ESRI International User Conference, 2006

Paper # 2044

Session Title: Economic Modeling, Analysis and Web Implementation

Paper Title

Geographic Information Systems (GIS) in Business and Industry Cluster Analysis: A Case Study of Indiana*

Prepared by

Christine Nolan, Senior Planner and Associate, Purdue Center for Regional Development
Indraneel Kumar, Spatial and GIS Analyst, Purdue Center for Regional Development

***** This paper is based on research supported by the U.S. Economic Development Administration: Unlocking Rural Competitiveness: The Role of Regional Cluster (EDA contract 11.312-EDA-7-15-2005)

Geographic Information Systems (GIS) in Business and Industry Cluster Analysis: A Case Study of Indiana

Abstract:
This research was funded by a grant from the Economic Development Administration (EDA) to study business and industry clusters in the U.S., with an emphasis on cluster development for rural areas. This paper explores the use of Geographic Information Systems (GIS) in business and industry cluster analysis, a regional and urban economic development tool. The objective is to develop a GIS database to visualize clusters over space and time and apply spatial analysis techniques. This GIS approach was instrumental to the research team in studying the geographic distribution of clusters in the USA.

Acknowledgements:
The authors would like to thank Fred Byon, Purdue Center for Regional Development (PCRD) and Rachel Justis, Indiana Business Research Center (IBRC) for research and mapping assistance. We are also grateful to the IBRC, our partners in the EDA cluster grant project, for providing data, GIS, and web-development support. We found *Hawth's Analysis Tools for ArcGIS* available at www.spatialecology.com very useful for expediting the necessary spatial analysis processes.

Introduction:
Traditional economic development methods focus primarily on individual industries or businesses. The objective is to retain or expand existing industries and businesses, and bring new firms into an area. Pursuing one firm at a time and addressing individual problems and needs consumes enormous efforts and resources.

In the last ten to fifteen years, business and industry cluster analysis has emerged as a new way of focusing on economic development by integrating regional differences in development and economic specialization. Cluster-based economic development calls for regional stewardship, cooperation, and capitalizing on synergies between existing industries, businesses, and institutions. An increasing number of states and regions in the U.S. and overseas have modified their economic development strategies to focus and capitalize on the business and industry clusters where they have, or would like to have, a competitive advantage.

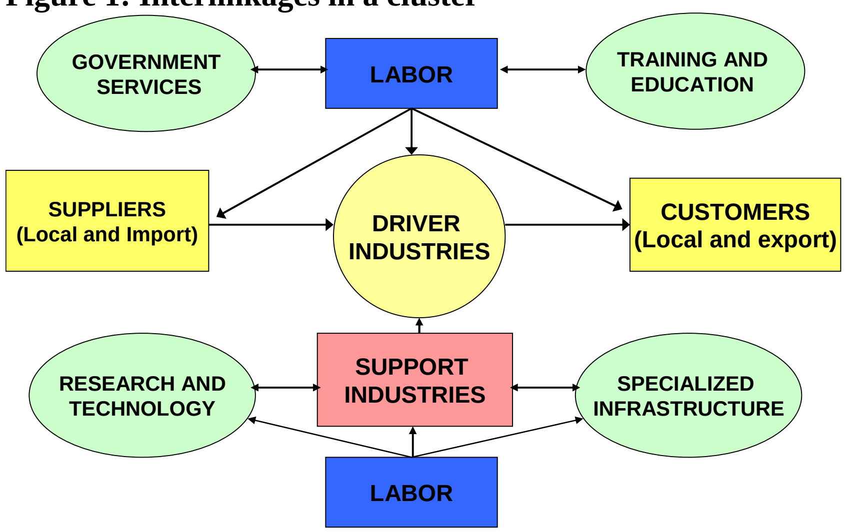
What are business and industry clusters?
"Industry clusters are *geographic* concentrations of competing, complementary, or interdependent firms and industries that do business with each other and/or have common needs for talent, technology, and infrastructure. The firms included in the cluster may be both competitive and cooperative. They may compete directly with some members of the

cluster, purchase inputs from other cluster members, and rely on the services of other cluster firms in the operation of their business."[1]

Compared to traditional economic development methods, business and industry clusters offer additional approaches:

- Solve groups of industry problems/needs
- Reveal groups of industries that have similar workforce needs
- Build sustained business-to-business connections
- Invest and assist groups of firms to build synergy and economic impact

Figure 1: Interlinkages in a cluster



Who are using business and industry cluster tools?

