Package: historydata (via r-universe)

October 15, 2024

Type Package

Title Datasets for Historians

Version 0.3.0

Description These sample data sets are intended for historians learning R. They include population, institutional, religious, military, and prosopographical data suitable for mapping, quantitative analysis, and network analysis.

Depends R (>= 3.0.2)

Suggests ggplot2 (>= 3.0.0)

License MIT + file LICENSE

LazyData true

URL https://docs.ropensci.org/historydata/,
https://github.com/ropensci/historydata

BugReports https://github.com/ropensci/historydata/issues

RoxygenNote 7.3.2

Encoding UTF-8

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

RemoteUrl https://github.com/ropensci/historydata

RemoteRef master

RemoteSha e24e0cccc447dceed3b259efc40587beb7903496

Contents

catholic_dioceses	*Roman Catholic dioceses in the United States, Canada, and Mexico*

Description

Dates when Roman Catholic dioceses and archdioceses in the United States were founded or made metropolitan sees, with geocoded locations. The sources cited indicate that none of these sees have been discontinued.

Usage

```
catholic_dioceses
```

Format

A data frame with 425 observations of 6 variables.

Variables

- diocese: name, and thus location, of the diocese.
- event, date: the date when the diocese was erected, that is first founded, or made a metropolitan see. Encoded as a date object.
- rite: the rite overseen by the diocese. Regions with ordinary jurisdiction but not episcopal character are not included.
- lat, long: latitude and longitude coordinates for the headquarters city of the diocese.

References

This data is compiled from several sources:

Joseph Bernard Code, *Dictionary of the American Hierarchy (1789-1964)* (New York: Joseph F. Wagner, 1964), 425-26.

For the United States since 1963, Canada, and Mexico: https://www.catholic-hierarchy.org/

Examples

```
head(catholic_dioceses)
```

---

dijon_prices   *Wholesale Market Prices in Dijon, France 1568–1630*

Description

These are wholesale market prices in the city of Dijon in Burgundy in central France from 1568 to 1630. They include the major cereal grains, different qualities of wine, dried legumes, oils used for cooking, seeds, and candle wax made from beef tallow. All prices were generally recorded by the city council at the same time of year on the first market day after the feast of St. Martin (November 11). All prices are in *sous tournois* / 20 *sous* = 1 *livre tournois*.

Usage

```
dijon_prices

dijon_prices_wide
```

Format

dijon_prices is a data frame with 1,110 observations of 6 variables. dijon_prices_wide is a data frame with 19 observations of 65 variables. dijon_prices_wide contains the data as it was transcribed; that data has been converted to a tidy format in dijon_prices.

An object of class tbl_df (inherits from tbl, data.frame) with 19 rows and 65 columns.

Variables

- commodity: the commodity for sale
- measure: the amount of the commodity for that price
- price: the price in *sous tournais*.
- citation, citation_date: citations and dates for documents in the Archives municipales de Dijon.

Author(s)

Mack Holt, George Mason University

References

All citations are to the Archives municipales de Dijon. See the columns citation and citation_date in dijon_prices for the documents from which each price was gathered.

---

early_colleges	*Early colleges in the United States*

Description

This dataset contains information about the founding of colleges established before 1848 in the United States of America.

Usage

```
early_colleges
```

Format

A data frame with 65 observations of 6 variables.

Variables

- college: The name of the college or university.
- original_name: The name under which the institution was founded, if different.
- city, state: The location of the institution.
- established: The year that the institution was founded.
- sponsorship: The sponsoring religious denomination, or secular if not founded by a denomination.

Author(s)

George Oberle, George Mason University

Daniel Walker Howe

References

This data was transcribed by George Oberle from the table "Some American Institutions of Higher Education Founded Before 1848, in Daniel Walker Howe, *What Hath God Wrought: The Transformation of America, 1815-1848* (New York: Oxford University Press, 2007), 460-461.

Examples

```
head(early_colleges)
if(require(ggplot2)) {
  ggplot(early_colleges, aes(x = established)) + geom_bar(binwidth = 5) +
  ggtitle("Founding Dates of Early American Colleges")
}
```

---

historydata	*Datasets for historians*

Description

This package provides sample datasets of interest to historians, analogous to the **datasets** package for R generally and the **histdata** package for datasets from the history of statistics.

