GeoVISTA

GeoViz Toolkit:

Coordinated analysis of multivariate relationships in geographic space

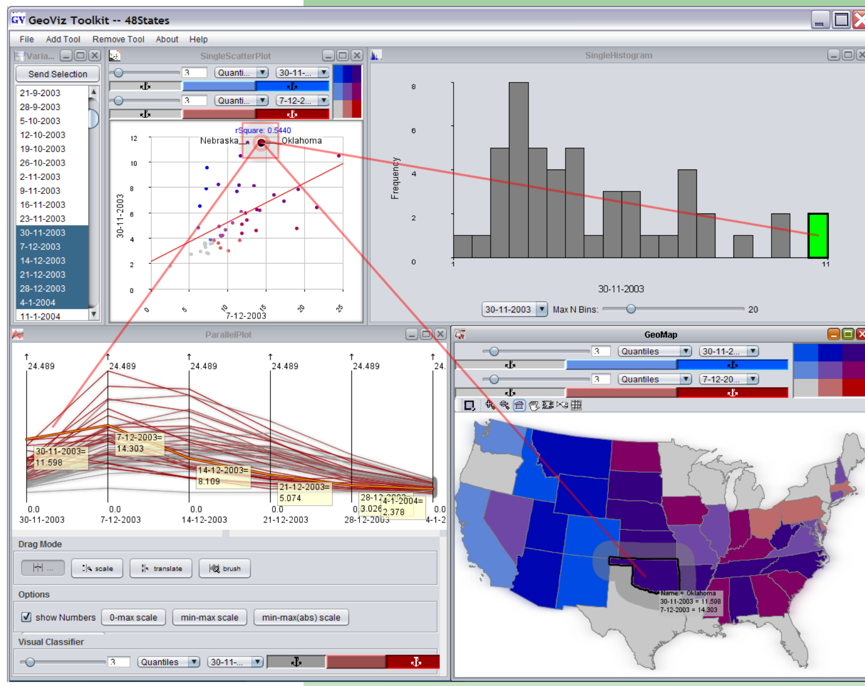
Frank Hardisty

http://www.geovista.psu.edu/geoviztoolkit/

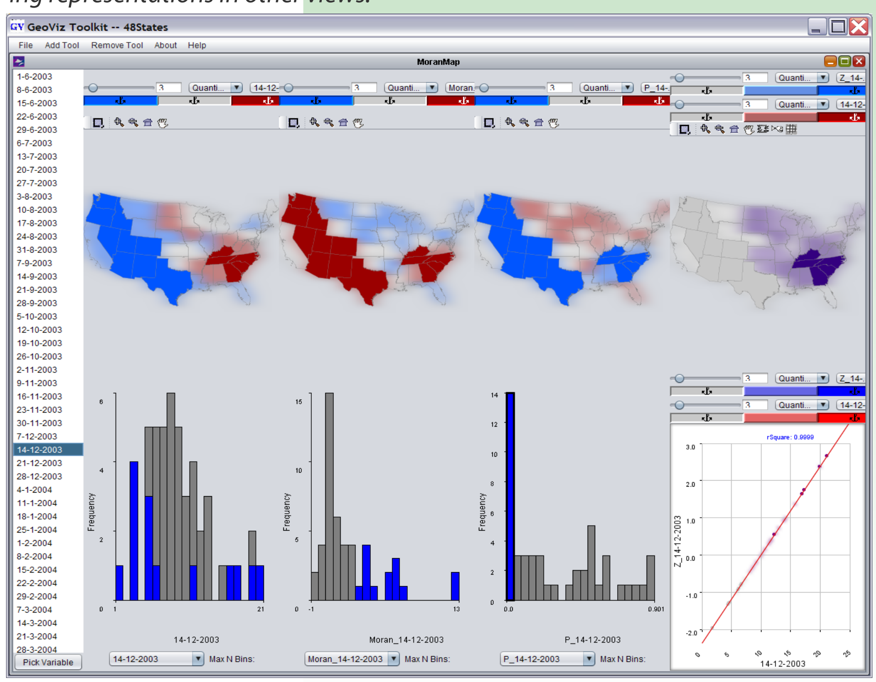
The GeoViz Toolkit supports systematic analysis of spatial, temporal, and attribute data sets. It allows analysts to discover previously hidden patterns in data, moving from spatial patterns to statistical patterns and back again by mixing and matching data visualization components to quickly construct custom analysis tools. Components are automatically coordinated across multiple views to enable insight into highly multivariate data, especially geographically organized data. Potential applications range from research in public health (e.g., infectious disease dynamics or cancer etiology, surveillance, and control), through analysis of socioeconomic and demographic data in support of both research and public policy, to exploration of patterns of incidents related to terrorism or crime.

The Toolkit provides a large selection of mapping and statistical graphing components for depicting univariate and multivariate data in dynamically linked views. Univariate data is represented in histograms, statistical maps, and cartograms (a map form that scales places to match their data value). Relationships among pairs of variables are explored with bivariate maps and graphs, while multivariate relationships are explored using matrixes of views (with many variables in pairwise views) and multivariate glyphs. The matrixes can combine other graphs in novel ways.

Additionally, the components within the variety of available spatio-statistical views are automatically coordinated. Data loaded into one of the components is reflected in the rest and interaction in one view produces relevant actions or highlighting in other views, enabling analysts to quickly explore relationships among many attributes geographically, e.g., identifying unusual places.



Leader lines help orient the user among multiple representations. Mousing over the values for Oklahoma in the scatterplot causes lines to be drawn to the corresponding representations in other views.



This local Moran's I tool shows spatial autocorrelation among flu incidence rates in the week of December 14-21 of 2003, as estimated by Google Flu Trends. The maps and statistical graphics show, in turn, the original rates, the local Moran's I rates, which rates are statistically significant using Monte Carlo simulation, and the relationship between the original rates and the local Moran's I rates. Regions highlighted are significant clusters. In the left map, blue is a cluster of low rates (flu rates were high here in earlier weeks) and red is a cluster of high rates currently.

PENNSTATE

Research described here is supported (in part) by The National Science Foundation, The National Cancer Institute, The Centers for Disease Control and by the the National Visualization and Analytics Center (NVAC), a U.S. Department of Homeland Security Program. NVAC is operated by the Pacific Northwest National Laboratory, a U.S. Department of Energy Office of Science laboratory.