

Mapping the Record of American Democracy (ROAD) 1984-1990



data.fas.harvard.edu/ROAD/

The Record Of American Democracy (ROAD) data includes election returns, socioeconomic summaries, and demographic measures of the American public at unusually low levels of geographic aggregation. The NSF-supported ROAD project covers every state in the country from 1984 through 1990 (including some off-year elections). One collection of data sets includes every election at and above State House, along with party registration and other variables, in each state for the roughly 170,000 precincts nationwide (about 60 times the number of counties). Another collection has added to these (roughly 30-40) political variables an additional 3,725 variables merged from the 1990 U.S. Census for 47,327 aggregate units (about 15 times the number of counties) about the size one or more cities or towns. These units completely tile the U.S. landmass. The collection also includes geographic boundary files so users can easily draw maps with these data.

Gary King
 Director, Institute for Quantitative Social Science
 David Florence Professor of Government
 Department of Government
 Harvard University

With Contributions from Project Leaders:

Micah Altman
 Associate Director
 Harvard-MIT Data Center
 Senior Research Scientist
 Institute for Quantitative Social Science

Jeffrey Lewis
 Assistant Professor
 Massachusetts Institute of Technology

Greg Adams
Kenneth Benoit

and **Bradley Palmquist**

Eric Reinhardt
 Associate Professor
 Department of Political Science
 Emory University

Claudine Gay
 Associate Professor
 Department of Political Science
 Stanford University

Russ Mayer



Poster Author: Anders N. Hopperstead, GIS Specialist

The ROAD Project Invented the "MCD-Group": The Smallest Scale Geographic Unit at which Precinct Boundaries and Census Boundaries Coincide

In most counties, eligible citizens in a geographically contiguous area are assigned to a single voting booth to cast their ballot. We refer to the geographical description of the unit of analysis as a precinct, although in some states it is called an election district, and occasionally several voting booths exist within one precinct. The number of precincts nationwide is approximately 170,000, with some variation over time. More detailed official electoral data than precinct returns cannot be created for the entire nation unless the laws governing the secret ballot are revoked.

To permit analysis of this electoral data in conjunction with publicly-available census data, the ROAD project devised the **MCD-Group**, a geographic unit composed of Census Bureau defined Minor Civil Divisions (MCDs). Individual MCDs correspond roughly to towns and cities. There are 25,973 of these MCD-groups in the U.S. outside of California and an additional 21,200 within California, and completely tile the U.S. land mass. For each of these units, the ROAD collection provides all electoral data aggregated from the precinct level, as well as **3,725 variables** from the U.S. Census long form. This is the smallest level of aggregation at which publicly-available census data and electoral data can be made to coincide.

