Package 'satscanMapper'

July 5, 2018

Type Package

Version 1.0.1

Date 2018-06-30

Title 'SaTScan' (TM) Results Mapper

Author Jim Pearson <jpearson@statnetconsulting.com>, with support from 'NCI'.

Maintainer Jim Pearson <jpearson@statnetconsulting.com>

Imports graphics, grDevices, utils, foreign, stringr, RColorBrewer,
stats, sp, methods, SeerMapper

Depends R (>= 3.3.0)

Description Supports the generation of maps based on the results from
'SaTScan' (TM) cluster analysis.
The package handles mapping of Spatial and Spatial-Time analysis using
the discrete Poisson, Bernoulli, and exponential models of case data generating
cluster and location ('GIS') records containing observed, expected and observed/expected
ratio for U. S. states (and DC), counties or census tracts of individual
states based on the U. S. 'FIPS' codes for state, county and census tracts
(locations) using 2000 or 2010 Census areas, 'FIPS' codes, and boundary data.
'satscanMapper' uses the 'SeerMapper' package for the boundary data and
mapping of locations. Not all of the 'SaTScan' (TM) analysis and models generate
the observed, expected and observed/expected ratio values for the clusters and
locations.
The user can map the observed/expected ratios for locations
(states, counties, or census tracts) for each cluster with a p-value less than 0.05
or a user specified p-value.
The locations are categorized and colored based on either the cluster's Observed/Expected
ratio or the locations' Observed/Expected ratio.
The place names are provided for each census tract using data from 'NCI', the 'HUD' crossover
tables (Tract to Zip code) as of December, 2013, the USPS Zip code 5 database for 1999,
and manual look ups on the USPS.gov web site.

License GPL (>= 2)

LazyData no

LazyLoad no

BuildResaveData best

ByteCompile yes

NeedsCompilation no

Repository CRAN

Date/Publication 2018-07-05 13:00:02 UTC

R topics documented:

---

satscanMapper-package *A graphics package to create U. S. maps from SaTScan (TM) analysis results data at the state, state/county or state/census tract levels.*

Description

The satscanMapper package provides a means of creating U. S. maps at the state, county within a state, or census tracts within a state level (locations) using the results data from a SaTScan (TM) Spatial and Spatial-Temporal analysis. The package categorizes and color the observed over expected rations for locations used in the analysis. The locations can be states, or counties or census tracts within a state. The package requires the location ID be the U. S. FIPS codes to be able to create a geo file and be able to map the locations (states, counties or census tracts within a state or DC.) This allows the package to link up the data to the US Census 2000 and 2010 boundary data.

The package uses the observed/expected ratios calculated by SaTScan (TM) to categorize and color the locations within each cluster. The package can use the ratios of each location or the ratio of the cluster being mapped.

The SaTScanMapper package uses the SeerMapper and Seer2010Mapper packages for the state, county and census tract boundaries and mapping functions. SaTScanMapper then adds the cluster labels and outlines as requested.

Details

Package:	satscanMapper
Type:	Package
Version:	1.0.1
Date:	2018-06-30
License:	GPL-2
LazyLoad:	no

The package contains a function to validate the population and case file location IDs against the contained boundary data and generate an appropriate geographic file for use with SaTScan (TM). The minimum SaTScan results data files are:

1. Saved Parameter File

2. Cluster Information File

3. Location Information File

4. Relative Risk Estimates File

and must be in DBF format.

The boundary data for the states, counties and census tracts and mapping functions from the *SeerMapper* package is used.

A state, county and place name (for census tracts) information is provided to help make the cluster reports more readable to the user. Only the needed boundary data are loaded into memory.

Author(s)

Jim Pearson <jpearson@statnetconsulting.com>
Maintainer: Jim Pearson <jpearson@statnetconsulting.com>
Package compiled by Jim Pearson <jpearson@statnetconsulting.com>

---

co99_data	*U.S. County data table for all U. S. counties containing information to support the county fips code validation and report generation in the satscanMapper package.*

Description

Contains all of the U. S. county information for all 50 states, the District of Columbia and Puerto Rico needed by the satscanMapper package to validate the user provided county level location IDs and support the generation of the text reports on the SatScan (TM) results.

Usage

```
data(co99_data)
```

Format

The co99_data structure is:

row.names - character - 5 digit State/County FIPS code.

stID - character - 2 digit - State FIPS code.

stName - character - state full name.

coName - character - county full name.

saID - character - Seer Registry Area identifier (abbreviation).

c_X_00 - integer - the X value of the cartesian coordinate for the center of the county in an equal area projection for the census year 2000.

c_Y_00 - integer - the Y value of the cartesian coordinate for the center of the county in an equal area projection for the census year 2000.

c_X_10 - integer - the X value of the cartesian coordinate for the center of the county in an equal area projection for the census year 2010.

c_Y_10 - integer - the Y value of the cartesian coordinate for the center of the county in an equal area projection for the census year 2010.

tracts_00 - integer - the number of census tracts in the county in census year 2000.

tracts_10 - integer - the number of census tracts in the county in census year 2010.

