Multiple Weather Station Analysis

Process

- We constructed econometric formulations for a few states with multiple weather stations using the data from last year
  - this provides for a comparison to last year's models
- This resulted in problems with model coefficients of the wrong sign and with variables that were not statistically significant

2014 TX Model (San Antonio)

Dependent Variable: ELECTRICITY_SALES
Method: Least Squares
Date: 03/13/15 Time: 15:42
Sample: 1990 2012
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	30527.38	33755.29	0.904373	0.3784
@MOVAV(REAL_ELECTRICITY_PRICE,5)	-4641.131	1375.351	-3.374506	0.0036
REAL_INCOME/POPULATION	3515.000	1454.605	2.416465	0.0272
REAL_GSP	0.145494	0.026098	5.574822	0.0000
CDD	16.16171	3.823118	4.227363	0.0006
HDD	20.84670	5.074099	4.108454	0.0007

R-squared	0.993627	Mean dependent var	307551.8
Adjusted R-squared	0.991752	S.D. dependent var	43037.33
S.E. of regression	3908.547	Akaike info criterion	19.59918
Sum squared resid	2.60E+08	Schwarz criterion	19.89539
Log likelihood	-219.3905	Hannan-Quinn criter.	19.67367
F-statistic	530.0731	Durbin-Watson stat	1.629527
Prob(F-statistic)	0.000000		

2014 TX (multiple stations)

Dependent Variable: ELECTRICITY_SALES
Method: Least Squares
Sample: 1990 2012
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-9940.844	34462.99	-0.288450	0.7784
@MOVAV(REAL_ELECTRICITY_PRICE,5)	-3619.431	1297.105	-2.790392	0.0176
REAL_INCOME/POPULATION	5110.873	1386.709	3.685613	0.0036
REAL_GSP	0.105381	0.026246	4.015128	0.0020
DALLAS_TX_C	9.074900	5.438765	1.668559	0.1234
DALLAS_TX_H	-5.838066	8.363243	-0.698063	0.4996
HOUSTON_TX_C	7.655999	6.386797	1.198723	0.2558
HOUSTON_TX_H	28.83526	12.32312	2.339931	0.0392
LUBBOCK_TX__C	2.291983	7.991036	0.286819	0.7796
LUBBOCK_TX__H	3.467606	7.266146	0.477228	0.6425
SAN_ANTONIO_TX_C	5.075477	6.074335	0.835561	0.4212
SAN_ANTONIO_TX_H	-4.522494	13.89882	-0.325387	0.7510

R-squared	0.997415	Mean dependent var	307551.8
Adjusted R-squared	0.994829	S.D. dependent var	43037.33
S.E. of regression	3094.724	Akaike info criterion	19.21866
Sum squared resid	1.05E+08	Schwarz criterion	19.81110
Log likelihood	-209.0146	Hannan-Quinn criter.	19.36766
F-statistic	385.7912	Durbin-Watson stat	1.950824
Prob(F-statistic)	0.000000		

HDD Correlations

	Dallas	Houston	Lubbock	San Antonio
Dallas	-	0.90	0.79	0.87
Houston	0.90	-	0.70	0.88
Lubbock	0.79	0.70	-	0.80
San Antonio	0.87	0.88	0.80	-

A Different Process

- We developed virtual weather stations using population-weighted CDD and HDD from multiple weather stations within the states
- Each state is divided into regions along county lines and a representative weather station is chosen for each region
- The CDD and HDD for a given year are determined using the population of each region and the CDD and HDD of the region's weather station

2014 TX (virtual station)

Dependent Variable: ELECTRICITY_SALES
Method: Least Squares
Date: 03/13/15 Time: 15:42
Sample: 1990 2012
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	18197.81	29241.15	0.622336	0.5420
@MOVAV(REAL_ELECTRICITY_PRICE,5)	-3792.856	1121.703	-3.381337	0.0035
REAL_INCOME/POPULATION	3605.803	1238.386	2.911695	0.0097
REAL_GSP	0.137162	0.022457	6.107792	0.0000
CDDMWS	20.02943	3.703811	5.407790	0.0000
HDDMWS	16.95795	3.897780	4.350669	0.0004

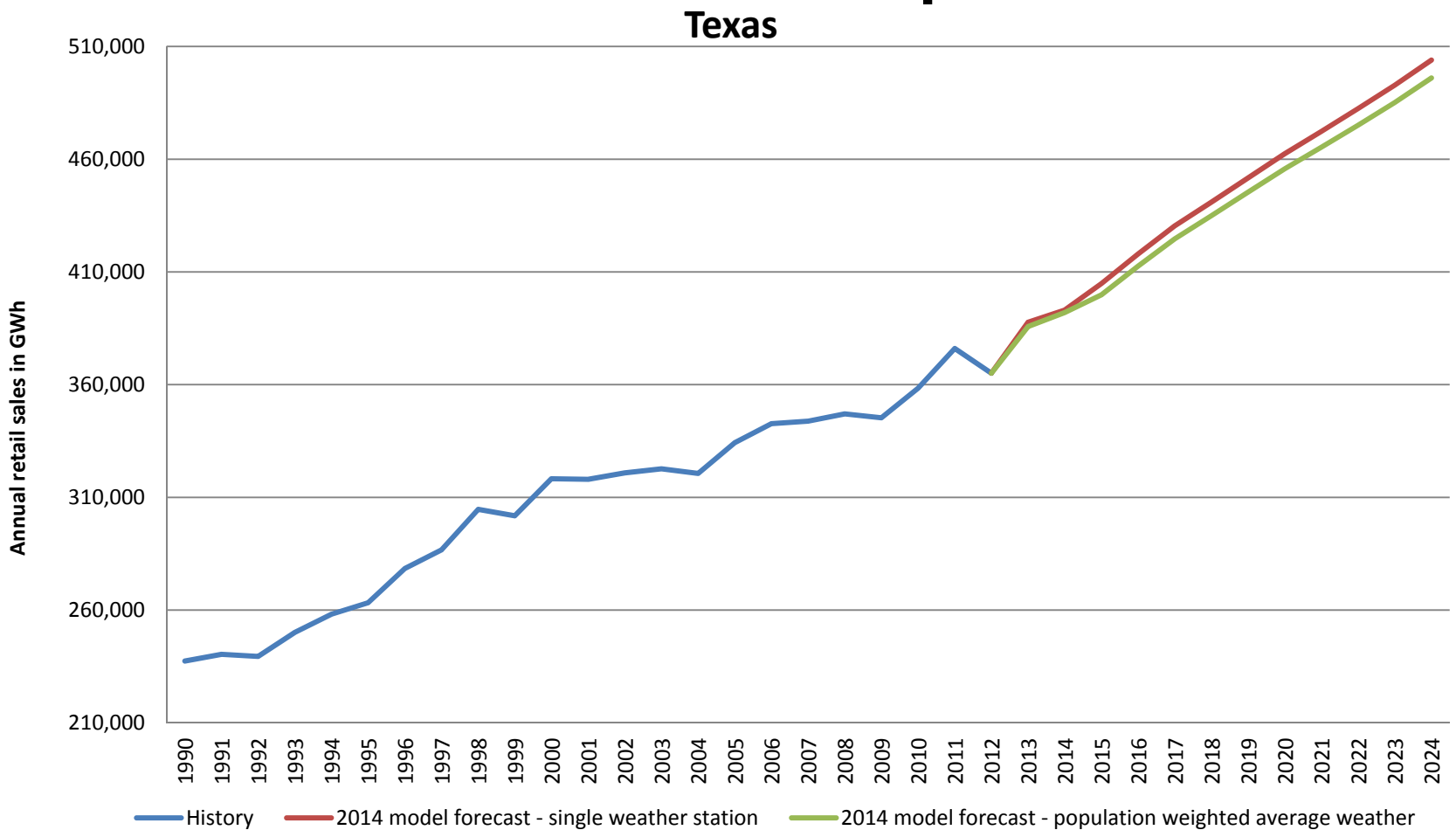
R-squared	0.995288	Mean dependent var	307551.8
Adjusted R-squared	0.993902	S.D. dependent var	43037.33
S.E. of regression	3360.723	Akaike info criterion	19.29716
Sum squared resid	1.92E+08	Schwarz criterion	19.59337
Log likelihood	-215.9173	Hannan-Quinn criter.	19.37166
F-statistic	718.1691	Durbin-Watson stat	1.598551
Prob(F-statistic)	0.000000		