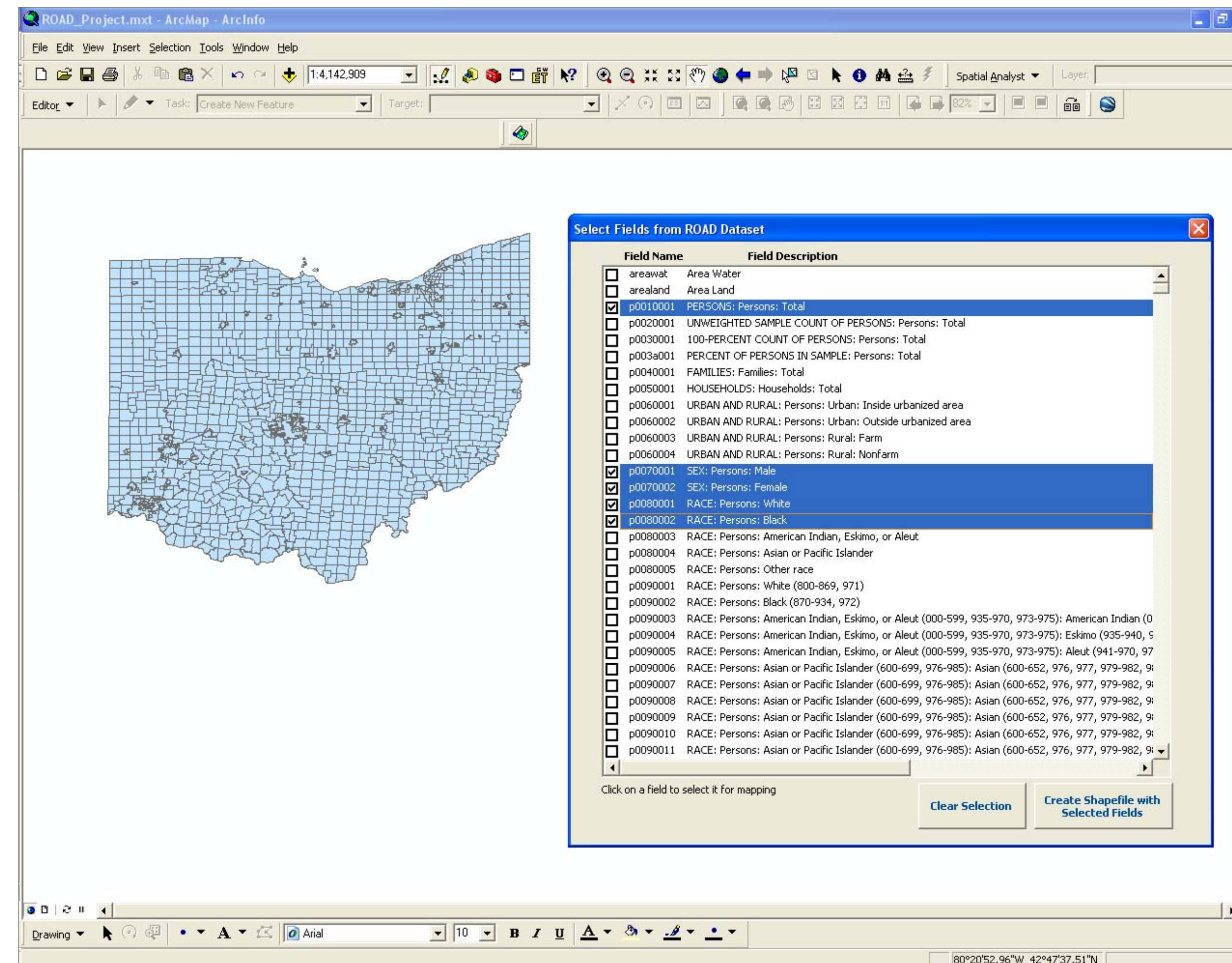
Making Maps Using the Virtual Data Center: The Challenge of Creating GIS Data Files from the ROAD Collection of Over 3,500 Variables

GIS Software limits the number of fields, or variables, in a single file. In ArcGIS 9.1, the limit is 255 fields. To use GIS software to analyze the 3,500+ variables in the ROAD collection, researchers must select a subset of fields and join them with the MCD-Group boundary files. In collaboration with the Harvard-MIT Data Center and the Murray Archives, the Center for Geographic Analysis has automated this process through the development of an ArcObjects application that allows researchers to select a subset of ROAD variables and generate a corresponding GIS data file. The application reads the data from files stored in the Virtual Data Center.

Figure 1. ROAD Data in SPSS File Format
 Statistical analysis software such as SPSS allows for large numbers of fields, or variables. All 3509 fields from the Ohio ROAD data set can be viewed at one time.

1:	areawat	arealnd	p0010001	p0020001	p0030001	p003a001	p0040001
1	49.00	83396.00	917.00	152.00	862.00	17.6	271.
2	18.00	107373.00	1271.00	190.00	1196.00	16.9	344.
3	0.00	140452.00	1092.00	203.00	1096.00	18.6	338.
4	1305.00	142782.00	716.00	181.00	704.00	25.7	215.
5	0.00	111686.00	946.00	166.00	1022.00	16.2	263.
6	19.00	112015.00	1478.00	234.00	1400.00	16.7	400.
7	53.00	5800.00	2994.00	1120.00	2393.00	46.8	660.
8	824.00	162867.00	3717.00	1232.00	3701.00	33.3	1010.
9	2013.00	70975.00	639.00	95.00	657.00	14.5	197.
10	98.00	75015.00	824.00	140.00	865.00	16.2	240.
11	5.00	96861.00	1974.00	642.00	1920.00	33.4	546.
12	196.00	113912.00	1934.00	271.00	1499.00	18.1	401.
13	219.00	120436.00	6178.00	860.00	5144.00	16.9	1413.
14	6.00	72532.00	1089.00	311.00	1147.00	27.1	308.
15	57.00	84627.00	1643.00	608.00	1764.00	34.5	455.
16	2378.00	86790.00	1767.00	305.00	1773.00	17.2	489.
17	42.00	63416.00	12414.00	2474.00	12407.00	19.9	3511.
18	63.00	93623.00	2793.00	623.00	2778.00	22.4	775.
19	2787.00	80259.00	10105.00	1420.00	10105.00	14.1	2521.
20	402.00	92708.00	2715.00	647.00	2737.00	23.6	800.
21	251.00	32818.00	45549.00	6702.00	45549.00	14.7	11008.
22	35.00	110250.00	6676.00	1097.00	6676.00	16.4	1845.
23	0.00	93784.00	21603.00	525.00	2065.00	25.1	636.
24	3.00	86644.00	3577.00	495.00	3577.00	14.0	953.