/br The county data information contains the information for all U. S. counties.

Source

CENSUS.GOV web site (boundary data), HUD.gov Census Tract to ZipCode crossover table, USPS Zip code database (1999), and R MAPS package (state and county names) for the U. S. 2000 Census.

| co99_M_data | *U. S. County data table for Alaska, Hawaii, and Puerto Rico counties containing the projected coordinates of the state's county centroids.* |

Description

Contains the U. S. county information for Alaska, Hawaii, and Puerto Rico needed by the satscan-Mapper package to create the coordinates file to be used by SaTScan (TM). The main coordinates and boundary files were modified to shift Alaska, Hawaii and Puerto Rico to better positions on the US map. To be able to create a coordinates file based on the location IDs in the *.pop and *.cas input files to SaTScan (TM), the coordinates must be based on the original Lat and Long and not the moved locations. This dataset contains the unmoved projected coordinates of the of the centroids for all of the counties in the three affected states. Once the analysis is completed, only the mapping coordinates are needed.

Usage

data(co99_M_data)

Format

The co99_M_data structure is:

row.names - character - 5 digit State/County FIPS code.

stID - character - 2 digit - State FIPS code.

stName - character - state full name.

coName - character - county full name.

saID - character - Seer Registry Area identifier (abbreviation).

c_X_00 - integer - the X value of the cartesian coordinate for the center of the county in an equal area projection for the census year 2000.

c_Y_00 - integer - the Y value of the cartesian coordinate for the center of the county in an equal area projection for the census year 2000.

c_X_10 - integer - the X value of the cartesian coordinate for the center of the county in an equal area projection for the census year 2010.

c_Y_10 - integer - the Y value of the cartesian coordinate for the center of the county in an equal area projection for the census year 2010.

tracts_00 - integer - the number of census tracts in the county in census year 2000.

tracts_10 - integer - the number of census tracts in the county in census year 2010.

y - integer - a bit string identifying the census years the census tract existed. If the value is 1 or 3, the tract existed in the 2000 census. If the value is 2 or 3, the tract existed in the 2010 census.

/br The county data information contains the information for the states of Alaska, Hawaii, and Puerto Rico.

Source

CENSUS.GOV web site (boundary data), HUD.gov Census Tract to ZipCode crossover table, USPS Zip code database (1999), and R MAPS package (state and county names) for the U. S. 2000 Census.

CreateGeo4SS	*Function reads the population and case SaTScan (TM) input files and creates a coordinates file based on the U. S. state, county, or census tract FIPS code as the location IDs.*

Description

The satscanMapper creates maps based on the information and statistics data provide by in the SaTScan (TM) Results Files. The package can map the 50 U.S. states, District of Columbia and the Puerto Rico territory at the state, county and census tract levels.

The satscanMapper requires the location IDs to be the U. S. FIPS codes for the States (and DC/Territories), State/Counties, or State/County/Census Tracts depending on the mapping level required for the data. The FIPS codes and boundary data used by this package is based on the 2000 and 2010 U. S. Census and were pulled from the Census Bureau web site. The *SeerMapper* package is used for the mapping and the source of all boundary data.

The *CreateGeo4SS* function creates a coordinates file based on the location IDs in the input population and case files to the SaTScan (TM) package. The *satscanMapper* function takes the results and generates maps. To make sure the analysis and the clustering coordinates matches the state, county and census tract boundary data used by *SeerMapper* to map the results, it is CRITICAL the initial input coordinates file matches the projection and coordinates used by *SeerMapper*.

CreateGeo4SS reads the population and case files, collects the locations IDs used, validates the Location IDs against the U. S. FIPS standards, and creates a coordinates (.geo) file for use as input to the SaTScan (TM) package.

Usage

```
CreateGeo4SS( path    = NULL,   # optional path to the case and population files.
                 pop      = NULL,   # filename of the population file
                 cas      = NULL,   # filename of the case file

               geo      = NULL,   # output filename for the coordinates file
             overwrite = FALSE,  # overwrite permission on coordinates file
                                 #   def = FALSE
             header   = TRUE,   # headers present on the pop and case files
                                 #   def = TRUE
               censusYear= NULL    # def = 2000.
           )
```