In the U.S., the states of Arizona, Louisiana, California, Oregon, Washington, New York, Massachusetts, Connecticut, Minnesota and Indiana (or localities and regions in these states) are among many who have commissioned cluster studies and who have implemented or are implementing cluster strategies. Internationally, Australia, India and many European Union countries are also conducting economic development initiatives based on cluster analysis. In addition to local governments, regional councils, economic development organizations, and non-governmental organizations are participating in economic development based on business and industry clusters. In developing nations, institutions such as the World Bank and the United Nations Industrial Development Organizations are assisting in developing cluster-based economic development approaches.

How is cluster analysis carried out?

Cluster analysis involves identifying those businesses and industries that form a cluster. Researchers have used various quantitative and qualitative methods to group businesses

---

[1] Industry Clusters: An Economic Development Strategy for Minnesota Preliminary Report. http://www.hhh.umn.edu/centers/slp/economic_development/econdev_strategy_industry_cluster.pdf; University of Minnesota Extension Service, 1999. Accessed June 19th, 2006.

and industries into clusters. These include analysis of input-output tables to determine value-chain linkages, surveys, location quotient analysis, shift-share methods, literature review, etc. In our case, the research team did an extensive literature review of the existing cluster studies and applied local knowledge of the area, experiences from past projects, and expertise to group the industries.

Table 1: List of business and industry clusters

Clusters and Sub-Clusters Studied in this Project	
1	Advanced Materials
2	Agribusiness, Food Processing and Technology
3	Apparel and Textiles
4	Arts, Entertainment, Recreation and Visitor Industries
5	Biomedical/Biotechnical (Life Sciences)
6	Business and Financial Services
7	Chemicals and Chemical-Based Products
8	Defense and Security
9	Education and Knowledge Creation
10	Energy (Fossil and Renewable)
11	Forest and Wood Products
12	Glass and Ceramics
13	Information Technology and Telecommunications
14	Transportation and Logistics
15	Manufacturing Supercluster
	Primary Metals
	Fabricated Metal Products
	Machinery
	Computer and Electronic Products
	Electrical Equipment, Appliance and Components
	Transportation Equipment
16	Mining
17	Printing and Publishing

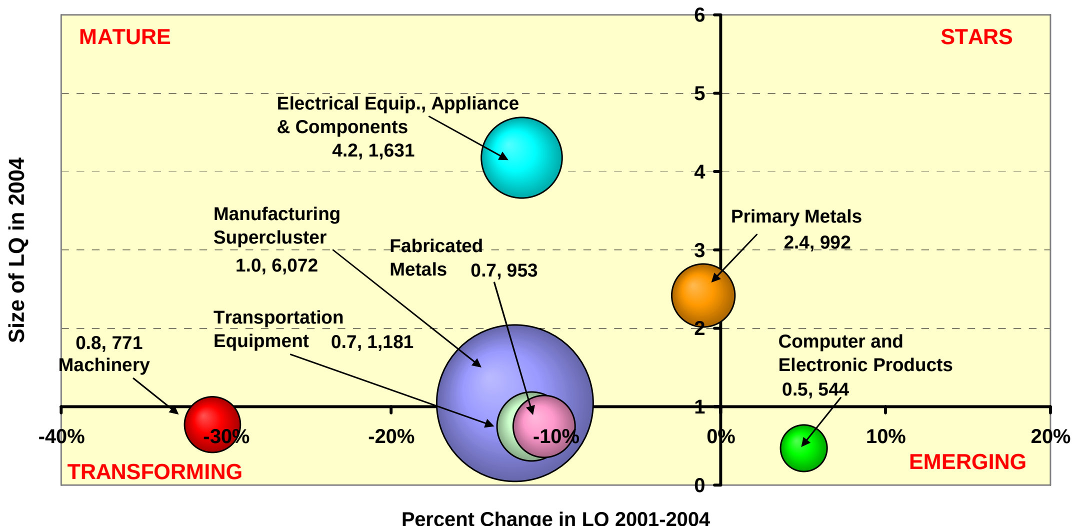
Measurement of growth rates and location quotients for each industry in a cluster, as well as the total cluster, is the first step towards determining where the region's comparative advantage lies.[2] Location quotients show where industry sectors in particular localities are more strongly represented than they are in the nation as a whole. *To the extent that a particular location quotient is greater than 1, the area is considered to be more* *specialized* *in that industry or cluster than the nation is, and industries in the cluster are assumed to be producing for export as well as local consumption.*