References

Citations to the sources for datasets are provided in the documentation for each dataset.

---

judges	*Federal judges in the United States of America*

Description

This dataset contains information about the appointments and careers of all federal judges in United States history since 1789. It includes judges who "judges presidentially appointed during good behavior who have served since 1789 on the U.S. District Courts, the U.S. Courts of Appeals, the Supreme Court of the United States, the former U.S. Circuit Courts, and the federal judiciary's courts of special jurisdiction." Some of the unnecessary information from the source has been excluded.

Format

Two data frames, judges_people and judges_appointments.

Data frames

The data frame judges_people contains information about the judges, such as names and vital information. The data frame judges_appointments contains information about their appointments, such as the name of the court, nominating president, and the dates of service.

References

This data is taken from the Biographical Directory of Federal Judges, 1789-present.

Examples

```
data(judges_people)
data(judges_appointments)
```

| methodists | *Methodist Minutes of the Annual Conferences, 1786–1834* |

Description

This dataset contains transcriptions of the membership figures from the annual meeting minutes of the Methodist Episcopal Church in the United States of America.

Usage

```
methodists
```

Format

A data frame with 20241 observations of 11 variables.

Author(s)

Lincoln Mullen, George Mason University

References

See the url field in each row for a link to the page in the Hathi Trust from which that row was transcribed.

| minneapolisfed_cpi | *Consumer Price Index, 1800-2024* |

Description

This dataset contains historic consumer price index (CPI) data including estimates before the modern U.S. CPI, retrieved from the Federal Reserve Bank of Minneapolis.

Usage

```
minneapolisfed_cpi
```

Format

A data frame with 225 observations of 3 variables.

Details

These details are taken from https://www.minneapolisfed.org/about-us/monetary-policy/inflation-calculator/consumer-price-index-1800- and edited.

Official U.S. data go back to 1913 for a consumer price index (CPI) comparable to what the U.S. Bureau of Labor Statistics (BLS) still calculates today. The table below reflects the following historical series as initially compiled by the BLS for the Handbook of Labor Statistics, with modern CPI data from 1913 to the present day:

- 1800 to 1851 - Index of Prices Paid by Vermont Farmers for Family Living
- 1851 to 1890 - Consumer Price Index by Ethel D. Hoover
- 1890 to 1912 - Cost of Living Index by Albert Rees
- 1913 to 1977 - Consumer Price Index (CPI)
- 1978 to present - Consumer Price Index for all Urban Consumers (CPI-U)

The dataset uses 1967 as the index (1967=100). With the caveat that data before 1913 should be considered estimates To find out how much a price in Year 1 would be in Year 2 dollars:

Year 2 Price = Year 1 Price x (Year 2 CPI/Year 1 CPI)

Variables

- `year`: date of CPI, or estimate of CPI.
- `annual_average_index`: average annual CPI, or estimate of average annual CPI.
- `annual_percentage_change`: annual percentage change of the CPI.

References

This data is compiled from the Federal Reserve Bank of Minneapolis:

https://www.minneapolisfed.org/about-us/monetary-policy/inflation-calculator/consumer-price-index-1

Examples

```
head(minneapolisfed_cpi)
```

---

naval_promotions	*Promotions of U.S. Navy officers, 1798-1849*

Description

This dataset contains records of the careers of officers of the line in the United States Navy who became midshipmen between 1798 and 1849. It lists the dates at which officers were promoted to the ranks of midshipment, lieutenant, master commandament, and captain.

Usage

```
naval_promotions
```

Format

A data frame with 5,705 observations of 4 variables.

Variables

- `id`: A unique numeric identifier for each officer.

- `name`: Name of the officer.

- `generation`: A numeric value to represent a judgment about the generations within the officer corp of the U.S. Navy. First generation officers entered the service before 1801, as a rank higher than midshipman. Second generation officers entered the service before the Peace Establishment Act (or by the end of 1801), but as midshipman. Thus, they essentially became adults in the service, and they learned their craft from the first generation. Third generation officers entered the service as midshipmen after the Peace Establishment Act but before the end of the War of 1812. Those officers in this generation who became captain rose to that rank in the 1830s and '40s. Fourth generation officers entered the service after the War of 1812 had ended. These officers saw almost no wartime service, and many of the ones who achieved captain found themselves having to decide whether to serve in the Union or the Confederacy during the Civil War.