Arguments

path	is the name of the directory containing the input population and case files. If present, it will also be used for the coordinates output file. The path= value will be prepended to the pop, case and geo filenames if needed to access the file.
pop	is the file name of the input population file. The value can be a full filename with a path or just a filename to be used with the path= call parameter. The format of the population file must be the Location ID followed by a white-space separator and any number of fields required. Only the location ID is evaluation. A valid pop or case file must be specified.
cas	is the file name of the input case file. The value can be a full filename with a path or just a filename to be used with the path= call parameter. The format of the case file must be the Location ID followed by a white-space separator and any number of fields required. Only the location ID is evaluation. A valid pop or case file must be specified.
geo	is the file name of the output coordinates file. If no extension is provided, ".geo" will be used. It the geo= call parameter is not specified, the base filename and path of the case and population will be used to create a coordinates file. The value can be a full filename with path or a filename to be used with the path= call parameter. The file specified is check to see if it already exists. If the overwrite= call parameter is set to TRUE, the existing file will be overwritten. If overwriting is not permitted and the file exists, no coordinates file is created and the caller informed . Note: A data.frame representing the coordinates file, is returned to the *CreateGeo4SS* caller and the functions value.

overwrite   is a logical value. The *overwrite* call parameter determines what will happen if the output coordinates filename already exists. If TRUE, the fill will be over-written. If FALSE, the file will not be overwritten and no output file will be generated. The default value is FALSE.

header   is a logical value. The *header* call parameter is passed through to the read.table functions to indicate if the input files have a header line as the first line of the file. If set to TRUE, the first line is processed as a field name line instead of data. If set to FALSE, the first line of the file is treated as data. The default value is TRUE. The *header* call parameter is also used to control whether or not to include a header line in the coordinate output file. Both the population and case files must have headers or not have headers.

censusYear   This is a character value of the census year of the location IDs in the population and case data records or the SatScan (TM) results data files. The acceptable values for the censusYear parameter are "2000" and "2010". The default value is "2000" for the 2000 census year.

Details

The *CreateGeo4SS* function creates a coordinates file for input to the SaTScan (TM) package. By using the *CreateGeo4SS* function, the caller an ensure the location IDs used in the population and case files match the FIPS codes for the specified census years and that the coordinates used in the analysis are equal-area projection coordinates that match the boundary data stored in the *SeerMapper* package (which draws the maps.) Once the coordinate file is created, the population, case and coordinates file can be used by the SaTScan (TM) package and the results mapped by the *satscanMapper* function in this package.

The locations IDs used in the population and case files must be the U. S. FIPS states (2-digit), state/counties (5-digit), or state/county/census tract (11-digit) codes based on the geographical units being mapped. Since the FIPS codes are different in the 2000 and 2010 census years, the caller needs to notify *CreateGeo4SS* and *satscanMapper* via the *censusYear* call parameter if the location identifiers are from the 2010 census years. The packages default to using the 2000 census FIPS codes and boundary data. If the location ID based on the 2010 census, set *censusYear* call parameter to "2010" and the packages adjust.

This package will only map the boundaries of states that are linked to data or have counties or tracts linked to data within their boundaries.

Value

None

Author(s)

Jim Pearson and Linda Pickle of StatNet Consulting, LLC, Gaithersburg, MD

Examples

```
######
##
```

```
## Example # 1 - creating a coordinates file for a SaTScan (TM) analysis
##
## This example access the population and case data files used in a
## SaTScan (TM) analysis:
##     a) read the files to build a list of all of the location IDs
##           involved in the analysis,
##     b) validates the location IDs against the U. S. Census web site's
##           list of Location IDs for that year and geographic level
##           (State, State/County or State/County/Census Tract), and
##     c) returns a list of centroid X, Y coordinates for each location ID.
##
## The coordinates file can be written to disk or returned as a data.frame
## to the caller as the return value of the CreateGeo4SS function.
##
## The coordinates values returned are based on an equal-area projection
## of the lat and long values of the area centroids based on the
## area boundary data used to map the areas.
##
## The example population and case data files distributed with the package
## are located in the installation directory for the satscanMapper package.
##
#
## return a DF of the coordination data
#
## Not run:

SSInstDir <- system.file(package="satscanMapper", "extdata")

PopFile   <- "inc_noadj_cir.pop"

CasFile   <- "inc_noadj_cir.cas"

GeoDF     <- CreateGeo4SS(path=SSInstDir,
                          pop=PopFile,

                          cas=CasFile,
                          censusYear="2010",
                          header=FALSE)

str(GeoDF)


## End(Not run)
#  Example removed due to time constraints
#
######


#####
#
#  This example #2 creates a coordinates file.
#
```

```
## Not run:

SSInstDir <- system.file(package="satscanMapper", "extdata")

PopFile   <- "inc_noadj_cir.pop"

CasFile   <- "inc_noadj_cir.cas"

GeoDF     <- CreateGeo4SS(path=SSInstDir,
                          pop=PopFile,
                          cas=CasFile,
                          geo="inc_noadj_cir.geo",
                          censusYear="2010",
                          header=FALSE)

# output file written to the extdata directory of the package installation.


## End(Not run)

####

print(" End of CreateGeo4SS examples.")

####
```

---

pl99_data | *A data.frame containing place name information used by the county and census tract data levels when creating the detailed run report in the satscanMapper package.*

Description

Contains the place name information table to support the text report on county and census tract place names.

