Historical Analysis of Census Figures, 1790-2000: A Preliminary Report on NAGIS (Native American Geographic Information System)



December 2007

Table of Contents

- 1 I. A short history of the US Decennial Census
- 1-2 II. Project Goals
- 2 III. Notes on Racial Classification
- 2-3 IV. Research Methodology
- 3-6 V. Beta Test Results
- 6-7 VI. Preliminary Map Analysis
- 7-8 VII. Future Directions
- 9-33 VIII. Maps
- 34-35 IX. Bibliography

Acknowledgements

There are so many people that I would like to mention that have helped me with this project. To LeeAnne, my lovely wife; you graciously put up with this immense undertaking despite it being what I termed a "fun diversion." I want to thank all the people at Warm Springs Middle School who showed interest in this project even before I knew what I was really doing.

The History Department at West Virginia University provided me a home while I perfected the computer and research skills which made this feasible. They are again my home, and I hope this will be something I can give back to the department and the university. I especially would like to thank Dr. Tyler Boulware and Dr. Kenneth Fones-Wolf for their interest and questions concerning this project. This project would still be a figment of my overactive imagination without the assistance of Dr. Gregory Elmes in the WVU Geography Department. You have shown great patience with my many questions concerning the grey areas of GIS.

I have also been assisted by Jeff Davis and Andrea Keller at the West Virginia State Historic Preservation Office, Jennifer Cosham at the Maryland Historical Trust, Dave Snyder at the Ohio State Historic Preservation Office, Mike Barber and Sara Leonard at the Virginia Department of Historic Resources, Sarah Briland at the Kentucky Heritage Council, Kurt Carr and Amanda Shafer at the Pennsylvania Historical and Museum Commission, Christina Pappas at the Kentucky Office of State Archaeology, and John Long at the Newberry Library's Scholl Center Atlas of Historical County Boundaries Project. To you and anyone I might have accidentally missed, thank you so much for you time and assistance.

Historical Analysis of Census Figures, 1790-2000: Native American Geographic Information System, NAGIS

This project began as a brief look at census numbers for a few counties in West Virginia but has expanded greatly over the past year. This project encompasses a very broad research question: What kinds of numerical data can be found and used to look at the Native American history of West Virginia and the surrounding region? This has produced some very basic concerns about using census data and the role of ethnicity and empirical data. This project has also provided an opportunity to see how effective technology can be in historical queries. This is the analysis of the initial attempt at creating a working GIS based on such wide ranging census data. My primary focus here is the technical details of the project, though in the last few sections I will provide a cursory analysis of the maps produced from the beta test.

I. A Short History of the US Decennial Census

The United States was the first country to systematically and periodically gather census data. The Constitution created decennial censuses to count the people of each state for the apportioning of Congressional representatives. Over time the schedules of questions grew and the types of questions began to reach beyond the most basic information. By the late nineteenth century, the US Census had developed to also include commerce and industry. It wasn't until the 1870 Census that Native Americans or "Indians" were included, but still only selectively. The only Indians enumerated were those living among the general population, not those "untaxed" Indians still living on reservations. From 1870 till the present, the census has counted Indians in ever increasing complexity. The 1970 Census initiated self-identification of race or ethnicity among other categories. All previous surveys had required Census officials to identify the race/ethnicity. The 2000 Census altered the way Indians are counted by allowing individuals to self-identify as more than one group. In the next two years, the Census Bureau will be preparing for the 2010 Census, marking 22 censuses, 12 of those presided over by the Census Bureau established in 1900.

II. Project Goals

As stated above, I was interested in what types of materials I could use to research Native Americans in the state of West Virginia. After surveying the records, settling on census data, and developing a basic GIS structure, I developed four goals.

A. Collect and Query Census Data Temporally and Regionally: This is the most basic structural goal, but provided many unforeseen complications. The collection of data was mostly complete by the time I began developing this project but the ability to query the data proved elusive. Regions are geographic in nature and thus easy to view in a GIS environment. Time was another matter entirely for the GIS. Time is a human construct that computers still have trouble understanding.

B. Analyze Population movements in West Virginia and surrounding states: By analyzing the census maps created by the GIS, I hoped to determine patterns of mobility of people in the region across time.

