

mapStats: An R Package for Geographic Visualization of Survey Statistics

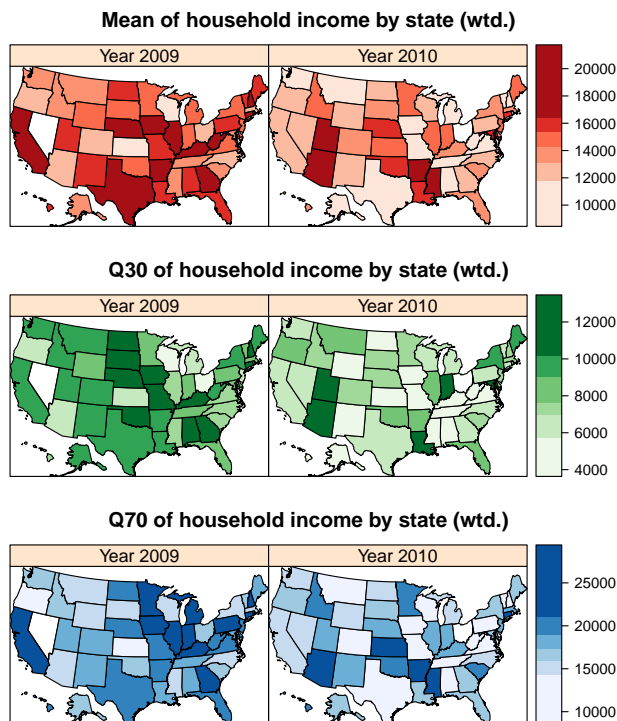
Samuel S. Ackerman*

Department of Statistics, Temple University

*Contact author: ackerman@temple.edu

Keywords: maps, lattice, shapefile

mapStats is an *R* package that produces color-coded maps of survey data using **lattice** graphics. It calculates variable statistics (mean, quantiles, total), either weighted or unweighted, from survey data where observations have a geographic element, and produces a map with colors representing the levels of the statistic. For example, one may want to visualize how the median and mean income of respondents by state in a survey changes over several years or differs by education level. Visualizing survey statistics geographically, which this package greatly simplifies, is an important way to understand characteristics of the survey data quickly. **mapStats** has flexible user control over appearance, including options such as overplotting maps with another shapefile or graphical element. Generating quality graphics without the package would require familiarity with **lattice** utilities and knowledge of manipulating shapefiles to merge them with data for plotting, which may be difficult for a casual user. Methods in *R* for calculation of summary statistics, such as quantile regression, require manual case-by-case manipulation of output, and do not necessarily perform well in cases where levels of class combinations are missing. The package includes functions for calculating variable statistics which deal well with missing data or class combinations, and that can be used for class variables in general (not necessarily geographies) or without plotting. An example of output on synthetic data is shown below:



References

Ackerman, Samuel (2013). “mapStats: An R package for geographic display of survey statistics.”