- `rank`: The ranks which officers obtained, including `midshipman`, `lieutenant`, `master_commandant`, and `captain`. Also included are the dates some officers `left_service`.

- `date`: The date as a character string on which an officer attained the associated rank.

Author(s)

Abby Mullen, Northeastern University

Edward W. Callahan

References

This data was compiled by Abby Mullen from Edward W. Callahan, *Officers of the Continental and U.S. Navy and Marine Corps, 1775-1900* (New York: L. R. Hamersly, 1901), made available via the Naval Historical Center, `https://www.ibiblio.org/hyperwar/NHC/Callahan/index.htm`. See Mullen's blog post about compiling the data, "Database of Officers of the Line," abbymullen.org, January 5, 2013.

Examples

```
head(naval_promotions)
```

| paulist_missions | *Records of missions held by the Paulist Fathers, 1851–1893* |

Description

This dataset contains transcriptions of records of missions held in by the Paulist Fathers, a Roman Catholic missionary order, in the nineteenth-century United States. This dataset contains only the most interesting data recorded in the full Paulist mission chronicles. The founding members of the Paulist Fathers began as members of the Redemptorist order. Their chronicles include both the missions held as Redemptorists and as Paulists. This transcription only includes missions up to the year 1893; the Paulist Chronicles contain more records than are transcribed here, and the Paulists continued holding missions for many more years than are recorded in the Chronicles.

Usage

```
paulist_missions
```

Format

A data frame with 841 observations of 17 variables.

Variables

- mission_number: The number assigned to the mission in the Paulist Mission Chronicles.
- church: The name of the church or cathedral at which the mission was held.
- city, state: The location of the mission.
- start_date, end_date: The start and end dates of the mission as date objects.
- year: The year of the mission.
- decade: The decade of the mission (useful for faceting).
- duration_days, duration_years: The duration of the mission in days (as an integer) and in weeks (as an ordered factor).
- confessions: The number of confessions heard by the Paulists at the mission, which is a rough proxy for the number of people who attended the mission.
- converts: The total number of converts made during the mission and people left under instruction for conversion after the mission was over.
- order: Whether the mission was held under the Redemptorist or Paulist order.
- lat, long: The latitude and longitude of the city where the mission was held.
- volume, page: The location of the mission record in the Paulist mission chronicles.

Author(s)

Lincoln Mullen, George Mason University

References

The Paulist missions are recorded in *Chronicle of the Missions Given by the Congregation of Missionary Priests of St. Paul the Apostle*, six manuscript volumes, Office of Paulist History and Archives, North American Paulist Center, Washington, DC. Data transcribed by Lincoln Mullen.

Examples

```
head(paulist_missions)
```

---

presbyterians	*Presbyterians Statistics Through One Hundred Years*

Description

This dataset contains transcriptions of annual denominational records for Presbyterians in the United States. This data was compiled Herman Carl Weber for the Presbyterian Church U.S.A. in 1927. For an explanation of the variables, see the book from which the records were transcribed.

Usage

```
presbyterians
```

Format

A data frame with 133 observations of 37 variables.

Details

Weber includes two sections, one of raw data compiled from the Minutes of the General Assembly, which apparently includes the foreign membership of the Presbyterian churches. The second section contains the data that Weber used for his visualizations, which excludes the foreign membership and makes various calculations on the data.

Not all of the data in Weber is reproduced here. The data comes from part 1, with the tables on financial information excluded.

One field that varies in title from members received by exam to members received by confession varies in title, but has been presumed to refer to the same process.

Author(s)

Lincoln Mullen, George Mason University

References

Weber, Herman C. *Presbyterian Statistics Through One Hundred Year, 1826–1926: Tabulated, Visualized, and Interpreted*. Philadelphia: The General Council, Presbyterian Church in the U.S.A., 1927. https://catalog.hathitrust.org/Record/007109885.

---

quasi_war	*Naval encounters during the Quasi War between France and the United States of America*

Description

Naval encounters during the Quasi War between France and the United States of America

Usage

```
quasi_war
```

Format

A data frame with 198 observations of 16 variables.

References

This data was gathered by Abby Mullen.

