A GIS-based investigation of the association between median household income and alcohol-sales locations in Pulaski County, Arkansas

Don Dailey

GEOG 6380: GIS Research Methodology

December 14, 2012

Problem Statement

Studies have shown in some areas of the United States that alcohol-selling establishments are disproportionately located in low-income and predominately African-American census tracts (Romley, Cohen, Ringel, & Sturmhttp, 2007). A high density of alcohol-selling outlets has been linked to a spectrum of negative effects from increased levels of violent crime to increased littering (Maxwell & Immergluck, 1997). The purpose of this study was to determine whether the location of alcohol-selling establishments is associated with median household income in Pulaski County, Arkansas. This study illuminated a pattern in the use of alcohol and the data could be used to inform public policy decisions that could control where alcohol is sold in the county. Public policy based on this data could be employed to control alcohol overuse in an effort to improve living conditions in those areas with higher density of alcohol outlets.

Specific objectives

Using the software functionality of ArcGIS, the locations of alcohol-sales permits were mapped and placed in the context of median household income in the county. The study determined the relationship between median household income and alcohol-sales locations. In addition, the study determined the number of people per alcohol permit in each of the Census tracts in the county.

Study area/Data layers

Pulaski County, Arkansas, was chosen as the study area because the county is the most populous and the most urban in the state (Figure 1). The state capital of Little Rock, the state's largest city, is located there. The county contains a full range of alcohol-selling establishments,

liquor stores, restaurants, bars, entertainment venues, etc. The county also is home to a full range of economic conditions. Additionally the county contains rural, suburban and urban areas.

Studying a larger area or the entire state was problematic because many counties in the state are completely dry, meaning no alcohol sales are permitted. Many other counties allow alcohol sales on a limited basis. These limits are very strict in some counties and quite liberal in others. To account for these differences to produce a valid statewide study was beyond the scope of this project.

The list of data layers used or created during the course of the study included:

- Address of every alcohol-sales establishment
- Census tract boundaries
- Census derived median household income data
- Ordinary least squares regression calculations
- Geographically weighted regression calculations
- Spatial autocorrelation calculations
- Alcohol permits per population

Methods, techniques, assumptions

Data on the location of alcohol sales and household income in Pulaski County were subjected to a series of statistical analyses to determine whether the level of household income predicts the location of alcohol permits. For the sake of brevity, familiarity with ESRI's ArcGIS geographic information systems software will be assumed.

In order to map the locations of all alcohol-sales permits, the permit addresses had to be assigned a location by using the address locator functionality of the software. The locator was provided by Pulaski Area GIS, the agency charged with creating and maintaining spatial data for

the various governments in the county. The locator contains the spatial parameters of assigned addresses in the county. With spatial data attached to the alcohol sales locations, the locations were mapped (Figure 2). Of the 849 alcohol permits active in the county, 842 were geocoded.

The county's Census tracts were added to the map (Figure 3). Census tracts are the smallest division for which population data, including median household income, is provided. A demographic table showing median household income for each tract was joined to the census block spatial data. This produced an attribute table for the spatial data that also showed the median household income in each block (Figure 4).

The next step was to associate each alcohol license location with the tract in which the license was located. A one-to-many location join was performed that added a column to each census block record showing the total number of liquor licenses in each Census tract.

A column was manually added to the attribute table and, using the field calculator, the number of people per alcohol permit was calculated for each tract. This gave an idea of the accessibility of alcohol-sales locations (Figure 5). Some tracts were left blank because they contained no alcohol permits. While this data gives a broad overview of the density of licenses, it's not very useful to this study's analysis.

Putting all the pertinent data into one data table gave the ability to use the software's statistical-analysis functions on the data. The purpose of using the statistical analysis was to show whether some demographic variable predicted the license locations. In this case the variable was median household income. For the sake of brevity, familiarity with spatial statistics will be assumed. A spatial autocorrelation test and two types of regression analyses were applied to the data. Spatial autocorrelation is the clustering of spatial data, meaning that something is located where it is because a similar something is nearby (Clark & Hosking, 1986). In the case of

the alcohol licenses, if spatial autocorrelation were found, it would indicate the locations of the licenses were linked in some way. Using the spatial autocorrelation functionality of the spatial statistics tools in ArcGIS, the Moran's I statistic on the count of liquor license locations in each tract was calculated.

Following the spatial autocorrelation test, a regression analysis was performed. Regression analysis attempts to provide answers on why a particular variable is what it is (Clark & Hosking, 1986). In the case of the liquor licenses, the knowledge sought was whether the location of liquor licenses is predicted by the median household incomes of the population of Pulaski County. By using ArcGIS software, the results of the regression analysis can be easily computed and then mapped. The liquor license data was put through two types of regression tools in the software: ordinary least squares, which is a globally even application, and geographically weighted regression, which is a localized regression technique that often works better for spatially autocorrelated data.