Model Comparison

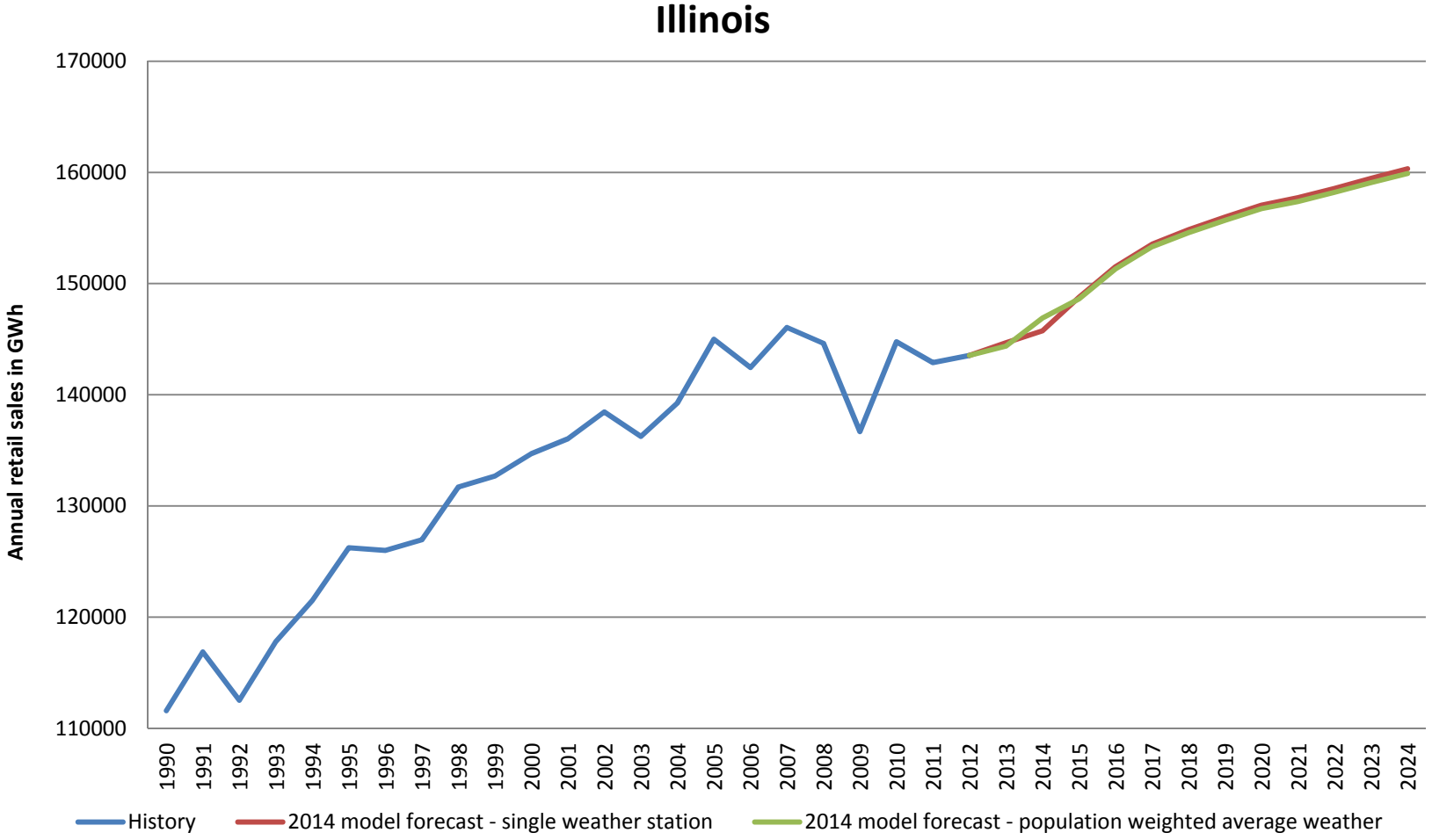
TX – Comparison across models

	Single weather station	Population-weighted average weather	Multiple weather stations
Adjusted R-squared	0.992	0.994	0.995
Prob(F-statistic)	0.000	0.000	0.000
Durbin-Watson stat	1.630	1.599	1.951
Mean absolute error	2685.298	2276.612	1696.387