C. Analyze Native American population centers in hopes of establishing continuity: By analyzing the census data concerning Native Americans, I hoped to establish the nature of Native American involvement in the region through time. My historical research into Native Americans has been an effort to showcase the continuity of Native communities in the region.

D. Analyze archaeological site distribution and NA populations: By overlaying archaeological site data over the existing census data, I hoped to determine any connections between precontact/early historical Indian settlements and clusters of modern Indian people. This is connected closely with the third goal, but was not realized by this initial project.

III. Notes on Racial Classification

I have selected the census data for three categories: Native Americans (Indians), African Americans (Blacks both free and enslaved), and Total Population at the county level. The identification of racial categories provided a major issue for this project. Historians have been dealing with the implication of racial classification for over 300 years. I have based my classifications on the standards set by each census' schedule defining racial categories, though even this proved confusing at times. The classification and identification of Native Americans in the Census has proven especially problematic. First one has to answer whether Native Americans or Indians are part of a race or should be considered an ethnicity. Historically, Native Americans were classified as a "race" though there are no scientific or biological reasons behind this. The race construction also breaks down through the process of sexual relationships between Indians and non-Indians. More recently, Native Americans have been labeled as an ethnicity, or groups with significant differences of culture. The problem here is that the term cannot be applied to all Native American people because of the wide variations in culture and history. I have chosen here to use the "race" construction for the purpose of initial research. Any individual with an ancestral link to the inhabitants of North America before European contact thus can be classified as Native American. The complexities of racial identification are a quagmire for most researchers.

IV. Research Methodology

There are five stages of research for NAGIS. Stage One was collecting data. The queries of Census data were made primarily online through the US Census website and the University of Virginia Library Historical Census Data Browser. Seeking to cast a very wide net, I decided to collect data for the states bordering West Virginia, Maryland, Virginia, Kentucky, Ohio and Pennsylvania. I narrowed down the counties to those closest to West Virginia. Though I am focused on Native American census figures, I decided to also record African American (Black) and the Total population figures at the county level. I collected this data from each census, 1790 to 2000. I created an Excel spreadsheet for Excel 2003 to include state and county level census figures with calculations for percentages of population and percentages of change. After collecting much of the census data, I also realized that I would need information about the county area changes through the census years. In addition, I collected county level total archaeological site numbers as of 2007. All these figures were placed in the spreadsheet. This stage is nearing completion with the exception of dealing with the complicated counting of Native Americans in the 2000 census and standardizing the county land area values across time. I have prepared to increase the data collected to include individual data from the 1870-1930 Censuses for those classified as Native American. These are the only available years that have Native American data listed. This will involve using the US Gen Website.

Stage Two is the research phase and was in progress very soon after data collection began. This background research included gathering historical information about the census, county line evolution, racial classification, as well as general information about using GIS in historical projects.



Stage Three involved evaluating the data collected, methods of interpretation, further data queries and overall error checking of the spreadsheet against existing Census data. At this time, I began the final preparations for entering the data into the GIS. As this is my first experience with GIS, I decided truncate my research to create a manageable GIS structure using just the West Virginia data. I have created a general flow chart of what I think the structure of the database will be. The yellow indicates imagery, and the blue indicates database files. After coming up with a plan of action and collecting all needed materials, I will move on to building.

Stage Four is the construction of a beta version of the GIS using ArcDesktop. One special concern I have is the creation/collection of base maps, especially for the historical counties. The basemaps were obtained from the National Historic Geographic Information System (NHGIS) at <u>www.nhgis.org</u>. There was a national map for each census year, and I extracted the counties from my prescribed research area. My data was in a spreadsheet, so I had to reformat it into a dbf file. This stage will be discussed in greater detail in the next three sections.

After the beta test, I produced a series of maps for analysis. I hope to answer a few questions. Are there any clusters in the data? Is there evidence of population movements through the region? What role does topological region play in settlement patterns of Native Americans?

Stage Five brings the whole process full circle by enacting the changes suggested by the beta test. This will eventually include an expansion of database materials, more specific archaeological data, historical imagery, and Indian trails. The grandest dream of this so far would be a one-stop shop for site level archaeological data coupled with census data, environmental data, historical data and imagery, which aid in the mapping and analysis of a wide history in the West Virginia region.