Results

Spatial autocorrelation: The software calculations determined that spatial autocorrelation was present in alcohol license locations (Figure 6). The locations were linked to each other and not randomly distributed through the county.

Regression: Using the ordinary least squares regression technique, the number of alcohol license locations per Census tract was identified as the dependent variable and median household income the explanatory variable. The OLS regression estimates what the value of the dependent variable should be based on the explanatory variable value and compares that estimate to the actual values observed in the dependent variable. The software returns several statistics that define how well the explanatory variable fits the dependent variable. In this case the explanatory

variable unambiguously did not fit the dependent. In other words, median household income did not predict alcohol license location using this regression technique. A map of the standard deviation of the residuals gives an idea where the alcohol permits are concentrated, which is in the tracts where the model underpredicted the number of permits by more than 2.5 standard deviations (Figure 7). Additionally, determining the Moran I statistic for the residuals from the OLS regression showed the residuals (the difference between the observed number of liquor licenses in each block and the predicted number of licenses) to be spatially autocorrelated, an indicator that the explanatory variable doesn't explain the dependent variable.

The geographically weighted regression, which often works better for spatial data because such data is often spatially autocorrelated also showed that alcohol license location was not associated with median household income. The map of the standard deviation of the residual values produced by this technique looked very similar to the ordinary least squares derived map, but it showed even more clearly how alcohol permits are spatially autocorrelated. Likewise, the Moran's I statistic on the residuals showed them to be spatially autocorrelated, again an indicator that the explanatory variable doesn't explain the independent variable.

Conclusions

Based on the results of the statistical analysis of alcohol license locations in Pulaski County, it is clear the licenses aren't disproportionately located in areas with low median household income. Indeed, median household income has no effect on the location of the licenses. However, the analysis does show that license locations are highly spatially autocorrelated. Alcohol license locations are spatially linked, but why? Additional statistical analysis would have to be undertaken to pinpoint the cause of the autocorrelation. I think it most likely that the

level of commercial activity and vehicle traffic in a given Census tract predicts the location of liquor licenses.

On a broad scale it's easy to see by looking at the mapped raw location data and the mapped regression analysis data that more licenses are located in the urban areas of the county. On a more narrow scale, the tracts within the urban areas with the highest concentration of licenses, downtown Little Rock and a portion of west Little Rock, are areas with very high levels of vehicle traffic and/or commercial activity. Those areas are shown in Figures 7 and 8 as the tracts with residuals more than 2.5 standard deviations from the mean. Downtown Little Rock is home to a major entertainment district that draws people from all over the county to eateries and drinking establishments. West Little Rock hosts many shopping and dining establishments and also draws clientele from a wide area.

Limitations

This study examined a single explanatory variable in an effort to determine if that variable could be ruled out as a predictor of alcohol license locations. Ideally, a researcher would examine many explanatory variables until at least one fit the regression models and could be considered a true explanation for the permit locations. I feel it would be worthwhile to gather data measuring levels of human traffic, consumer spending, number of shoppers in the stores, traffic counts on streets, etc. and applying the same statistical analysis to see if there is association of those factors with liquor license locations.

Additionally, this study was confined to the total number of outlets permitted to sell alcohol. A finer-grained study could be conducted by testing the various permit types to see if different demographic variables predict the location of different types of permits. For example, it would be worthwhile to use only those locations permitted for retail sale of beer and alcohol for off-

premises consumption and test those data against the household income data. This would limit the study to mainly liquor stores, which studies have shown elsewhere are disproportionately located in predominately low-income areas (LaVeist, T., & J. Wallace, 2000).

References

LaVeist, T., & J. Wallace. (2000). Health risk and inequitable distribution of liquor stores in African American neighborhoods. *Social Science and Medicine*, 51(4), 613–617.

Maxwell, A., & D. Immergluck. (1997). Liquorlining: liquor store concentration and community development in lower-income Cook County (IL) neighborhoods. Chicago, IL: Woodstock Institute.

Romley, J., Cohen, D., Ringel, J., & Sturmhttp, R. (2007). Alcohol and Environmental Justice: The density of liquor stores and bars in urban neighborhoods in the United States. *Journal of Studies on Alcohol and Drugs*. 68(1): 48–55.

Bluthenthal, R., Cohen, D., Farley, T., Scribner, R., Beighley, C., Schonlau, M., & Robinson, P. (2008). Alcohol Availability and Neighborhood Characteristics in Los Angeles, California and Southern Louisiana. *The Journal of Urban Health*. 85(2): 191-205

Clark, W. A. V., & Hosking, C. L. (1986). *Statistical Methods for Geographers*. New York City, NY: John Wiley & Sons, Inc.