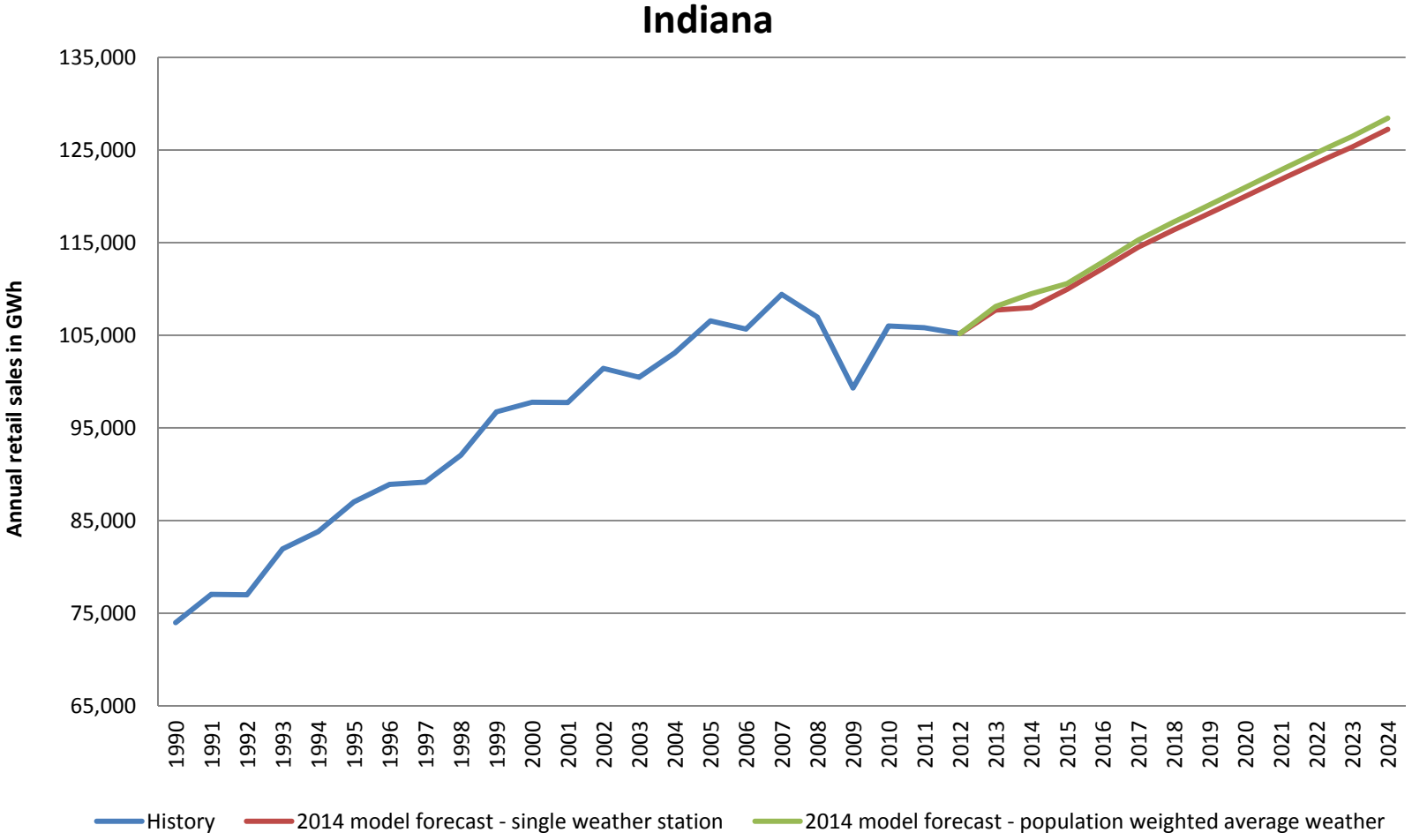
TX Model Comparison



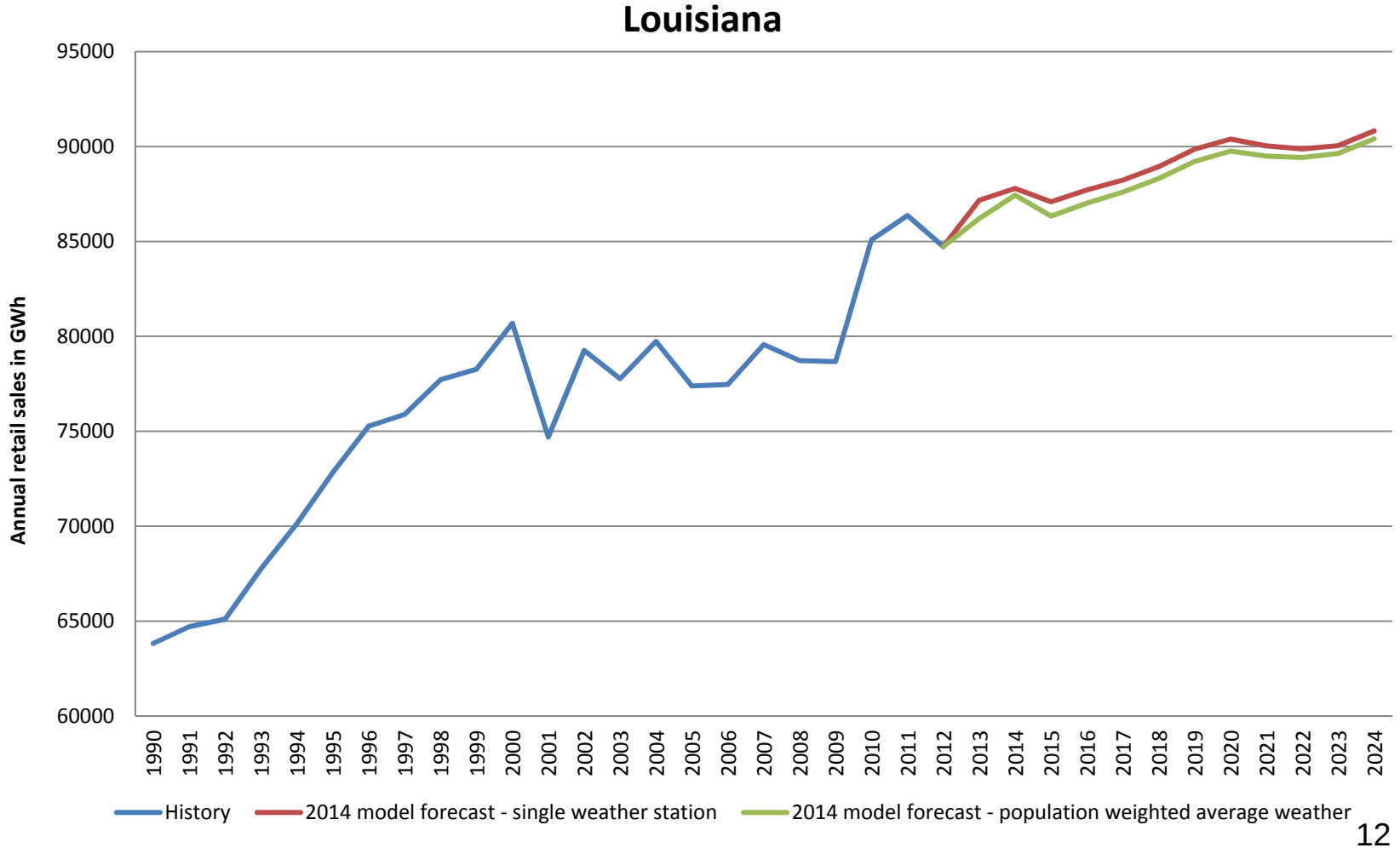
IL Model Comparison



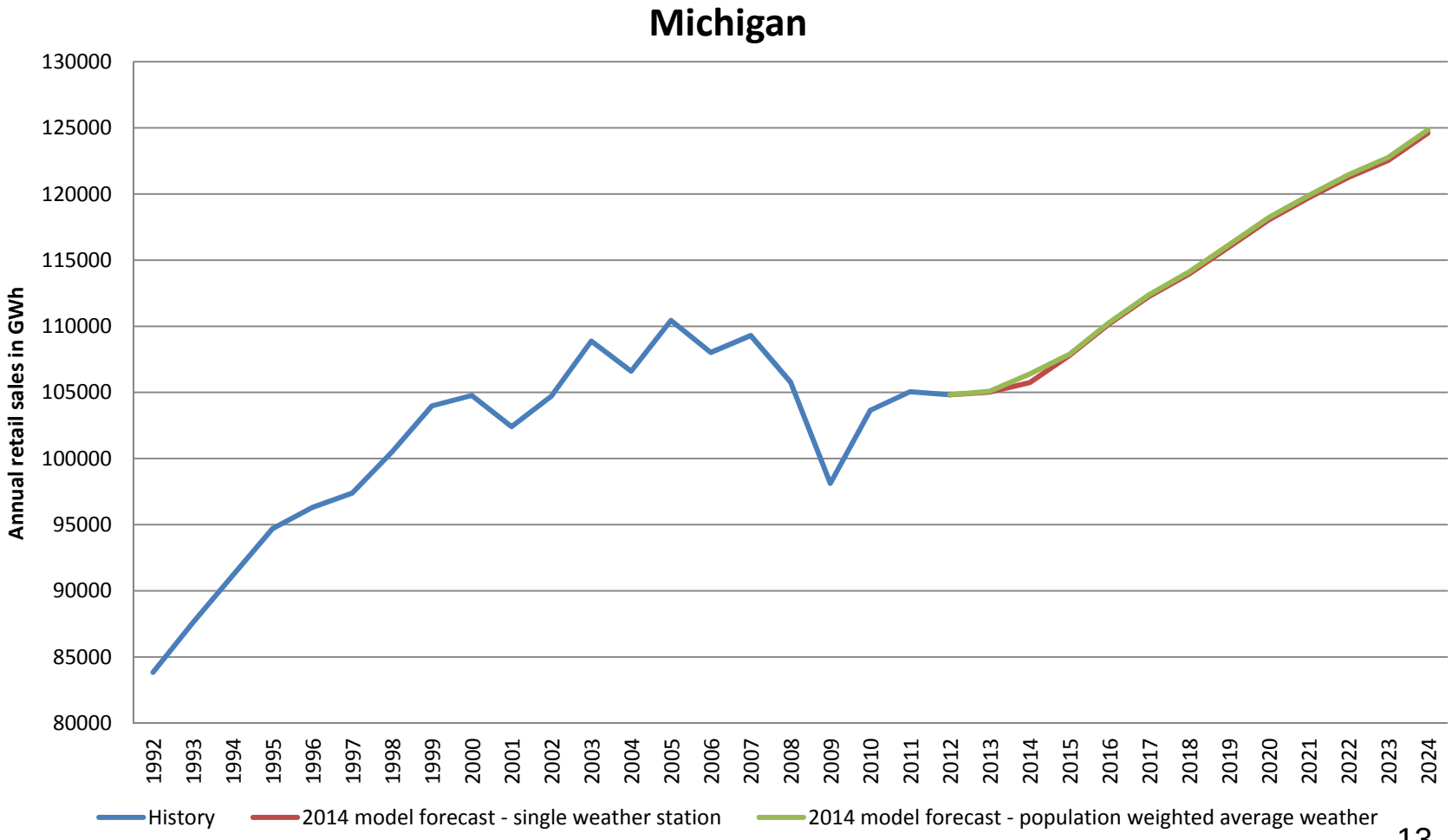
IN Model Comparison



LA Model Comparison



MI Model Comparison



Results

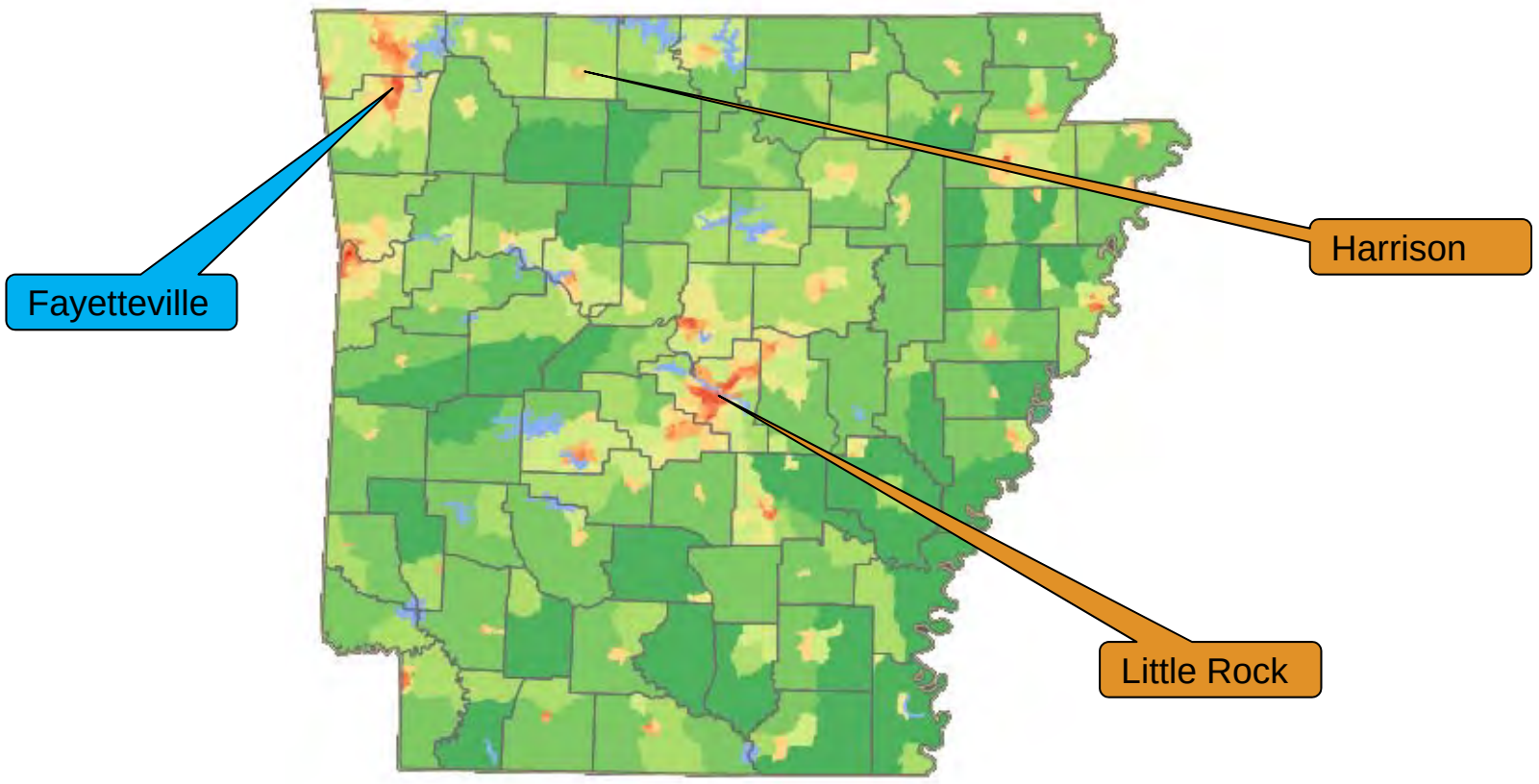
- We examined the possibility of using multiple weather stations (each as a separate variable) and found it infeasible
- We believe that the population-weighted virtual station works well but has little impact on the forecast results
- We propose to use the population-weighted virtual station method