V. Beta Test Results

As with learning any new skill, there were both successes and failures. The initial articulation of the Native American GIS came with many issues to be resolved. Below is a discussion of both the resolved and unresolved issues. The most general and fundamental issues of design were resolved but left many of the smaller problems to be fixed after closer inspection. Overall, though, this beta test was a success in that it produced a series of maps depicting the census data with which I began. I can see from the maps included in this project that my first three goals are feasible and will produce meaningful results.

A. Resolved Issues

I began this process using base maps with county outlines that required sophisticated time constraints. The Newberry Library has been developing interactive web-based GIS viewing of the historical boundaries of many states at http://www.newberry.org/ahcbp/state_index.html. There was a polygon for each time the boundary of a county changed. The SQL query for building a single map was inordinately complicated, in fact I still have not figured out this method. In addition to this problem, I was unable to find similar shapefiles for Pennsylvania and Maryland. This prompted me to use only West Virginia for my beta test. After some more practice using ArcGIS, I was able to find some base maps which better suited my needs. The National Historic Geographic Information System (NHGIS) is a clearinghouse of data and maps for use in GIS environments at www.nhgis.org. I was able to find county level data for all censuses for the entire United States. I painstakingly processed each census map to extract only the 6 states of my research area and saved them to my personal database. I placed each map as a layer and overlaid state outlines, which also changed through time. It became clear that it would be just as easy to use all the region, and more importantly all my data, rather than just looking at West Virginia. There were some minor problems with this method as will be discussed below.

As with the base maps, my beginning database structure was overly complicated because I ignored many of the analyses which ArcGIS can do with raw data. I provided databases with pre-calculated population densities and percent change. I also decomposed the databases needlessly so each state had nine smaller databases, for each ethnic group's population, population densities, and population changes. This seemed necessary in order to utilize the Newberry shapefiles. The discovery of the NHGIS shapefiles changed the way I needed to handle database management. I combined all of my states into one spreadsheet with all figures included and no unnecessary calculations and then broke this down to three separate sheets for Native American Population, Black Population, and Total Population. Then I converted them into Access Database files, where I cleaned them up in preparation for conversion into dbf (IV) files. These dbf files were then imported into my geodatabase through ArcCatalog and placed on my multilayered map.

The final step of joining the databases proved a challenge in many ways. I had first attempted to join access formatted files to the basemaps, and this failed. I learned that the best way to format the databases is as dbf files. This done, I attempted to join the database to each basemap but ran into similar errors. This seems to be connected to the unresolved issues with the basemaps themselves. The joins were successful, but I was unable to view the attribute tables for each layer now. After some trial and error, I discovered that the joins were successful and I was able to symbolize the corresponding census year to the map to which it was joined. Though extremely cumbersome, joining each map layer to each of the three databases produced the ability to view the data at the most basic level.

B. Unresolved Issues

Though I declared this beta test a success, I have to admit there were more unresolved issues than resolved ones. The most general issue remaining for this project is how to best query across time. As I designed this GIS, each census was snapshot or layer which could in essence stand by itself. Analyzing the data between two censuses remains very cumbersome and artificial. Temporal queries will require testing separate from display of raw data as done here.

More detrimental to the flow of NAGIS as a major basemap zoom error. I discovered a drawing error when attempting to zoom in to just view West Virginia. At a certain point,

currently unknown, ArcGIS was unable to draw the shapefile. This was not due an established visible scale range, but rather to an internal file flaw. I circumvented this by only zooming in as close as it would let me and leaving it there for printing. This may also be connected to some difficulty in joining the census database to the maps. After successfully joining any of the three databases to a layer map, I would be unable to view the new attribute table. This seems to be a viewing problem as the data was joined enough to be used for symbolizing the map. This inability to view the attribute table complicated a problem verifying the appropriate connection of census data to the correct counties. There were instances, notably with Fauquier County and the Independent Cities in Virginia, where I had data in my database but no information appears on the map for the shapefile. In addition, some counties that I know I did not collect census data for appeared to have data. Since I was unable to view the attribute tables, I was unable to locate the exact issue. This, therefore, limits the utility of the maps to a degree.

