Labor Markets Versus Trade: Identifying a Roadmap to Defining New Regions to Measure OCS Effects

J. Matthew Fannin, Ragu Kongari, Arun Adhikari, & Ashok Mishra
Louisiana State University

Background

- As shown by Alan Pulsipher and Kathy Perry, there can be measurable diversity of oil and gas industrial concentration between on shore area regions
- As co-PIs in the cooperative agreement with Alan Pulsipher and BOEMRE, our responsibility was to generate and evaluate alternative regional classifications based on theories and methods from regional science

Conceptual Framework

- Defining Regions
  - Homogenous – similarity of place
    - E.g. Corn Belt, Delta, Piney Woods
  - Nodal – functional relationships
    - E.g. metropolitan areas, BEA regions
  - Administrative – for management purposes
    - E.g. Federal Reserve System Regions, DOTD Regions, LSU Ag. Center Regions
- BOEMRE regions serve a hybrid purpose: a functional region for administrative purposes

Alternative Functional Regions

- Labor markets are one of the most popular methods to identify economic regions
  - Labor considered one of the most important inputs in the production process
- However, cluster advocates such as Michael Porter from Harvard and others argue spatially proximate inputs can have equal or greater influence in defining a region

Alternative Functional Regions

- We develop alternative regional clustering frameworks based on two approaches to defining economic regions
  - Traditional labor market regions (based on commuting patterns of laborers)
  - Industrial trade (based on physical trade of goods and services between buyer and seller)

New Datasets for Regional Classification

- New datasets represent vectors of economic relationships between regions
  - Includes a magnitude of economic linkage as well as a direction (similar to velocity in physics)
- New datasets represent combining statistical techniques, survey data and administrative data

New Datasets for Regional Construction

- Local Employment Dynamics (LED) Project http://lehd.did.census.gov/led/
  - The successor to journey-to-work commuting data
  - Generates annual county-to-county (and sub-county) commuting linkages
- IMPLAN™ Version 3.0 estimates inter-county trade flows by industry for over 400 industrial sectors

Strategy

- We use the county-to-county relationship between incommuters and outcommuters and imports and exports between counties to calculate strength of linkages
- We then apply a clustering methodology where the strongest counties are aggregated together into new multi-county regions

Strategy Overview

- Clustering Methodology
- Maps for 100 cluster regions
- Clusters selected for comparison based on BOEMRE regions
- Variables comparison

Clustering Methodology Details

Table 1: Exports matrix - E: Thousands of dollars exported from row parish to column parish

E_{ij}	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
A1	0	1.422	0.527	1.293	0.35	0.674	1.624	1.377	0.46	0.297
A2	0.821	0	0.569	0.162	0.349	0.34	1.2	0.837	0.161	0.139
A3	0.777	1.452	0	0.092	0.241	3.614	0.484	0.726	0.656	0.308
A4	2.497	0.181	0.103	0	0.333	0.051	0.16	0.955	0.051	0.112
A5	0.344	0.463	1.179	0.377	0	0.654	0.147	3.719	0.131	0.695
A6	1.043	0.745	3.421	0.07	0.143	0	0.267	0.273	0.527	0.05
A7	4.735	1.572	1.008	0.221	0.388	0.524	0	0.846	0.179	0.298
A8	2.019	1.684	0.906	1.133	5.262	0.907	0.798	0	2.155	16.32
A9	0.455	0.434	0.371	0.098	0.249	0.334	0.271	1.878	0	0.424
A10	0.185	0.48	0.211	0.092	0.732	0.108	0.081	10.1	0.177	0

Clustering Methodology Details

Table 2: Transaction matrix: Thousands of dollars of imports and exports between the parishes.

T_{ij}	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	ToP*
A1	0	2.243	1.304	3.79	0.695	1.717	6.359	3.396	0.915	0.482	20.9
A2	2.243	0	2.021	0.343	0.812	1.085	2.772	2.521	0.595	0.62	13.01
A3	1.304	2.021	0	0.195	1.419	7.036	1.492	1.632	1.026	0.519	16.64
A4	3.79	0.343	0.195	0	0.711	0.121	0.381	2.087	0.149	0.204	7.981
A5	0.695	0.812	1.419	0.711	0	0.797	0.535	8.982	0.38	1.427	15.76
A6	1.717	1.085	7.036	0.121	0.797	0	0.791	1.18	0.861	0.158	13.75
A7	6.359	2.772	1.492	0.381	0.535	0.791	0	1.644	0.45	0.379	14.8
A8	3.396	2.521	1.632	2.087	8.982	1.18	1.644	0	4.033	26.42	51.89
A9	0.915	0.595	1.026	0.149	0.38	0.861	0.45	4.033	0	0.601	9.01
A10	0.482	0.62	0.519	0.204	1.427	0.158	0.379	26.42	0.601	0	30.81

The transaction matrix is given by $T = E + E^T$

Clustering Methodology Details

- We calculate strength of trade linkages by
  - $C_{ij} = 1 - \frac{T_{ij}}{ToP_i + ToP_j - T_{ij}}$

    Where,

    C_{ij} = Coefficient of linkage between parishes A_i and A_j

    T_{ij} = Imports and exports trade between parishes A_i and A_j

    ToP_i = Total trade by parish A_i

    ToP_j = Total trade by parish A_j

Clustering Methodology Details

Table 3: Coefficient / distance matrix: Linkage between parishes.

C_{ij}	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
A1	0	0.9292	0.964	0.8489	0.9807	0.9479	0.7833	0.9511	0.9685	0.9906
A2	0.9292	0	0.9269	0.9834	0.971	0.9577	0.8893	0.9596	0.9722	0.9857
A3	0.964	0.9269	0	0.992	0.9542	0.6988	0.9502	0.9756	0.9583	0.9889
A4	0.8489	0.9834	0.992	0	0.9691	0.9944	0.983	0.9639	0.9911	0.9947
A5	0.9807	0.971	0.9542	0.9691	0	0.9722	0.9822	0.8469	0.9844	0.9684
A6	0.9479	0.9577	0.6988	0.9944	0.9722	0	0.9715	0.9817	0.9607	0.9964
A7	0.7833	0.8893	0.9502	0.983	0.9822	0.9715	0	0.9747	0.9807	0.9916
A8	0.9511	0.9596	0.9756	0.9639	0.8469	0.9817	0.9747	0	0.9291	0.5306
A9	0.9685	0.9722	0.9583	0.9911	0.9844	0.9607	0.9807	0.9291	0	0.9847
A10	0.9906	0.9857	0.9889	0.9947	0.9684	0.9964	0.9916	0.5306	0.9847	0

Clustering Methodology Details

- Since SAS treats distance (coefficient) as a measure of strength of linkage between parishes, the smaller the distance, the stronger the linkage between parishes.
- For example, if we want a group of 9 out of the 10 counties, then we need to find one pair of parishes which are strongly bonded. Those include A8 and A10 with a linkage coefficient of 0.5306. We name the group CL9. The resulting matrix results from taking the average method as shown in Table 4.

Clustering Methodology Details

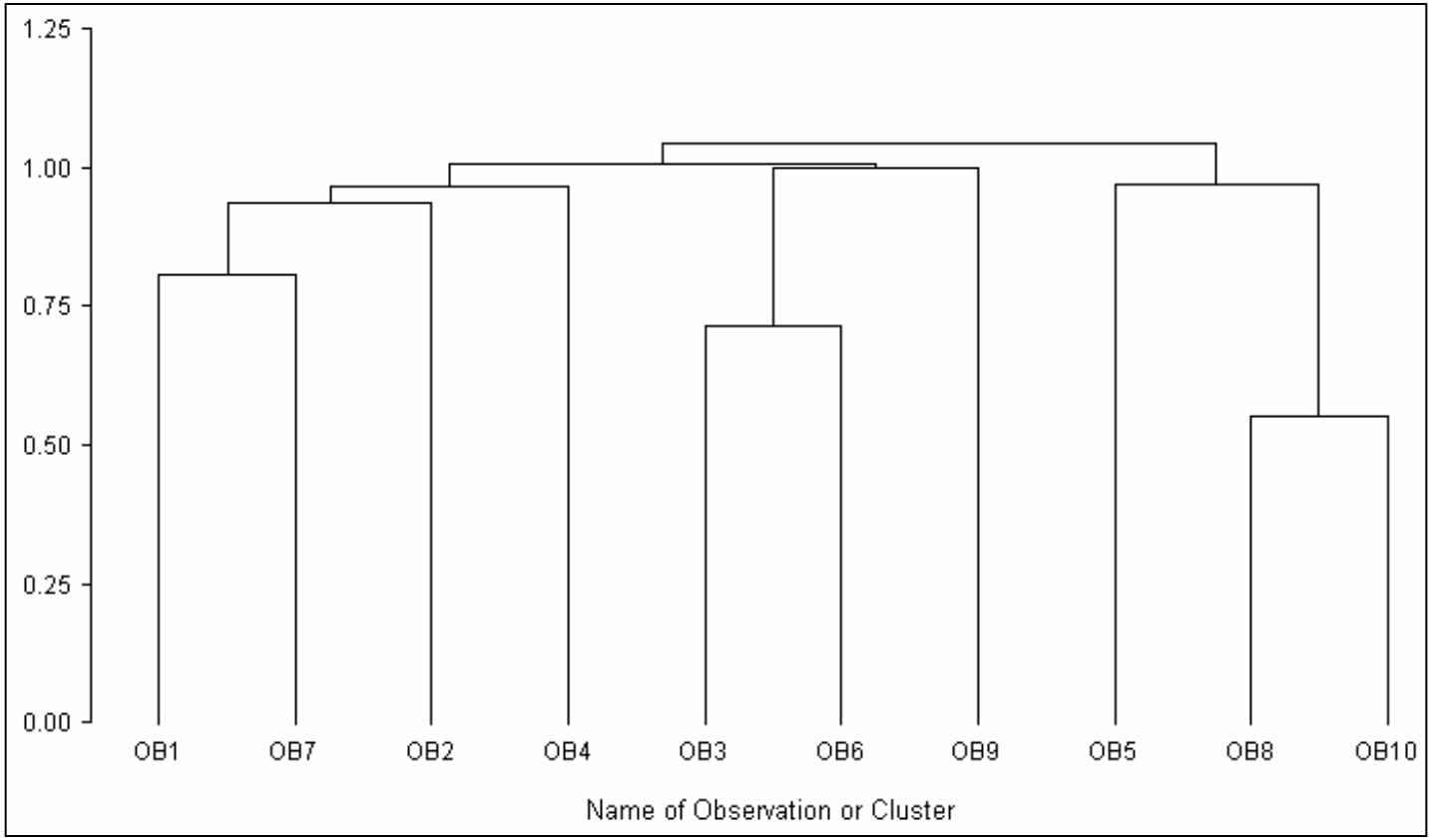
Table 4: Distance / coefficient matrix after 10 counties grouped into 9.

C_{ij}	A1	A2	A3	A4	A5	A6	A7	A9	CL9
A1	0	0.9292	0.964	0.8489	0.9807	0.9479	0.7833	0.9685	0.9708
A2	0.9292	0	0.9269	0.9834	0.971	0.9577	0.8893	0.9722	0.9726
A3	0.964	0.9269	0	0.992	0.9542	0.6988	0.9502	0.9583	0.9823
A4	0.8489	0.9834	0.992	0	0.9691	0.9944	0.983	0.9911	0.9793
A5	0.9807	0.971	0.9542	0.9691	0	0.9722	0.9822	0.9844	0.9077
A6	0.9479	0.9577	0.6988	0.9944	0.9722	0	0.9715	0.9607	0.9891
A7	0.7833	0.8893	0.9502	0.983	0.9822	0.9715	0	0.9807	0.9832
A9	0.9685	0.9722	0.9583	0.9911	0.9844	0.9607	0.9807	0	0.9569
CL9	0.9708	0.9726	0.9823	0.9793	0.9077	0.9891	0.9832	0.9569	0

From the above table the second strongly bonded parishes are A3 and A6 with a linkage coefficient of 0.6988, we then name the group CL8. We would then aggregate those regions and continue the iterative process until we reaches the number of regions we requested.

Clustering Methodology Details

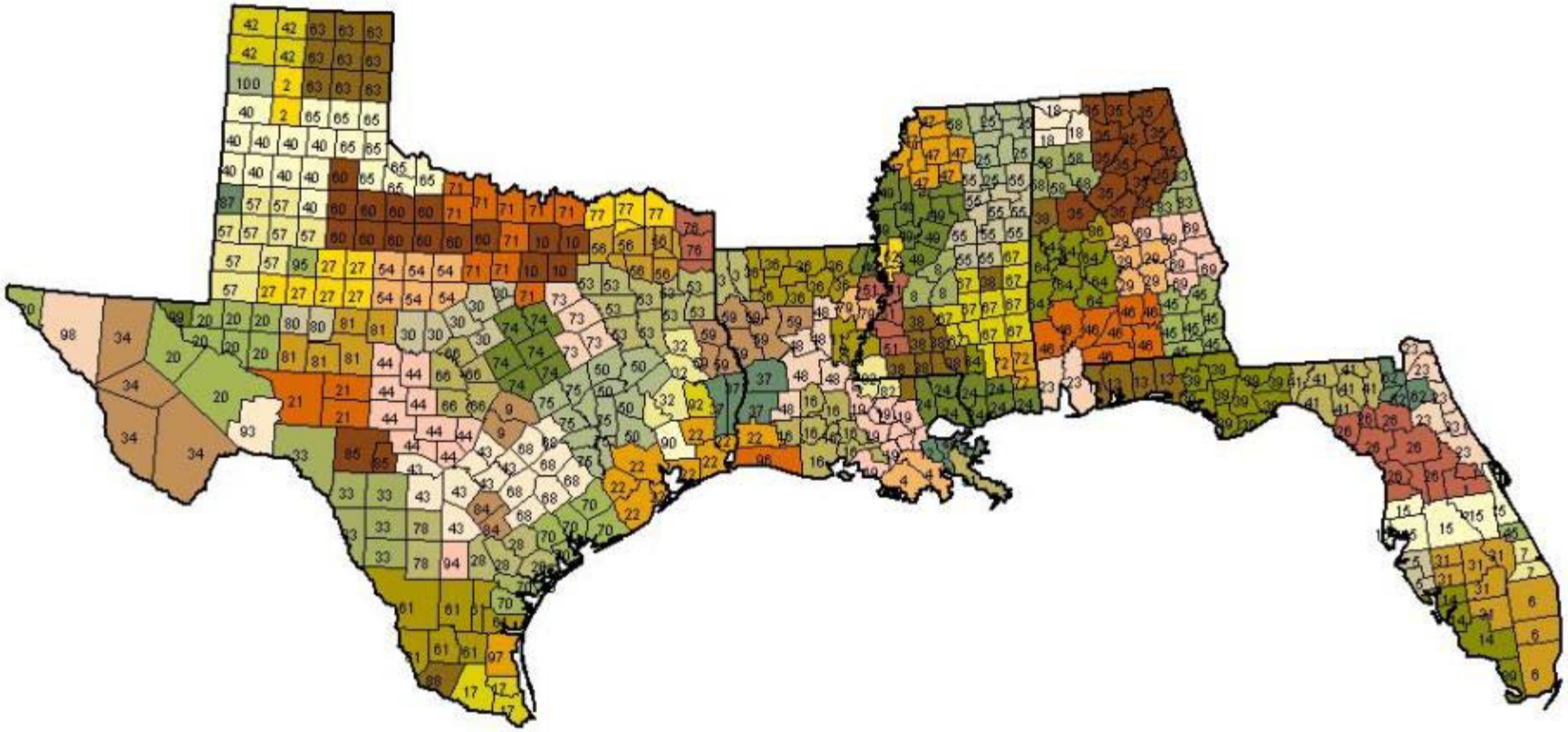
Figure1: SAS cluster tree.