Population Weighted Virtual Weather Stations

Orange indicates a weather station that is used

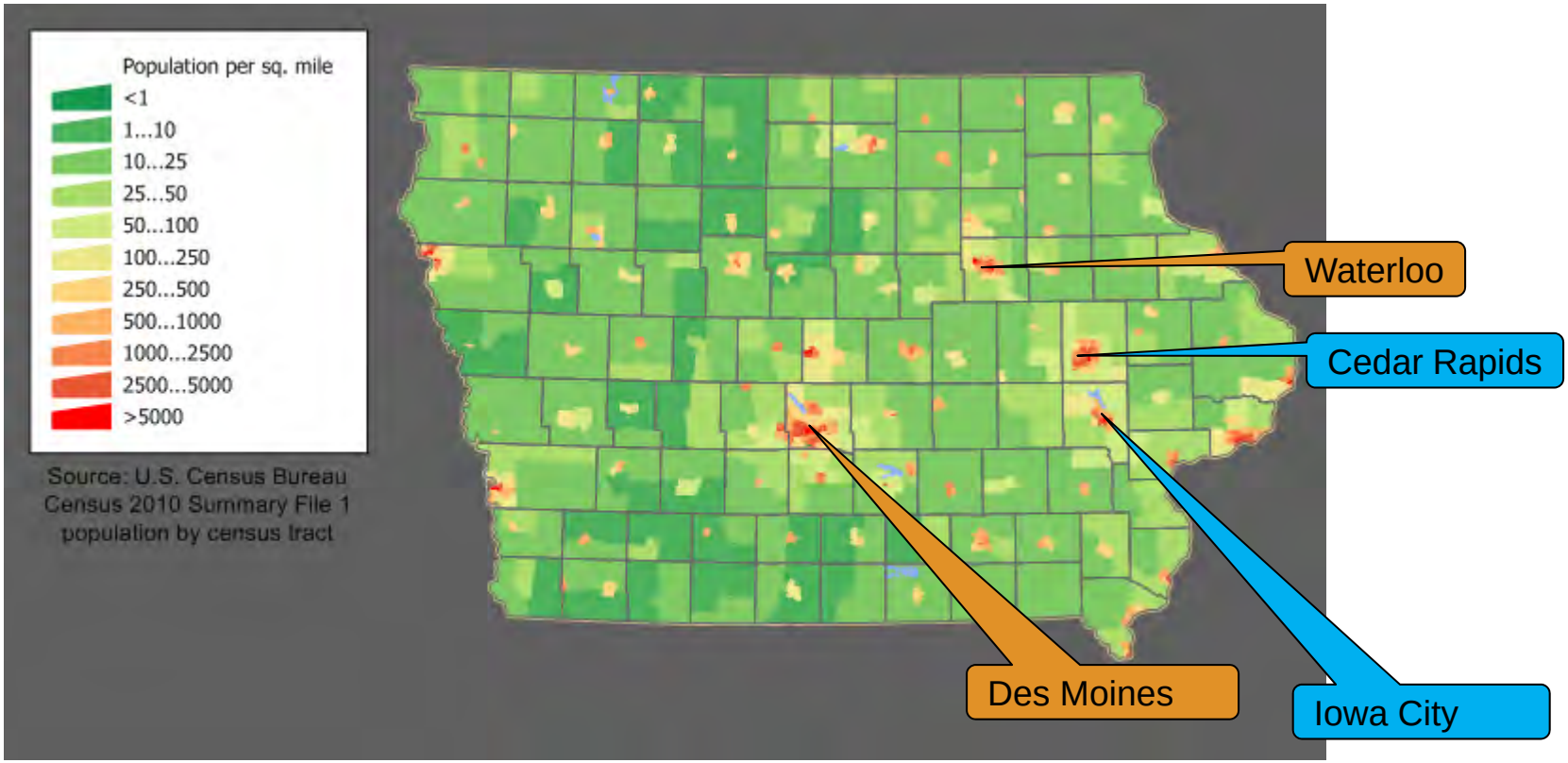
Blue indicates a weather station that could not be used because of data problems