Usage

```
data(pl99_data)
```

Format

The pl99_data data.frame contains:

row.names - character - a search key containing the county 5 digit fips code and the place name within the county.

stcoID - character - 5 digit state/county FIPS code containing the place name.

plName - character - the place name.

tracts_00 - integer - the number of census tracts in the county in census year 2000.

tracts_10 - integer - the number of census tracts in the county in census year 2010.

/br The place name index contains the information for all counties and census tracts. /br The data used to provide place names for the census tracts are derived the Missouri Census Data Center correlation tables, the HUD.gov Census Tract to ZipCode crossover tables for 2000 and 2010, and the USPS.gov zip code database for the 2000 and 2010 census years. Additional place name information was provided by NCI and via location research in the Internet.

Source

CENSUS.GOV web site (boundary data), HUD.gov Census Tract to ZipCode crossover table, USPS Zip code database (1999), and R MAPS package (state and county names) for the U. S. 2000 Census.

---

sa99_data	*Coordinates data table for the 20 U. S. NCI Seer Registry areas.*

Description

Contains the required information and centroid coordinates for the 20 U. S. NCI Seer Registry areas.

Usage

```
data(sa99_data)
```

Format

```
sa99_data structure:
    row.names - character - Seer Registry abbreviation.
    saName    - Character - Registry Name
    tracts    - integer   - Number of census tracts within state
    c_X_00    - integer   - centroid X in 2000 census
    c_Y_00    - integer   - centroid Y in 2000 census
    c_X_10    - integer   - centroid X in 2010 census
    c_Y_10    - integer   - centroid Y in 2010 census
```

/br The Seer Registry data set the information and centroid coordinates for the 20 NCI Seer Registries in the U. S.

Source

CENSUS.GOV web site (boundary data), HUD.gov Census Tract to ZipCode crossover table, USPS Zip code database (1999), and R MAPS package (state and county names) for the U. S. 2000 Census.

---

satscanMapper	*Function to map Results data produced by a SaTScan (TM) cluster analysis*

Description

The satscanMapper creates maps based on the information and statistics data provide by in the SaTScan (TM) Results Files. The package can map the 50 U.S. states, District of Columbia and the Puerto Rico Territory at the state, county and census tract levels.

The locations are identified using the U. S. FIPS codes for the State (and DC/Territories), Counties with in a state, and Census Tracts with a state. The boundary data used by this package is based on the 2000 and 2010 U. S. Census and were pulled from the Census Bureau web site. The *SeerMapper* package is used for the mapping and the source of all boundary data.

Usage

```
satscanMapper( resDir     = NULL,           # no default, must be provided
                 prmFile    = NULL,            # no default, must be provided
               outDir     = NULL,          # default = NULL, uses value in resDir

                 censusYear  = NULL,            # default = census year 2000.
                 categ       = NULL,          # default = 5
                 title       = NULL,          # no default, optional
                 pValue      = NULL,          # default = 0.05
                 report      = NULL,          # default = TRUE
                 runId       = NULL,          # default = ""
                 bndyCol     = NULL,          # default = grey(0.9)
                 label       = NULL,          # default = TRUE
                 outline     = NULL,          # default = TRUE
                 locO_EMap   = NULL,          # default = TRUE
                 clusO_EMap  = NULL           # default = TRUE
              )
```

Arguments

resDir   is the name of the directory containing the SaTScan (TM) saved session parameter file (.PRM). This directory will be used to hold the output graphics (PDF) and text output report (.txt) files.

prmFile   is the file name of the SaTScan (TM) saved session parameter file (.PRM). The saved session parameter file contains all of the information about the SaTScan analysis to be mapped, including the location of the results cluster, location and RR files. The base of the .PRM filename (without extension) is used to create the filenames for the graphic and text output files. If the output filename already exists, a new name based on the original name is created and used. or you can specify the 'runId' to prepend a few characters to the filenames to make them unique for each *satscanMapper* run.

outDir   is the name of the directory to write the PDF graphic map file and the TXT sum-
                  mary report file. Normally, the output files are written to the directory specified
                  in the 'resDir' parameter to keep the input and output information in the same
                  location. However, if the output files need to be written to a different directory,
                  the 'outDir' parameter can be used to specify the different directory. If the
                  'outDir' parameter is missing or *NULL*, the value of the 'resDir' parameter
                  is used as the output directory. If parameter is set to "", the current working
                  directory will be used when the output files are written. If the parameter is set
                  to a character string, it must be a valid/existing directory and will be used as the
                  destination when the output files are written. The default value is *NULL*.

censusYear   This is a character value of the census year of the location IDs in the population
                  and case data records or the SatScan (TM) results data files. The acceptable
                  values for the censusYear parameter are "2000" and "2010". The default value
                  is "2000" for the 2000 census year.

