Package: spDataLarge (via r-universe)

May 20, 2026

Title Large datasets for spatial analysis

Version 2.2.0

Description Large datasets for spatial analysis. The data from this package could be retrived using the spData package.

Depends R (>= 3.5)

License CC0

RoxygenNote 7.3.3

LazyData true

LazyDataCompression bzip2

URL https://github.com/Nowosad/spData

BugReports https://github.com/Nowosad/spDataLarge/issues

Encoding UTF-8

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

Date/Publication 2026-01-27 14:53:52 UTC

RemoteUrl https://github.com/nowosad/spDataLarge

RemoteRef HEAD

RemoteSha d75b765711931463de59f8c8431814bff565fc3e

Contents

`bristol_ways` *Datasets providing a snapshot of Bristol's transport system*

Description

Data used in the transport chapter in Geocomputation with R. See https://r.geocompx.org/transport.html for details.

Usage

```
bristol_ways
```

Format

sf data frame objects

Source

https://wicid.ukdataservice.ac.uk/ and other open access sources

Examples

```
## Not run:
library(sf)
bristol_ways
bristol_od
bristol_region
bristol_ttwa
bristol_zones
bristol_stations

## End(Not run)
```

census_de *Datasets providing building blocks for a location analysis*

Description

Data used in the geomarketing chapter in Geocomputation with R. See https://r.geocompx.org/location.html for details.

Usage

```
census_de

metro_names

shops
```

Format

census_de A data.frame with census variables

metro_names A data.frame with metropolitan area information

shops An sf data.frame with POINT geometry

An object of class `data.frame` with 10 rows and 3 columns.

An object of class `sf` (inherits from `data.frame`) with 180035 rows and 3 columns.

Details

census_de A data.frame with German census data at 1km resolution. Contains columns: x, y (EPSG:3035 coordinates), pop (population count, 2022), women (percentage of women, 2011), mean_age (mean age in years, 2022), hh_size (average household size, 2022). Missing values are coded as -1.

metro_names A data.frame with metropolitan area names derived from Census 2022 population data. Contains columns: city, town, state. Generated by aggregating 1km population grid to 20km resolution, filtering cells with >500,000 inhabitants, and reverse geocoding centroids.

shops An sf data.frame of shop locations from OpenStreetMap for the identified metropolitan areas. Contains columns: osm_id, shop, geometry.

Examples

```
data("census_de", package = "spDataLarge")
head(census_de)

data("metro_names", package = "spDataLarge")
metro_names

data("shops", package = "spDataLarge")
head(shops)
```

comm *Community matrix of the Mt. Mongón*

Description

A community matrix with species as columns and sites as rows. The rownames correspond to the id which can be also found in [random_points]. Please note that in fact 100 sites have been visited but in 16 of them no species could be found (see again [random_points]). The data is used in the "Ecology" chapter in Geocomputation with R. See https://r.geocompx.org/eco.html for details.

Format

A dataframe with 100 sites (rows) and 69 species (columns). Species presence is given in percentage points (between 0-100 site). Due to overlapping cover between individual plants, the total cover per site can be >100%.

References

Muenchow, J., Bräuning, A., Rodríguez, E.F. & von Wehrden, H. (2013): Predictive mapping of species richness and plant species' distributions of a Peruvian fog oasis along an altitudinal gradient. Biotropica 45, 5, 557-566, doi: 10.1111/btp.12049.

Examples

```
data("comm", package = "spDataLarge")
```

dem.tif *Digital elevation model (DEM) of the Mongón study area.*

Description

A raster geotiff (EPSG:32717) representing altitude (ASTER GDEM, LP DAAC 2012) with 117 rows and 117 columns:

dem Altitude in m asl.. For more details, please refer to Muenchow et al. (2013).

. The data is used in the "Ecology" chapter in Geocomputation with R. See https://r.geocompx.org/eco.html for details.