Arkansas Population Density Map

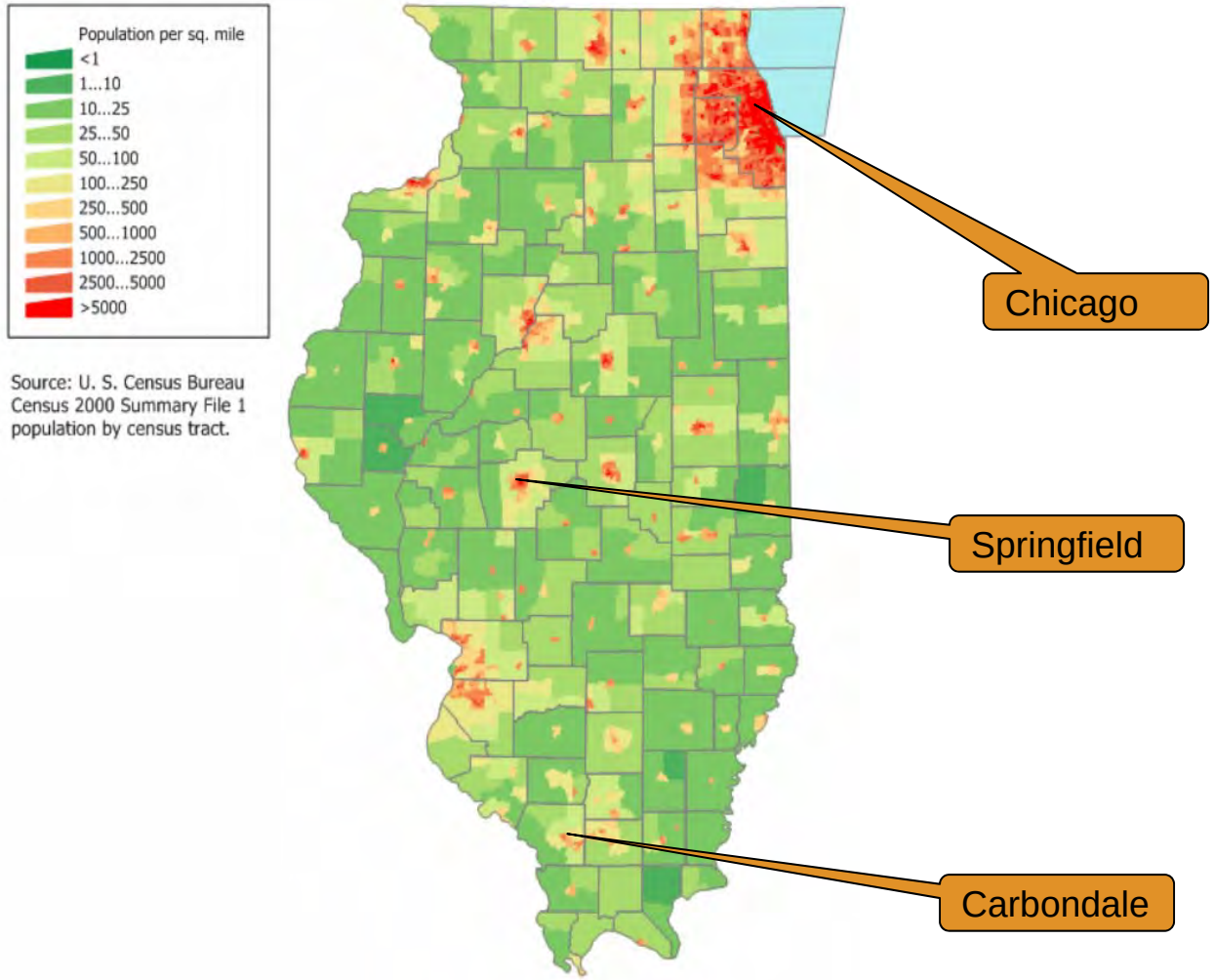


Source: US Census Bureau

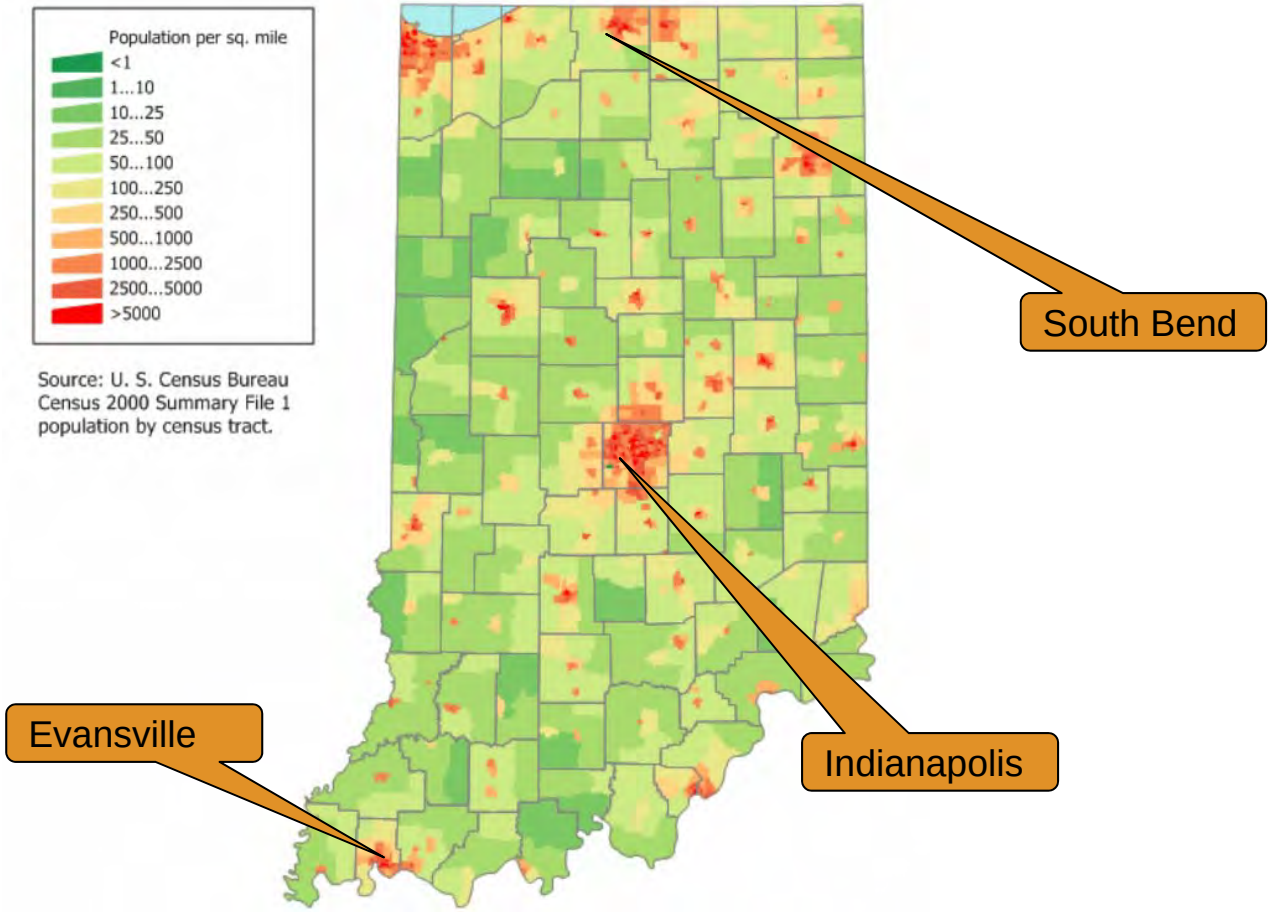
Iowa Population Density Map



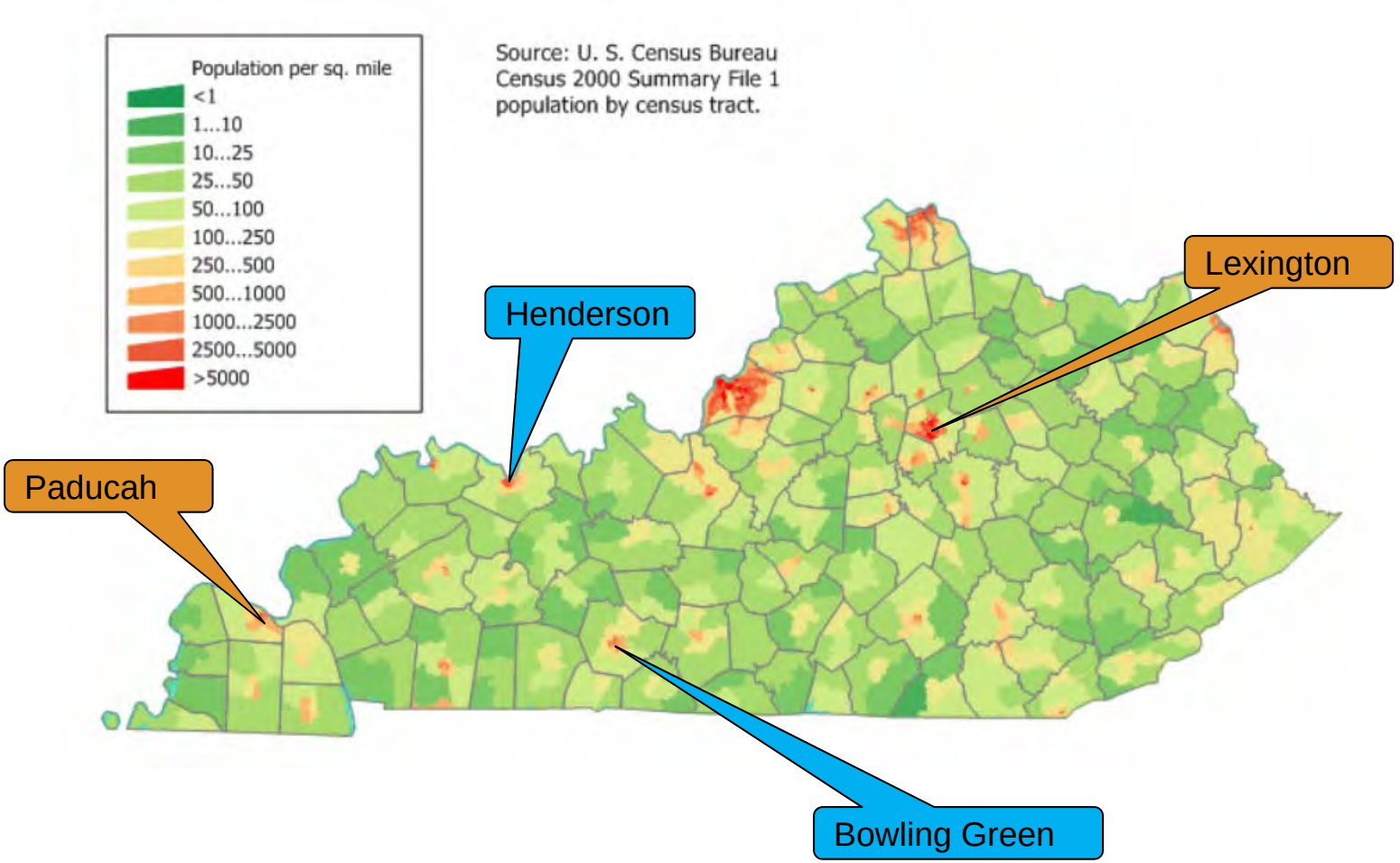
Illinois Population Density Map



