The National Historical Geographic Information System

Pétra Noble and Catherine A. Fitch

Minnesota Population Center
University of Minnesota

The census is the primary source of statistical information about growth and change of the American population since 1790. Aggregate data tables, in print or electronic format, are the principal means of describing the characteristics of states, metropolitan areas, cities, counties, minor civil divisions and neighborhoods. Approximately 670 gigabytes of U.S. Census summary data covering the period 1790 through 2000 exist but they are scattered across dozens of archives and are stored in incompatible formats on magnetic tape, CD-ROMs, or on paper.

Even if this massive body of aggregate census data were readily accessible, historical analysis would be complicated by changes in the geographic units they describe. Summary counts of the population characteristics of places are meaningful only if those places are clearly defined. In every census year, the boundaries of the geographic units described by the census have been modified. This makes the analysis of geographic change in the American population exceedingly difficult. When researchers do assess change, they must develop their own maps at great expense. Not surprisingly, the number of such studies is small and their geographic and chronological scope is limited (e.g., Duncan 1957, Leiberson 1961, Denton and Massey 1991, Massey and Eggers 1990; Massey and Denton 1988, Alba et al. 1995, Logan et al. 1996). To allow systematic analysis of change over time, we need a compatible set of electronic maps that describe the location of each geographic unit tabulated by the Census. Before NHGIS, such maps existed only for the 1990 and 2000 census years.

The National Historical Geographic Information System (NHGIS) is designed to make the rich body of aggregate census data accessible within a GIS framework for historical

population research. The five-year project began in April 2001 with funding from the National Science Foundation Social Science Infrastructure Program. The goals were to gather together all surviving census data from 1790 to 2000; format them consistently; develop comprehensive standardized machine-readable documentation; create high-precision historical electronic boundary files describing census tracts and counties; and develop innovative web-based tools for disseminating both microdata and metadata over the Internet. As we reach the conclusion of this phase of the project we would like to illustrate what we have been done to achieve these goals, highlight research that has been conducted as a result of NHGIS, demonstrate the data access system (available at http://nhgis.org), and discuss the next steps for the project.