Figure 2. Selecting ROAD Variables for GIS Analysis and Mapping
 The application developed by the CGA allows researchers to browse and select up to 255 fields from the thousands of census and electoral variables in the ROAD collection. The application outputs a shapefile with the proper variable names and geographic boundaries.



Geographic Analysis of ROAD Data: Investigating Spatial Patterns and Analysis of the Correlates of Voting Behavior

One of the most significant contributions of the ROAD project is that it provides appropriate data for examining the relationship between the local environment and actual political behavior. While, political scientists have recognized that local context affects political behavior, the kind of detailed information need to test hypotheses regarding this relationship has heretofore been unavailable. An analysis of this type using the ROAD project files might consider a variety of electoral variables, including voter turnout, the vote for various candidates, registration rates, and partisan differentials in voting or registration. The explanatory variables in such an analysis would consist of census variables, including education levels, racial composition, and urbanicity.

Establishing a means for researchers to create GIS data files from the ROAD data collection is the first step towards supporting detailed geographic analysis of ROAD data. The Center for Geographic Analysis is working with political scientists in the Institute for Quantitative Social Science to develop tools and techniques for manipulation and investigation of ROAD data in GIS and spatial statistics software. Shown here are a few examples of thematic maps that were generated from ROAD data.

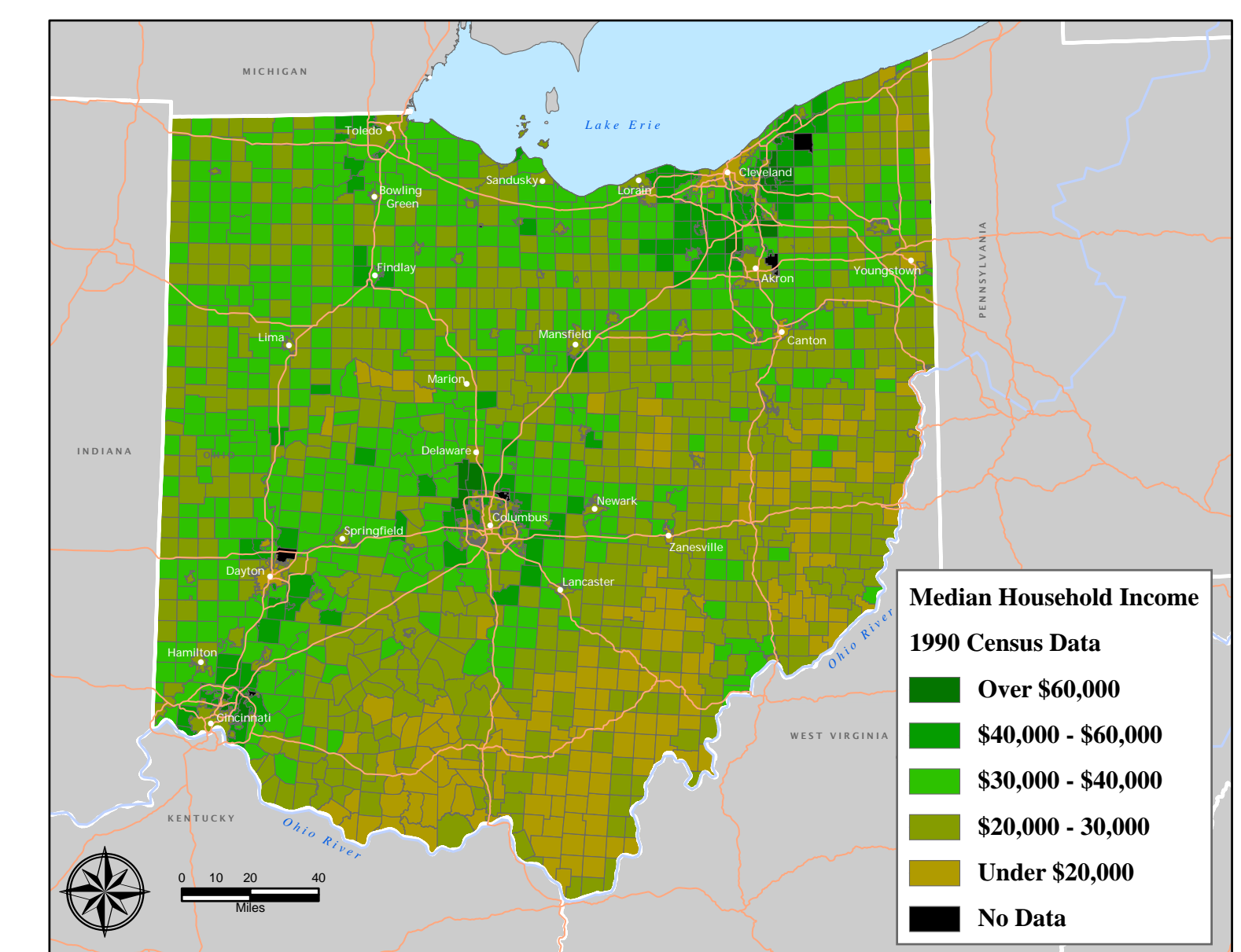
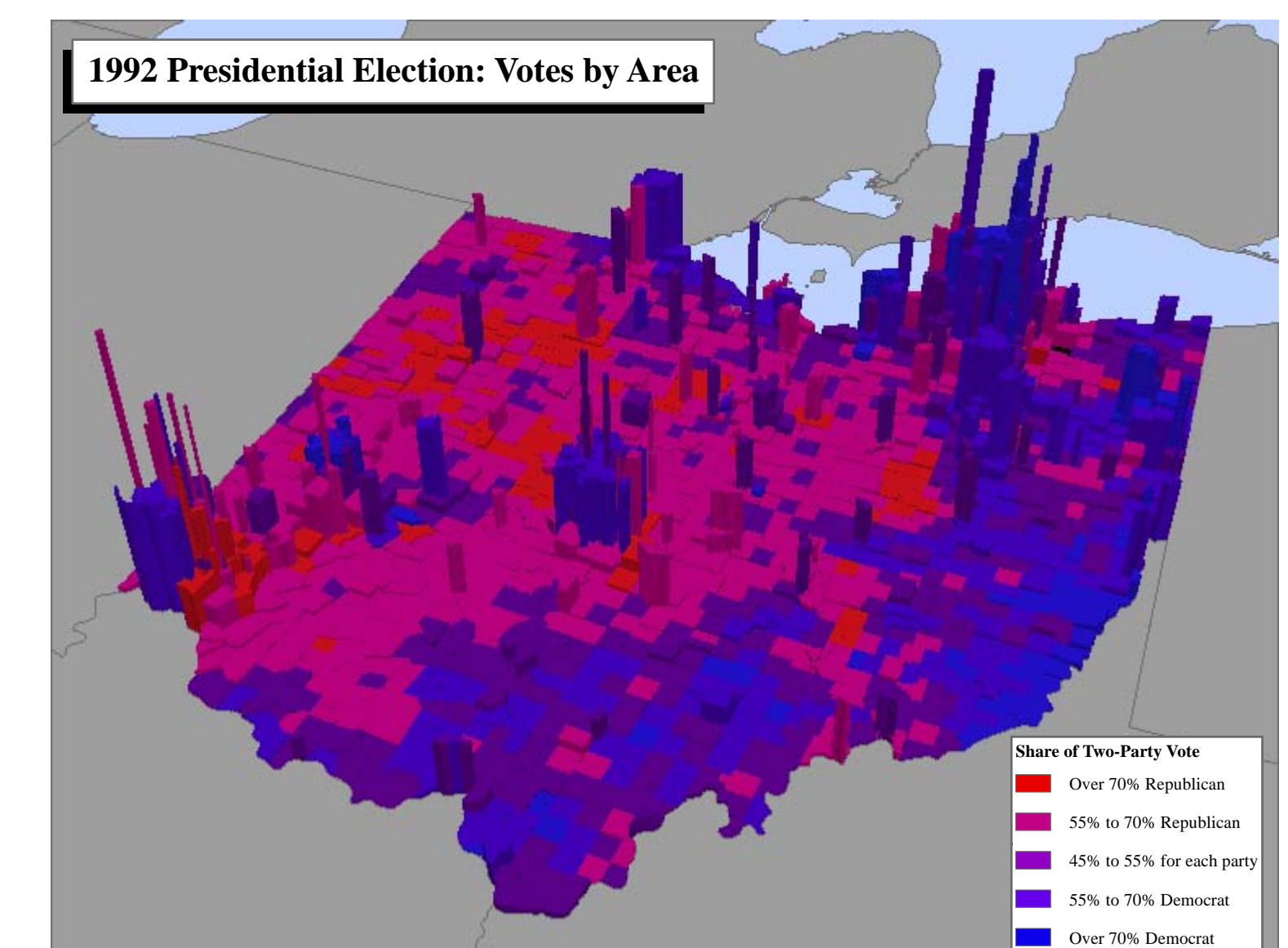
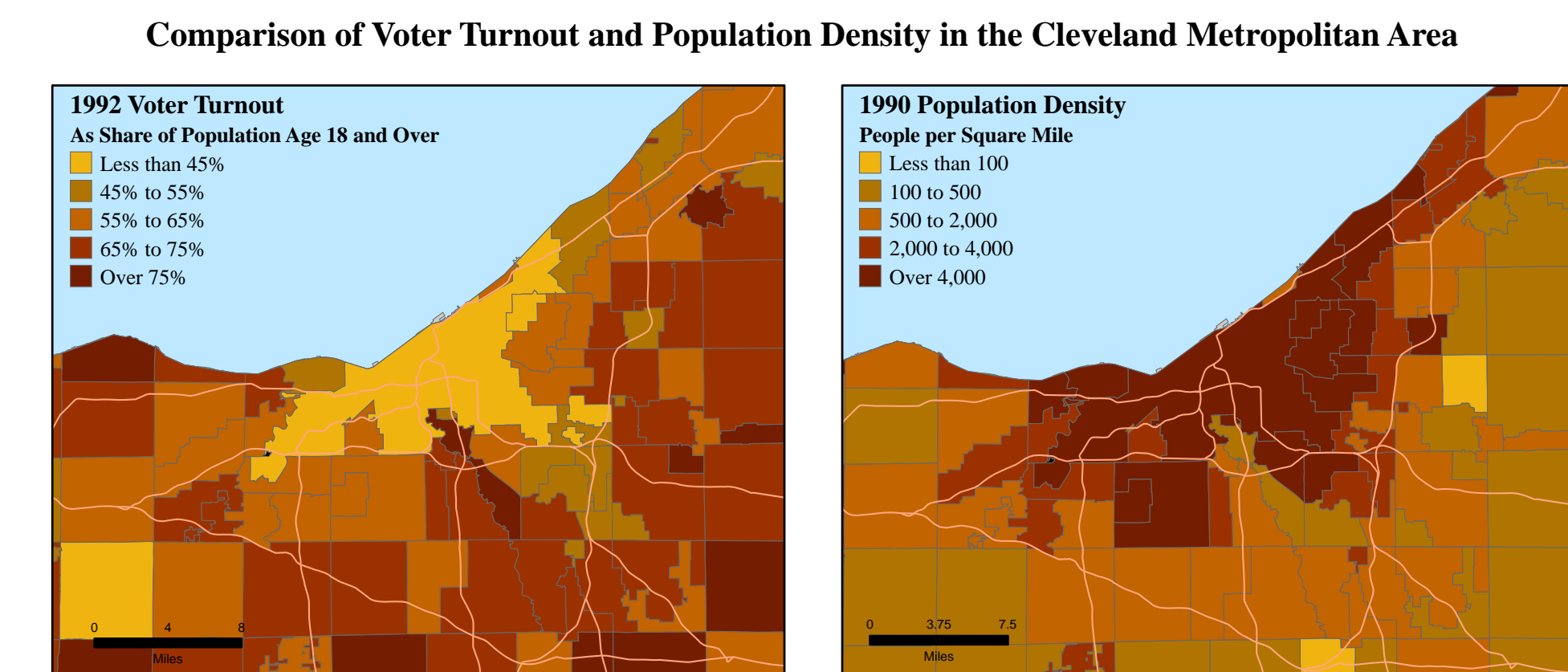


Figure 3. Performing Geographic Analysis of ROAD Data
 The ROAD data collection is primarily count data from the 1990 Census and elections. In order to perform effective geographic analysis, additional fields for normalization and percentage calculation must be created.