We turn now to the maps themselves and the maximization of the display of the census data. In an effort to conserve time and produce maps, I committed the novice error of symbolizing the raw census data in choropleth form. It was suggested by more knowledgeable cartographers that this data should either be visualized as graduated symbols or calculated as population density. Choropleth depictions of raw census data are commonly understood by historians despite their limited analytical value. To show population growth, I used the same symbol levels for each series of maps, using different colors for Native American, Black, and Total population. There were some problems with the color schemes used for each series. The middle oranges of the Native American scale are hard to distinguish. The color scheme for Black populations is bad for color-blind historians. Lastly, the color scheme and symbol levels for Total population were poorly planned and can be misleading with the two colors. Despite these deficiencies, the shared symbol levels does allow for viewing population increases, especially on the Native American series.

Database issues consumed much of the time on this project. In converting the Access database to a dbf file, all null values were turned into zeros. This is problematic for the Native American table, especially because while most places did collect Native American populations, some counties did not. Thus their values would be "null" not zero. This could also mess up the joins between the databases and the census year maps. The collection of Native American population numbers also proved difficult for the period of 1900-1950 when I could only find "All other races" as a category. Therefore, the maps for those years may be skewed higher. I am still seeking the Native American numbers at the county level for this period. The figures for Native American population in the 2000 census proved equally frustrating for this project. For the first time, individuals were able to classify themselves as multiple races and ethnicities. I only included single race identification here, but will eventually include the much larger numbers. This does complicate the display and analysis of racial categories, though rightfully so. Lastly, the database is prepared for an analysis of the project identified regions and state level, but I was constrained by time for this level query. At the state level I would also have to include the numbers gathered from the census general figures and look at how the two sometimes do not match. Though mainly dealing with the complicated practices involved in collecting Native American census figures, these database issues suggest much wider issues in database management in a GIS environment.

There are two relatively minor issues remaining. First, time and greater concerns have precluded the writing and revision of metadata, though its importance can not be understated. Lastly, the collection and inclusion of additional map layers such as hydrography, archaeological

site map and the digital elevation model (DEM) remain to be completed. I was able to find the hydrography for West Virginia, but there were problems with the projection that have yet to be resolved. Finding suitable maps in these categories for the entire region may take some time.

Many of these unresolved problems are minor but must be resolved to smooth out the workflow of this project. The next version of this GIS will begin handling these issues.

VI. Preliminary Map Analysis

There were three series of raw population maps produced from this project. The N series depicts the Native American populations 1870-2000 at the county level. B series maps depict African Americans and the T series the total populations for the years 1790, 1870, 1900, 1950, and 2000. These years are correlated with significant historical events. The 1790 census was the first national census. The 1870 census is the first year that Native Americans were counted. 1900 was the first year for the permanent US Census Bureau. 1950 was a major peak of the coal industry in West Virginia. Finally, the 2000 census is the most current census data. The following preliminary analyses are focused specifically on establishing the population trends of West Virginia, the first goal of this project. Further analyses will produce maps depicted population change, percentages and other similar queries.

A. N Series: Native American Populations

The Native American population series begins in 1870 with the first enumeration of Indians not living on reservations. There are few Native Americans in any county of West Virginia until 1900. There is no discernable pattern till the turn of the twentieth century. There seems to be higher concentrations along the major river valleys, especially the Ohio, Kanawha, and Tygart Rivers. Kanawha County is consistently one of the highest populations of native people. It is not surprising that southern West Virginia has the highest concentrations of Indians after the 1930s and the booming coal industry. What was more surprising was that the highest populations seem to be focused in more urban counties as opposed to more rural counties. The Native American population rises dramatically. This might be due to the peaking coal industry but the widespread nature of the increase could also be due to a change in the reporting. I am particularly interested in when individuals began self-identification of race. The period from 1900-1950 may be skewed higher because this depicts the more general category "All Other Races." The 1970 census is the first time that all counties of West Virginia had Indian residents. The highest concentrations continued to follow the river systems, particularly the Kanawha and Tygart with stronger populations in the south. Greenbrier County, a rural county, has a notably high Native American population in the 1970 census for unknown reasons. There is little change in this pattern until the 2000 census when individuals were allowed to self-identify as multiple races. This data only represents individuals identifying as Native American alone. Again, the river valleys, south and urban counties have the highest populations of Indians.