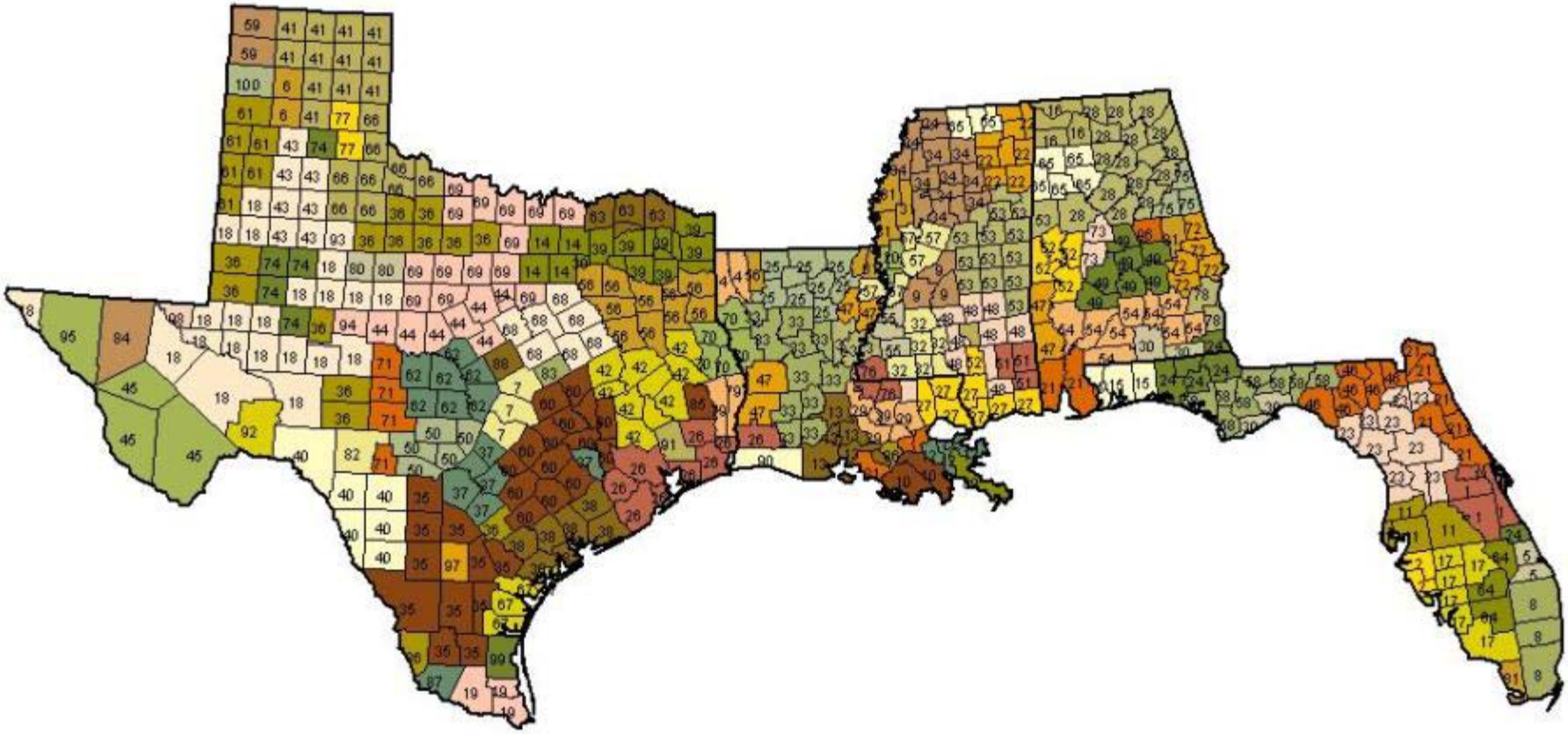
Alternative Regional Classificatons

- Commuting Regions (2008 LED data)
- All Industrial Trade (2008 IMPLAN trade data)
- Oil and Gas Industrial Trade (2008 IMPLAN trade data using Saha, Manik and Phillips 2005 OCS oil and gas supply chain IMPLAN sector subset (approximately 40 industries))
- Compare to BOEMRE on shore area baseline regions

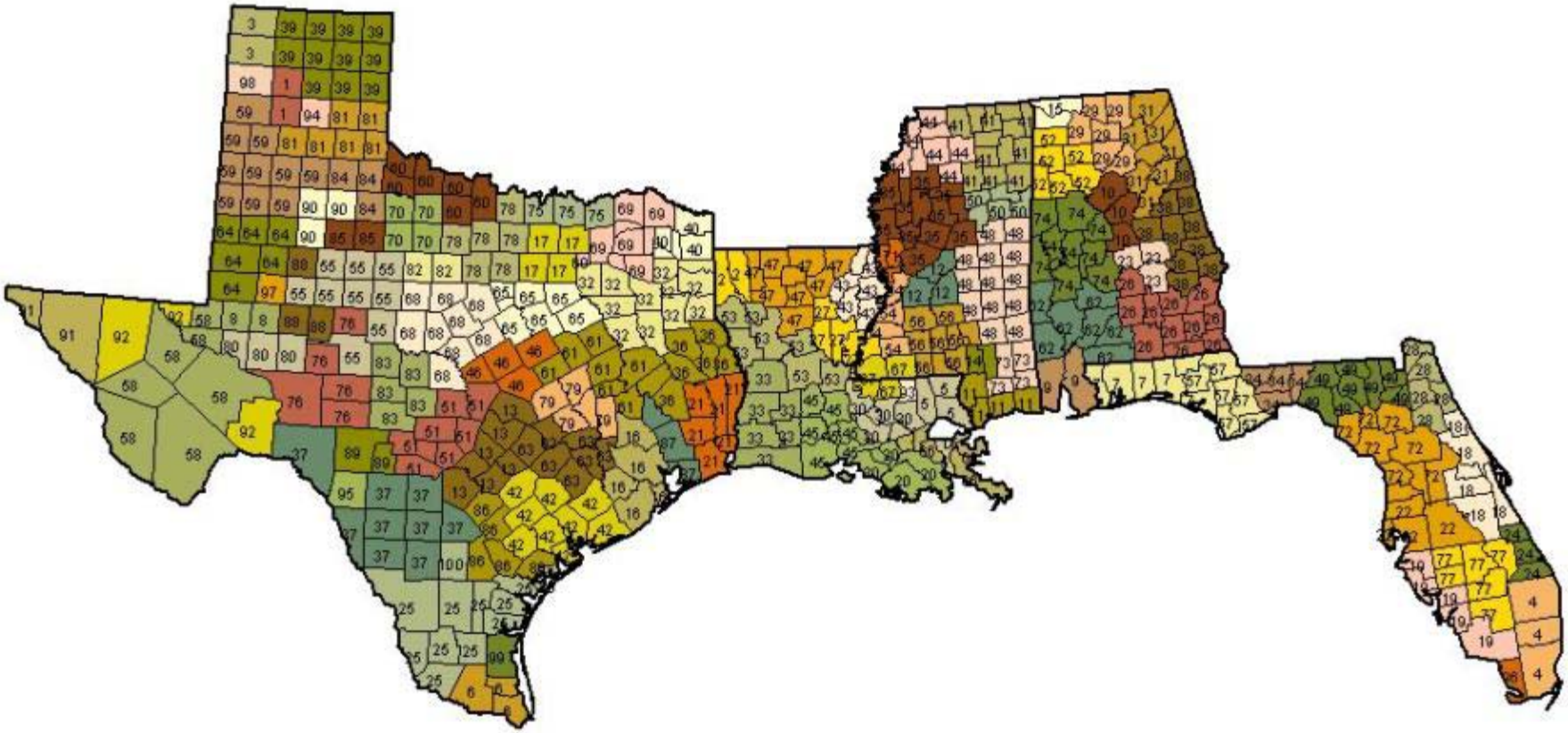
All Industry Trade Clusters 100



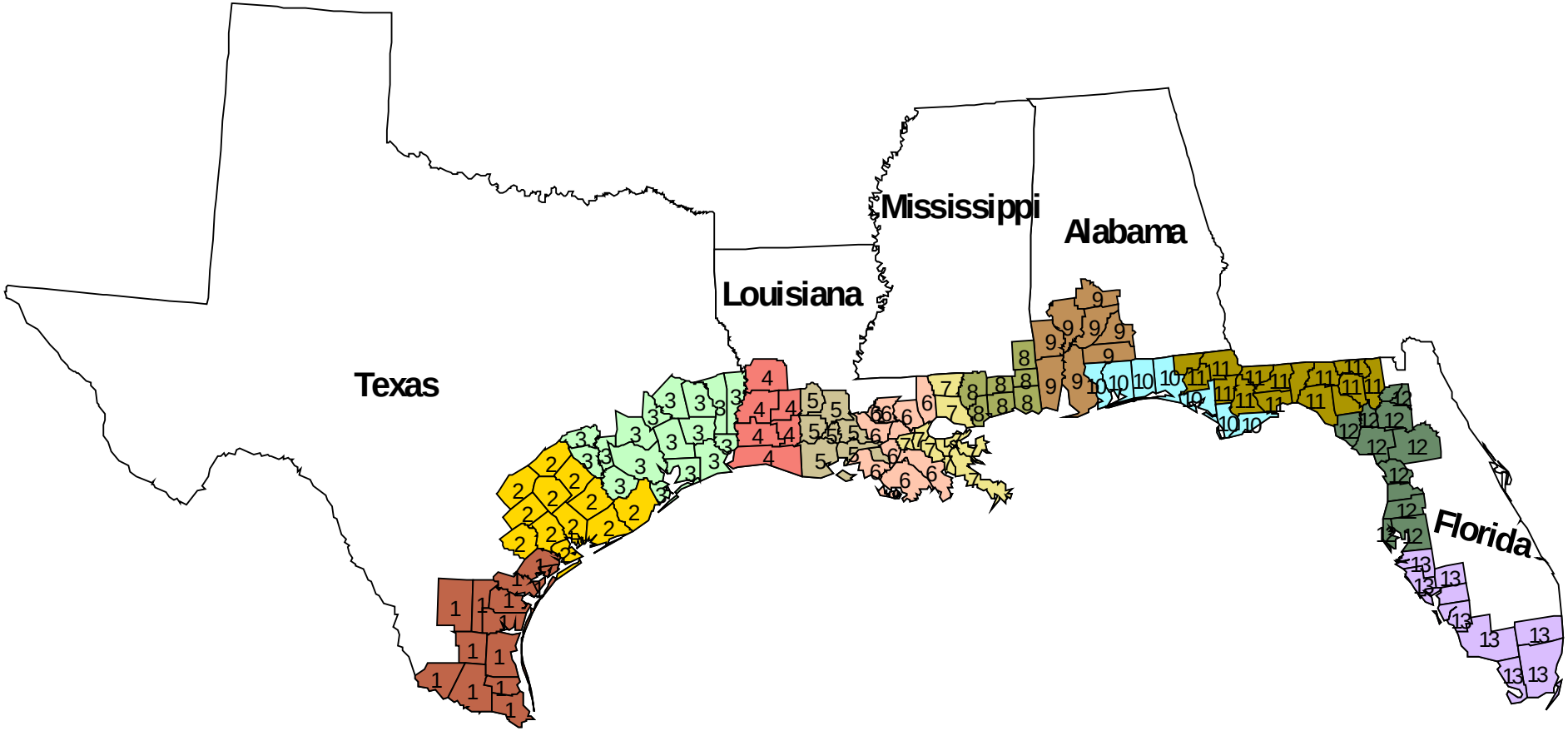
Oil and Gas Industry Clusters 100



Commuter Clusters 100



BOEMRE Regions



Regions for Analysis

- BOEMRE regions as base
- Select clusters with at least one BOEMRE county
- Socio-economic variables for comparison

Table 5: Number of counties and regions for the analysis

	BOEMRE	Commuter Cluster	All Trade Cluster	Oil and Gas Cluster
Number of Regions	13	30	37	39
Number of Counties	132	159	188	211

Variables Comparison

- Demographic statistics
- Industry Location Quotient of NAICS 211, 213111, 213112, 333132, 336611, 541360, Oil and gas (indicated as 444), and Pipeline (indicated as 555)

Statistical Analysis

- Apply a fixed effects type estimator where the dependent variable is regressed against only dummy variables representing the individual regions for each cluster
- Compare R-squares to determine best fitting model to demographic and location quotient data

Variables Comparison (cont.)

Table 6: R-square values for several socio-economic attributes within different regions

Variables	MMS R-square	Commuting R-square	All Trade R-square	Oil and Gas Trade R-square
Popest2008	0.139	**0.5540**	0.466	0.441
Povertyratio2007	0.431	0.652	**0.658**	0.515
Percapinc2006	0.365	0.596	**0.620**	0.480
Educationratio2007	0.134	**0.533**	0.356	0.315
Poppersqmile2000	0.148	0.425	0.410	**0.440**
Lqn211y08	0.154	0.267	**0.518**	0.372
Lqn213111y08	0.320	0.294	**0.411**	0.234
Lqn213112y08	0.221	0.313	**0.461**	0.275
Lqn333132y08	0.196	0.301	**0.518**	0.500
Lqn336611y08	0.465	**0.485**	0.319	0.319
Lqn541360y08	0.156	0.336	**0.427**	0.259
Lqn444y08	0.249	0.338	**0.483**	0.377
Lqn555y08	0.229	0.352	0.541	**0.550**
Lqn211percent	0.144	0.135	**0.189**	0.149
Lqn213111percent	0.127	0.571	**0.761**	0.654
Lqn213112percent	0.174	**0.295**	0.276	0.129
Lqn333132percent	0.087	**0.983**	0.144	0.148
Lqn336611percent	0.660	0.664	0.489	**0.689**
Lqn541360percent	0.491	0.708	0.661	**0.728**
Lqn444percent	0.210	0.292	**0.317**	0.259
Lqn555percent	0.118	0.152	**0.225**	0.160
Best R-square value in Total	**0**	**5**	**12**	**4**

Variables comparison (cont.)

Table 7: Adjusted R-square values for several socio-economic attributes within different regions

Variables	MMS Adjusted R-square	Commuting Adjusted R-square	All Trade Adjusted R-square	Oil and Gas Trade Adjusted R-square
Popest2008	0.052	**0.454**	0.339	0.317
Povertyratio2007	0.373	0.574	**0.576**	0.407
Percapinc2006	0.301	0.505	**0.529**	0.365
Educationratio2007	0.047	**0.428**	0.203	0.164
Poppersqmile2000	0.062	0.295	0.269	**0.317**
Lqn211y08	-0.001	0.001	**0.322**	0.148
Lqn213111y08	**0.174**	0.001	0.155	-0.069
Lqn213112y08	0.091	0.102	**0.281**	0.055
Lqn333132y08	-0.153	-0.324	**0.141**	0.120
Lqn336611y08	**0.292**	0.086	-0.164	-0.224
Lqn541360y08	-0.116	-0.185	**-0.068**	-0.306
Lqn444y08	0.146	0.139	**0.318**	0.209
Lqn555y08	0.139	0.185	0.419	**0.436**
Lqn211percent	**-0.032**	-0.212	-0.183	-0.183
Lqn213111percent	-0.173	0.270	**0.557**	0.417
Lqn213112percent	0.030	**0.067**	0.035	-0.128
Lqn333132percent	-0.369	**0.969**	-0.578	-0.581
Lqn336611percent	**0.516**	0.392	0.127	0.362
Lqn541360percent	0.222	0.299	0.332	**0.439**
Lqn444percent	**0.091**	0.062	0.088	0.037
Lqn555percent	**-0.054**	-0.206	-0.142	-0.216
Best R-square value in Total	**6**	**4**	**8**	**3**

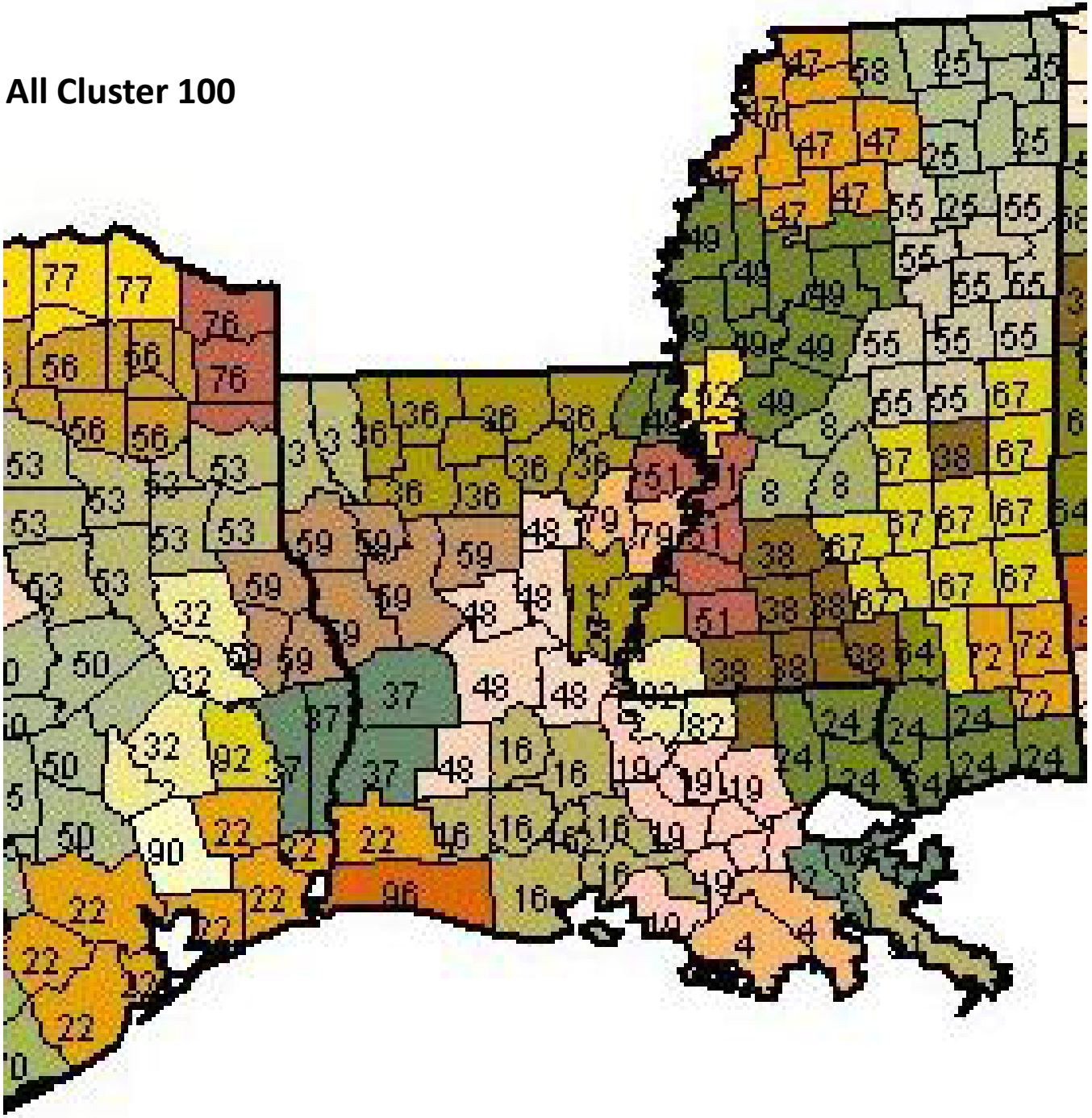
Discussion

- The all-industrial trade classification appears to combine both goodness of fit to the oil- and gas-related location quotient variables as well as the demographic variables
- Original BOEMRE on-shore areas also show reasonable fit on a portion of the LQ variables
- Commuting and Oil and Gas Industry trade clusters underperformed