categ   This is an integer value specifying the number of categories to be used when
                  coloring the locations in the cluster maps. A preset of Observed/Expected value
                  break points is used based on the number of categories specified in the 'categ'
                  argument. 'categ' can range from *3* to var10. The default 'categ' value is
                  5. The breakpoints for each categ value from *3* to *10* are hard coded into the
                  package as follows:

```
categ    breakpoint list
 3     -Inf,                        0.9, 1.1,                    Inf
 4     -Inf,                        0.9, 1.1,           3.0, Inf
 5     -Inf,            0.5,        0.9, 1.1,           3.0, Inf
 6     -Inf,            0.5,        0.9, 1.1,    2.0, 3.0, Inf
 7     -Inf, 0.33,      0.5,        0.9, 1.1,    2.0, 3.0, Inf
 8     -Inf, 0.33,      0.5,        0.9, 1.1, 1.5, 2.0, 3.0, Inf
 9     -Inf, 0.33,      0.5, 0.67, 0.9, 1.1, 1.5, 2.0, 3.0, Inf
10     -Inf, 0.33, 0.4, 0.5, 0.67, 0.9, 1.1, 1.5, 2.0, 3.0, Inf
```

                  Future plans include allowing the caller to specify their own set of break points.

title   A character string(vector) used as the main title for the cluster maps. A *title* may
                  be one or two lines. If two lines, code title=c("line 1", "line 2"). for the
                  title call parameter. If no title is provided (*NULL* or ""), the prmFile argument
                  is used for the main title. The default is *NULL*, the package uses the prmFile
                  argument as the title.

pValue   The package uses a default P Value of *0.05*. Only cluster with a P Value < *0.05*
                  are colored on the resulting map by default. If the P Value needs to be a different
                  value, then the user can specify the override P Value using the PValue argument.
                  The PValue argument must be between var0.001 and var0.5 to be valid.

report   This is logical value. If *TRUE* a text report of the statistics from each cluster and
                  location in a tier format by State, County, Place, and Location (Census Tract)
                  will be created. If varFALSE, no report is generated. The default is *TRUE*.

runId   If a user is creating multiple mappings of the same SaTScan (TM) Results Data
                  Files, but with different arguments and options, they will want to save the PDF
                  and report text files for each run. Normally the PDF and report text files would

be overlaid. The 'runId' argument provides a way to modify the output PDF and text filenames to make a unique filename for the output. The 'runId' string is inserted near the end of the output filename. This allows multiple mapping runs against the same data and a collection of resulting maps. If 'runId' is not specified, the package will append to the '.prm' filename the string of "-Rnn", where 'nn' is a number from 01 to 99. If an output set already exists, 'nn' is increased by 1 until an unused output file name is discovered.

label   This is a logical argument that indicates whether or not the outlines of each cluster on a map is labeled with the cluster number from the SaTScan (TM) data. If a cluster spans multiple years, the number of years is added to the label. For example: "1-2", this is cluster number 1 and the cluster existing over a two year period including the current year being mapped. The default value for 'label' is varTRUE.

outline   This is a logical value specifying whether or not draw the outline of the cluster on the cluster map. If TRUE, an outline is draw. If *FALSE*, no outlines are drawn. The default value is *TRUE*.

bndyCol   is a character string that defines the color to use for the area boundaries. The default color value is *grey50*. The value provided is verified against the values returned from the *colors()* function.

locO_EMap   a logical variable. This parameter enables the use of the location Obs/Exp ratio to categorize and color the data areas. If *TRUE*. the location ratios are used. If *FALSE*, the location Obs/Exp ratio is not used. The default value is *TRUE*.

clusO_EMap   a logical variable. This parameter enables the use of the cluster Obs/Exp ratio to categorize and color the data areas. If *TRUE*. the cluster ratios are used. If *FALSE*, the cluster Obs/Exp ratio is not used. The default value is *FALSE*.

Details

The satscanMapper function creates one or more maps for the SaTScan (TM) cluster analysis results data. The following must be done to facilitate the use of SaTScan (TM) cluster analysis and the satscanMapper package:

1. The coordinates file must be generated by the *CreateGeo4SS* function in the package to ensure the area centroids used in the analysis match the boundary projection used in the area mapping.

2. The Coordinates type must be set to "Cartesian". The centroid values returned by the *CreateGeo4SS* function are based on an equal-area projection and not Lat/Long coordinates.

3. Time Precision must be set to "Years". The mapping package only supports time units of years.

4. Since mapping is spatial, only the Retrospective Purely Spatial and Space-Time and the Prospective Space-Time analysis are supported.

5. There restrictions on the Probability Model used are imposed by the type of analysis.

6. Time Aggregation is support for units of "years" for any number of years. The default is 1 year. This parameter is generally not used.