[2] Location quotients may be calculated on the basis of an industry's or cluster's employment, its number of establishments, or certain other measures of economic activity. This study computed LQs based on employment.

In this project, the research team used data from the Covered Employment and Wages (CEW) annual data sets for Indiana and the nation, with a base year of 2001 and comparison year of 2004. The base year of 2001 was chosen because this is the first year at which CEW data are available on the basis of NAICS codes rather than SIC codes.

After calculation of LQs for each cluster and industries within each of the clusters, the specialization of clusters and changes in specialization were shown by using a method developed by the Boston Consulting Group.[3] Figure 2 shows LQs and percent change in LQ in the manufacturing supercluster for Economic Growth Region (EGR) 8, one of the 11 economic growth regions in Indiana.

Figure 2: Subclusters in Economic Growth Region 8's manufacturing supercluster



Note: The supercluster represents the sum of all 6 subclusters

Source: Purdue Center for Regional Development, using data provided by Indiana Business Research Center, 2006

Here, the values associated with each cluster are the LQ and size of the cluster as measured by employment. For example, primary metals cluster has an LQ of 2.4 and 992 employees. The classification of Stars, Emerging, Mature, and Transforming means the following:

- Specialized, and becoming increasingly specialized ("Stars")
- Not specialized, but becoming increasingly specialized ("Emerging")
- Specialized, but decreasing in degree of specialization ("Mature")
- Not specialized and decreasing in degree of specialization ("Transforming")

This initial analysis of a region's clusters begins to give policymakers and other stakeholders some idea of what is happening in the local economy, and which industries and clusters might need support to mitigate decline or to give an extra boost to growth.

---

[3] Boston Consulting Group; http://www.netmba.com/strategy/matrix/bcg/.

What is the role of Geographic Information Systems (GIS) in business and industry cluster analysis?

Geographic Information Systems (GIS) has many uses in economic development and particularly in business and industry cluster analysis. Of foremost importance is the capability to map the distribution of industry and business clusters over space and time, providing insights into their distribution patterns and temporal changes. A mapping at two different points of time sheds light on changes in the distribution pattern in a given cluster. This kind of mapping is a useful tool in policy-making or analyzing impacts of a policy on clusters.

The research team used ArcGIS to develop a GIS database for business and industry clusters of the whole U.S. at the county level; Indiana at the level of economic growth regions; and county level mapping for a pilot study area, Economic Growth Region 8 (EGR 8) in southern Indiana. GIS helped in mapping employment and establishment data for each of the clusters. Following are the examples of point-pattern mapping and spatial analysis used in this research.

Mapping:

- Dot-density mapping at the county level for whole USA
- Point-symbol mapping at the county level for whole USA
- Graduated symbol mapping for economic growth regions and county level in Indiana
- Geocoding for Economic Growth Region 8 in Indiana

Spatial analysis:

- Nearest-neighbor

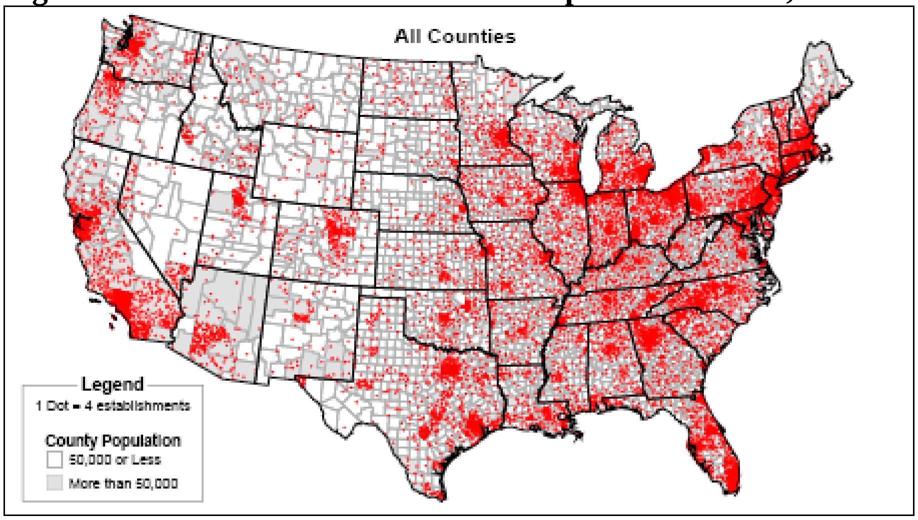
Mapping:

Dot-density Mapping

Dot-density maps were created for employment and establishments for each of the 17 business and industry clusters. These maps (reflecting 2004 data) show a snapshot of distribution of clusters over geographic regions. Dots are generated randomly within the county boundaries, the lowest level of geography. These maps provide useful insights at higher geographies such as states or census regions. Regardless of randomness in distribution of dots within counties, researchers can infer patterns of distribution and geographic clustering from the dot-density maps.

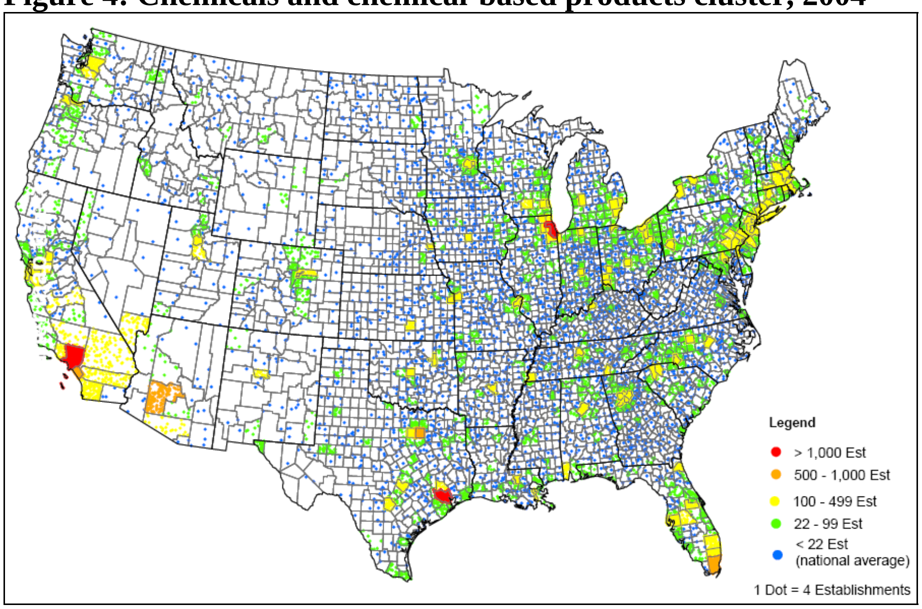
Figure 3 shows overall patterns of the distribution of establishments in chemical and chemical-based products cluster in 2004. The researcher can summarize that establishments in this cluster are concentrated in the midwestern, northeastern, and a few southern and western states.

Figure 3: Chemicals and chemical-based products cluster, 2004



In addition to distribution patterns, dot-density maps can show qualitative information by changing colors or symbols. The colors of the dots can be varied depending on the total number of establishments in a county.