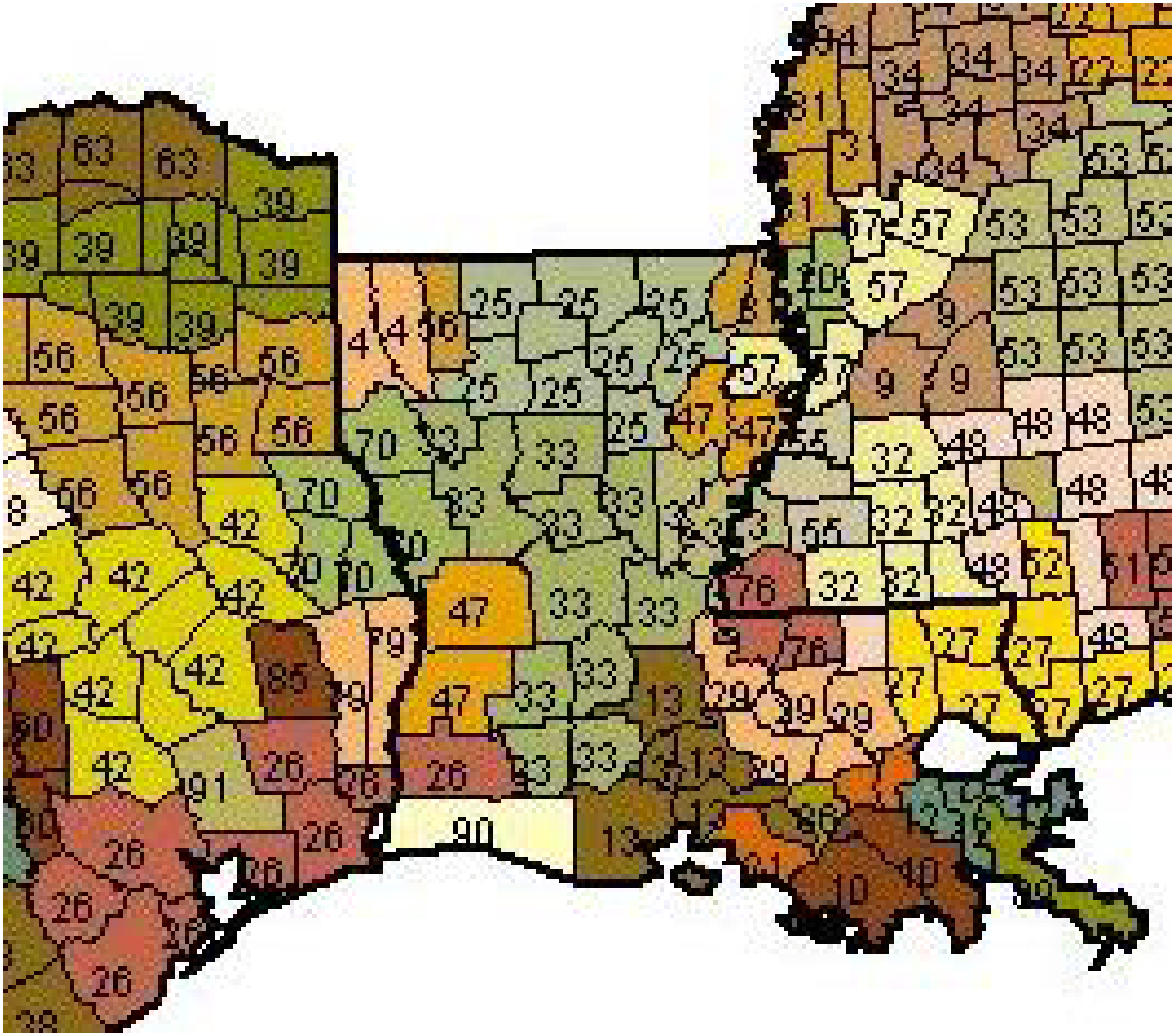
Next Steps

- Final sensitivity analysis to multi-state regions and additional counties
- Make final subjective calls on marginal counties with BOEMRE project officers
- Apply new multi-county regional classifications to portfolio of socio-economic BOEMRE research projects