7. Both the circular and elliptic Spatial Window Shapes are supported.

8. The Minimum Temporal Cluster Size is supported for any number of years. When specified, maps are generated on a sliding time line for group of years.

9. The main results file and all Column Output files use the same "Main Results File" name. This information is stored in the saved session *.PRM file. Any value may be specified.

10. The Cluster Information dbase box must be checked.

11. The Location Information dbase box must be checked.

12. The Risk Estimate for Each Location dbase box must be checked.

13. IMPORTANT: Just before executing the SaTScan (TM) analysis, save the session *.prm file in an appropriate directory. Menu -> File -> Save Session or Save Session as. If you make changes to the analysis parameters make sure to re-save the session file.

The saved session *.PRM file provides the main mapping function with all of the information to generate the maps of the analysis results. The saved session parameters are checked to ensure the mapping can be done and any errors reported to the user.

The *CreateGeo4SS* function should be used to evaluate the location IDs in the population and/or case data files and generate the coordinates(geo) file for use in the SaTScan (TM) analysis.

The Location IDs used in the original population, case and coordinates(geo) input files and the output results files must be in one of three formats based on the mapping level desired: the 2 digit U. S. fips code for state data, the 5 digit U. S. fips code for county data, or 11 digit U. S. fips code for census tract data. If the Location IDs for the census year being mapped, do not match the boundary data, warnings will be generated to inform the user of the issue. Check to make sure the correct census year was specified. The boundary data and FIPS codes were pulled from the CENSUS.GOV web site for the 2000 and 2010 census years and should be complete.

If 2 digit fips code IDs are used, the SaTScan (TM) results information is mapped at the state, district and territory level. When states are mapped, the boundaries for all of the states in the US will be drawn.

If 5 digit fips code IDs are used, the SaTScan (TM) results information is mapped at the county level for the states containing the counties with data. State boundaries are drawn around the counties as a reference. Only the county boundaries in states with data will be drawn. Therefore, is the analysis results contain counties in multiple states, all of the boundaries for the counties in these states are drawn.

If 11 digit fips code IDs are used, the SaTScan (TM) location information is mapped at the census tract level for the states containing the census tracts with data. State and county boundaries are drawn for any state contain census tracts referenced by an analysis.

The area categorization and coloring are based on either the location and cluster ODE ratios reported in the results files. The ratio used is controlled by the *locO_EMap* and *clusO_EMap* parameters.

The 'bndyCol' parameter specifies the color used for the data area boundaries (state, county or census tract). It's value must be a colors that is either from the *colors()* name list or is a *#hhhhhh* or *#hhhhhhhh* format as a character vector.

Only areas with data are mapped. To drawn all of the data level boundaries up to the next level, specify 'fill' = *TRUE*. For example: All tract boundaries will be drawn up to their county level for tract data. For county data, all counties will be drawn up to the state boundary.

At the current time, the boundary data is based on the boundary data in the SeerMapper and Seer2010Mapper packages. An additional parameter is required to specify which census year is to be used: 2000 or 2010.

If a text report is required ('report' = *TRUE*), the package provides a listing of the Cluster Information, Location Information and Risk data in a tabular format for the user's review. The location information is listed for each cluster in a tier format with the following levels:

```
State
   County
      PlaceName
         Census Tract
```

If the locations are counties, county names are provided, but not the Place names or the Census Tract IDs levels.

The critical SaTScan (TM) run options the must be set are:

```
StartDate = yyyy/mm/dd
EndDate = yyyy/mm/dd

AnalysisType = 4   (default)
    Valid Types: 1 = Purely Spatial,
                 3 = Retrospective Space-time,
                 4 = Prospective Space-Time.
    Not-Supported Types:
                 2 = Purely Temporal,
                 5 = Spatial Variation in Temporal Trends,
                 6 = Prospective Purely Temporal


IncludeRelativeRisksCensusAreasDBase         = y     # *.rr.dbf
CensusAreasReportedClustersDBase             = y     # *.col.dbf
MostLikelyClusterEachCentroidDBase           = y     # *.gis.dbf
MostLikelyClusterCaseInfoEachCentroidDBase = y     # *.sci.dbf
```

The *satscanMapper* package uses the *SeerMapper* package to generate all of the maps. Other R packages used by *satscanMapper* are: base, utils, graphics, grDevices, stats, RColorBrewer, stringr, sp, and foreign.