Burt, J. E., Barber, G., & Rigby, D. (2009). *Elementary Statistics for Geographers* (3rd. ed.). New York City, NY: The Guilford Press.

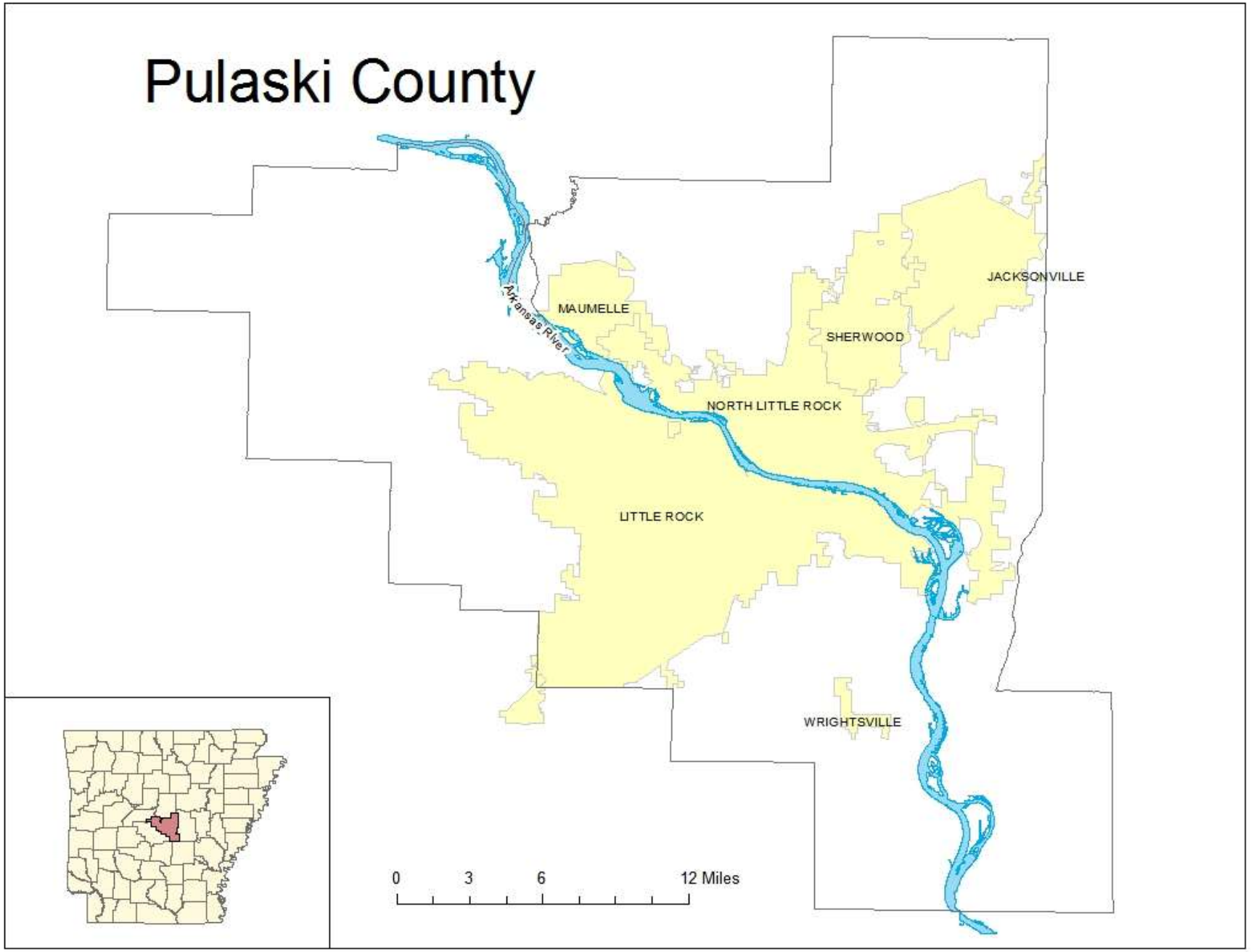