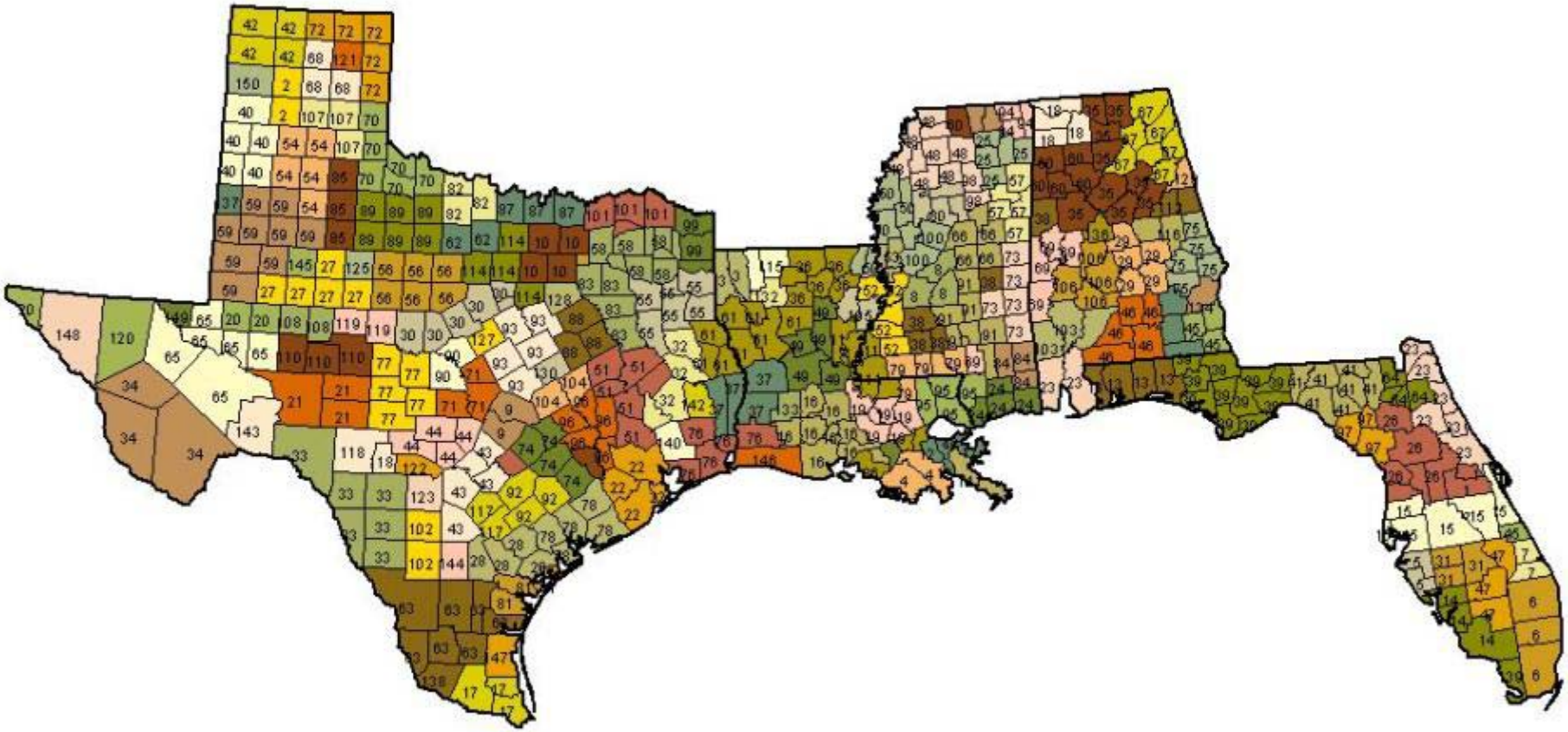
All Cluster 100



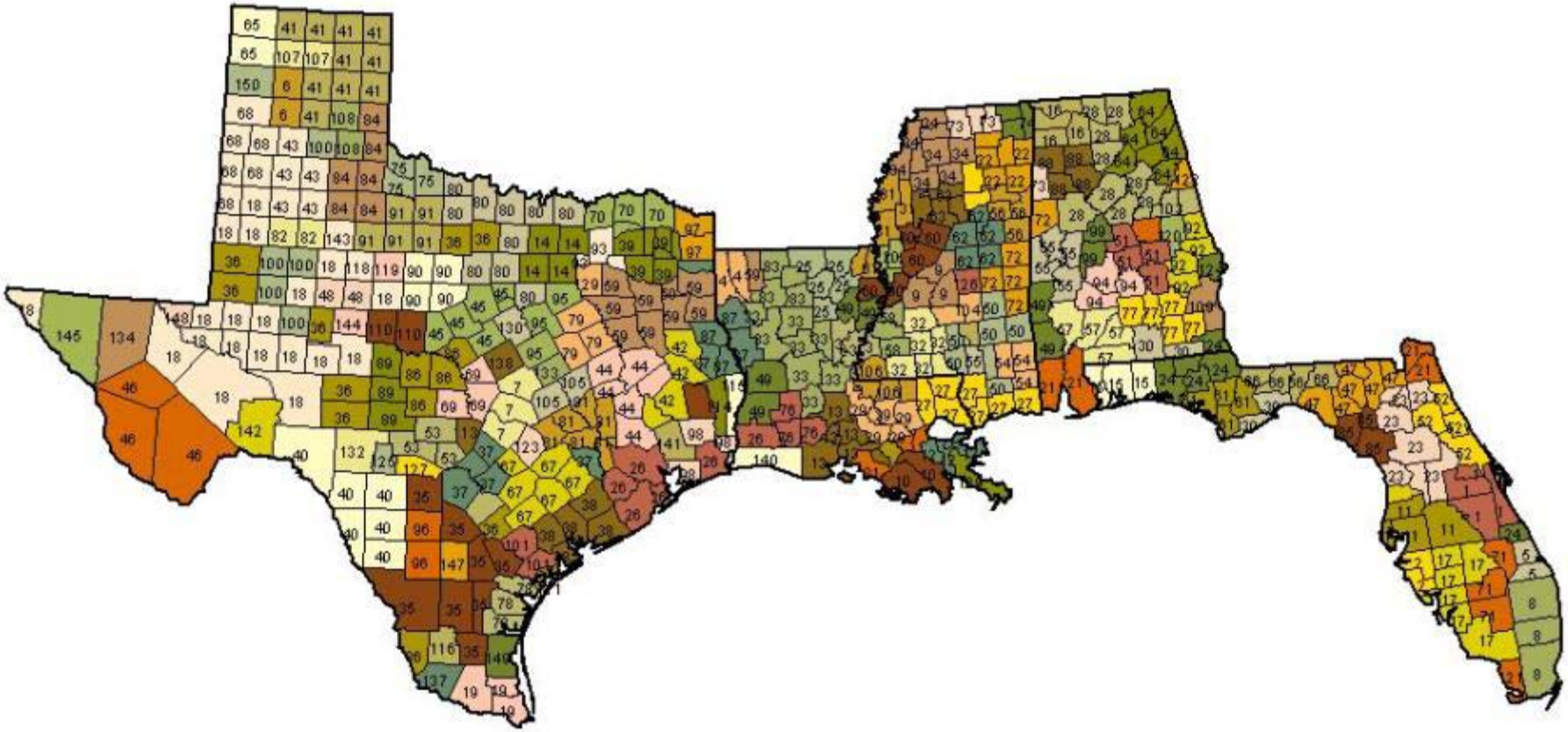
OGCluster 100



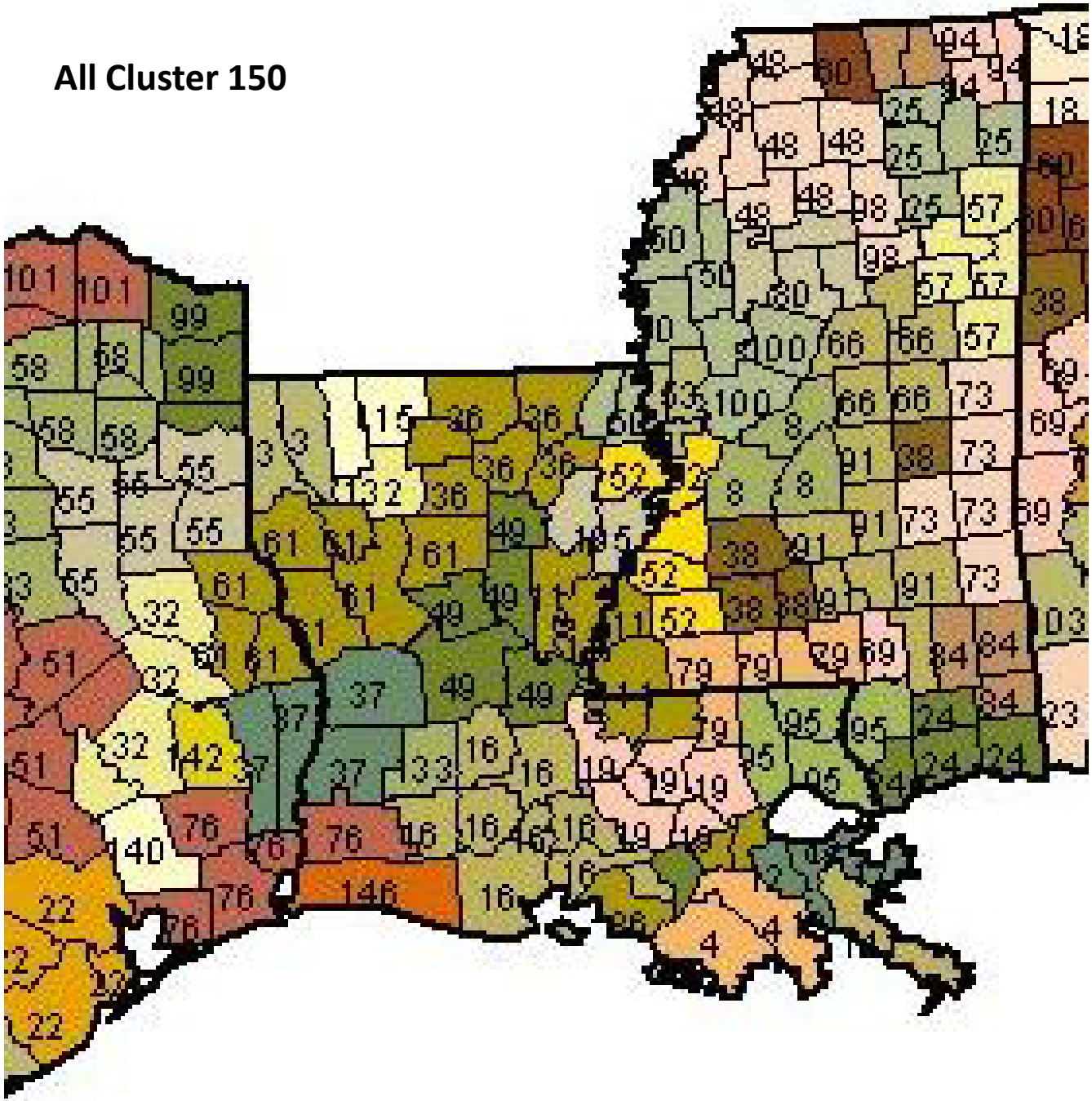
All Cluster 150



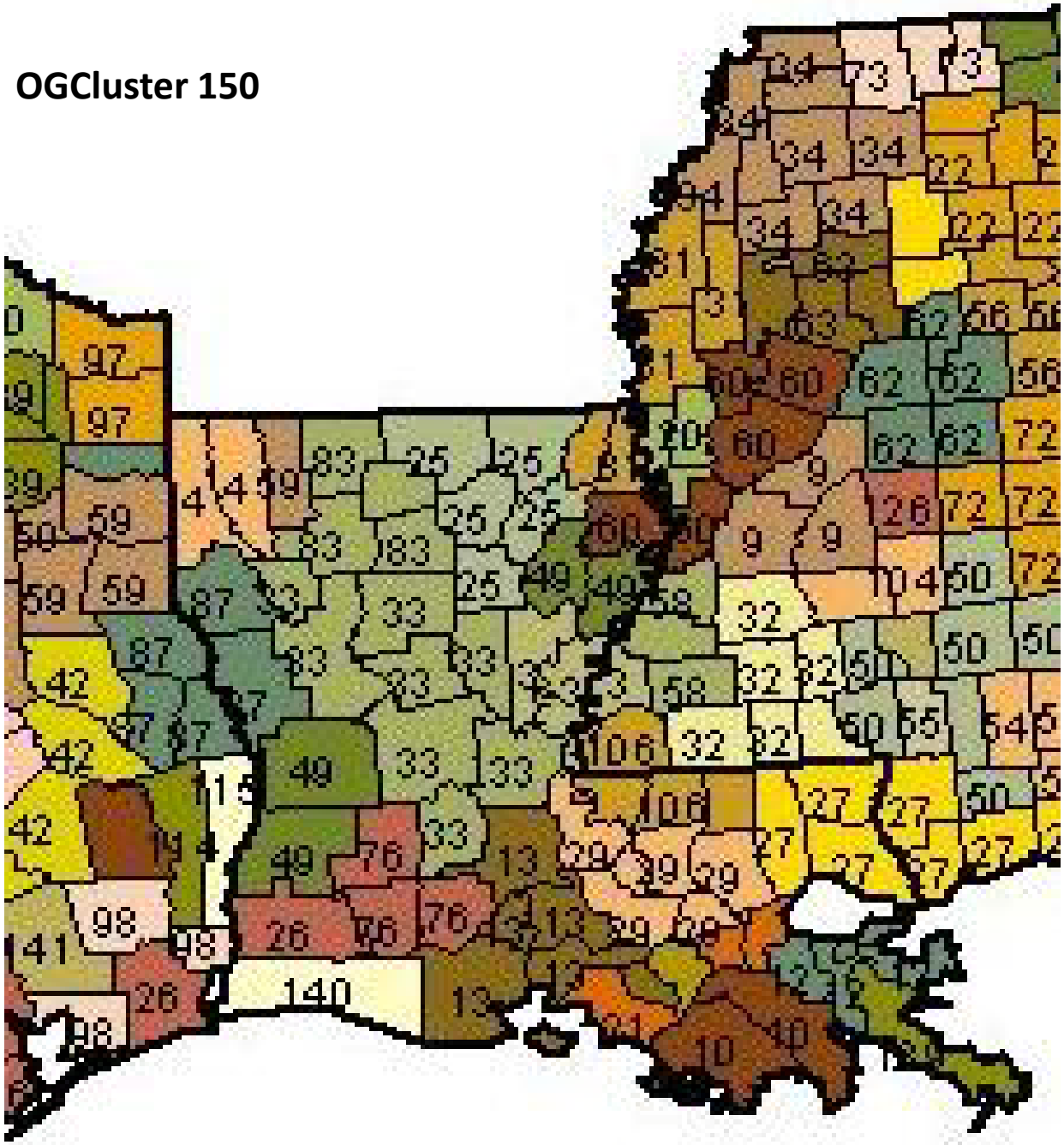
OGCluster 150



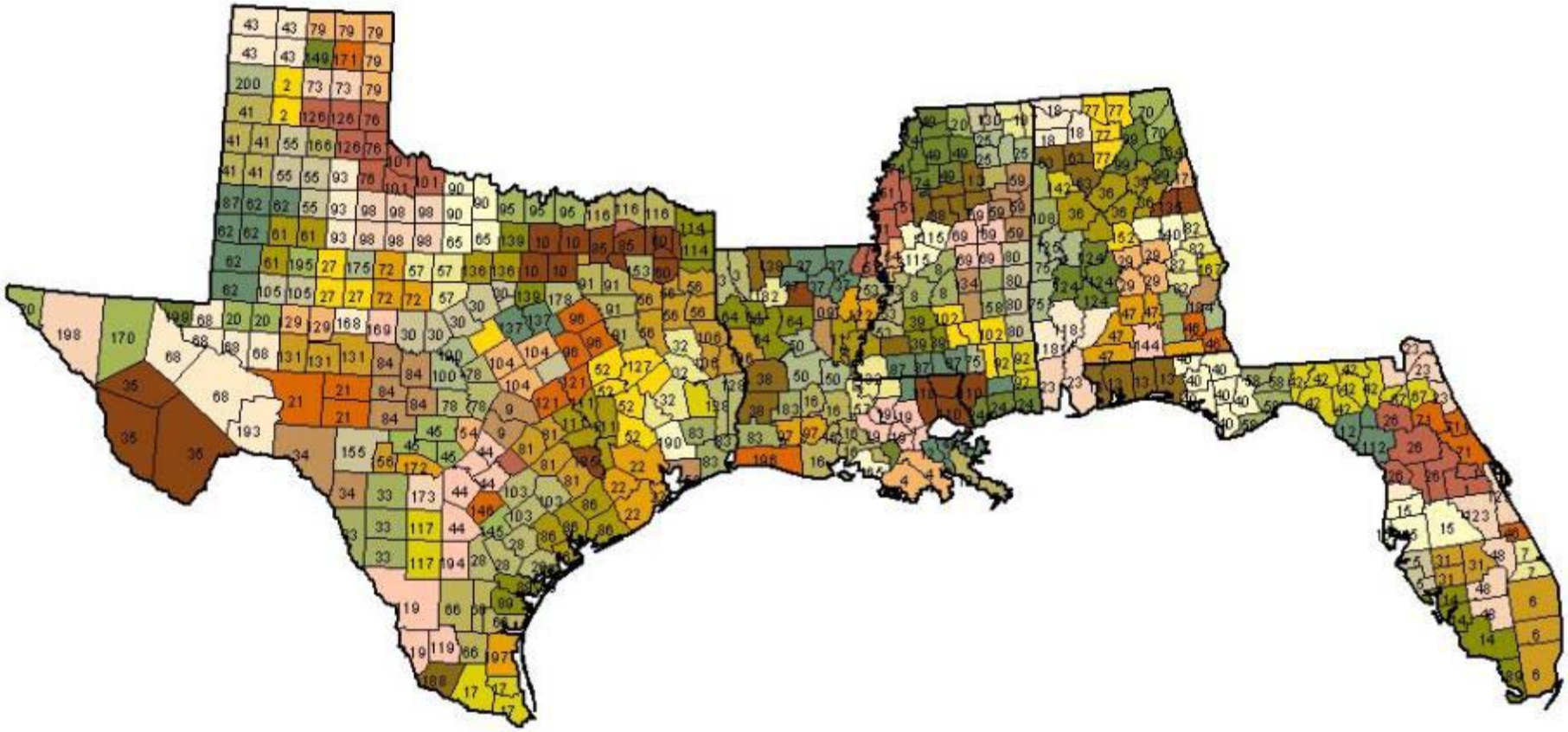
All Cluster 150



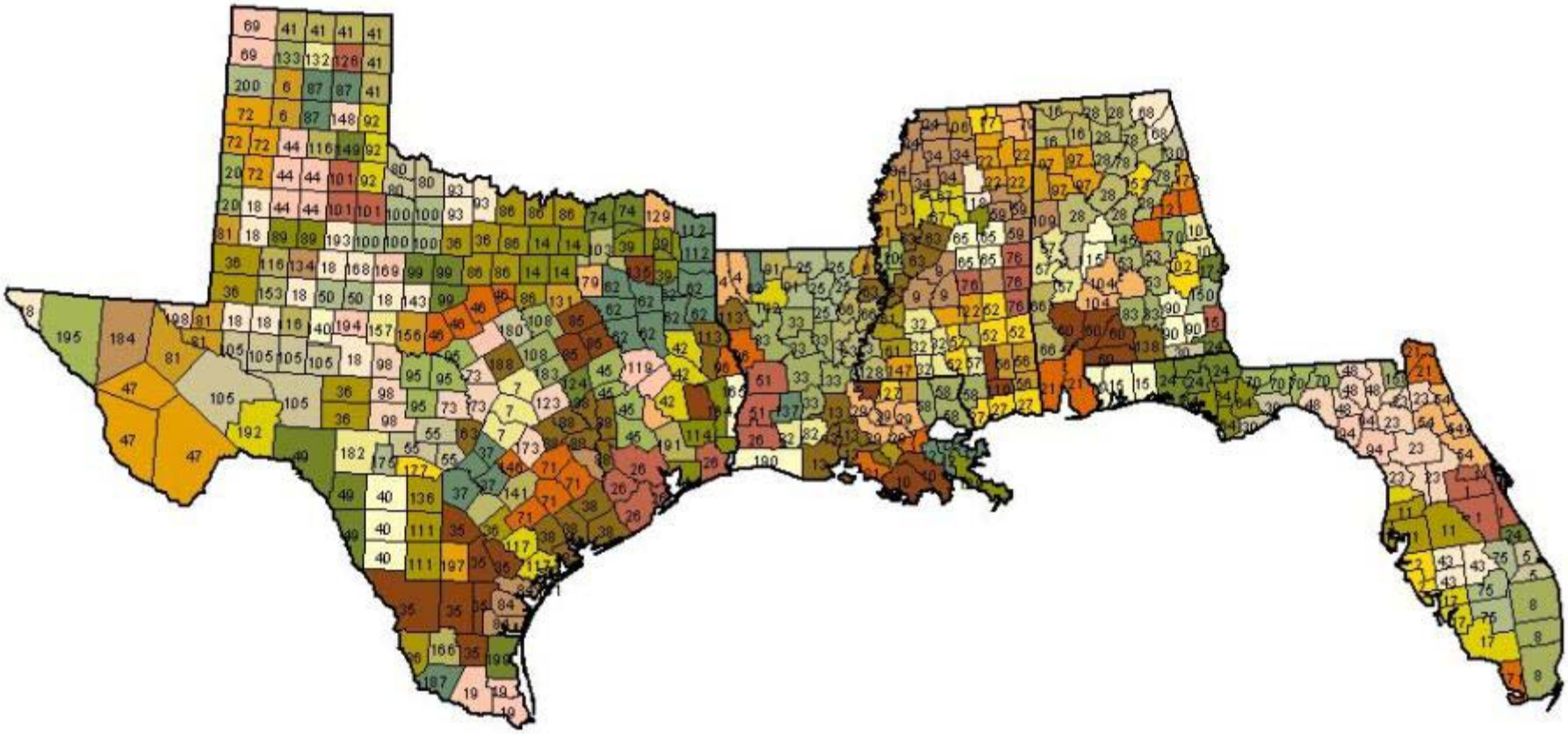
OGCluster 150



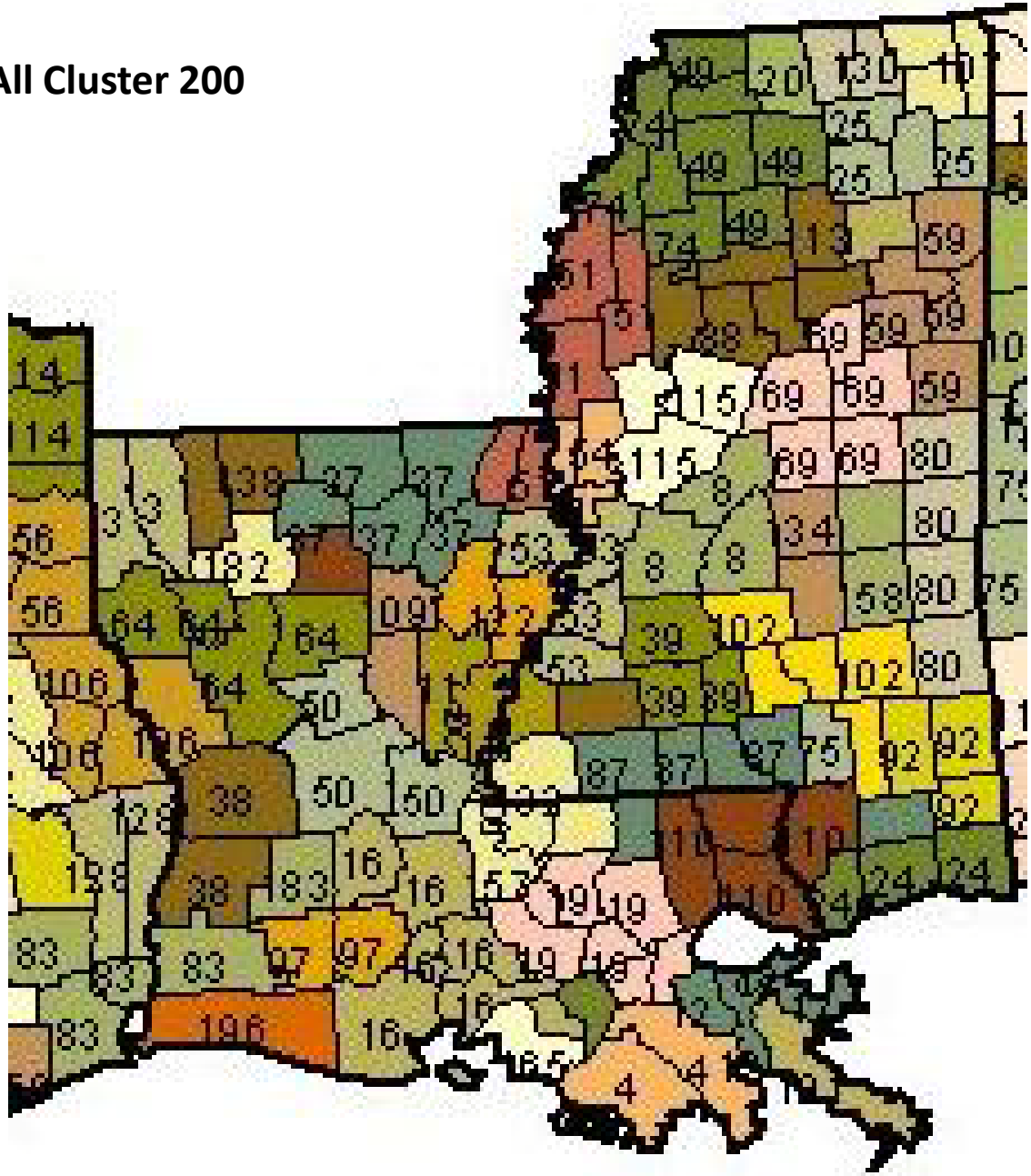
All Cluster 200



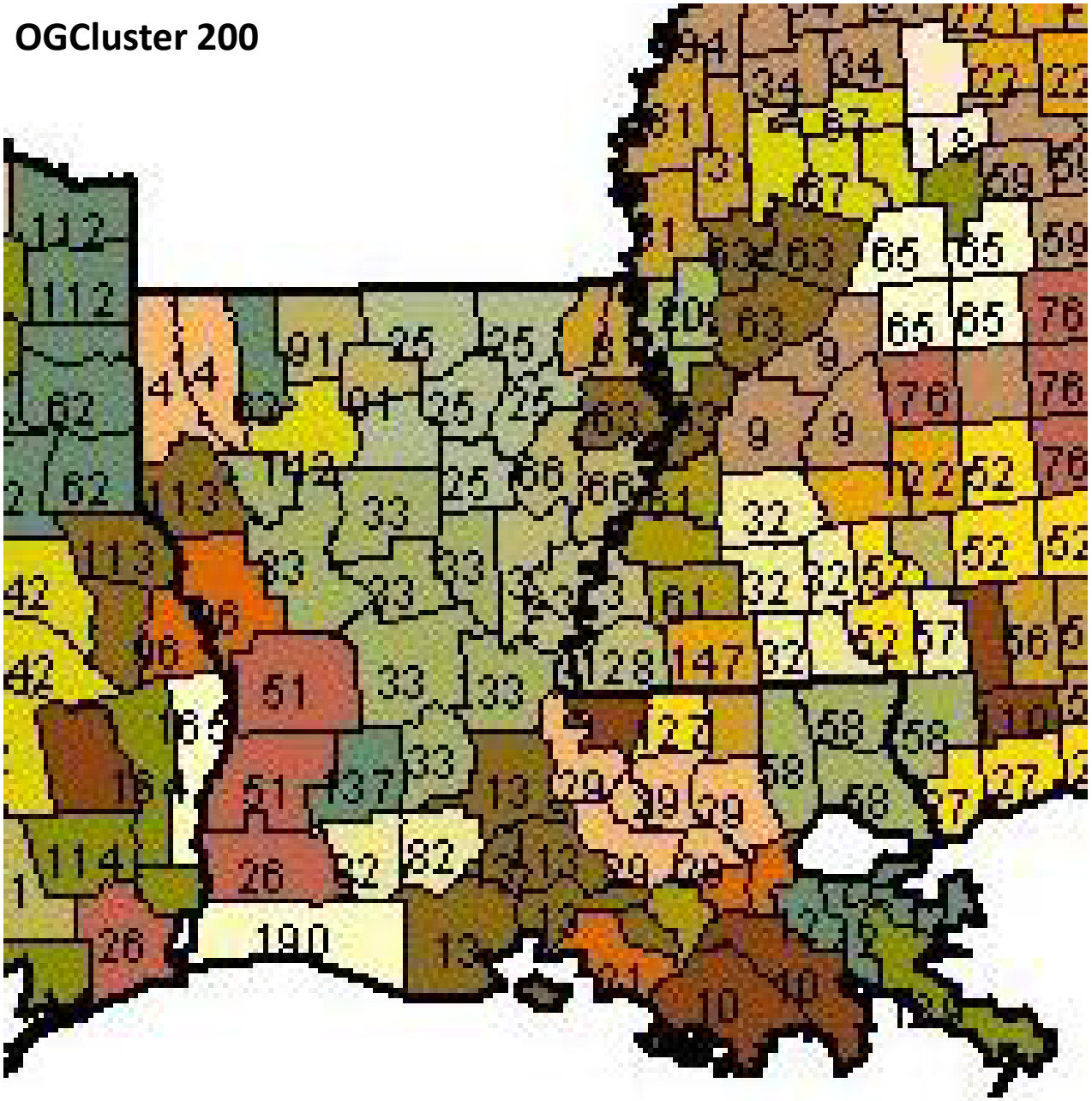
OGCluster 200



All Cluster 200



OGCluster 200



Clustering Methodology – Details

- We decomposed regions that were separate islands and crossed state boundaries
- Procedure chooses the maximum linkage of counties to create regions
- If a county 500 miles away has only ½ of 1% greater trade linkage to a reference county than an adjacent county, then the procedure will link the reference county with the 500 mile distant county as a region

Clustering Methodology – Details

- Issue with methodology
  - Do we want one half of one percent driving the creation of disjointed regions?
- We suggest an alternative to modify the procedure to weigh both trade linkage and geographic proximity

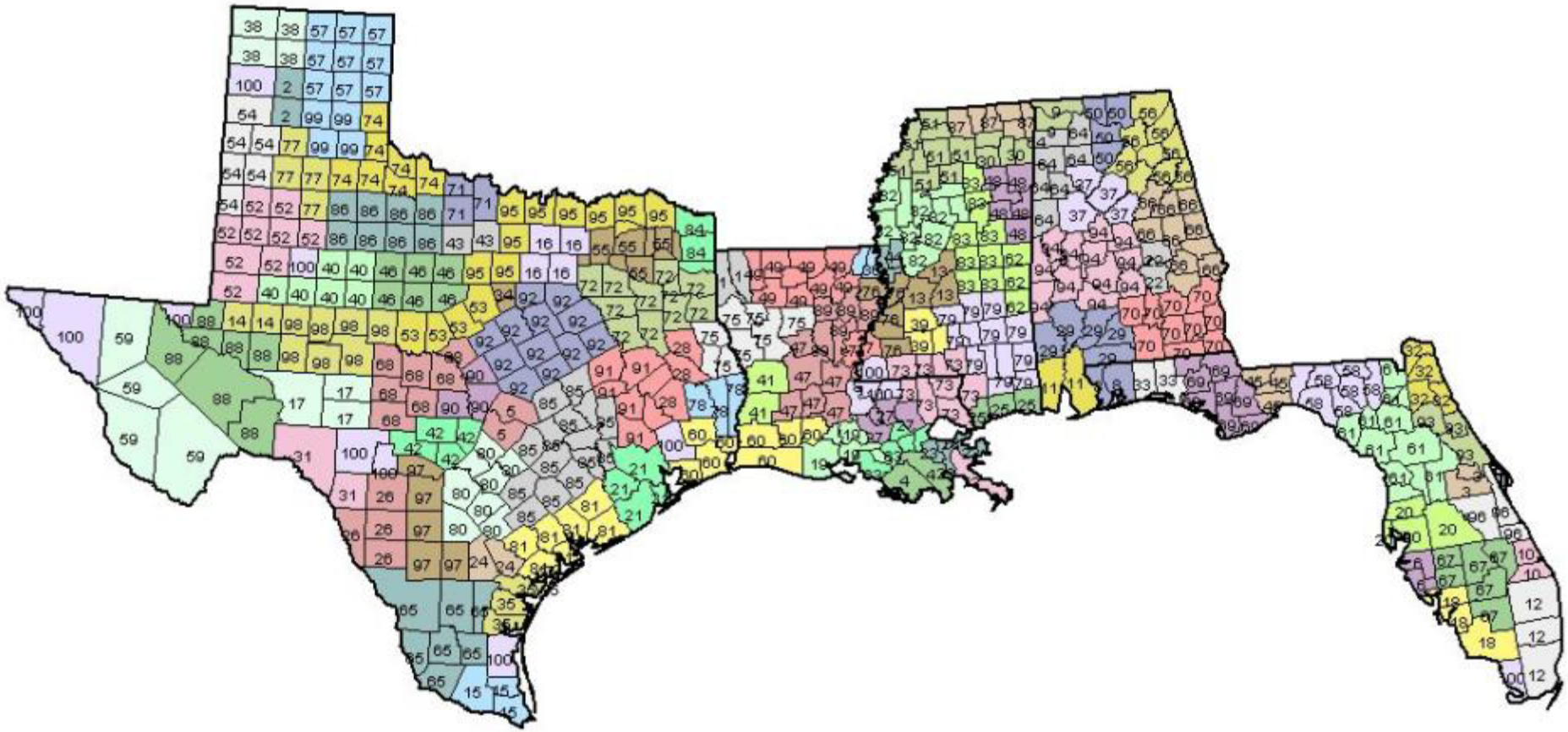
Weighted Clustering Methodology – Details