Value

None

Author(s)

Jim Pearson and Linda Pickle of StatNet Consulting, LLC, Gaithersburg, MD

Examples

```
######
#
#  These examples focus on mapping the SaTScan (TM) data after the analysis
#  has been completed.
#  See the section on the CreateGeo4SS function for an example of how
#  to build a coordinates file (.geo) that matches the satscanMapper
#  boundary data.
#
######
```

```
######
#
#  Example # 1 - Mapping existing SaTScan (TM) results - "USStateLung" result files
#     Make sure SaTScan (TM) results files are still located in the directories
#     documented in the session saved (.prm) file.
#
#     This example maps data for the US States containing circular clusters.
#
#   Get location of .prm file and location to write mapping output files
#
#\dontrun{
#tempdirOut <- tempdir()
#cat("tempDirOut:",tempdirOut,"\n")

SSMInstDir <- system.file(package="satscanMapper", "extdata")
cat("SSMInstDir:",SSMInstDir,"\n")

TT =c("Contiguous US States All Lung Cancer Mortality, 2004")

satscanMapper(resDir  = SSMInstDir,              # path to .prm file location for output files.
              prmFile = "stateLung.prm",
              outDir  = "",
              title   = TT
             )
#}
#
##  end of example # 1  Removed due to time constraints.
##
######

######
#
#  Example # 2 - Mapping existing SaTScan (TM) results - "USCountyLung" result files
#     named "inc_noadj_cir" files
#     Make sure SaTScan (TM) results files are still located in the directories
#     documented in the session saved (.prm) file.
#
## Not run:
#tempdirOut <- tempdir()
#cat("tempDirOut:",tempdirOut,"\n")

SSMInstDir <- system.file(package="satscanMapper", "extdata")

TT = "Contiguous US County Female Breast Cancer Incidence, 2009-2013"

satscanMapper(resDir=SSMInstDir,            # path to .prm file location for output files.
              prmFile    = "inc_noadj_ellip_hilo_10_nosp.prm",
              outDir     = "",
              categ      = 7,
              title      = TT,
              censusYear = "2010"
             )
```

```
## End(Not run)
#
#  end of example # 2
#
#####

print("end of examples")
```

---

satscanMapper.Version | *Function returns the release date and time of the satscanMapper package.*

Description

The satscanMapper.Version function is a support function that returns a string set at the time the *satscanMapper* package was built and tested for release to CRAN. The string contains the package name, version number and the date and time the package was built.

"SaTScanMapper V1.0.1 - built 180630 at 6:08pm"

Usage

```
satscanMapper.Version()
```

Value

A character vector contain the package name, version, and build data and time

Author(s)

Jim Pearson and Linda Pickle of StatNet Consulting, LLC, Gaithersburg, MD

---

st99_data | *Data and index table for the state boundary information.*

Description

Contains the required state information for satscanMapper to be able to validate state location IDs in the SatScan (TM) data and produce the text reports.

Usage

```
data(st99_data)
```

Format

st99.ind00 structure:

row.names - character - 2 digit - State FIPS code.

abbr - character - 2 character - State US PS Abbreviation.

stName - character - State Name.

rgID - integer - region ID number.

rgName - character - region name. (NorthEast, South, MidWest, West, Offshore)

dvID - integer - division ID number.

dvName - character - division name

loc - character - postfix file string identifying the directory/package containing the census tract boundary datasets for this state. (Regs, East, West)

county_00 - integer - number of counties in state during census year 2000.

county_10 - integer - number of counties in state during census year 2010.

tracts_00 - integer - number of census tracts in state during census year 2000.

tracts_10 - integer - number of census tracts in state during census year 2010.

change10 - logical - indicates if the state had different counties and/or boundaries in census year 2010. If FALSE, use the 2000 boundary data. If TRUE, use the supplemental 2010 boundary data. Only 3 states required a 2010 set of county boundary datasets (Alaska, Colorado, and Virginia)

c_X and c_Y - integers - the X and Y cartesian coordinates of the center of the state.

The state data table contains the information for all of the U. S. states, DC and Puerto Rico. This dataset is load for state, state/county, and state/county/census tract level data.

Source

CENSUS.GOV web site (boundary data), HUD.gov Census Tract to ZipCode crossover table, USPS Zip code database (1999), and R MAPS package (state and county names) for the U. S. 2000 Census.

---

st99_M_data	*U. S. state centroid data for the states of Alaska, Hawaii, and Puerto Rico.*

Description

Contains the U. S. state information for Alaska, Hawaii, and Puerto Rico needed by the satscan-Mapper package to create the coordinates file to be used by SaTScan (TM). The main coordinates and boundary files were modified to shift Alaska, Hawaii and Puerto Rico to better positions on the US map. To be able to create a coordinates file based on the location IDs in the *.pop and *.cas input files to SaTScan (TM), the coordinates must be based on the original Lat and Long and not the moved locations. This dataset contains the unmoved projected coordinates of the of the centroids for the three moved states. Once the analysis is completed, only the mapping coordinates are needed.