Figure 1. Source: Arkansas Geographic Information Office

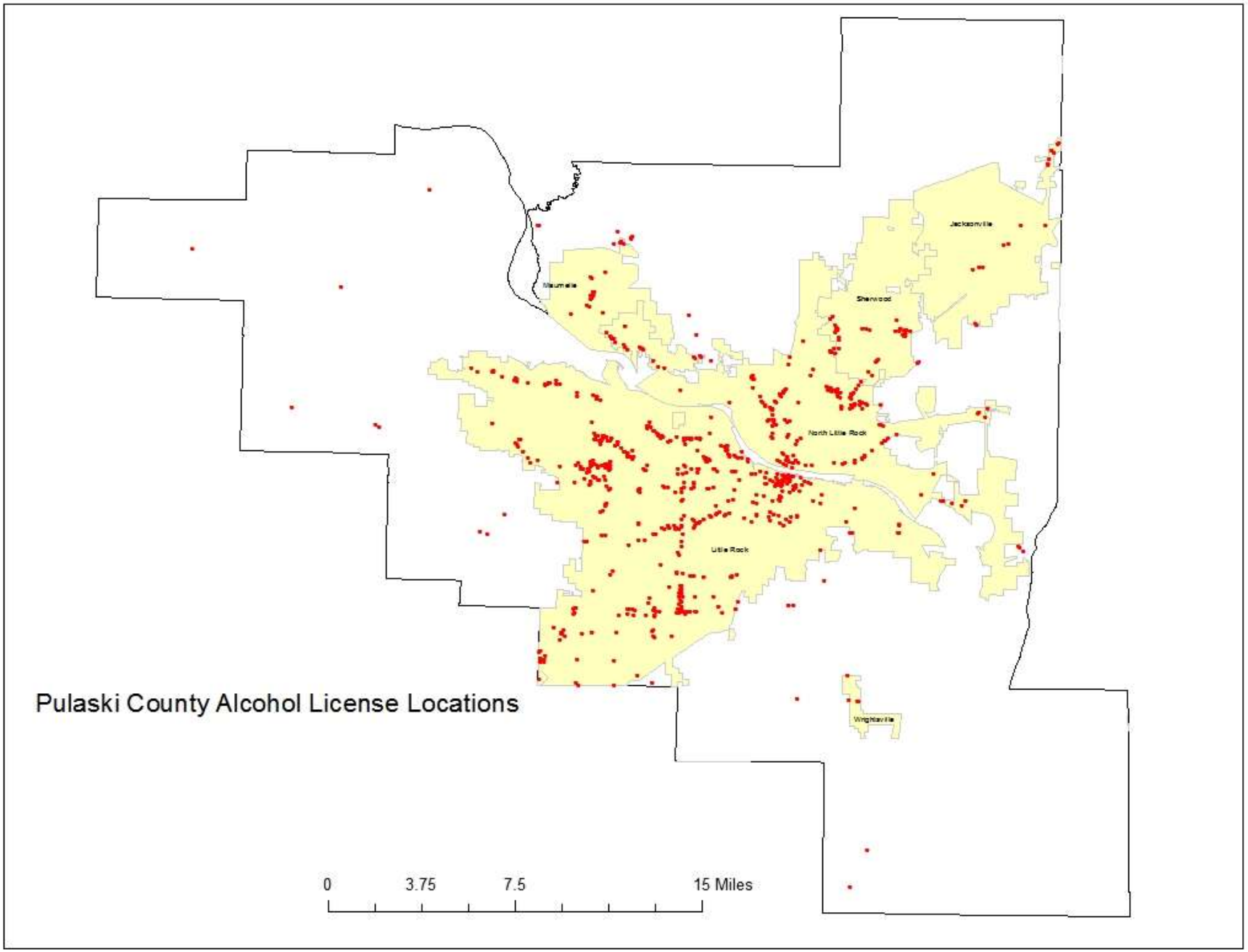

Figure 2. Source: Arkansas Geographic Information Office; Arkansas Alcoholic Beverage Control

