

README: Spatial Copyright Registrations Dataset

This dataset includes U.S. county-level information used to analyze the geography of copyright registrations between 1790 and 1870. The dependent variable 'log_copyright' is computed using scraped data from the Library of Congress. All other variables are taken from Haines (2010), Historical Statistics of the United States.

Variable Descriptions:

- stateicp: State identifier (ICPSR code).
- countyicp: County identifier (ICPSR code).
- totpop_1850: Total population in 1850.
- urb850_1850: Urban population in places >850 people in 1850.
- urb25_1850: Urban population in places >2,500 people in 1850.
- colleges_1850: Number of colleges in 1850.
- pubsch_1850: Number of public schools in 1850.
- totillit_1850: Total illiterate population in 1850.
- schlib_1850: Presence of school libraries in 1850 (binary or count).
- log_copyright: Log of total copyright registrations 1790-1870 (computed from Library of Congress data).

Data Sources:

Haines, Michael R. and Inter-university Consortium for Political and Social Research. (2010). Historical, demographic, economic, and social data: The United States, 1790-2002. Inter-university Consortium for Political and Social Research.