B. B Series: African American Populations

This truncated series of maps shows similar patterns to the Native American populations. The most striking correlation seems to be higher numbers along the larger rivers valleys, first in the Shenandoah in 1790 and then in the Kanawha and Tygart in 1870 and 1900. The Shenandoah remains higher than most of West Virginia through each of the selected census years. The dramatic rise of McDowell County from 1-10 African Americans in 1870 to 5,001-10,000 in 1900 is probably due to coal mining that was affecting all of the southern counties. The highest

populations of African Americans was Kanawha County in 1950 and 2000, which closely resembles the Native American population trend towards urban counties over rural. The height of African American population occurred in 1900 which was marked by a slighter higher overall population. Beginning in the 1950 census and continuing in the 2000 census there appears to be slightly lower numbers of African Americans in West Virginia. This may be a result of the great migration of minorities in the south towards urban centers in the north that severely decreased West Virginia's total population in the 1940s and 1950s.

C. T Series: Total Populations

West Virginia's total population maps show initially few numbers in 1790 and 1870, but a marked increase and then leveling off from 1900 till 2000. This selection has a major flaw in that it does not show the significant decline of West Virginia populations in the early and late twentieth century. There is an obvious bias towards higher populations in urban counties like Kanawha and Fayette in the south; Monongalia in north-central; and Jefferson and Berkeley in the eastern panhandle. The counties in the highest elevations of the Appalachian Mountains remain relatively poorly populated. Pocahontas County is a solid example of this trend. A majority of West Virginia counties seem to be between 10,001 and 50,000 people with Kanawha County reporting 100,000-500,000 total people. Again the Kanawha and Tygart River valleys loosely hold the highest population number. The south, though, remains slightly higher in total populations. The increasingly urban counties, Monongalia and Marion, are clustered in the north; but Berkeley and Jefferson Counties also had elevated populations.

VII. Future Directions

A. Poster Session Presentation: I will be representing the History Department with a poster concerning this project and its results at the West Virginia University Spring Poster Fair.

B. Refinement of GIS structure: This basically requires finding solutions for the major unresolved issues detailed above. The most important issues for this project are finding a better way to perform temporal queries and finding a more compatible and flexible database using SQL.

C. Broaden the types of data included: This will allow for the completion of the fourth goal dealing with archaeological site information. In addition to the archaeological database, I would like to add a hydrographic layer and a DEM. This could eventually include images, photographs, and oral history interviews, though obviously this would not be for spatial analysis.

D. Refine the Census Data: My original exhausting collection of Census data was flawed in many ways. I hope to go back and re-gather the Census data differently. First, I would place the data in a database instead of a spreadsheet. In addition to my original county-level data, I would collect census block-level data. Lastly, I would gather all Native American data at the individual-level. This is only possible for the 1870-1930 censuses; the 1940 census won't be released until 2015.

E. Web-based interface for wider use: Much of this data has been to ignored or remained too difficult for the public to view, so I would like to create a web-based interface to allow wider distribution. This will serve many purposes. First, the story of West Virginian Native Americans

will be brought to a wider public. Second, this could be developed as a tool for public school history and geography teachers in the use of Census data, historical maps, demography, GIS and computer technology. Lastly, this would provide for feedback on new avenues for queries of historical data, especially census data from interested parties both the public and academic.

F. Production of various paper versions of this project: There are three main products from this GIS. The most obvious is a book of maps depicting the entire series of raw and processed data. I will produce two articles, one about the technical aspects of this project, and the other a historical article developing the story of race and ethnicity in West Virginia and the surrounding regions. This last article could also be later developed into a monograph.