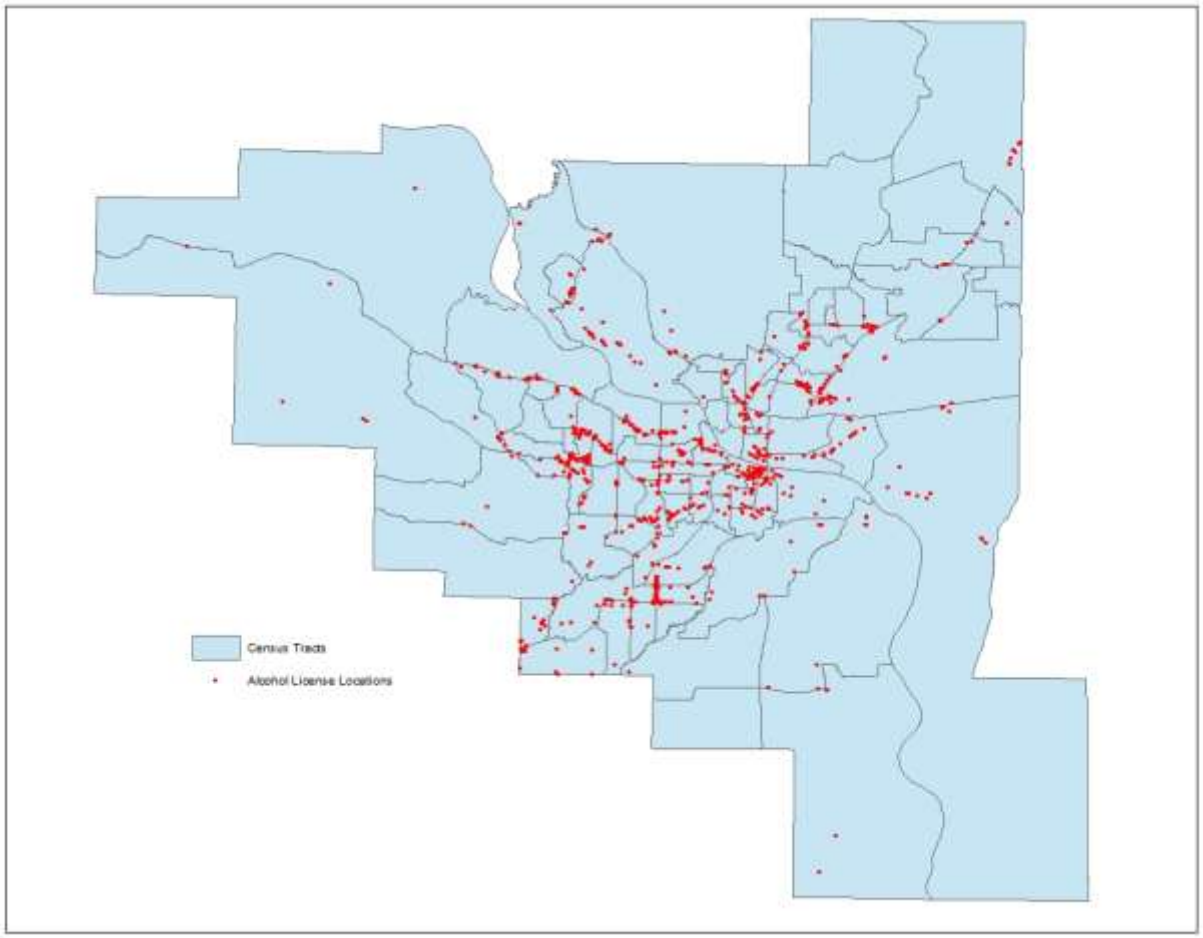

Figure 3. Source: U.S. Census Bureau; Arkansas Alcoholic Beverage Control