Usage

```
data(st99_M_data)
```

Format

st99.ind00 structure:

row.names - character - 2 digit - State FIPS code.

abbr - character - 2 character - State US PS Abbreviation.

stName - character - State Name.

rgID - integer - region ID number.

rgName - character - region name. (NorthEast, South, MidWest, West, Offshore)

dvID - integer - division ID number.

dvName - character - division name

loc - character - postfix file string identifying the directory/package containing the census tract boundary datasets for this state. (Regs, East, West)

county_00 - integer - number of counties in state during census year 2000.

county_10 - integer - number of counties in state during census year 2010.

tracts_00 - integer - number of census tracts in state during census year 2000.

tracts_10 - integer - number of census tracts in state during census year 2010.

change10 - logical - indicates if the state had different counties and/or boundaries in census year 2010. If FALSE, use the 2000 boundary data. If TRUE, use the supplemental 2010 boundary data. Only 3 states required a 2010 set of county boundary datasets (Alaska, Colorado, and Virginia)

c_X and c_Y - integers - the X and Y cartesian coordinates of the center of the state.

This data.frame contains the centroid projected coordinates for Alaska, Hawaii, and Puerto Rico prior to their relocation on the map for better presentation. It has a duplicate structure and contains to the *st99_data* dataset to allow the state data to be easily overlaid on the normal mapping data when the coordinates file is being generated.

Source

CENSUS.GOV web site (boundary data), HUD.gov Census Tract to ZipCode crossover table, USPS Zip code database (1999), and R MAPS package (state and county names) for the U. S. 2000 Census.

---

tr99_data	*Census Tract data table is used by the satscanMapper package to process census tract level location IDs for the census years 2000 and 2010.*

Description

Contains the census tract information for census years 2000 and 2010 to assist in the validation of the census tract level location IDs in user data and the generation of the text reports by the package on the SatScan (TM) cluster results.

Usage

```
data(tr99_data)
```

Format

The information contained in the tr99_data data.frame is:

row.names - character - 11 digit State/County/Census Tract FIPS code.

stcoID - character - 5 digit state/county FIPS code containing the census tract.

plKey - character - a placename key associated with the census tract. The key contains the state/county 5 digit fips code and the placename.

c_X_00 and c_Y_00 - integers - the X/Y cartesian coordinates of the center of the census tract based on an equal area projection in meters for the census year of 2000.

c_X_10 and c_Y_10 - integers - the X/Y cartesian coordinates of the center of the census tract based on an equal area projection in meters for the census year of 2010.

y - integer - a bit string identifying the census years the census tract existed. If the value is 1 or 3, the tract existed in the 2000 census. If the value is 2 or 3, the tract existed in the 2010 census.

/br The tr99 data contain the information for all census tracts. /br The place names in the tr99_data structure are derived from the HUD.gov Census Tract to ZipCode crossover tables and the USPS.gov zip code database for the 2000 and 2010 census. Additional place name information was provided by NCI as a supplement.

Source

CENSUS.GOV web site (boundary data), HUD.gov Census Tract to ZipCode crossover table, USPS Zip code database (1999), and R MAPS package (state and county names) for the U. S. 2000 Census.

---

tr99_M_data	*U.S. census tract data table for Alaska, Hawaii, and Puerto Rico census tracts containing the projected coordinates of the state's census tract centroids.*

Description

Contains the U. S. census tract information for Alaska, Hawaii, and Puerto Rico needed by the satscanMapper package to create the coordinates file to be used by SaTScan (TM). The main coordinates and boundary files were modified to shift Alaska, Hawaii and Puerto Rico to better positions on the US map. To be able to create a coordinates file based on the location IDs in the *.pop and *.cas input files to SaTScan (TM), the coordinates must be based on the original Lat and Long and not the moved locations. This dataset contains the unmoved projected coordinates of the of the centroids for all of the census tracts in the three affected states. Once the analysis is completed, only the mapping coordinates are needed.

Usage

```
data(tr99_M_data)
```

Format

The information contained in the tr99_data data.frame is:

row.names - character - 11 digit State/County/Census Tract FIPS code.

stID - character - 2 digit state FIPS code containing the census tract.

plKey - character - a placename key associated with the census tract. The key contains the state/county 5 digit fips code and the placename.

c_X_00 and c_Y_00 - integers - the X/Y cartesian coordinates of the center of the census tract based on an equal area projection in meters for the census year of 2000.

c_X_10 and c_Y_10 - integers - the X/Y cartesian coordinates of the center of the census tract based on an equal area projection in meters for the census year of 2010.

y - integer - a bit string identifying the census years the census tract existed. If the value is 1 or 3, the tract existed in the 2000 census. If the value is 2 or 3, the tract existed in the 2010 census.

/br The tr99 M data contain the coordinates for all of the Alaska, Hawaii and Puerto Rico census tracts. /br The place names in the tr99_M_data structure are derived from the HUD.gov Census Tract to ZipCode crossover tables and the USPS.gov zip code database for the 2000 and 2010 census. Additional place name information was provided by NCI as a supplement.

Source

CENSUS.GOV web site (boundary data), HUD.gov Census Tract to ZipCode crossover table, USPS Zip code database (1999), and R MAPS package (state and county names) for the U. S. 2000 Census.

Index