Format

A geotiff file

References

Muenchow, J., Bräuning, A., Rodríguez, E.F. & von Wehrden, H. (2013): Predictive mapping of species richness and plant species' distributions of a Peruvian fog oasis along an altitudinal gradient. Biotropica 45, 5, 557-566, doi: 10.1111/btp.12049.

LP DAAC (2012): Land Processes Distributed Active Archive Center, located at the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center. Available at: https://lpdaac.usgs.gov/ (last accessed 25 January 2012).

Examples

```
system.file("raster/ndvi.tif", package = "spDataLarge")
```

`ep.tif` *Environmental predictors*

Description

A geotiff file (CRS: UTM zone 17S; EPSG:32717) with 117 rows and 117 columns:

1. `dem` Digital elevation model (ASTER GDEM, LP DAAC 2012), see also [spDataLarge::dem].
2. `ndvi` Normalized Differenced Vegetation index.
3. `carea` Catchment area.
4. `cslope` Catchment slope

The data is used in the "Ecology" chapter in Geocomputation with R. See https://r.geocompx.org/eco.html for details.

Format

A geotiff file

References

LP DAAC (2012): Land Processes Distributed Active Archive Center, located at the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center. Available at: https://lpdaac.usgs.gov/ (last accessed 25 January 2012).

Muenchow, J., Brauning, A., Rodriguez, E.F. & von Wehrden, H. (2013): Predictive mapping of species richness and plant species' distributions of a Peruvian fog oasis along an altitudinal gradient. Biotropica 45, 5, 557-566, doi: 10.1111/btp.12049.

Muenchow, J., Schratz, P., and A. Brenning. 2017. RQGIS: Integrating R with QGIS for Statistical Geocomputing. The R Journal 9, 2, 409-428. https://rjournal.github.io/archive/2017/RJ-2017-067/RJ-2017-067.pdf.

Examples

```
system.file("raster/ep.tif", package = "spDataLarge")
```

landsat.tif *Dataset landsat*

Description

This is a dataset containing the four bands (2, 3, 4, 5) of the Landsat 8 image for the area of Zion National Park. A scale factor should be applied before using the data: https://www.usgs.gov/faqs/how-do-i-use-a-scale-factor-landsat-level-2-science-products

Format

A multilayer geotiff file

Source

http://landsat-pds.s3.amazonaws.com/L8/038/034/LC80380342015230LGN00/

Examples

```
system.file("raster/landsat.tif", package = "spDataLarge")
```

london_streets *Streets of london downloaded from OSM*

Description

Data used in the "Bridges to GIS" chapter in Geocomputation with R. See https://r.geocompx.org/gis.html for details.

Usage

```
london_streets
```

Format

An sf-object with one attribute (`osm_id`) and one `geometry` column.

Source

OpenStreetMap (see https://www.openstreetmap.org/).

Examples

```
## Not run:
library(sf)
library(osmdata)
library(spData)
library(dplyr)
data(cycle_hire)
points = cycle_hire[1:25, ]
b_box = sf::st_bbox(points)
london_streets = opq(b_box) %>%
        add_osm_feature(key = "highway") %>%
        osmdata_sf() %>%
        `[[`("osm_lines")
london_streets = dplyr::select(london_streets, 1)

## End(Not run)
```

`lsl` *Landslide dataset from Southern Ecuador*

Description

Data used in the "Statistical learning for geographic data" chapter in Geocomputation with R. See https://r.geocompx.org/spatial-cv.html for details.

Usage

```
lsl
```

Format

The landslide dataset consists of two objects (CRS: UTM zone 17S; EPSG:32717):

1. `lsl` A `data.frame` object representing the coordinates of landslide initiation points with 350 rows and 8 columns.
2. `study_mask` An `sf`-object delineating the natural part of the study area.

Source

Landslide dataset of the RSAGA package: `data("landslides", package = "RSAGA")`.

