

Using the StateLevelForest package and dataset

Gates Dupont

November 21, 2023

Contents

| | |
|---|----------|
| Introduction to StateLevelForest Dataset | 1 |
| Data Source | 1 |
| Caveats and Considerations | 2 |
| Loading and Exploring the Data | 2 |
| Example Analyses | 2 |

`library(StateLevelForest)`

Introduction to StateLevelForest Dataset

The `StateLevelForest` dataset provides a unique view of total forest area across all states in the United States, spanning over a century. It includes state-level data from 1907 to 2017, sampled on 11 occasions. It also includes estimated forest cover in 1630, included for reference, “at the time of European Settlement.” This dataset could be vital for understanding changes in forest cover over time, how this impacts ecosystems, ecological communities, and species, and can be instrumental in research, policy-making, and education.

Data Source

This dataset is transcribed from the data provided in the following report:

Oswalt, S. N., Smith, W. B., Miles, P. D., & Pugh, S. A. (2019). Table 3. In *Forest Resources of the United States, 2017: A Technical Document Supporting the Forest Service 2020 RPA Assessment* (pp. 77-78). U.S. Department of Agriculture, Forest Service. (doi: 10.2737/WO-GTR-97)

Information on the report is available at this link: <https://www.fs.usda.gov/research/treesearch/57903#>
And a PDF of the report is available at this link: https://www.fs.usda.gov/research/publications/gtr/gtr_wo97.pdf

Caveats and Considerations

While the `StateLevelForest` dataset is a valuable resource, users should be aware of several caveats:

- **Adjustments and Definitions:** The data for years 2007 and later are adjusted for forest definition change to a minimum of 10% cover and the removal of chaparral as a forest type.
- **Historical Data Adjustments:** Data for 1909-1997 have been adjusted for the removal of the chaparral type and the addition of historically unproductive forest areas in west Texas and west Oklahoma. 1630 Estimates: The 1630 data are estimated based on the original forest area and historic land clearing information, primarily for reference purposes.
- **Rounding:** Some data may not add up exactly due to rounding in the original calculations. Recommendations for Users

Before using the `StateLevelForest` dataset for scientific research or dissemination, we strongly recommended that users familiarize themselves with the original report by Oswald et al. (2019). Understanding the context, methodology, and limitations of the original data collection is crucial for accurate interpretation and use of this dataset.

Loading and Exploring the Data

```
# Load the StateLevelForest dataset
data("StateLevelForest")

# View the first few rows of the dataset
head(StateLevelForest)
#>   year      state forest_thousands_of_acres
#> 1 2017 Connecticut           1808
#> 2 2017  Delaware             361
#> 3 2017   Maine           17579
#> 4 2017 Maryland            2463
#> 5 2017 Massachusetts         3025
#> 6 2017 New Hampshire         4758
```

Example Analyses

Visualizing Forest Cover Change in Massachusetts

This example focuses on exploring the change in forest cover in Massachusetts over time, relative to its estimated forest cover in the year 1630. By visualizing these changes, we can gain insights into the environmental shifts and land use changes in Massachusetts.

Data Preparation and Transformation

We begin by filtering the `StateLevelForest` dataset to include only the data for Massachusetts. The goal is to compare the forest cover in different years with the forest cover in the year 1630, which we use as a baseline. The following steps are taken in the data transformation process:

- **Filtering for Massachusetts:** We select only the rows where state equals “Massachusetts”.

- Identifying the 1630 Baseline: We create a new column X1630 that holds the value of forest cover in 1630 for all rows.
- Excluding the 1630 Data Point: We exclude the year 1630 from our plot as it serves as a baseline for comparison.

Creating the Plot

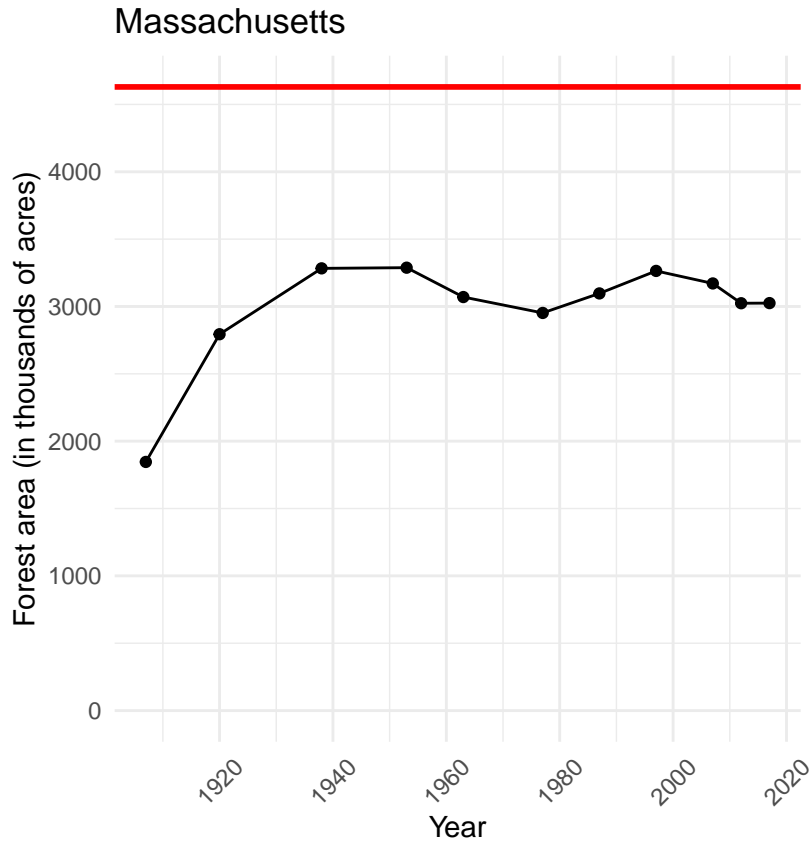
Using ggplot2, we create a line plot to visualize these changes:

- X-axis: Represents the years.
- Y-axis: Shows the forest area in thousands of acres.
- Red Horizontal Line: Indicates the forest cover in the year 1630, serving as a reference point.
- Points and Line: Display the change in forest cover across the years.

```
library(dplyr)
library(ggplot2)

MA_cover_relative_to_1630 <- StateLevelForest %>%
  filter(state == "Massachusetts") %>%
  mutate(X1630 = if_else(year == min(year), forest_thousands_of_acres, 0)) %>%
  mutate(X1630 = max(X1630)) %>%
  ungroup() %>%
  filter(year != min(year))

ggplot(MA_cover_relative_to_1630, aes(x = year, y = forest_thousands_of_acres)) +
  geom_hline(aes(yintercept = 0), color = NA, linewidth = 0) +
  geom_hline(aes(yintercept = X1630), color = "red", linewidth = 1) +
  geom_point() +
  geom_line() +
  scale_x_continuous(breaks = scales::pretty_breaks()) +
  scale_y_continuous(breaks = scales::pretty_breaks()) +
  labs(x = "Year", y = "Forest area (in thousands of acres)", title = "Massachusetts") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, vjust = 0.5),
        aspect.ratio = 1)
```



Interpretation of the Plot

The plot provides a clear visualization of how forest cover in Massachusetts has changed over time, compared to the baseline year of 1630. The red line representing the 1630 forest cover serves as a constant reference for these changes. By visualizing these data, we can observe trends, identify significant changes, and make informed hypotheses about the factors influencing these changes, or about how forest cover change might be influencing other dynamics.