector of variables, specified by name to remove

See Also

To create the conection please see Connect_to_db

Examples

```
## Not run:
Remove_vars(con, vars)
## End(Not run)
```

Rename_cols *Rename columns to match 2018 survey*

Description

Columns with different names between the surveys must be made to fit the last survey before merge

Usage

```
Rename_cols(con, csv)
```

Arguments

con	Connection the database
csv	CSV file with the relevant column name mappings

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Rename_cols(con, '/data/LUCAS_harmo/data/mappings/columnRename.csv')
## End(Not run)
```

UF_Consistency_check *User-friendly consistency check*

Description

Perform consistency checks on newly created UF fields to ensure conformity in terms of column order and data types

Usage

```
UF_Consistency_check(con)
```

Arguments

con	Connection to db

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
UF_Consistency_check(con)
## End(Not run)
```

Update_rd   *Update Record descriptor*

Description

Updates Record descriptor by adding a field (year) showing the year for which the variable exists
and removing variables listed in Remove_vars function from RD

Usage

```
Update_rd(con, rd, years)
```

Arguments

con	Connection to db
rd	Path to record descriptor csv
years	Character vector of the years of survey

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Update_rd(con, rd, years)
## End(Not run)
```

Upload_exif   *Upload_exif*

Description

Upload_exif

Usage

```
Upload_exif(con, exif)
```

Arguments

con	Connection to database
exif	the actual EXIF cvs located in mappings

Functions

- `Upload_exif`: to DB the exif information of LUCAS

See Also

To create the conection please see Connect_to_db

Examples

```
## Not run:
Add_photo_fields_2006n(con)
## End(Not run)
```

Upload_to_db	*Update csv to database*

Description

Upload to the DB all the 2009-2018 lucas csv downloaded from : https://ec.europa.eu/eurostat/web/lucas/data/primary-data there should be: EU_2012_20200213.CSV EU_2018_20200213.CSV OutScope_2015_20200225.CSV EU_2009_20200213.CSV EU_2015_20200225.CSV For 2006 it first combines them into one dataset for the entire year comprising of: BE_2006_0.xls CZ_2006_0.xls DE_2006_0.xls ES_2006_0.xls FR_2006_0.xls HU_2006_0.xls IT_2006_0.xls LU_2006_0.xls NL_2006_0.xls PL_2006_0.xls SK_2006_0.xls

Usage

```
Upload_to_db(data_dir, con)
```

Arguments

data_dir	Character. Folder where you saved all the micro data downloaded from EUROSTAT
con	PosGresSQLConnection Object.

Value

Boolean. True if the update to the DB worked FALSE otherwise

See Also

To create the conection please see lucas]Connect_to_db

To assert that you have the files Assert_files

Examples

```
## Not run:
Upload_to_db('/data/LUCAS_harmo/data/input', con)
## End(Not run)
```

Upper_case   *Upper case columns*

Description

Convert values in designated columns (lc1, lc1_spec, lu1, lu1_type, lc2, lc2_spec, lu2, lu2_type, cprn_lc) to uppercase for consistency's sake

Usage

```
Upper_case(con, years)
```

Arguments

con	Connection to db
years	Numeric vector of years to be harmonised

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
Upper_case(con,c(2009, 2012, 2015, 2018))
## End(Not run)
```

User_friendly   *User-friendly LUCAS harmonized*

Description

Creates columns with labels for coded variables and decodes all variables where possible to explicit labels

Usage

```
User_friendly(con, data_dir, years)
```

Arguments

con	Connection to db
data_dir	Directory where files are stored
years	Numeric vector of years to be harmonised

See Also

To create the conection please see [Connect_to_db](#)

Examples

```
## Not run:
User_friendly(con, '/data/LUCAS_harmo/data/mappings', c(2006, 2009, 2012, 2015, 2018))
## End(Not run)
```

Index