Landslide Data:

Muenchow, J., Brenning, A., Richter, R. (2012): Geomorphic process rates of landslides along a humidity gradient in the tropical Andes, Geomorphology 139-140, 271-284. DOI: 10.1016/j.geomorph.2011.10.029.

Stoyan, R. (2000): Aktivitaet, Ursachen und Klassifikation der Rutschungen in San Francisco/Suedecuador. Unpublished diploma thesis, University of Erlangen-Nuremberg, Germany.

See Also

```
?ta.tif
```

Examples

```
data("lsl", "study_mask", package = "spDataLarge")
```

---

ndvi.tif *Normalized difference vegetation index for the Mongón study area.*

Description

NDVI raster geotiff (EPSG:32717) computed from a Landsat scene (path 9, row 67, acquisition date: 09/22/2000; USGS 2013) with 117 rows and 117 columns:

ndvi Normalized difference vegetation index.. For more details, please refer to Muenchow et al. (2013).

. The data is used in the "Ecology" chapter in Geocomputation with R. See https://r.geocompx.org/eco.html for details.

Format

A geotiff file

References

Muenchow, J., Bräuning, A., Rodríguez, E.F. & von Wehrden, H. (2013): Predictive mapping of species richness and plant species' distributions of a Peruvian fog oasis along an altitudinal gradient. Biotropica 45, 5, 557-566, doi: 10.1111/btp.12049.

USGS (2013): U.S. Geological Survey. Earth Explorer. Available at: http://earthexplorer.usgs.gov/ (last accessed 1 March 2013).

Examples

```
system.file("raster/ndvi.tif", package = "spDataLarge")
```

nlcd.tif *Dataset nlcd*

Description

This is a dataset containing a simplified version of the National Land Cover Database 2011 product for the Zion National Park area

Format

A geotiff file

Source

https://www.mrlc.gov/nlcd2011.php

Examples

```
system.file("raster/nlcd.tif", package = "spDataLarge")
```

nlcd2011.tif *Dataset nlcd2011*

Description

This is a dataset containing the National Land Cover Database 2011 product for the Zion National Park area

Format

A geotiff file

Source

https://www.mrlc.gov/nlcd2011.php

Examples

```
system.file("raster/nlcd2011.tif", package = "spDataLarge")
```

nz_elev.tif *New Zeleand elevation raster data*

Description

Elevation raster data of the New Zealand area from the Mapzen Terrain Service. For teaching purposes only

Format

A geotiff file

Source

https://registry.opendata.aws/terrain-tiles/

See Also

See the elevatr package: https://cran.r-project.org/web/packages/elevatr

Examples

```
system.file("raster/nz_elev.tif", package = "spDataLarge")
```

pol_pres15 *Polish election data 2015*

Description

Polish Presidential election 2015 data by gminy and Warsaw borough areal units

Usage

```
pol_pres15
```

Format

sf data frame object with 2495 areal units and 65 variables

- TERYT, TERYT0, gm0TERYT areal unit IDs
- name0original areal unit names
- namecleaned areal unit names
- typesfactor with levels "Rural", "Urban", "Urban/rural" and "Warsaw borough"
- I_turnoutFirst round turnout proportion

- II_turnoutRunoff round turnout proportion
- I_Duda_shareWinner first round share
- II_Duda_shareWinner runoff round share
- I_Komorowski_shareIncumbent first round share
- II_Komorowski_shareIncumbent runoff round share
- I_*First round aggregated counts of all polling station data
- II_*Runoff round aggregated counts of all polling station data

Note

"PVE" in variable names means "postal voting envelopes"; voters requesting a postal voting package are expected to return a postal voting envelope with a declaration, and a sealed voting envelope to be placed in the ballot box.