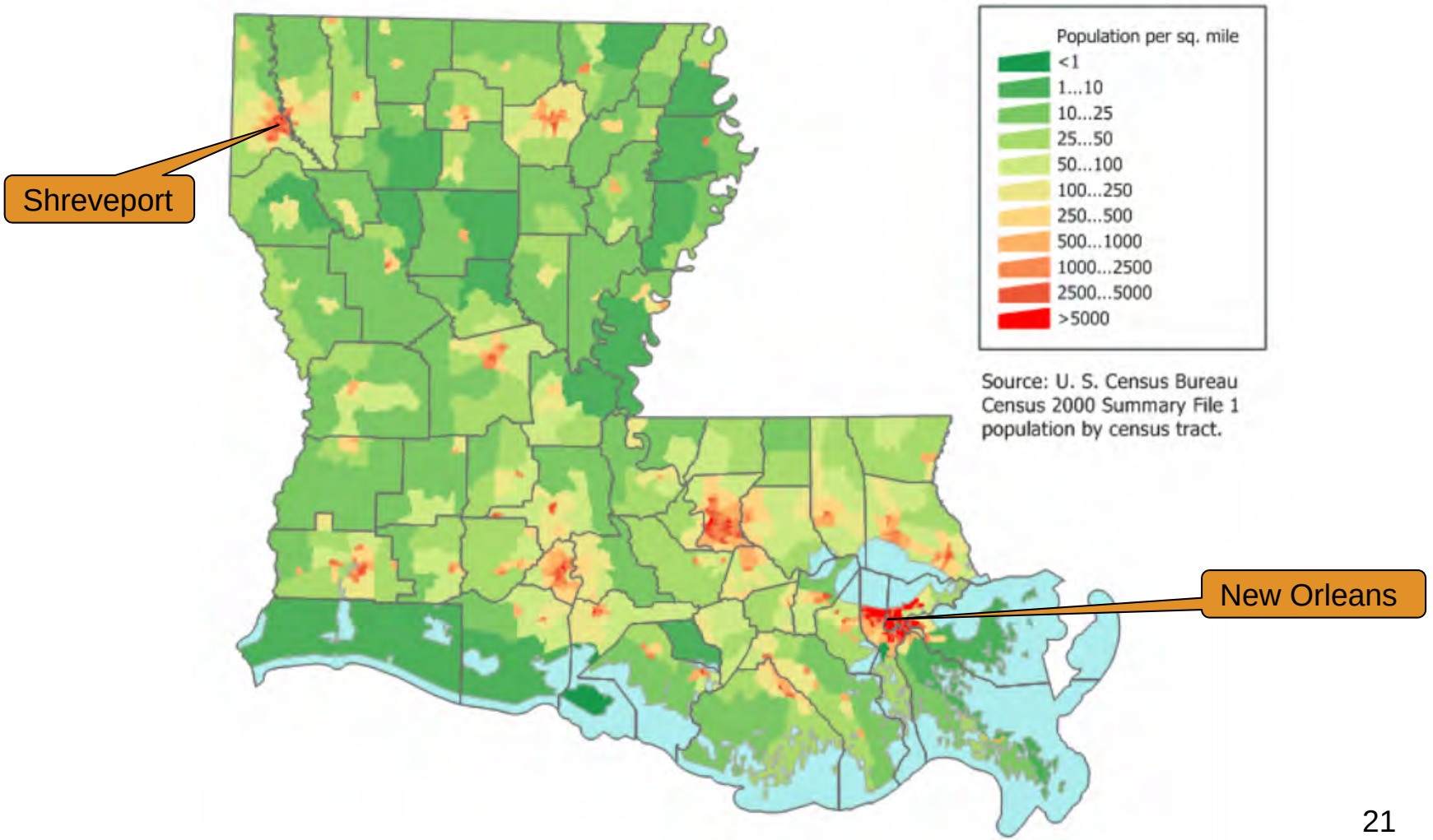
Indiana Population Density Map



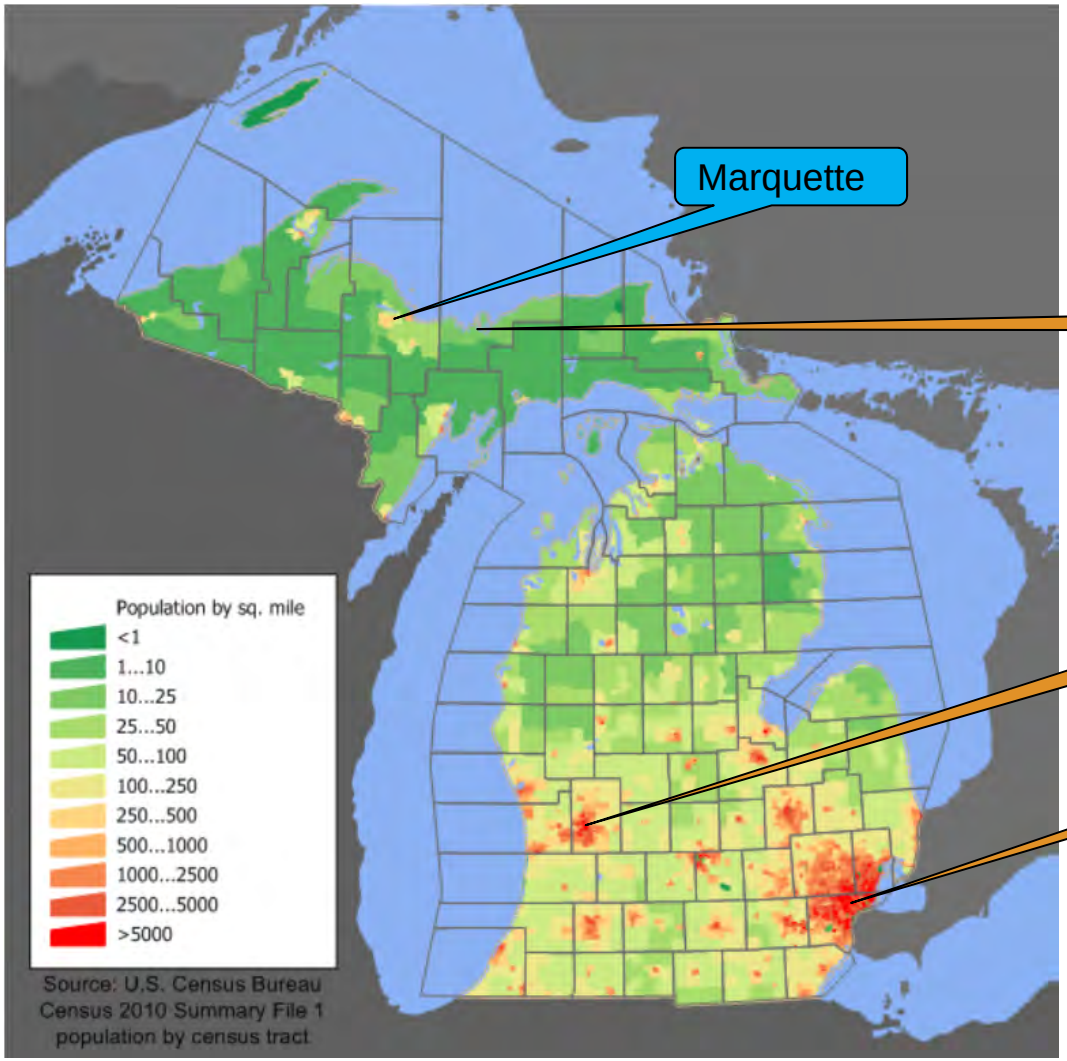
Kentucky Population Density Map



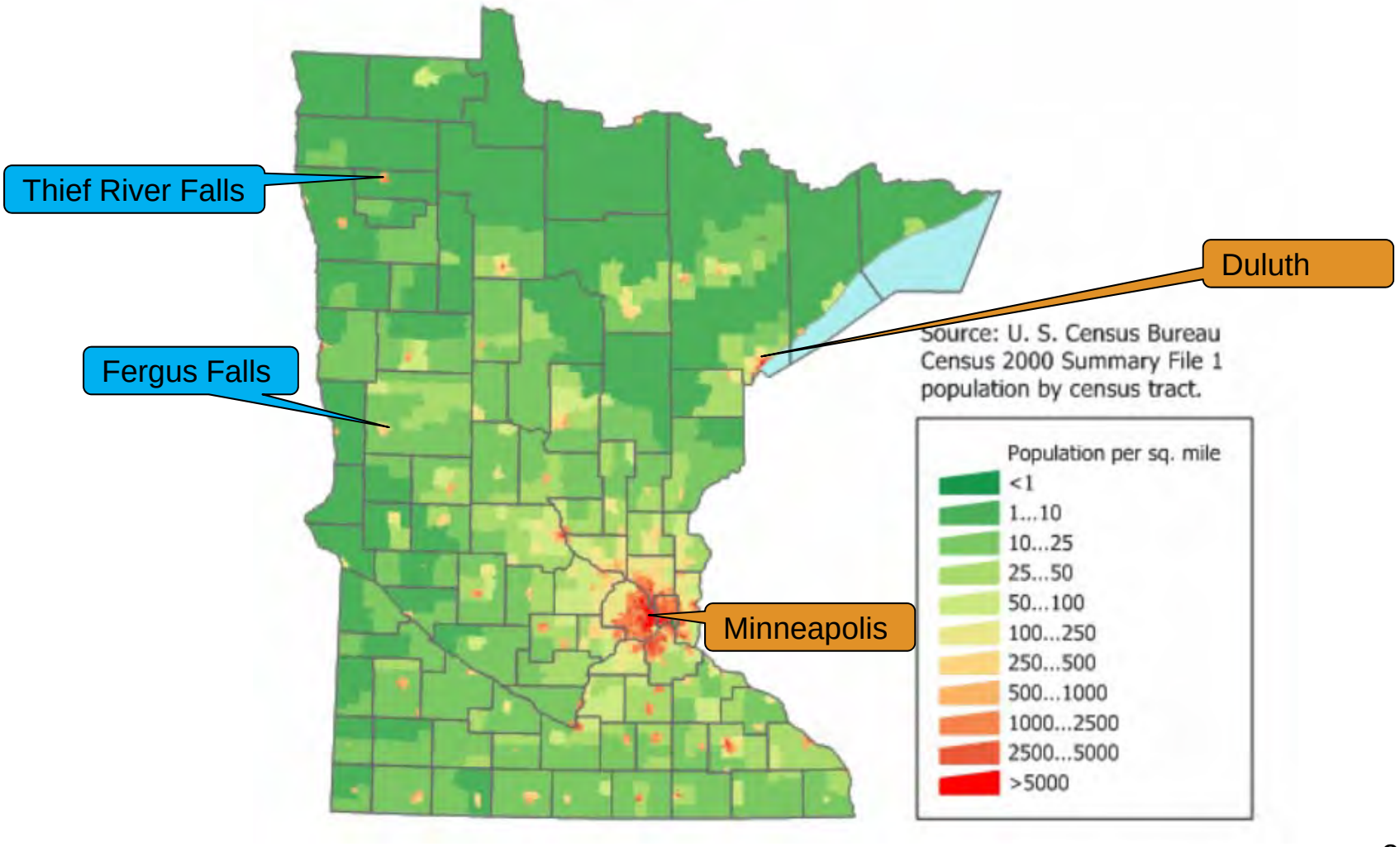
Louisiana Population Density Map



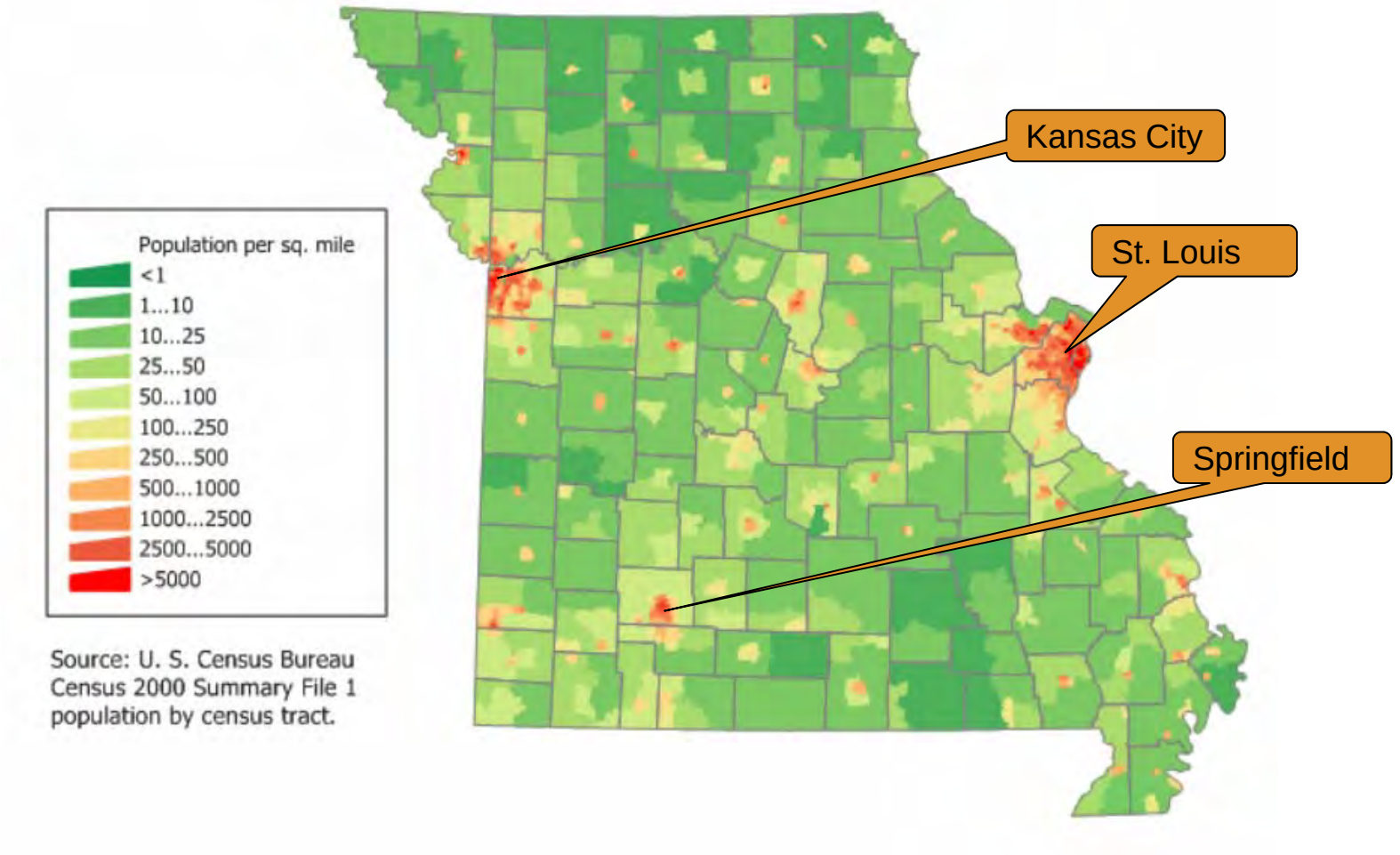