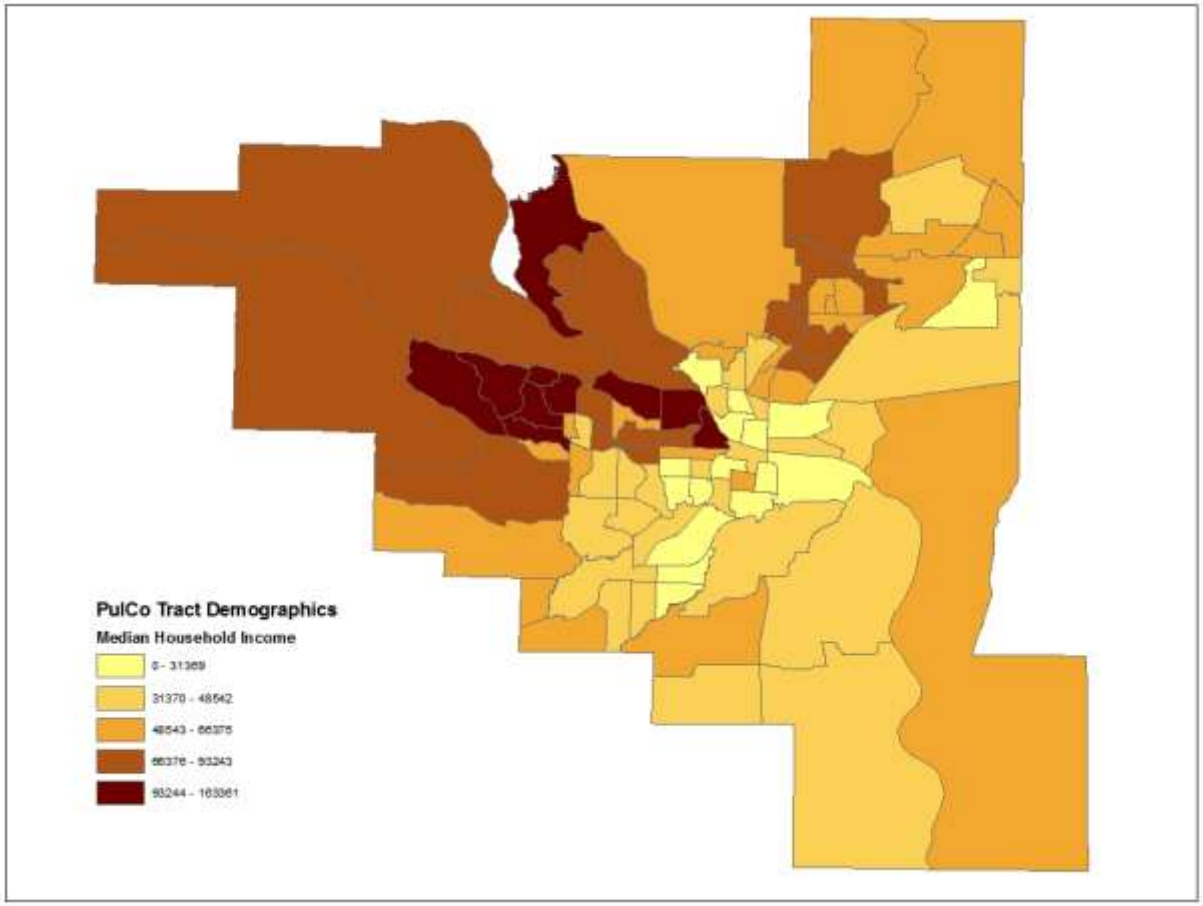

Figure 4. Source: U.S. Census Bureau

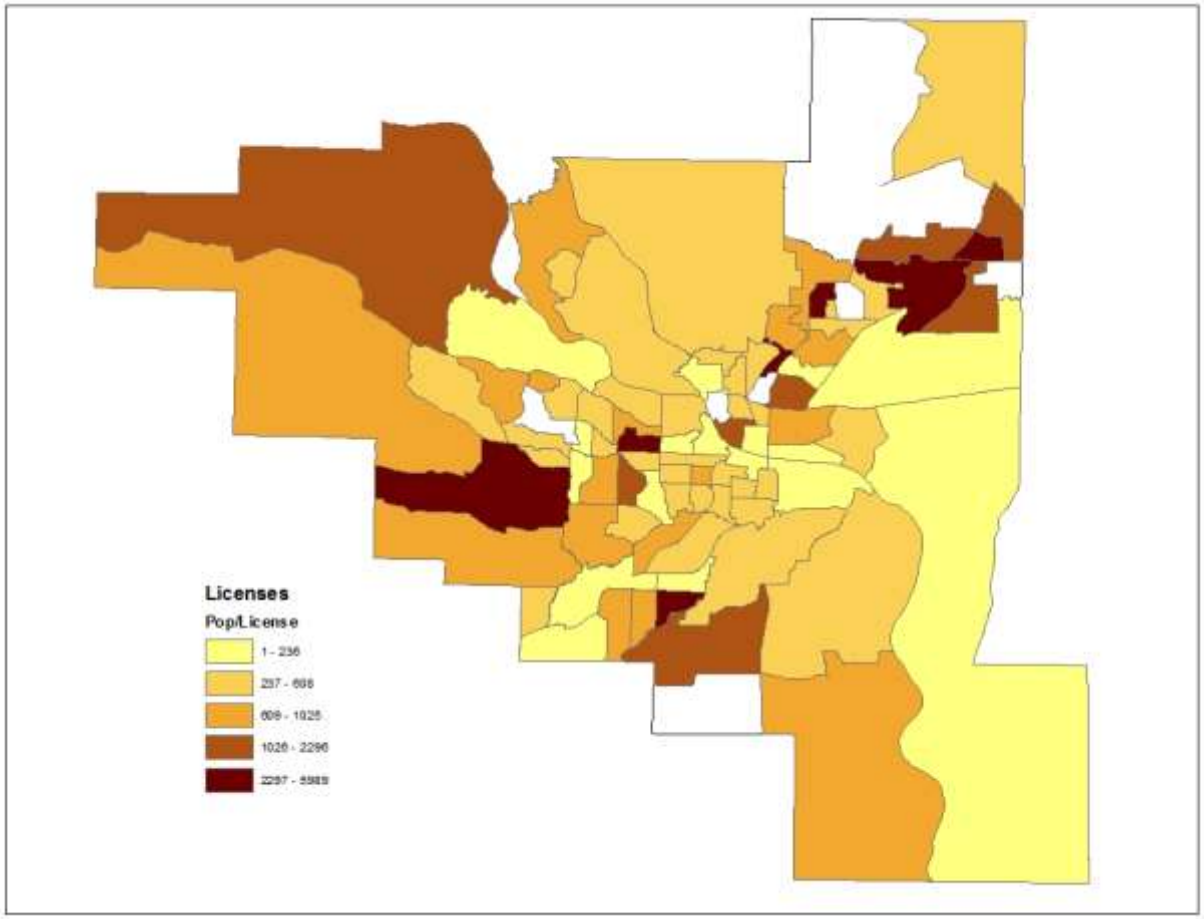