Author(s)

Roger Bivand

Source

http://prezydent2015.pkw.gov.pl/319_Pierwsze_glosowanie/, http://prezydent2015.pkw.gov.pl/325_Ponowne_glosowanie/, https://www.gov.pl/web/gugik

Examples

```
## Not run:
data("pol_pres15", package = "spDataLarge")
wd = aggregate(pol_pres15$I_entitled_to_vote, list(pol_pres15$types), sum)$x
boxplot(I_turnout ~ types, data = pol_pres15, width = wd)

## End(Not run)
```

---

random_points *Random points.*

Description

An [sf] (EPSG:32717) object with 100 randomly sampled points (stratified by altitude) on the Mt. Mongón (Peru). For more details, please refer to Muenchow et al. (2013). The data is used in the "Ecology" chapter in Geocomputation with R. See https://r.geocompx.org/eco.html for details.

Format

An [sf] object with 100 rows and 3 variables:

id Plot ID.

spri Number of vascular plant species per plot (species richness).

geometry Simple feature point geometry.

References

Muenchow, J., Bräuning, A., Rodríguez, E.F. & von Wehrden, H. (2013): Predictive mapping of species richness and plant species' distributions of a Peruvian fog oasis along an altitudinal gradient. Biotropica 45, 5, 557-566, doi: 10.1111/btp.12049.

Examples

```
data("random_points", package = "spDataLarge")
```

`srtm.tif` *Dataset srtm*

Description

This is a dataset containing the elevation raster data from SRTM of the Zion National Park area

Format

A geotiff file

Source

https://srtm.csi.cgiar.org/SRT-ZIP/SRTM_V41/SRTM_Data_GeoTiff/srtm_14_05.zip

Examples

```
system.file("raster/srtm.tif", package = "spDataLarge")
```

`study_area` *Mask of the study area on the Mount Mongón*

Description

An [sf] (EPSG:32717) object of geometry class polygon.

Format

An [sf] object with 1 row and 2 variables:

name Name.

geometry Simple feature polygon geometry.

References

Muenchow, J., Bräuning, A., Rodríguez, E.F. & von Wehrden, H. (2013): Predictive mapping of species richness and plant species' distributions of a Peruvian fog oasis along an altitudinal gradient. Biotropica 45, 5, 557-566, doi: 10.1111/btp.12049.

Examples

```
data("study_area", package = "spDataLarge")
```

`ta.tif` *Terrain attributes of the landslide dataset from Southern Ecuador*

Description

Data used in the "Statistical learning for geographic data" chapter in Geocomputation with R. See https://r.geocompx.org/spatial-cv.html for details.

Format

A geotiff file with the five terrain attribute layers: slope, plan curvature, profile curvature, elevation and catchment area.

Source

DEM dataset of the RSAGA package: `data("landslides", package = "RSAGA")`.

DEM:

Ungerechts, L. (2010): DEM 10m (triangulated from aerial photo - b/w). Available online:

'http://www.tropicalmountainforest.org/data_pre.do?citid=901'

Jordan, E., Ungerechts, L., Caceres, B. Penafiel, A. and Francou, B. (2005): Estimation by photogrammetry of the glacier recession on the Cotopaxi Volcano (Ecuador) between 1956 and 1997. *Hydrological Sciences* 50, 949-961.

See Also

`?lsl`

Examples

```
system.file("raster/ta.tif", package = "spDataLarge")
```

zion.gpkg *Dataset zion*

Description

This is a dataset containing the borders of Zion National Park

Format

A geopackage file

Source

https://irma.nps.gov/DataStore/DownloadFile/580617

Examples

```
system.file("vector/zion.gpkg", package = "spDataLarge")
```

zion_points *Point vector data*

Description

Dataset containing 30 randomly located points in the Zion National Park

Usage

```
zion_points
```

Format

A sf object

zion_points.gpkg *Point vector data*

Description

Dataset containing 30 randomly located points in the Zion National Park

Format

A geopackage file

Source

https://irma.nps.gov/DataStore/DownloadFile/580617

Examples

```
system.file("vector/zion_points.gpkg", package = "spDataLarge")
```

Index