- We “weight” the trade linkage by the centroid distance between counties
- What is the result?
  - A county would need to have a greater % trade difference than the greater % distance difference of the reference counties’ next largest county trading partner

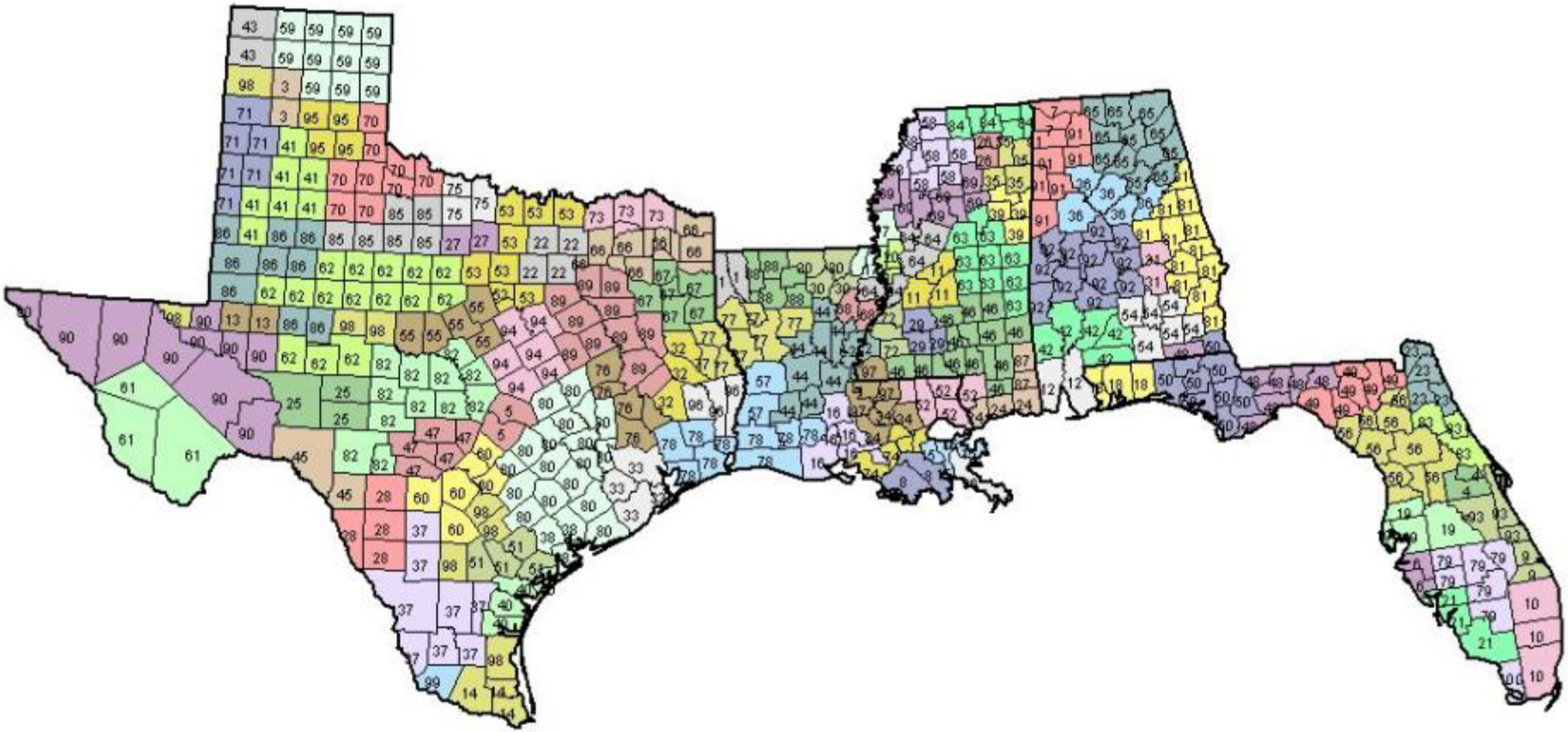
Weighted Clustering Methodology – Details

- Example
  - County A largest trading partner is with Counties B and C
  - 50% of A's total trade is with B and 30% is with C. County B is 100 miles from A and C is 50 miles from A.
  - Since B has only a 67% trade advantage over C but C has a 100% distance advantage over B, then A would first be linked with C.