VIII. Maps

	-	Table of Maps
10	NAGIS-N1870	Native American Population by County, 1870
11	NAGIS-N1880	Native American Population by County, 1880
12	NAGIS-N1890	Native American Population by County, 1890
13	NAGIS-N1900	Native American Population by County, 1900 *
14	NAGIS-N1910	Native American Population by County, 1910 *
15	NAGIS-N1920	Native American Population by County, 1920 *
16	NAGIS-N1930	Native American Population by County, 1930 *
17	NAGIS-N1940	Native American Population by County, 1940 *
18	NAGIS-N1950	Native American Population by County, 1950 *
19	NAGIS-N1960	Native American Population by County, 1960
20	NAGIS-N1970	Native American Population by County, 1970
21	NAGIS-N1980	Native American Population by County, 1980
22	NAGIS-N1990	Native American Population by County, 1990
23	NAGIS-N2000	Native American Population by County, 2000 **
24	NAGIS-B1790	African American Population by County, 1790
25	NAGIS-B1870	African American Population by County, 1870
26	NAGIS-B1900	African American Population by County, 1900
27	NAGIS-B1950	African American Population by County, 1950
28	NAGIS-B2000	African American Population by County, 2000 **
29	NAGIS-T1790	Total Population by County, 1790
30	NAGIS-T1870	Total Population by County, 1870
31	NAGIS-T1900	Total Population by County, 1900
32	NAGIS-T1950	Total Population by County, 1950
33	NAGIS-T2000	Total Population by County, 2000

* This set of data was created from a tabulation of "people of other races." ** This set of data only includes only people who identified as one race.











































0
1 - 5
6 -10
11 - 25
16 - 50
51 - 100
101 - 500
501 - 1000
1001 - 2500
2501+





















































IX. Bibliography

A. Census Data Census data for each state collected on November 13, 2006 from <u>http://fisher.lib.virginia.edu/collections/stats/histcensus/php/state.php</u>, 1790-1960 for all states.

Corraborating Census data collected on November 06, 2006, from http://www.census.gov/population/www/censusdata/hiscendata.html, 1790-2000 for all states.

B. Base Maps Census Map shapefiles 1790-2000 from <u>www.nhgis.org</u>.

Historical County shapefiles for Kentucky, Ohio, Virginia, West Virginia from <u>http://www.newberry.org/ahcbp/state_index.html</u>.

Map Data and outlines from http://www2.census.gov/geo/maps/general_ref/stco_outline/cen2k_pgsz/.

Thorndale, William and William Dollarhide. *Map Guide to the U.S. Federal Censuses, 1790-1920.* Baltimore: Genealogical Publishing Co. Inc. 1987.

C. Additional Research

Adivsory Committee on Problems of Census Enumeration. *America's Uncounted People*. Washington D.C.: Division of Behavioral Sciences, National Research Council, National Academy of Sciences. 1971.

Alonso, Willam and Paul Starr, eds. *The Politics of Numbers*: for the National Committee for Research on the 1980 Census. New York: Russell Sage Foundation. 1987.

Alterman, Hyman. *Counting People, The Census in History*. New York: Harcourt, Brace and World Inc. 1969.

Anderson, Margo J. *The American Census: A Social History*. New Haven: Yale University Press. 1988.

Anderson, Margo J. and Stephen E. Fienberg. *Who Counts? The Politics of Census-Taking in Contemporary America*. New York: Russell Sage Foundation. 1999.

Anderson, Margo J. "Encyclopedia of the U.S. Census". Washington, DC: CQ Press, 2000.

Bureau of the Census of the U.S. Department of Commerce. *The Statistical History of the United States from Colonial Times to the Present*. Stamford, CT: Fairfield Publishers. 1965.

Eckler, A. Ross. The Bureau of the Census. New York: Praeger Publishers. 1972.

Moore, Donald S.; Jake Kosek and Anand Pandian, eds. *Race, Nature, and the Politics of Difference*. Durham: Duke University Press. 2003.

Paisano, Edna L., et al, *We the...First Americans*. For the US Department of Commerce Economics and Statistics Administration Bureau of the Census on November 14, 2006, <u>http://www.census.gov/pubinfo/www/NEWamindML1.html</u>, September 1993.

Peters, Alan and Heather MacDonald. *Unlocking the Census with GIS*. Redlands, CA: ESRI Press. 2004.

Scott, Ann Herbert. *Census, U.S.A. Fact Finding for the American People, 1790-1970.* New York: Seabury Press. 1968.

Skerry, Peter. *Counting on the Census? Race, Group Identity and the Evasion of Politics*. Washington D.C.: Brookings Institution Press. 2000.

Wright, Carroll D. *The History and Growth of the United States Census*. Washington D.C.: Government Printing Office. 1900.

Yates, Frank, F.R.S. *Sampling Methods for Censuses and Surveys*. 3rd ed. London: Charles Griffin and Company Limited. 1965.