Package ‘Guerry’

February 19, 2015

Type Package

Title Maps, data and methods related to Guerry (1833) "Moral
Statistics of France"

Version 1.6-1

Date 2014-09-26

Author Michael Friendly [aut, cre], Stephane Dray [aut]

Maintainer Michael Friendly <friendly@yorku.ca>

Encoding latin1

Depends R (>= 2.10)

Suggests sp, shapefiles, spdep, ade4, maptools

Description This package comprises maps of France in 1830, multivariate data from A.-
M. Guerry and others, and statistical and
graphic methods related to Guerry’s "Moral Statistics of France". The goal is to facilitate the ex-
ploration and
development of statistical and graphic methods for multivariate data in a geo-
spatial context of historical interest.

License GPL

URL http://guerry.r-forge.r-project.org/

LazyLoad yes

LazyData yes

Repository CRAN

Repository/R-Forge/Project guerry

Repository/R-Forge/Revision 25

Repository/R-Forge/DateTimeStamp 2014-09-26 12:11:25

Date/Publication 2014-09-26 19:03:40

NeedsCompilation no
**Description**

Andre-Michel Guerry (1833) was the first to systematically collect and analyze social data on such things as crime, literacy and suicide with the view to determining social laws and the relations among these variables. He provided the first essentially multivariate and georeferenced spatial data on socially important questions, e.g., Is the rate of crime related to education or literacy? How does this vary over the departments of France? Are the rates of crime or suicide within departments stable over time?

In an age well before the idea of correlation had been invented, Guerry used graphics and statistical maps to try to shed light on such questions. In a later work (Guerry, 1864), he explicitly tried to entertain larger questions, but with still-limited statistical tools: Can rates of various crimes be related to multiple causes or predictors? Are the rates and ascribable causes in France similar or different to those found in England?

The Guerry package comprises maps of France in 1830, multivariate data from A.-M. Guerry and others, and statistical and graphic methods related to Guerry's *Moral Statistics of France*. The goal of providing these as an R package is to facilitate the exploration and development of statistical and graphic methods for multivariate data in a geo-spatial context.

**Details**

- **Package:** Guerry
- **Type:** Package
- **Version:** 1.6-0
- **Date:** 2014-09-23
- **License:** GPL
- **LazyLoad:** yes

Data from Guerry and others is contained in the data frame Guerry. Because Corsica is often considered an outlier both spatially and statistically, the map of France circa 1830, together with the Guerry data is provided as SpatialPolygonsDataFrames in two forms: gFrance for all 86 departments, and and gFrance85, for the 85 departments excluding Corsica.
Author(s)

Michael Friendly and Stephane Dray
Maintainer: Michael Friendly <friendly AT yorku.ca>

References


Angeville

Data from d’Angeville (1836) on the population of France

Description

Adolph d’Angeville (1836) presented a comprehensive statistical summary of nearly every known measurable characteristic of the French population (by department) in his Éssai sur la Statistique de la Population francaise. Using the graphic method of shaded (choropleth) maps invented by Baron Charles Dupin and applied to significant social questions by Guerry, Angeville’s Essai became the first broad and general application of principles of graphic representation to national industrial and population data.

The collection of variables in the data frame Angeville is a small subset of over 120 columns presented in 8 tables and many graphic maps.

Usage

data(Angeville)
Format

A data frame with 86 observations on the following 16 variables.

department  a numeric vector
Department  Department name: a factor with levels Ain Aisne ... Vosges Yonne
Mortality  Mortality: Number of births to give 100 people at age 21 (T1:13)
Marriages  Number of marriages per 1000 men aged 21 (T1:15)
Legit_births  Annual no. of legitimate births (T2:17)
Illeg_births  Annual no. of illegitimate births (T2:18)
Recruits  Number of people registered for military recruitment from 1825-1833 (T3:32)
Conscripts  Number of inhabitants per military conscript (T3:33)
Exemptions  Number of military exemptions per 1000 all of physical causes (T3:47)
Farmers  Number of farmers during the census in 1831 (T4:65)
Recruits_ignorant  Average number of ignorant recruits per 1000 (T5:69)
Schoolchildren  Number of schoolchildren per 1000 inhabitants (T5:71)
Windows_doors  Number of windows & doors in houses per 100 inhabitants (T5:72). This is sometimes taken as an indicator of household wealth.
Primary_schools  "Number of primary schools (T5:74)
Life_exp  Life expectancy in years (T1:9a,9b)
Pop1831  Population in 1831

Details

ID codes for department were modified from those in Angeville's tables to match those used in Guerry. Angeville's variables are recorded in a variety of different ways and some of these were calculated from other columns in his tables not included here. As well, the variable names and labels used here were often shortened from the more complete descriptions given by d'Angeville. The notation "(Tn:k)" indicates that the variable used here came from Table n, Column k.