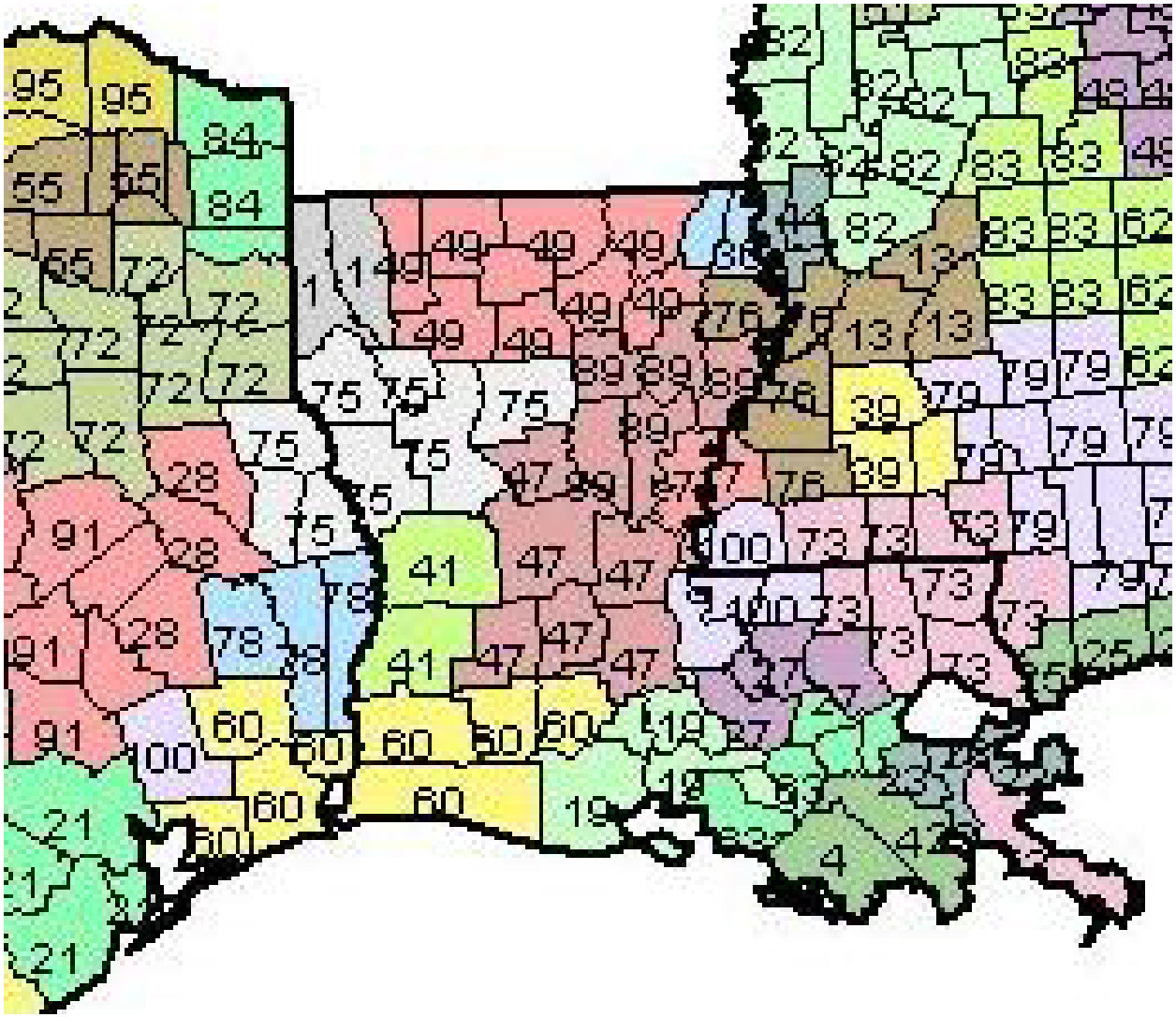
All Cluster Weighted 100



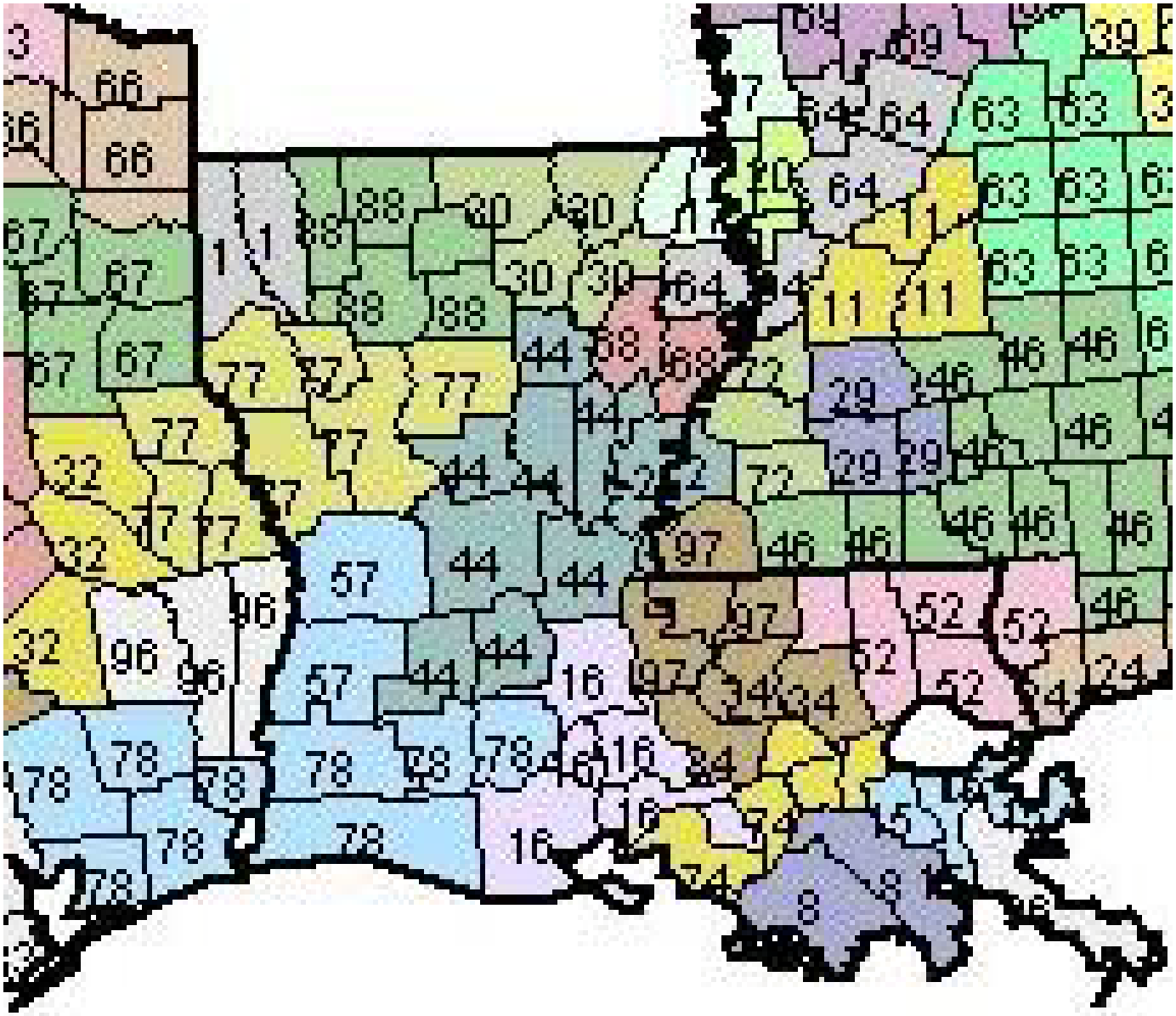
OGCluster Weighted 100



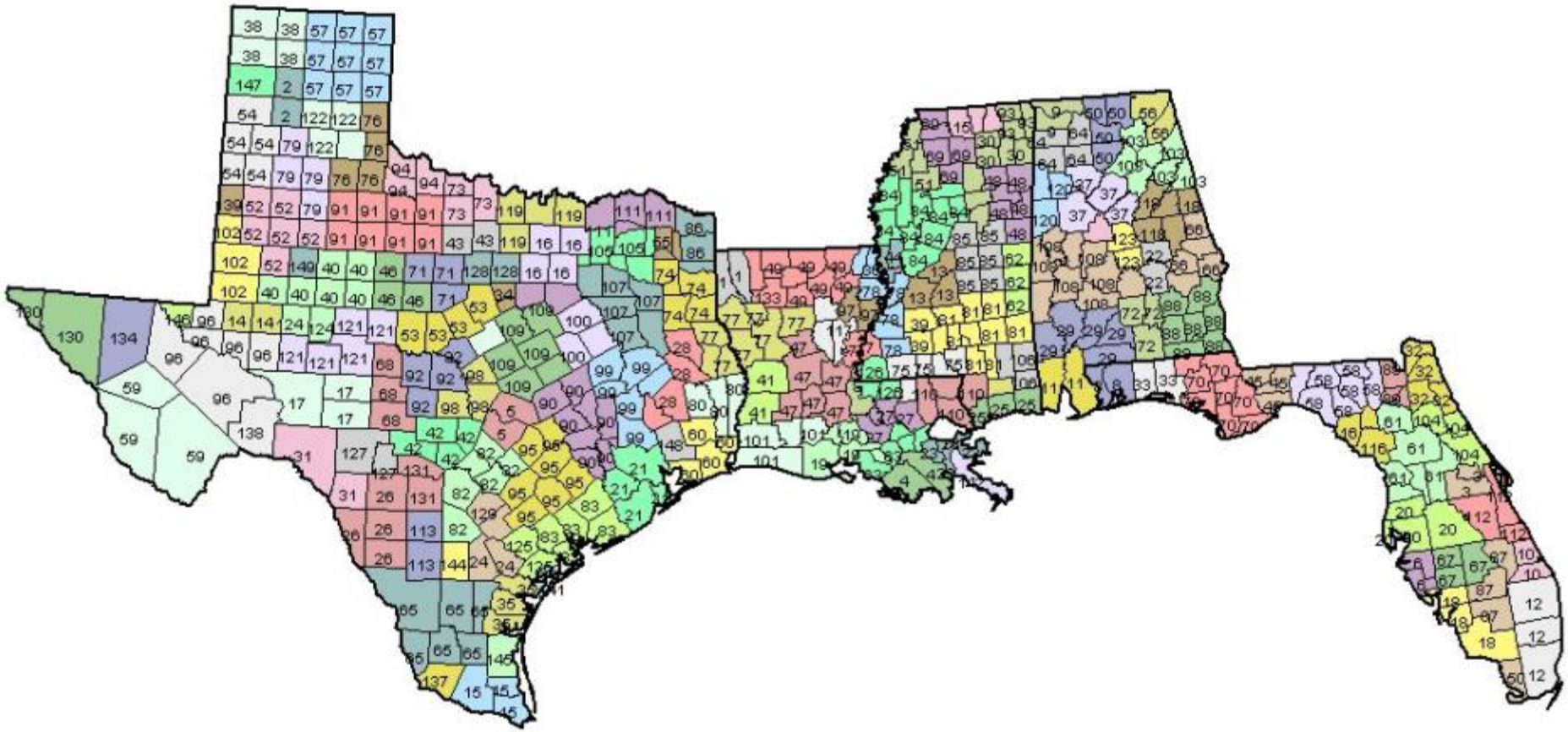
All Cluster Weighted 100



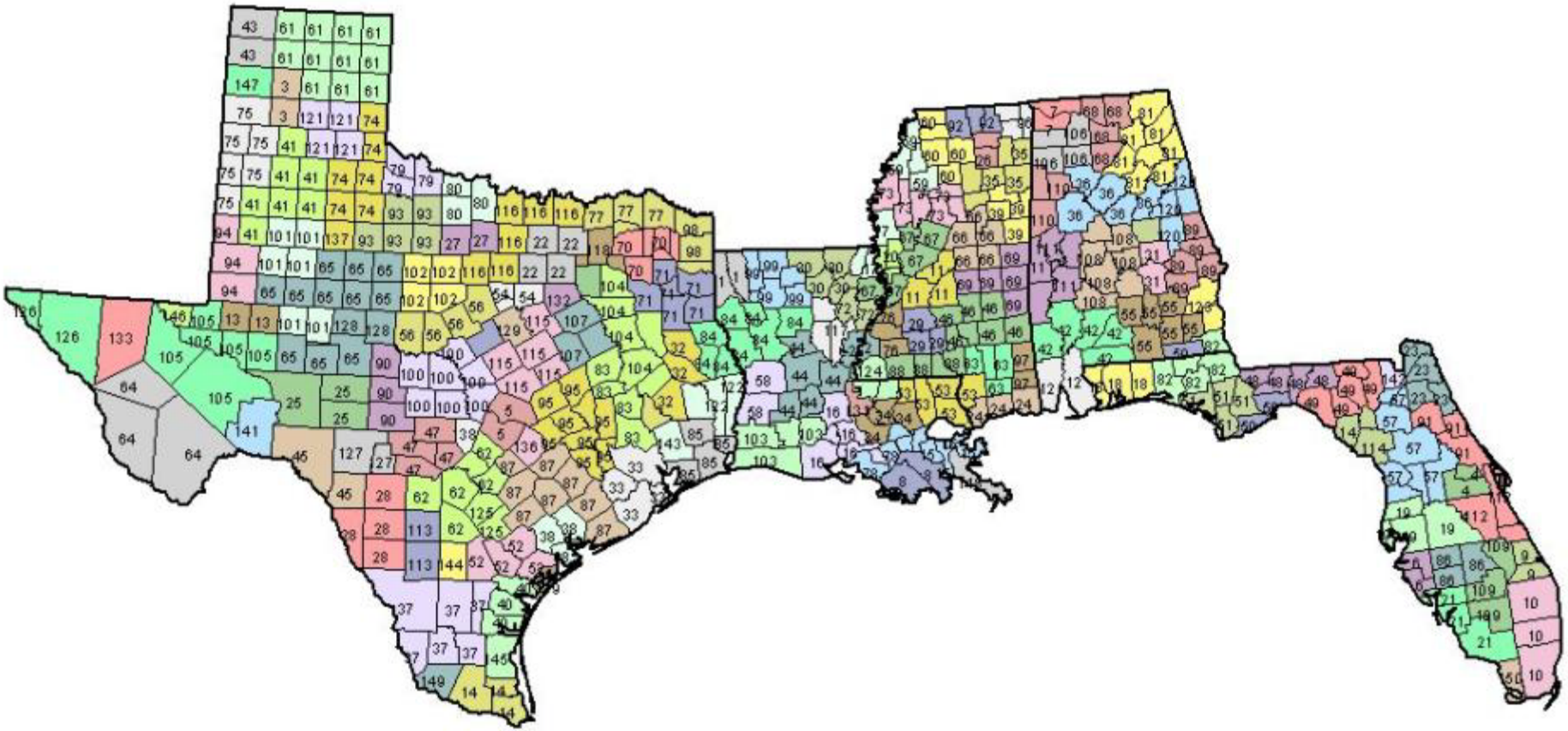
OGCluster Weighted 100



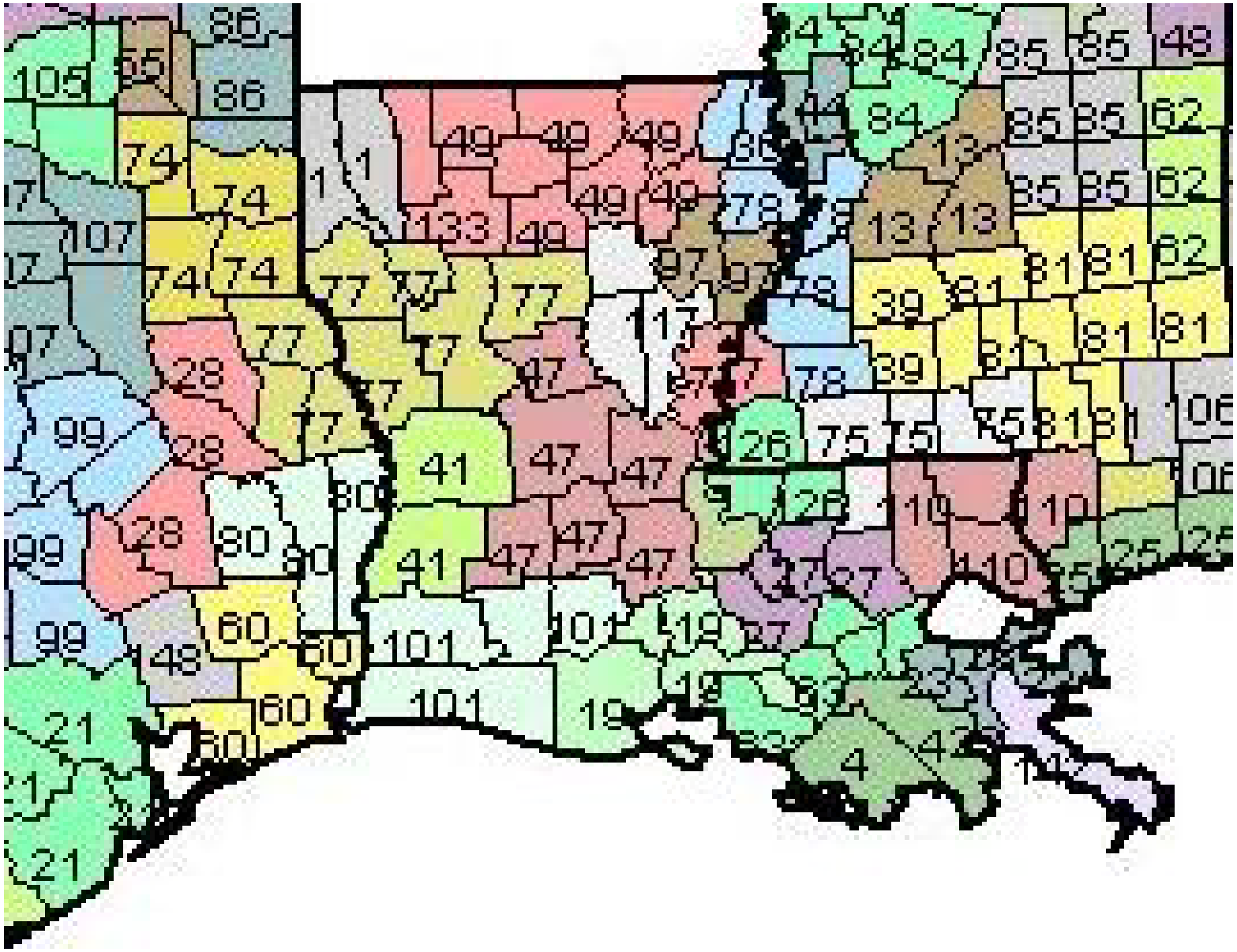
All Cluster Weighted 150



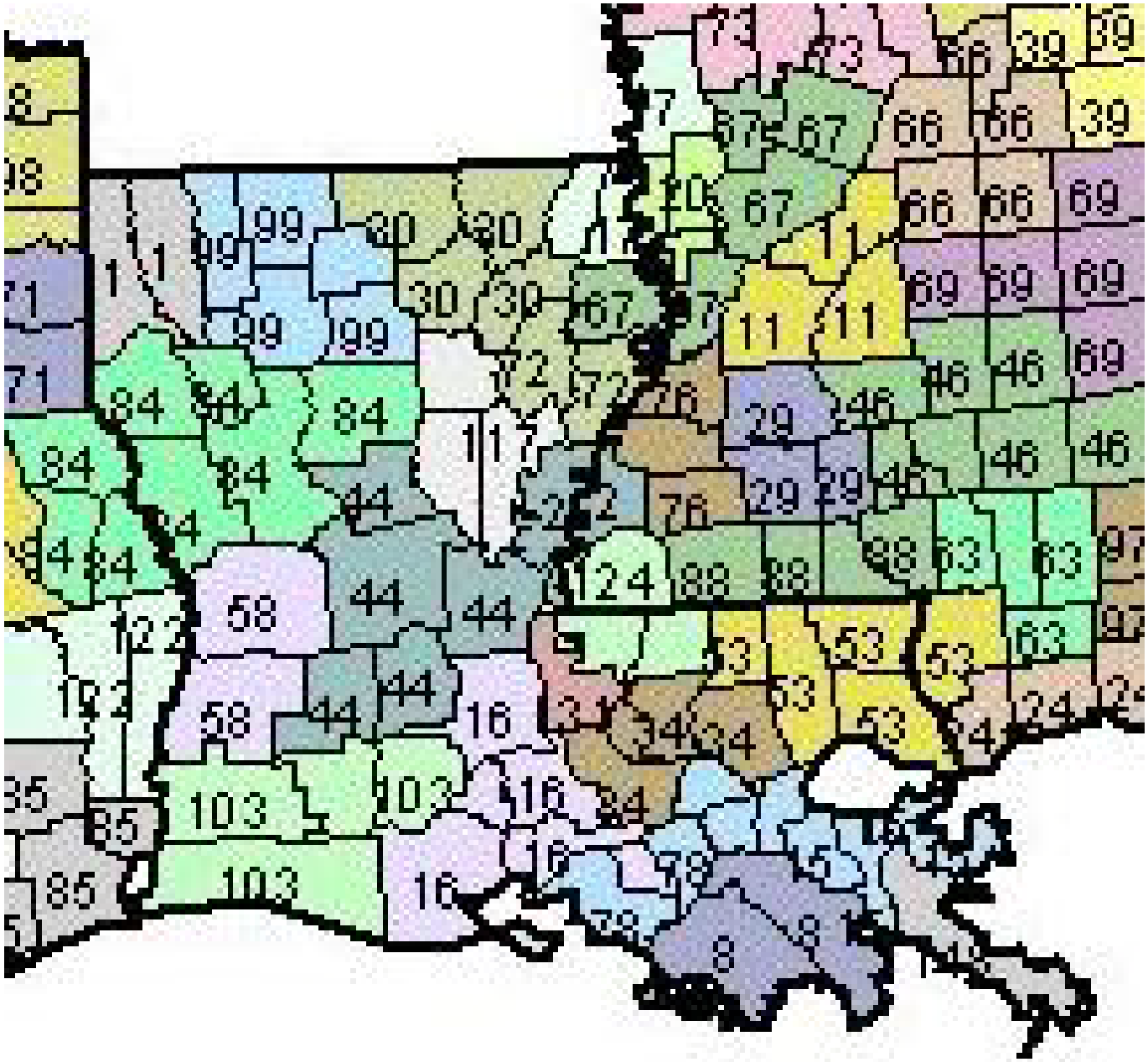
OGCluster Weighted 150



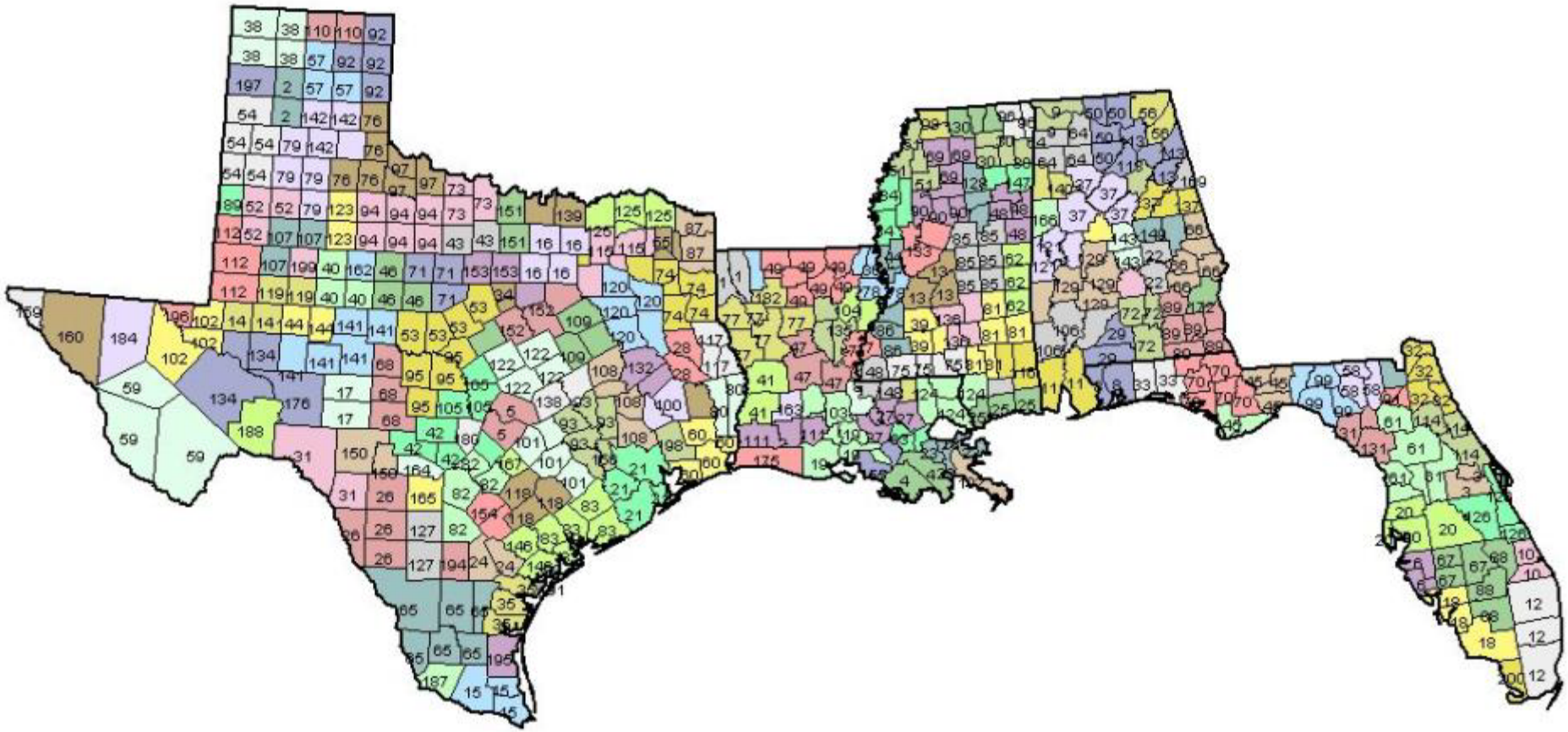
All Cluster Weighted 150