Michigan Population Density Map



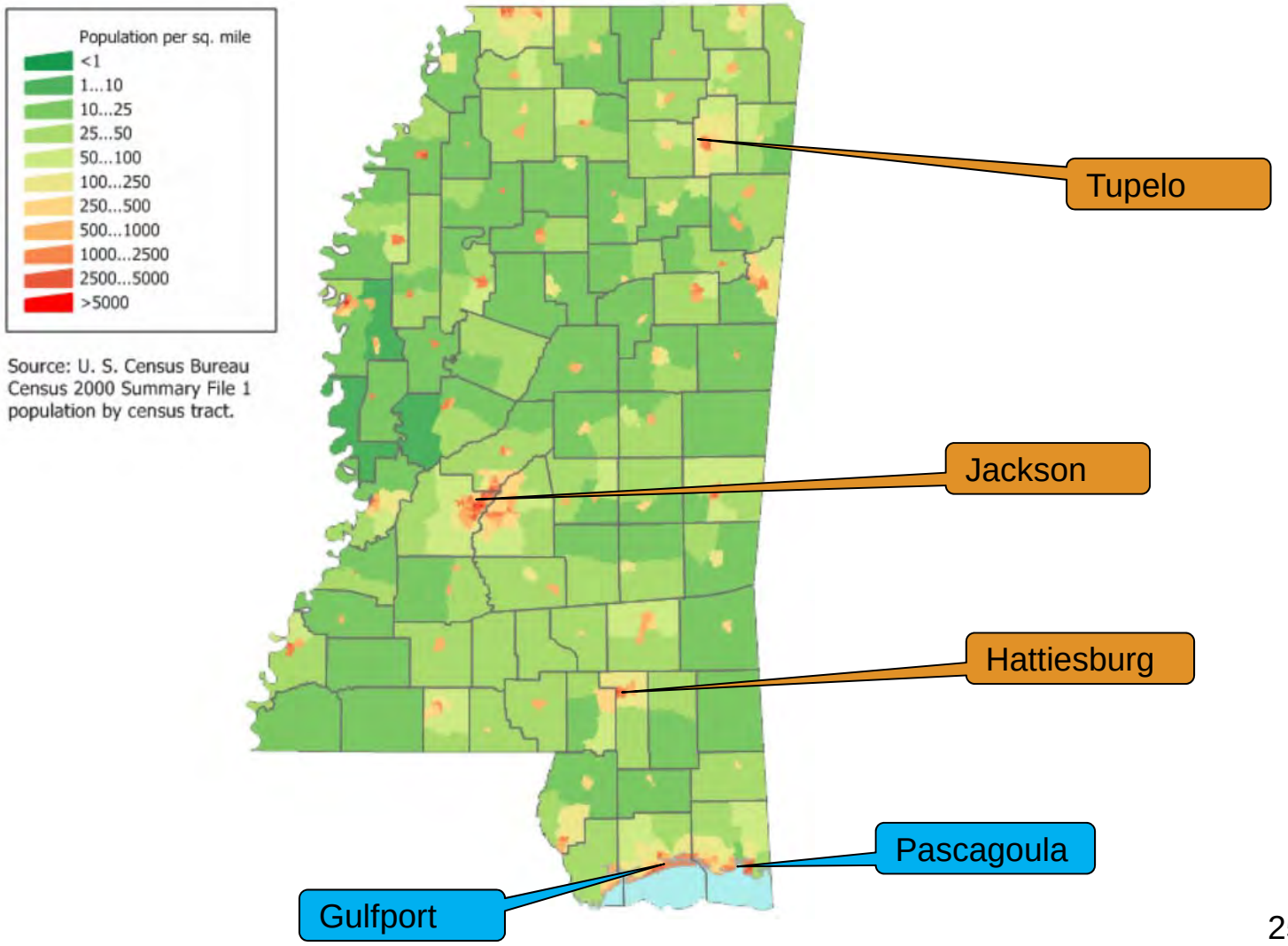
Minnesota Population Density Map



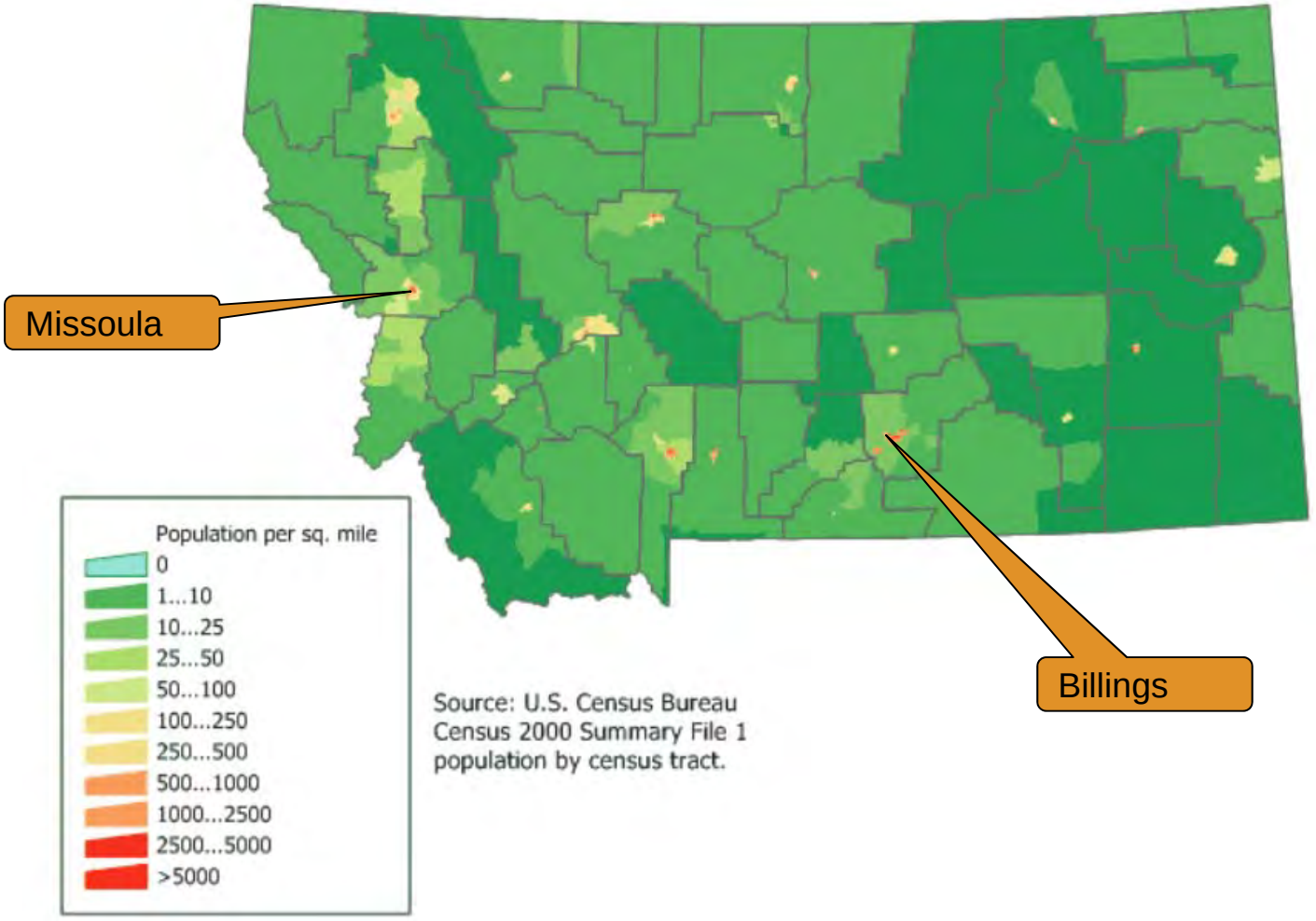
Missouri Population Density Map



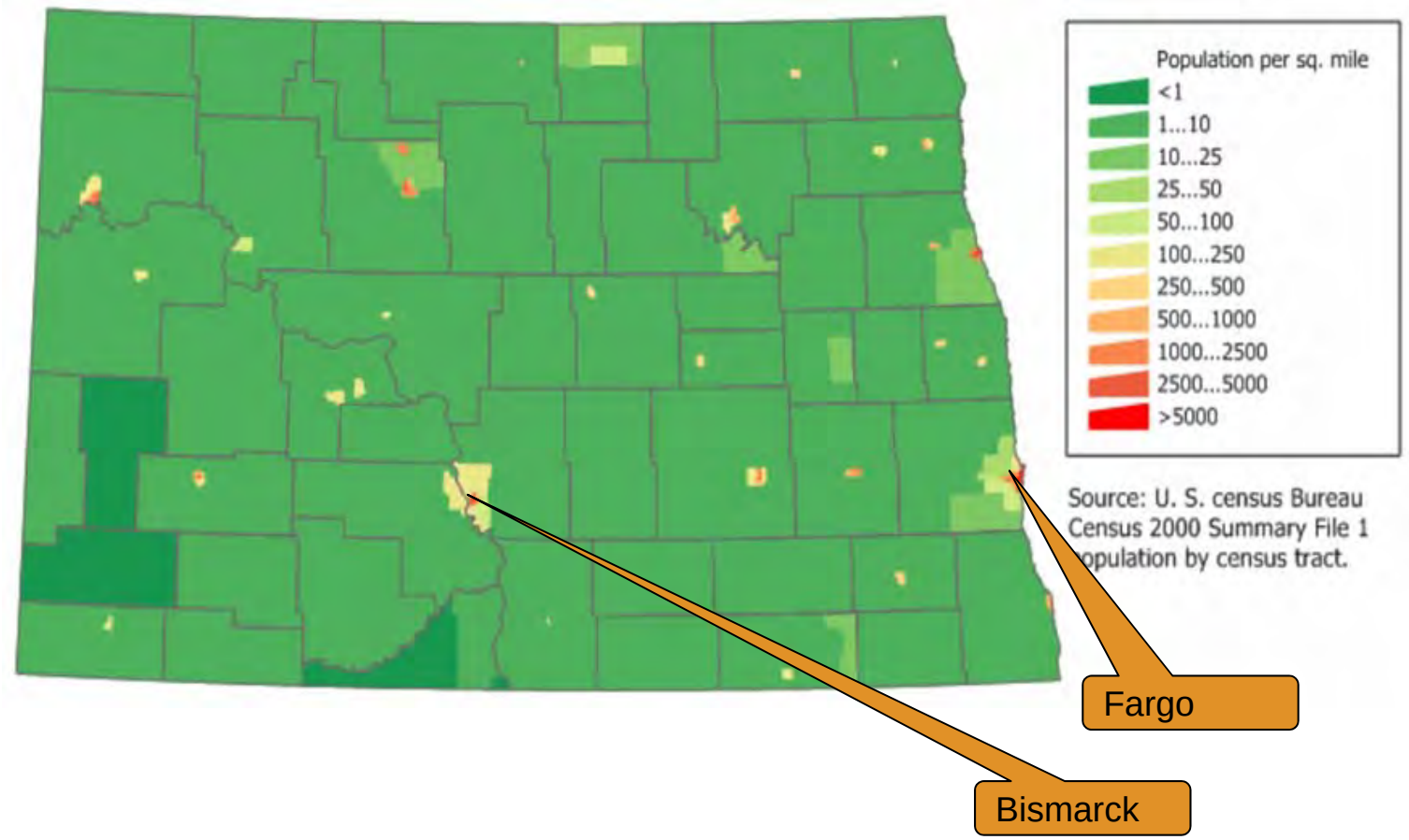
Mississippi Population Density Map