Source

The data was digitally scanned from Angeville's tables using OCR software, then extensively edited to correct obvious errors and finally subjected to some consistency checks using the column totals and ranked values he provided.

References


Examples

data(Angeville)
## maybe str(Angeville) ; plot(Angeville) ...
Description

gfrance is a SpatialPolygonsDataFrame object created with the sp package, containing the polygon boundaries of the map of France as it was in 1830, together with the Guerry data frame.

Usage

data(gfrance)

Format

The format is: Formal class 'SpatialPolygonsDataFrame' [package "sp"] with 5 slots: gfrance@data, gfrance@polygons, gfrance@plotOrder, gfrance@bbox, gfrance@proj4string. See: SpatialPolygonsDataFrame for descriptions of some components.

The analysis variables are described in Guerry.

Details

In the present version, the PROJ4 projection is not specified.

Source


References


See Also

Guerry for description of the analysis variables Angeville for other analysis variables

Examples

data(gfrance)
names(gfrance)  ## list @data variables
plot(gfrance)  ## just show the map outline

# Show basic choropleth plots of some of the variables
spplot(gfrance, "Crime_pers")
spplot(gfrance, "Crime_prop")

# Note that spplot assumes all variables are on the same scale for comparative plots
# transform variables to ranks (as Guerry did)

```r
# Not run:
local({
gfrance$Crime_pers <- rank(gfrance$Crime_pers)
gfrance$Crime_prop <- rank(gfrance$Crime_prop)
gfrance$Literacy <- rank(gfrance$Literacy)
gfrance$Donations <- rank(gfrance$Donations)
gfrance$Infants <- rank(gfrance$Infants)
gfrance$Suicides <- rank(gfrance$Suicides)

spplot(gfrance, c("Crime_pers", "Crime_prop", "Literacy", "Donations", "Infants", "Suicides"),
      layout=c(3,2), as.table=TRUE, main="Guerry’s main moral variables")
})
# End(Not run)
```

---

**gfrance85**

---

### Map of France in 1830 with the Guerry data, excluding Corsica

**Description**

`gfrance85` is a `SpatialPolygonsDataFrame` object created with the sp package, containing the polygon boundaries of the map of France as it was in 1830, together with the Guerry data frame. This version excludes Corsica, which is an outlier both in the map and in many analyses.

### Usage

```r
data(gfrance85)
```

### Format

The format is: Formal class `SpatialPolygonsDataFrame` [package "sp"] with 5 slots: `gfrance85$data, gfrance85$polys, gfrance85$plotOrder, gfrance85$bbox, gfrance85$proj4string`. See: `SpatialPolygonsDataFrame` for descriptions of some components.

The analysis variables are described in Guerry.

### Details

In the present version, the PROJ4 projection is not specified.

### Source

References


Examples

data(gfrance85)
require(sp)
plot(gfrance85)  # plot the empty outline map

# extract some useful components
df <- data.frame(gfrance85)[7:12]  # main moral variables
xy <- coordinates(gfrance85)        # department centroids
dep.names <- data.frame(gfrance85)[6]
region.names <- data.frame(gfrance85)[5]
col.region <- colors()[c(149,254,468,552,26)]

if (require(spdep)) {
  lw <- nb2listw(poly2nb(gfrance85))  # neighbors list
}

# plot the map showing regions by color with department labels
op <- par(mar=rep(0.1,4))
plot(gfrance85,col=col.region[region.names])
text(xy, labels=dep.names, cex=0.4)
par(op)

Guerry

Data from A.-M. Guerry, "Essay on the Moral Statistics of France"

Description

Andre-Michel Guerry (1833) was the first to systematically collect and analyze social data on such things as crime, literacy and suicide with the view to determining social laws and the relations among these variables.

The Guerry data frame comprises a collection of 'moral variables' on the 86 departments of France around 1830. A few additional variables have been added from other sources.