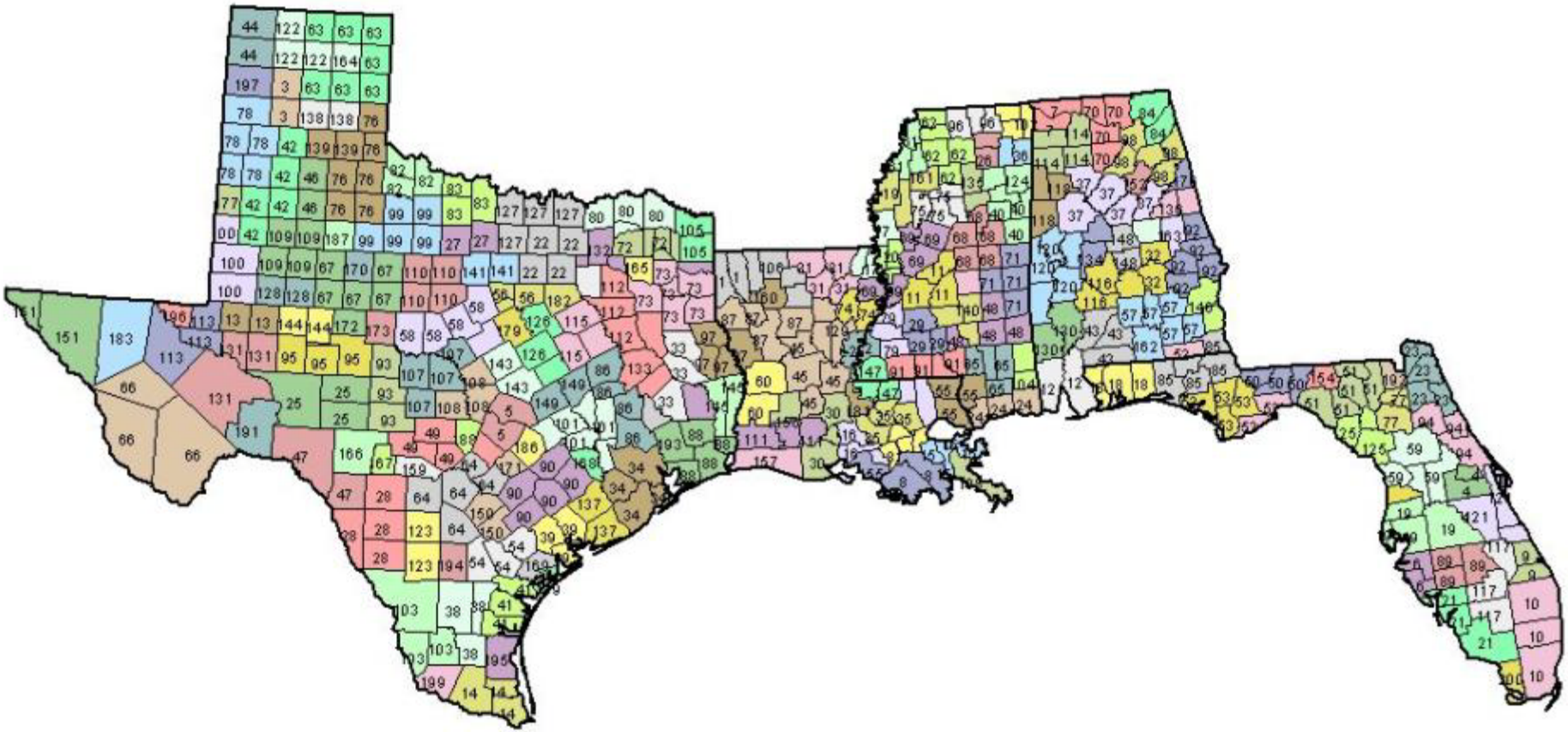
OGCluster Weighted 150



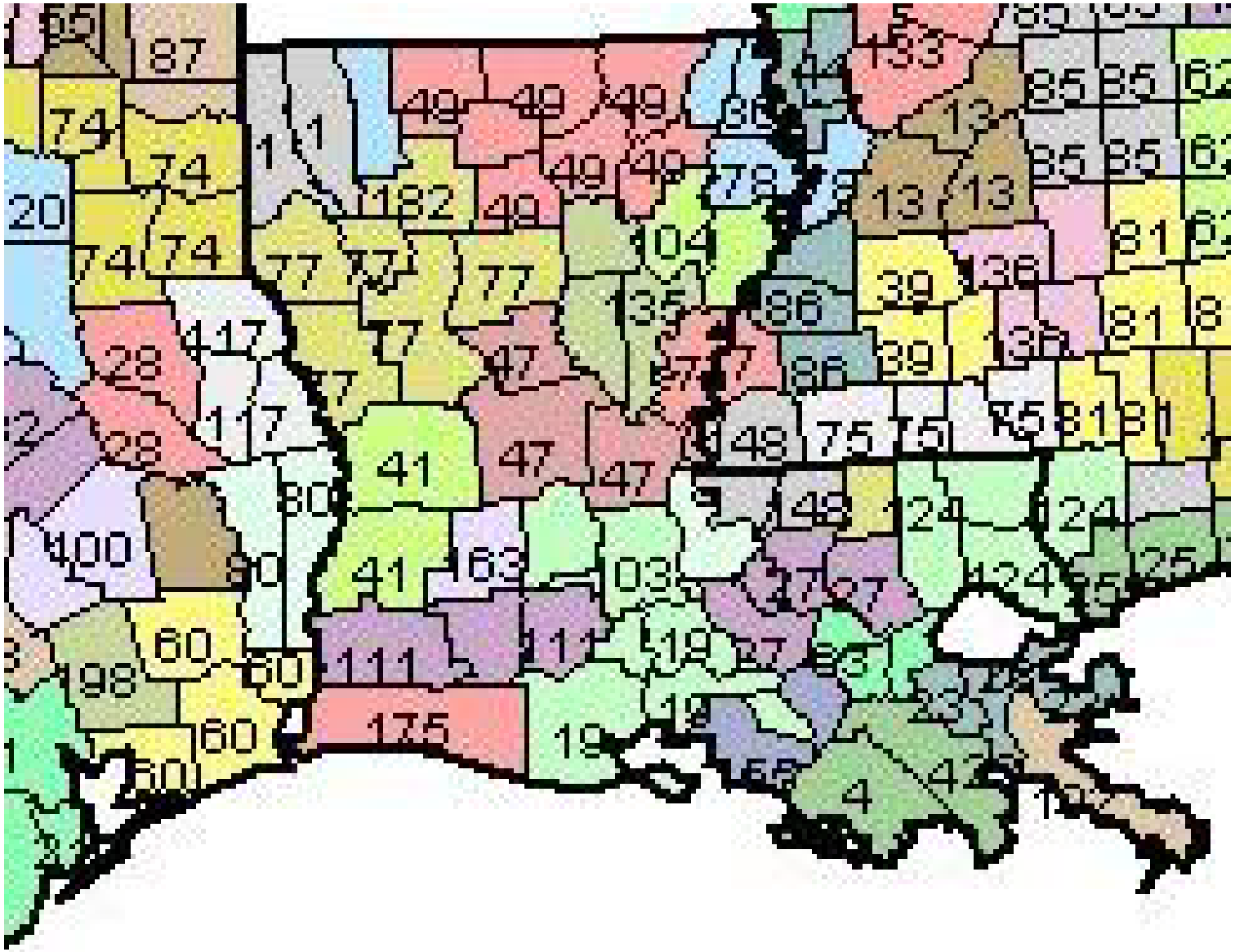
All Cluster Weighted 200



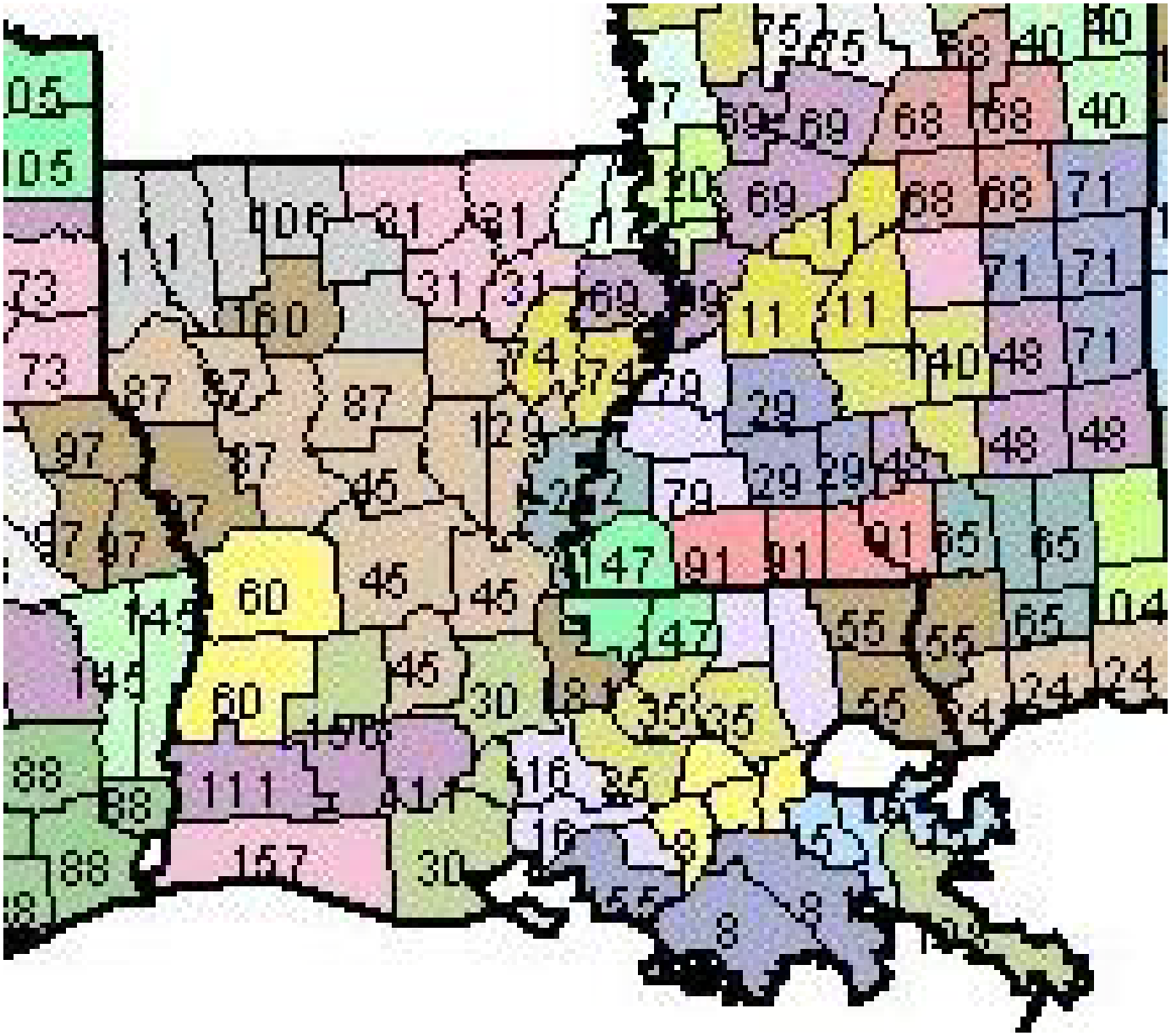
OGCluster Weighted 200



All Cluster Weighted 200



OGCluster Weighted 200

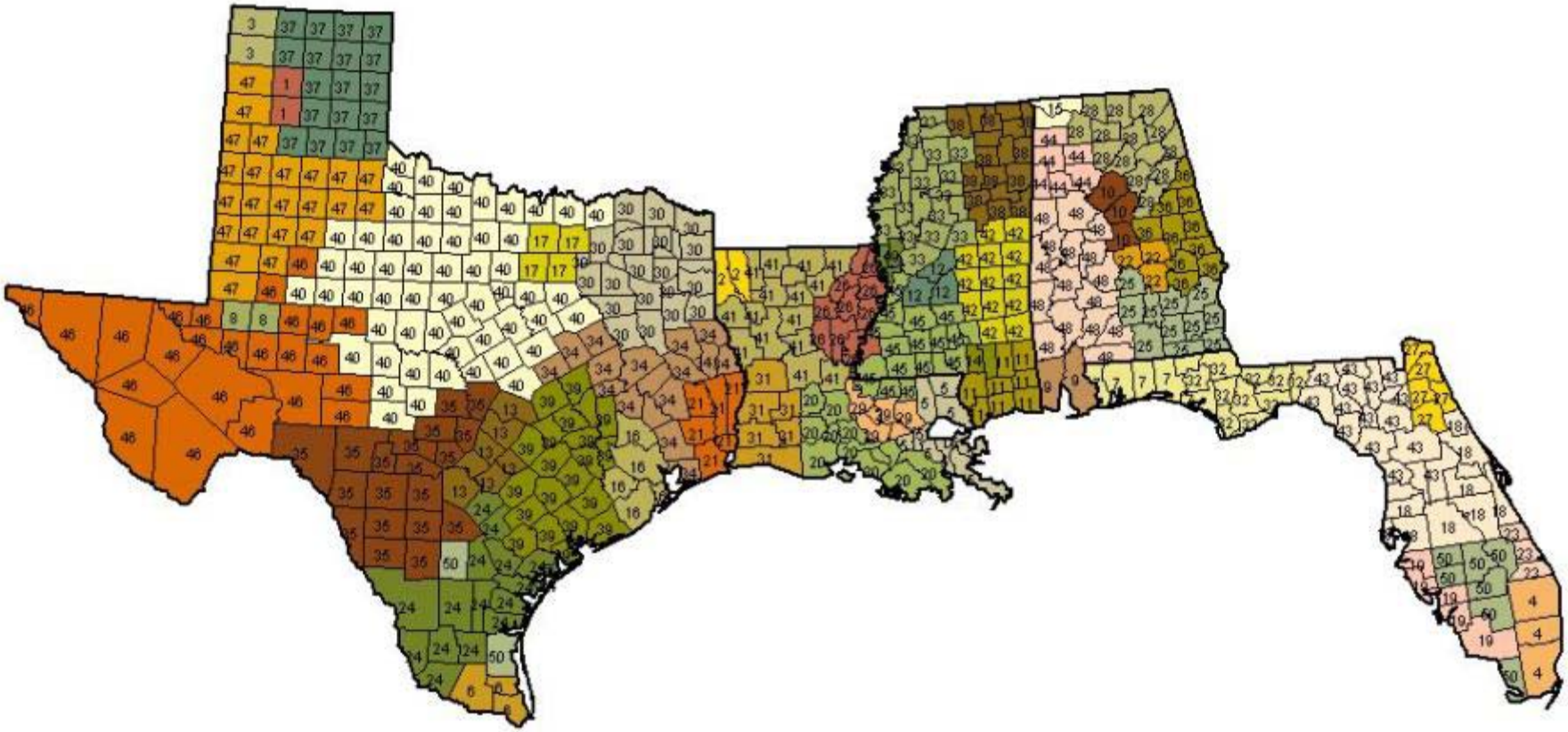


Next Steps

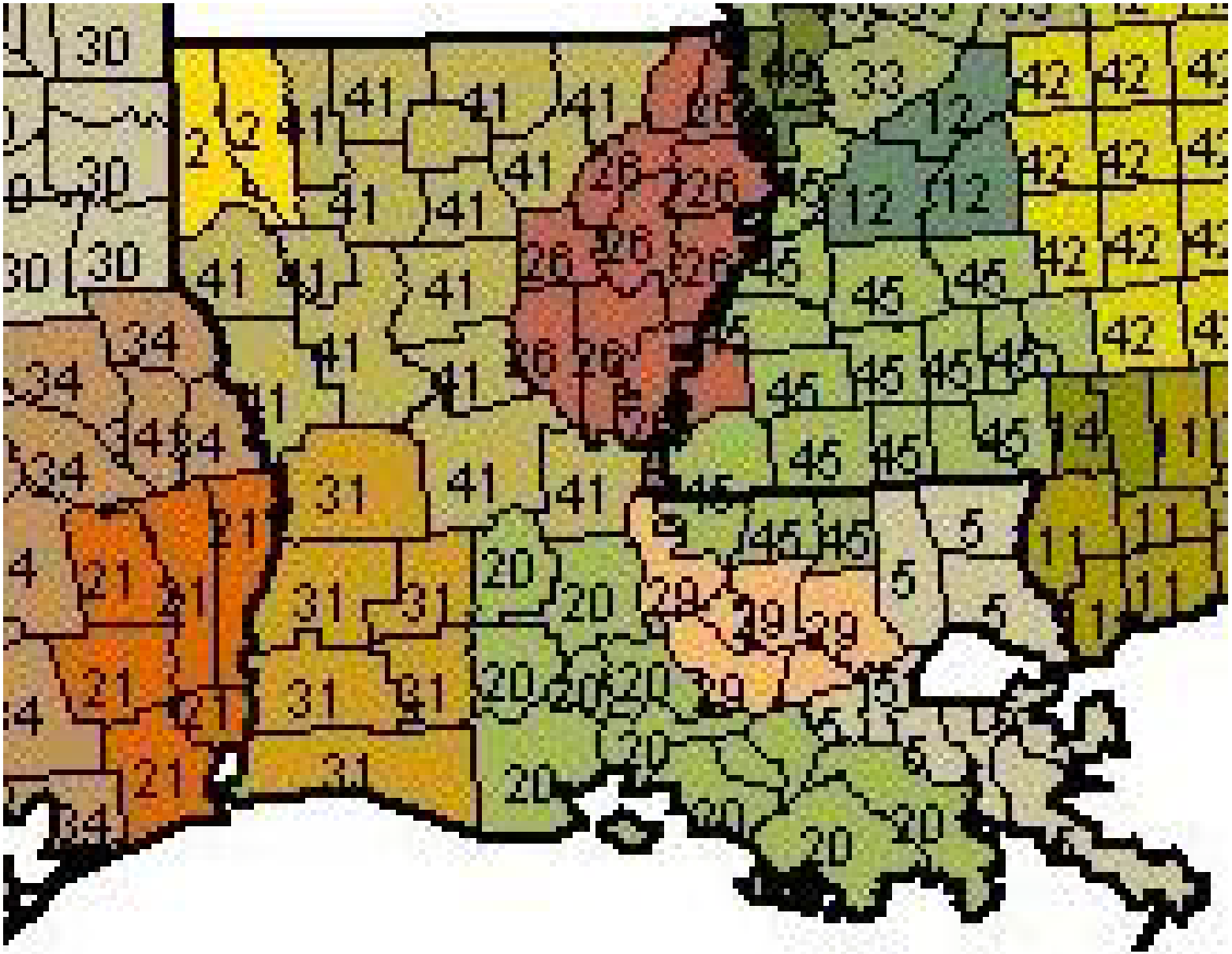
- Calculate key descriptive statistics for alternative regional classifications for use in subjective determination
- Other requests from project officers