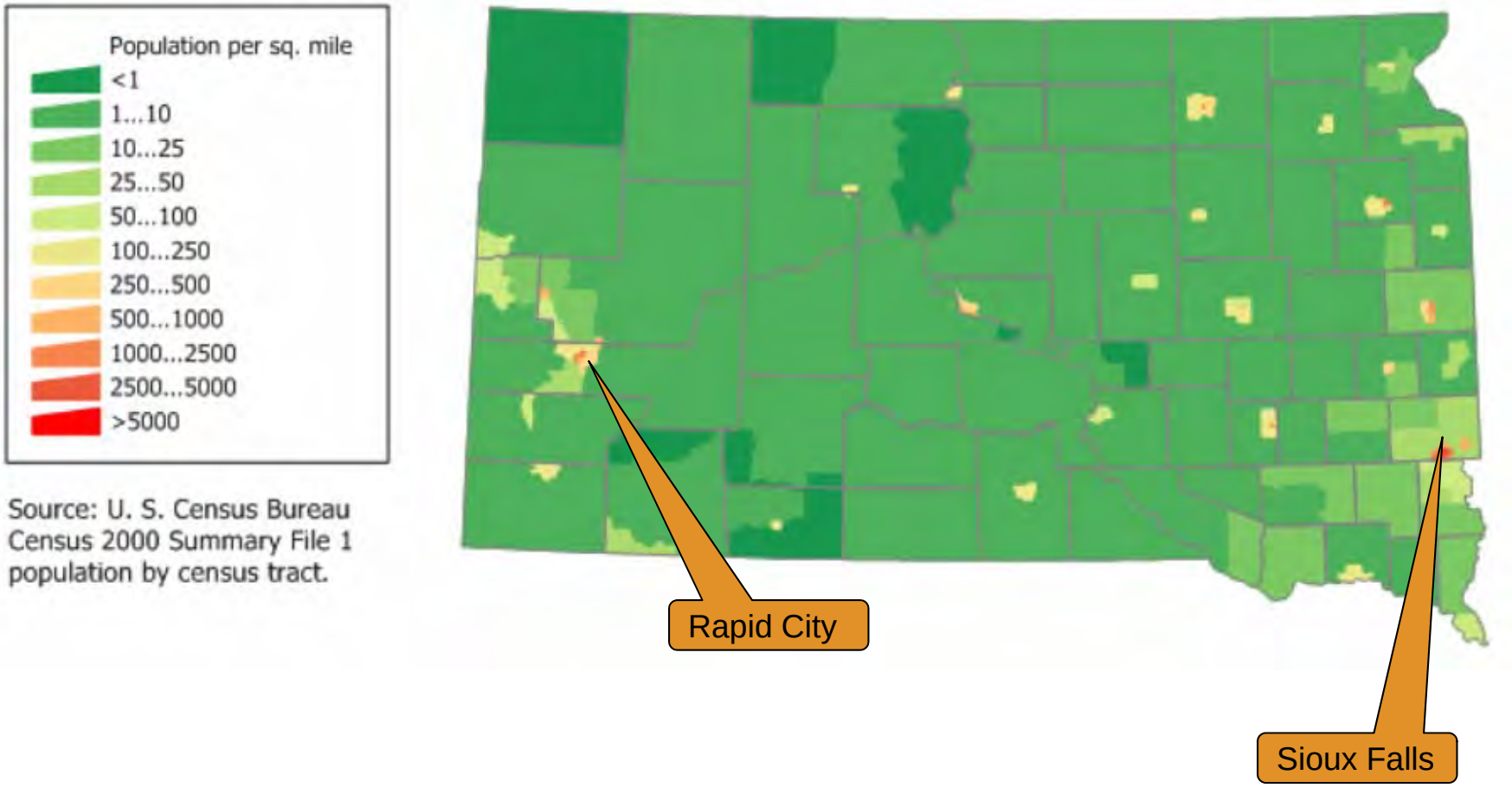
Montana Population Density Map



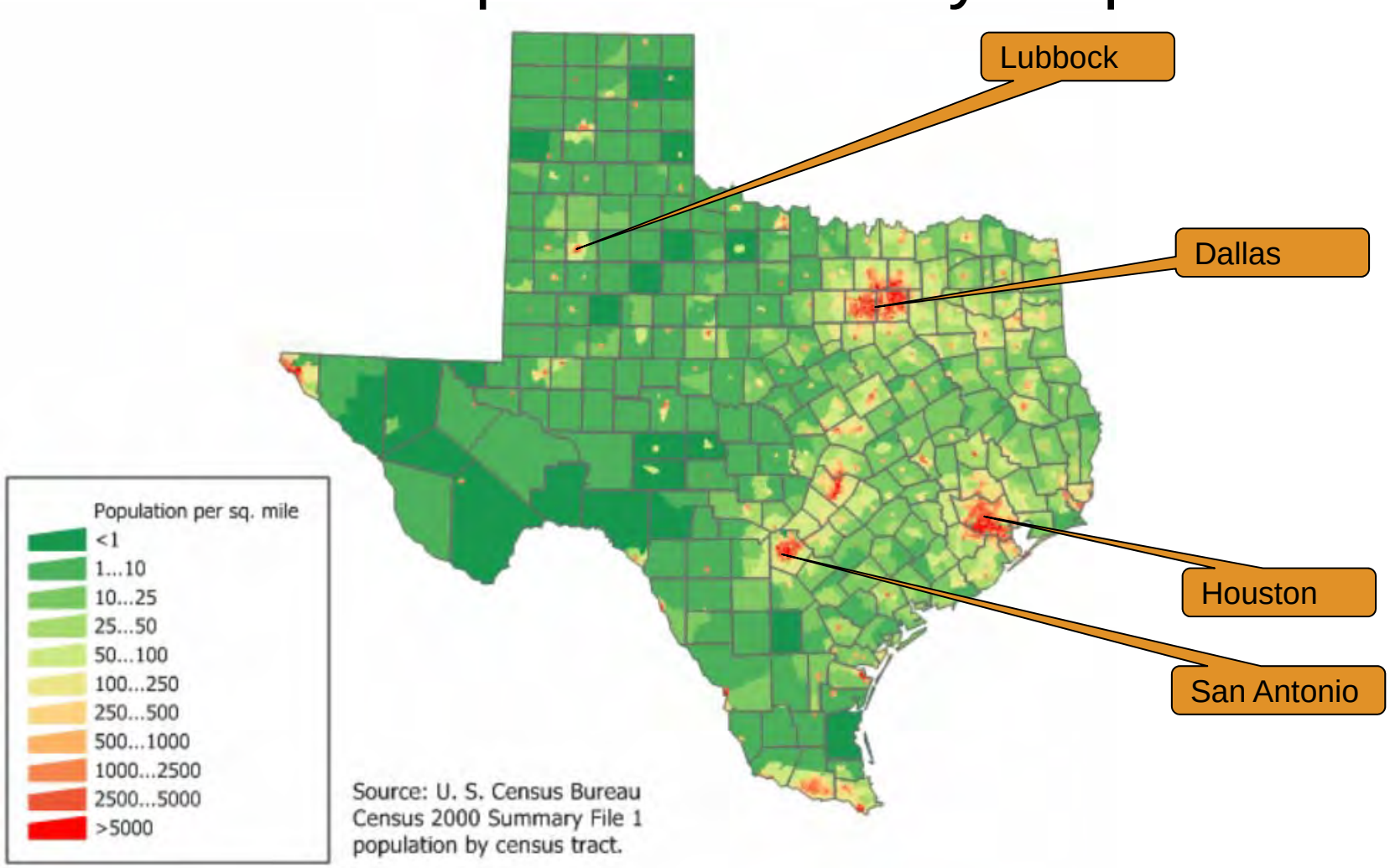
North Dakota Population Density Map



South Dakota Population Density Map



Texas Population Density Map



Wisconsin Population Density Map