Figure 5. Source: U.S. Census Bureau; Arkansas Alcoholic Beverage Control

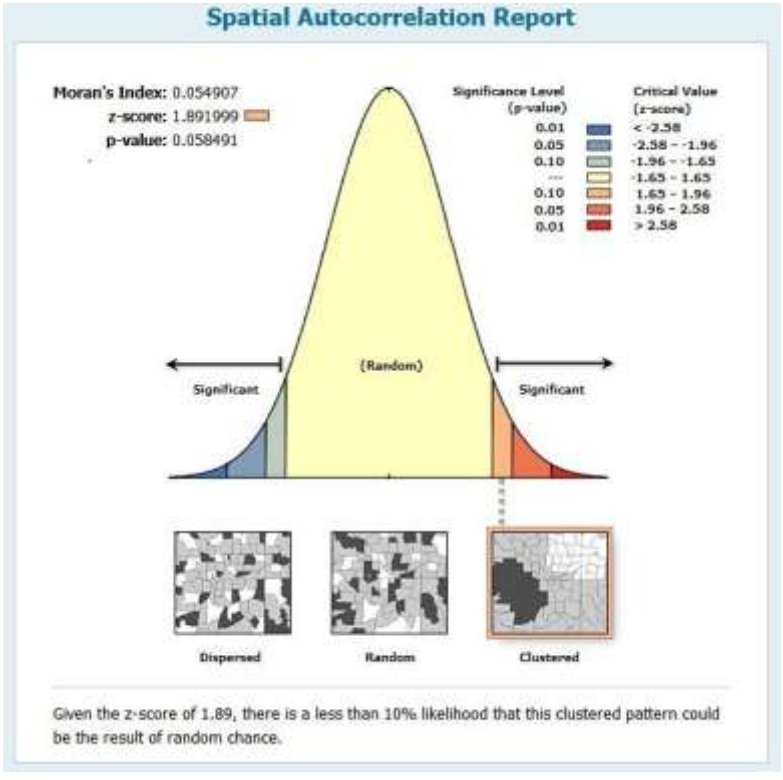

Figure 6.