Appendix:
Commuting Clusters

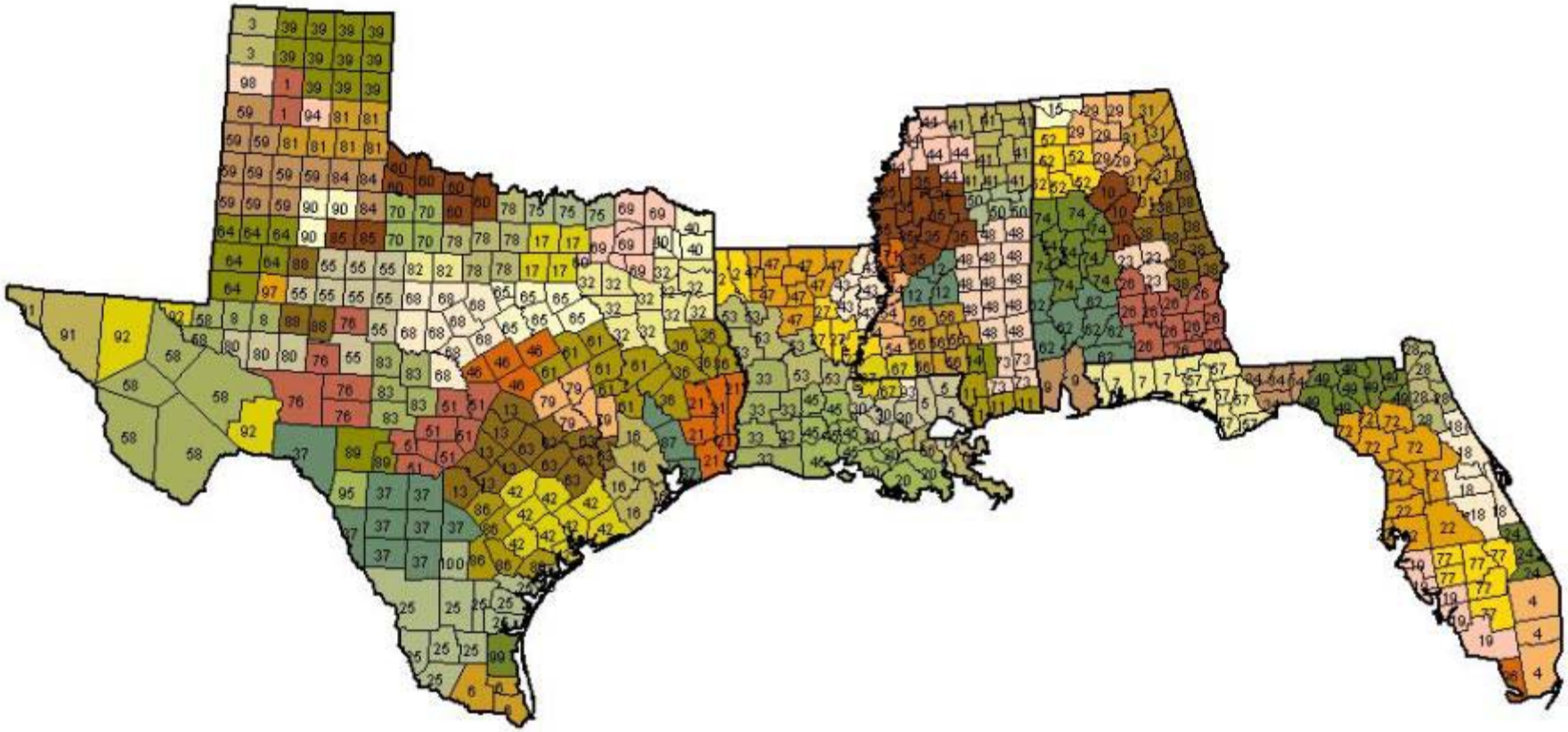
Commuter Cluster 50



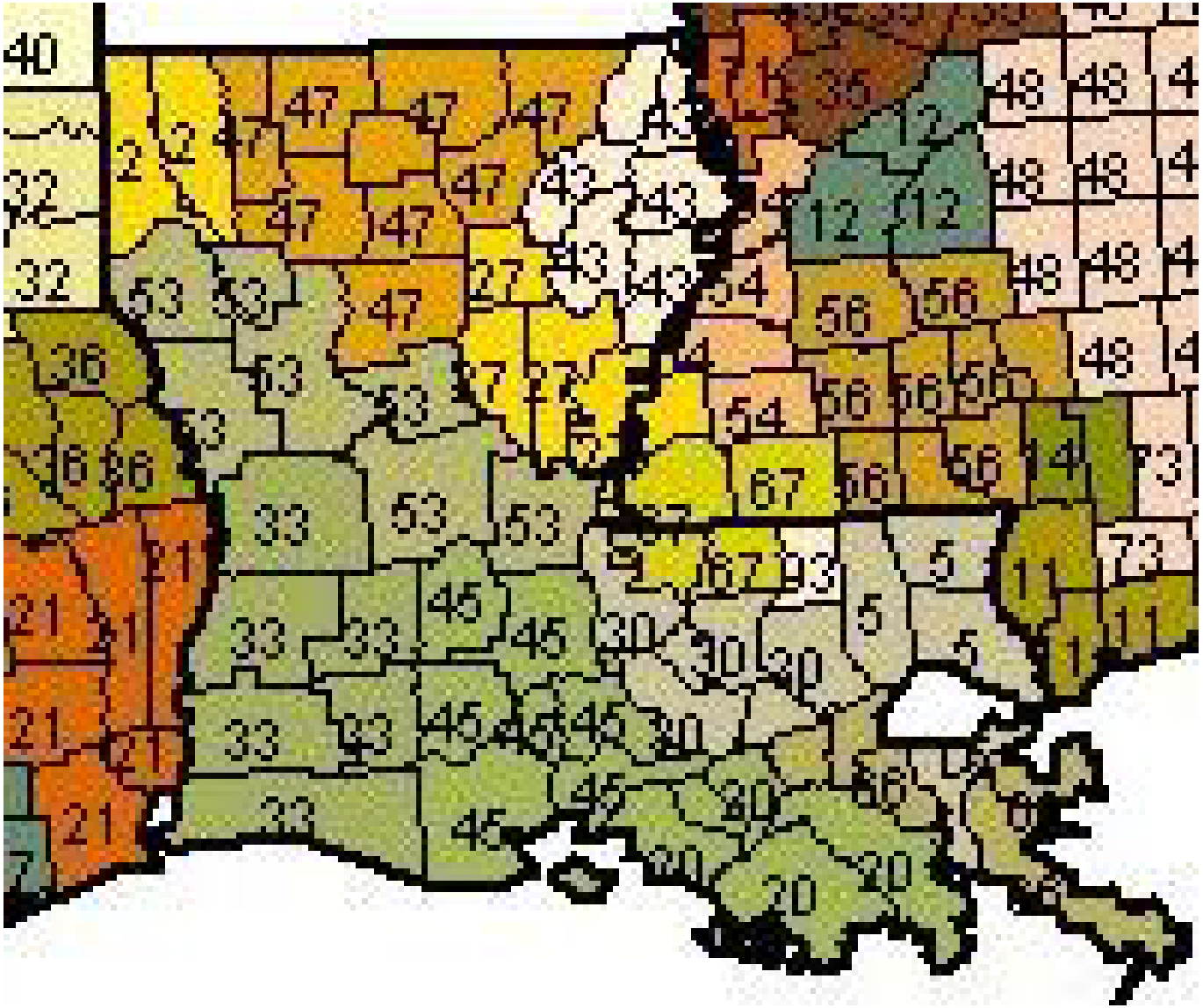
Commuter Cluster 50



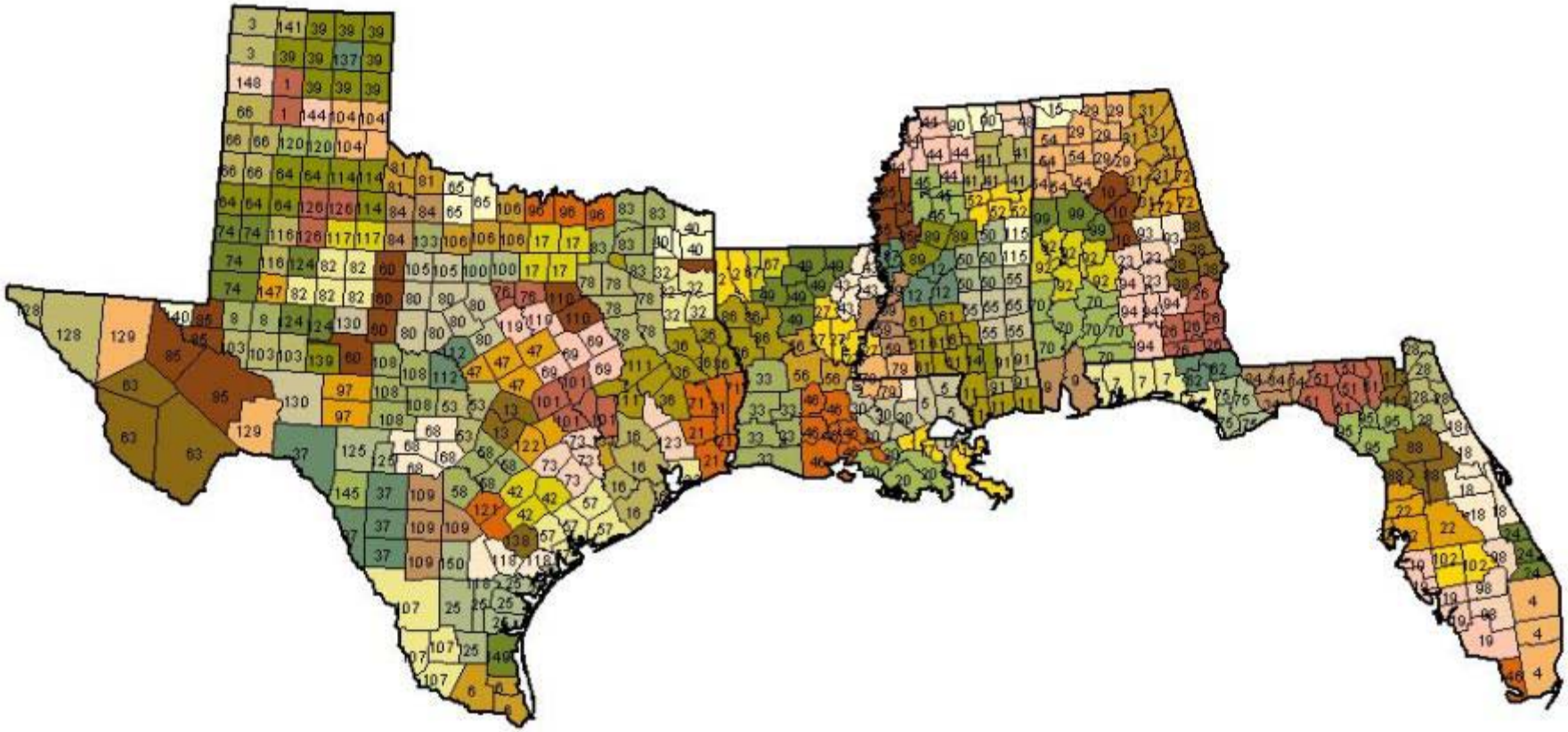
Commuter Cluster 100



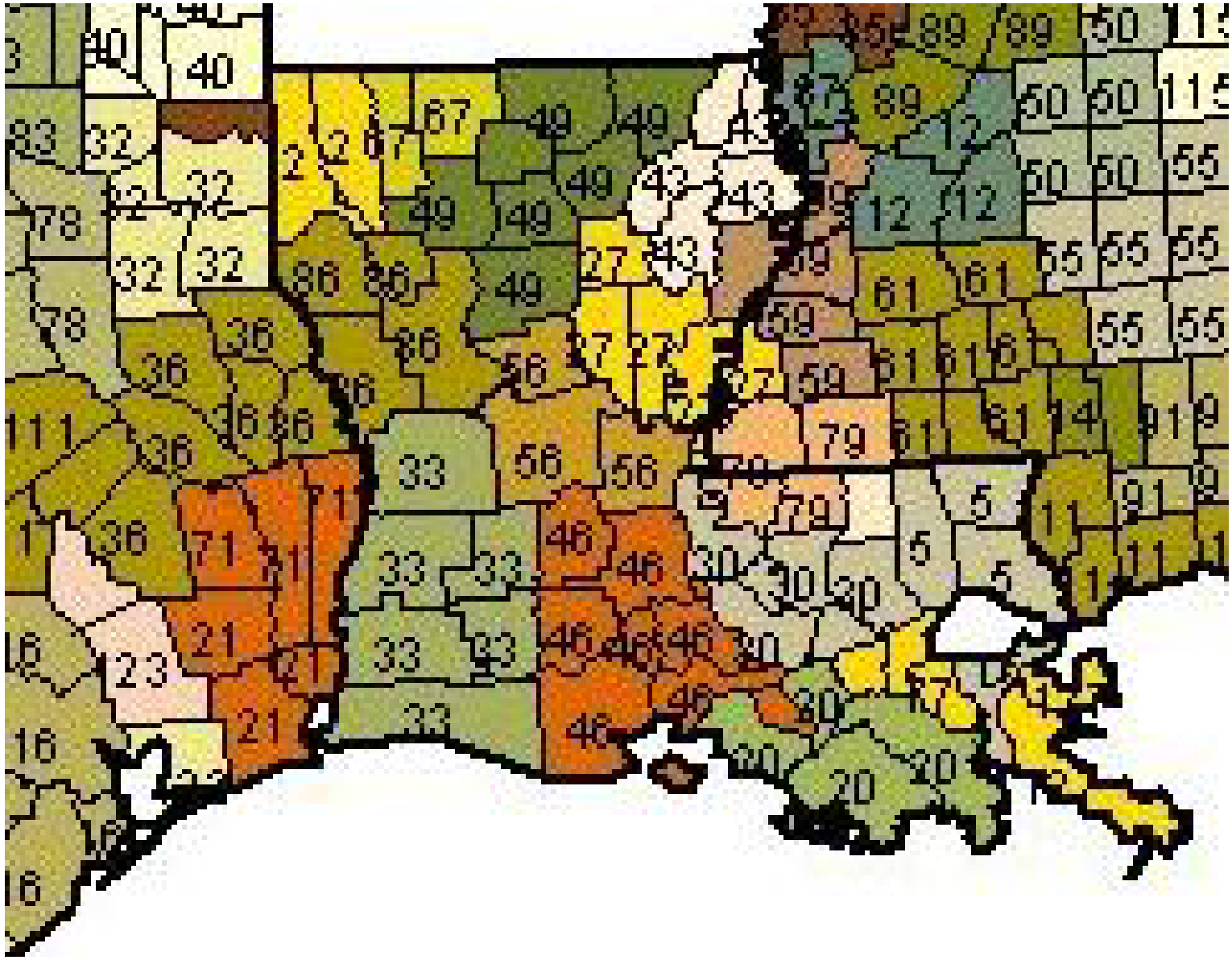
Commuter Cluster 100



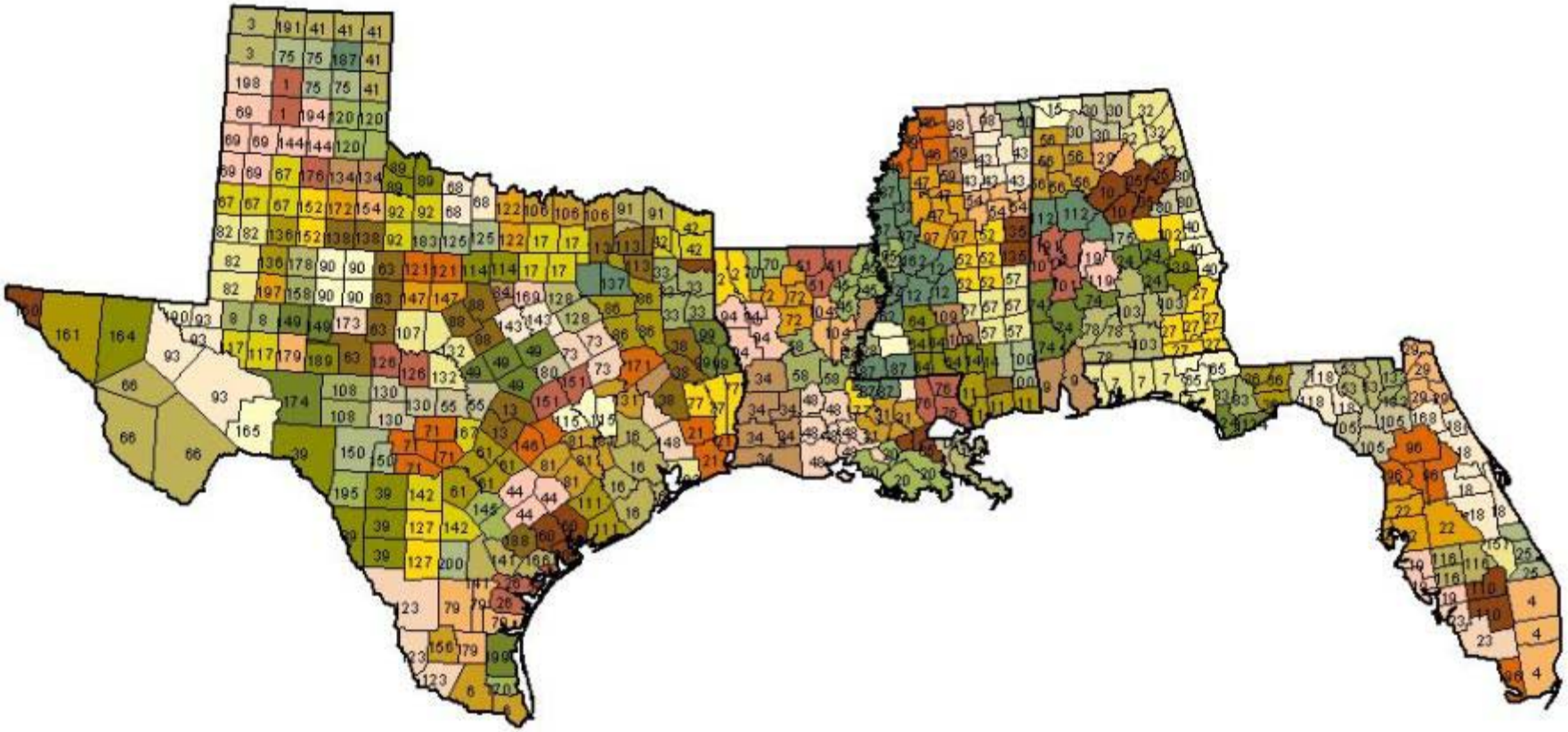
Commuter Cluster 150



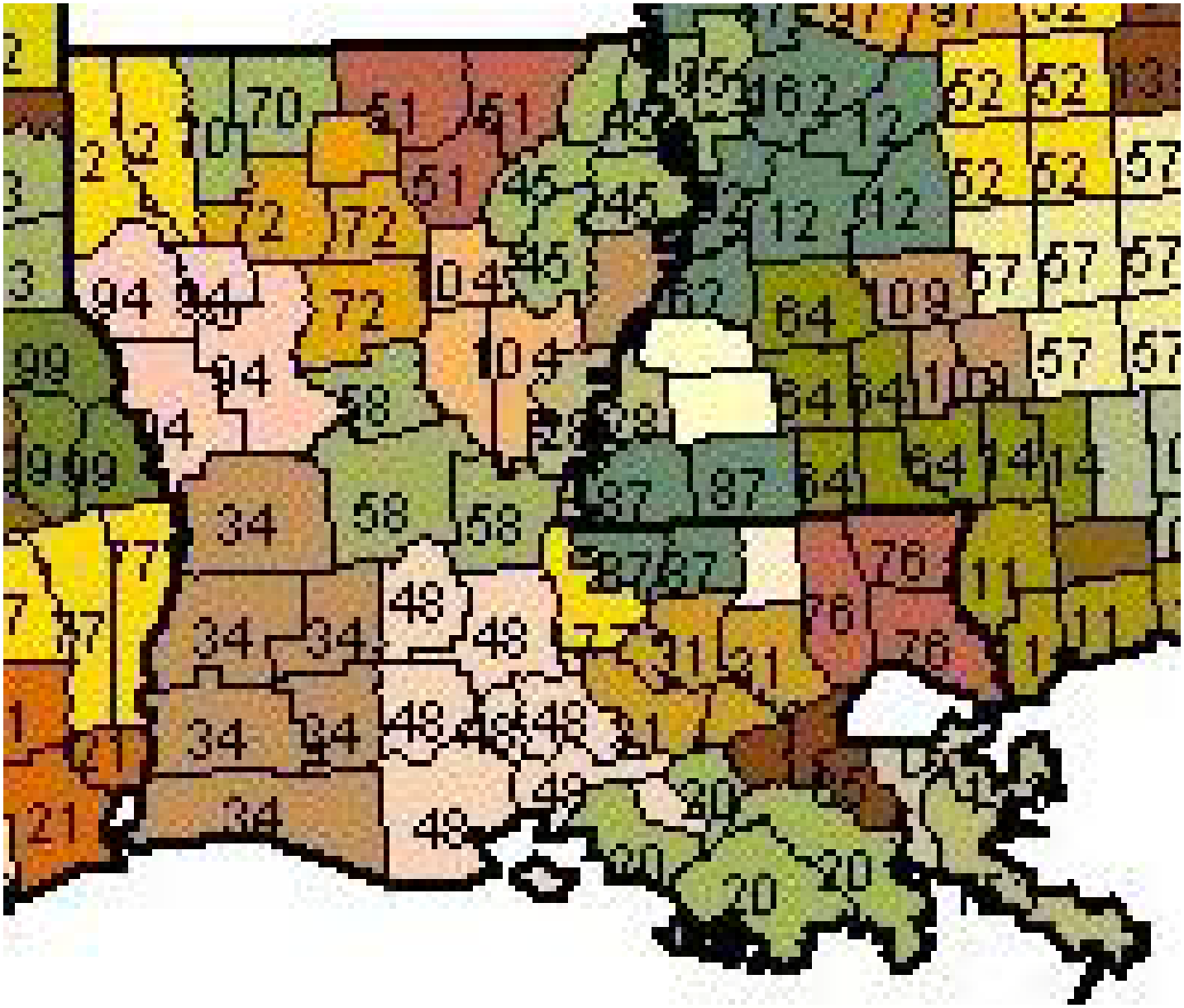
Commuter Cluster 150



Commuter Cluster 200



Commuter Cluster 200



References

Minnesota IMPLAN Group. 2008. IMPLAN Version 3 Blog. Online. www.implan.com, accessed March 10, 2010.

Saha, B., J. Manik, and M. Phillips. 2005. Upgrading the Outer Continental Shelf economic impact models for the Gulf of Mexico and Alaska (MAG-PLAN Study Report). U.S. Dept. of the Interior, Minerals Management Service, Herndon, VA. OCS Report MMS 2005-048. 164 pp.

U.S. Census Bureau. 2010. Local employment dynamics. Online. http://lehd.did.census.gov/led/led/led.html, accessed March 10, 2010.