Usage

data(Guerry)
A data frame with 86 observations (the departments of France) on the following 23 variables.

- **dept**  Department ID: Standard numbers for the departments, except for Corsica (200)
- **Region** Region of France (‘N’='North', 'S'='South', 'E'='East', 'W'='West', 'C'='Central'). Corsica is coded as NA
- **Department** Department name: Departments are named according to usage in 1830, but without accents. A factor with levels Ain Aisne Allier ... Vosges Yonne
- **Crime_pers** Population per Crime against persons. Source: A2 (Compte général, 1825-1830)
- **Crime_prop** Population per Crime against property. Source: A2 (Compte général, 1825-1830)
- **Literacy** Percent Read & Write: Percent of military conscripts who can read and write. Source: A2
- **Donations** Donations to the poor. Source: A2 (Bulletin des lois)
- **Infants** Population per illegitimate birth. Source: A2 (Bureauau des Longitudes, 1817-1821)
- **Suicides** Population per suicide. Source: A2 (Compte général, 1827-1830)
- **MainCity** Size of principal city (‘1:Sm’, ‘2:Med’, ‘3:Lg’), used as a surrogate for population density. Large refers to the top 10, small to the bottom 10; all the rest are classed Medium. Source: A1. An ordered factor with levels 1:Sm < 2:Med < 3:Lg
- **Wealth** Per capita tax on personal property. A ranked index based on taxes on personal and movable property per inhabitant. Source: A1
- **Commerce** Commerce and Industry, measured by the rank of the number of patents / population. Source: A1
- **Clergy** Distribution of clergy, measured by the rank of the number of Catholic priests in active service / population. Source: A1 (Almanach officiel du clergy, 1829)
- **Crime_parents** Crimes against parents, measured by the rank of the ratio of crimes against parents to all crimes— Average for the years 1825-1830. Source: A1 (Compte général)
- **Infanticide** Infanticides per capita. A ranked ratio of number of infanticides to population— Average for the years 1825-1830. Source: A1 (Compte général)
- **Donation_clergy** Donations to the clergy. A ranked ratio of the number of bequests and donations inter vivios to population— Average for the years 1815-1824. Source: A1 (Bull. des lois, ordunn. d’autorisation)
- **Lottery** Per capita wager on Royal Lottery. Ranked ratio of the proceeds bet on the royal lottery to population— Average for the years 1822-1826. Source: A1 (Compte rendus par le ministre des finances)
- **Desertion** Military disertion, ratio of the number of young soldiers accused of desertion to the force of the military contingent, minus the deficit produced by the insufficiency of available billets— Average of the years 1825-1827. Source: A1 (Compte du ministere du guerre, 1829 etat V)
- **Instruction** Instruction. Ranks recorded from Guerry’s map of Instruction. Note: this is inversely related to Literacy (as defined here)
- **Prostitutes** Prostitutes in Paris. Number of prostitutes registered in Paris from 1816 to 1834, classified by the department of their birth Source: Parent-Duchatelet (1836), *De la prostitution en Paris*
Distance  Distance to Paris (km). Distance of each department centroid to the centroid of the Seine (Paris) Source: calculated from department centroids

Area  Area (1000 km^2). Source: Angeville (1836)

Pop1831  1831 population. Population in 1831, taken from Angeville (1836), *Essai sur la Statistique de la Population française*, in 1000s

Details

Note that most of the variables (e.g., Crime_pers) are scaled so that 'more is better' morally. Values for the quantitative variables displayed on Guerry’s maps were taken from Table A2 in the English translation of Guerry (1833) by Whitt and Reinking. Values for the ranked variables were taken from Table A1, with some corrections applied. The maximum is indicated by rank 1, and the minimum by rank 86.

Source


References


See Also

Angeville for other analysis variables

Examples

data(Guerry)

## maybe str(Guerry); plot(Guerry) ...
Index

*Topic datasets
  Angeville, 3
  gfrance, 5
  gfrance85, 6
  Guerry, 7
*Topic package
  Guerry-package, 2
*Topic spatial
  gfrance, 5
  gfrance85, 6
  Guerry-package, 2
*Topic
  gfrance, 5
  gfrance85, 6
  Guerry-package, 2

Angeville, 3, 5, 9

gfrance, 2, 5
gfrance85, 2, 6
Guerry, 2, 4–6, 7
Guerry-package, 2

SpatialPolygonsDataFrame, 5, 6