Examples

```
head(quasi_war)
```

---

sarna	*Population estimates of American Jews*

Description

This dataset contains estimates of the population of American Jews.

Usage

```
sarna
```

Format

A data frame with 92 observations of 3 variables.

Variables

- year: date of estimate.
- estimate: the type of estimate. population_low and population_high are the lower and upper bounds on the population of American Jews; percentage_high and percentage_low are the lower and upper bounds on the percentage of Jews among the United States population.
- value: the value of the estimate.

References

This data is taken from the appendix in Jonathan D. Sarna, *American Judaism: A History* (New Haven: Yale University Press, 2004), 375-376.

Examples

```
head(sarna)
```

---

tudors	*Tudor dynasty*

Description

Spousal and parent/child relationships among selected members of the Tudor dynasty, suitable for network analysis.

Usage

```
tudors
```

Format

A data frame with 35 observations of 3 variables.

Variables

- person_1, person_2: The two people in the relationship.
- relationship: The type of relationship.

Examples

```
data(tudors)
```

---

us_cities_pop	*United States Historical City Populations, 1790-2010*

Description

A dataset of cities mentioned in the U.S. Census and other data sources. In general, cities are included once they reach a population of 2,500. Data is recorded every ten years. See the dataset's GitHub page for details on how the data was gathered.

Usage

```
us_cities_pop
```

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 62224 rows and 17 columns.

References

Erik Steiner and Jason Heppler, "United States Historical City Populations, 1790-2010," Spatial History Project, Center for Spatial and Textual Analysis, Stanford University, `https://github.com/cestastanford/historical-us-city-populations/`

`us_military_strengths`   *US Military Strengths*

Description

Active Duty US Military Personnel by Regional Area and by Country: Worldwide Manpower Distribution by Geographical Area

Usage

`us_military_strengths`

Format

A data frame with 50 observations of 27 variables.

Details

This data subset was extracted and transcribed in February 2018 from 48 individual DoD military troop strength yearly reports which varied in composition over time according to requirements mandated by Congress. The compiled spreadsheet covers the period 1950-1999, less the years 1951-1952 which are not available. It includes total military personnel, the numbers in the U.S. and overseas, the totals by military Service, and the figures for some twenty countries in which the United States had a significant military presence for much or all of the period. The original DoD reports are a mix of .pdf and .xls formats downloaded in one.zip file from the section "Historical Publications" at the bottom of the DMDC reports page. The country column headings in this spreadsheet use the ISO 3166 Alpha-2 country codes from the ISO Online Browsing Platform (OBP) at `https://www.iso.org/obp/ui/#search`.

Clarke Bursley put this data together.

References

Defense Manpower Data Center, "Active Duty Military Personnel by Regional Area and by Country: Worldwide Manpower Distribution by Geographical Area (M05), Historical Reports - Military Only - 1950, 1953 - 1999," DoD Personnel, Workforce Reports & Publications, `www.dmdc.osd.mil/appj/dwp/dwp_reports.jsp`

Examples

```
head(us_military_strengths)
```

---

us_national_population
Population of the United States, 1790-2010

Description

Population figures for the entire United States of America from the decennial census.

Usage

```
us_national_population
```

Format

A data frame with 23 observations of 2 variables.

Variables

- year: date of the census.

- population: population of the state or territory.

References

This dataset has been gathered by the NHGIS. Minnesota Population Center, *National Historical Geographic Information System: Version 2.0* (Minneapolis: University of Minnesota, 2011).

Examples

```
head(us_national_population)
if(require(ggplot2)) {
  ggplot(us_national_population,
         aes(x = year, y = population)) +
  geom_line() +
  ggtitle("Population of the United States, 1790-2010")
}
```

---

us_state_populations   *Populations of US states and territories, 1790-2010*

Description

Population figures for US states and territories from the decennial census.

Usage

```
us_state_populations
```

Format

A data frame with 983 observations of 4 variables.

Variables

- year: date of the census.
- state: name of the state or territory.
- population: population of the state or territory.
- GISJOIN: a unique identifier for joining NHGIS data to spatial data.

References

This dataset has been gathered by the NHGIS. Minnesota Population Center, *National Historical Geographic Information System: Version 2.0* (Minneapolis: University of Minnesota, 2011).

Examples

```
head(us_state_populations)
```

Index