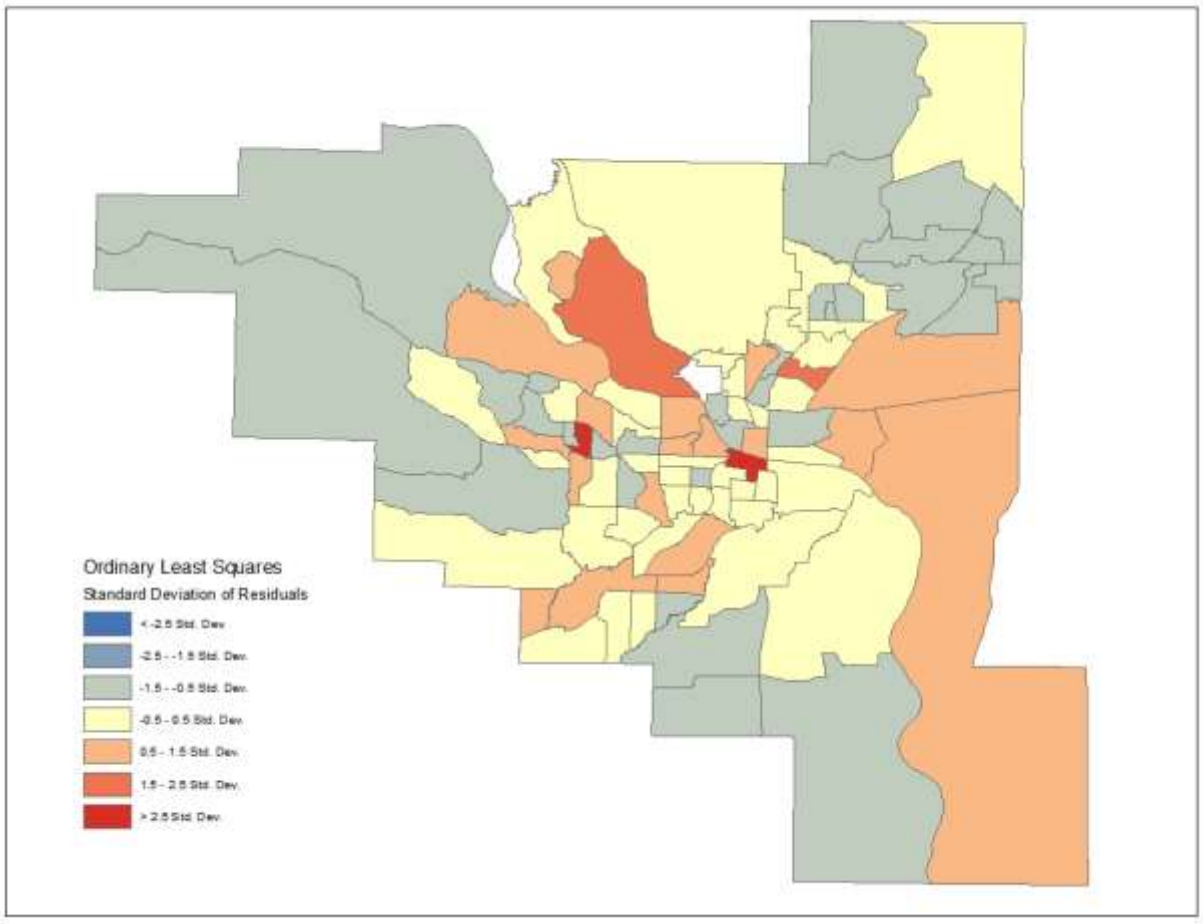

Figure 7. Source: U.S. Census Bureau; Arkansas Alcoholic Beverage Control

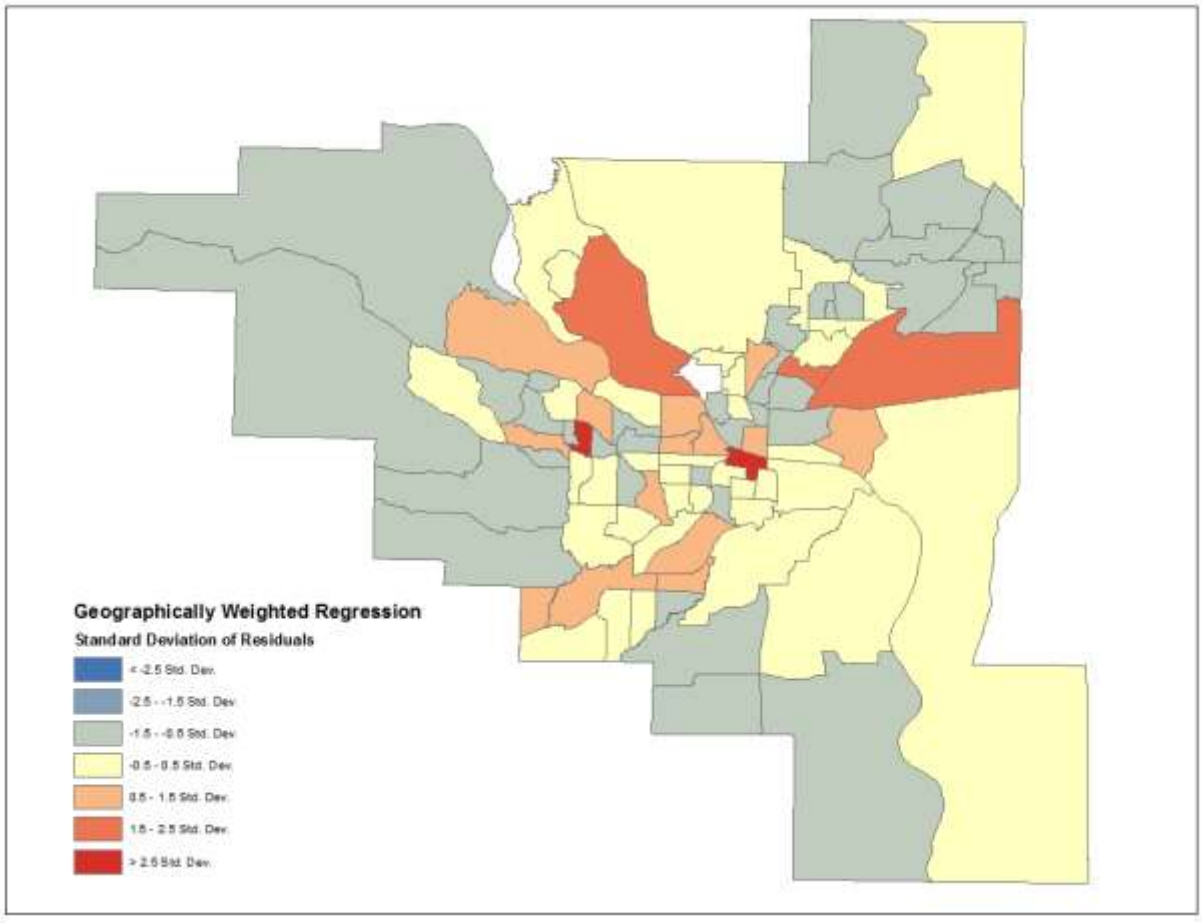

Figure 8. Source: U.S. Census Bureau; Arkansas Alcoholic Beverage Control