Figure 4: Chemicals and chemical-based products cluster, 2004



Note: Dots are classified by total number of establishments in a county

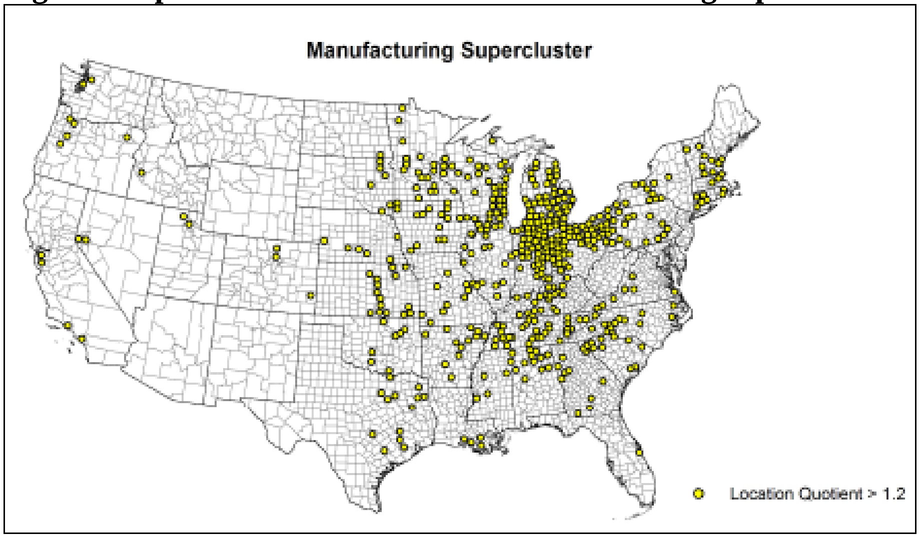
Compared to Figure 3, Figure 4 shows distribution as well as classifying counties according to the number of existing establishments. Researchers can use additional information for a sub-state level analysis. In Figure 4, the Chicago, Los Angeles, and Houston-Galveston regions emerge as large concentrations of chemical and chemical-based products cluster establishments.

Point-symbol Mapping

Point-symbol maps were used to locate counties that were identified as **hot-spots** of cluster activities. The hot-spots of cluster distribution are determined by using Location Quotient (LQ) analysis mentioned previously. This analysis compares the proportion of a county's cluster employment to the nation. If the LQ is more than one, it indicates that the region is exporting or specialized in the particular business and industry cluster. Exporting clusters include those kinds of businesses and industries that bring dollars into the community by selling their products and services to consumers outside of the region. These are also known as *Basic Industries*. For our research project, a benchmark LQ value of 1.2 was used to determine the specialized locations.

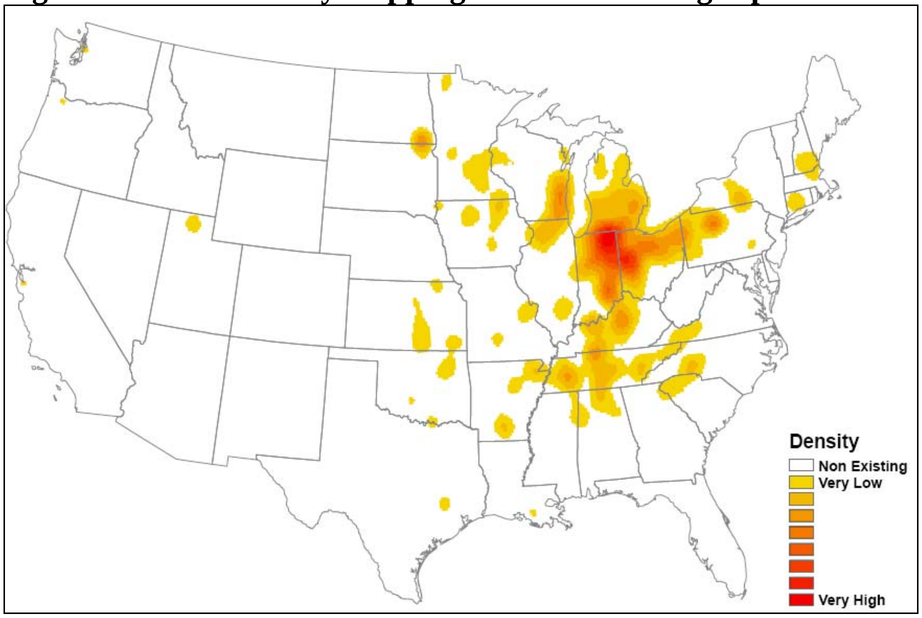
Figure 5 shows specialized locations of the manufacturing supercluster with counties having an LQ of 1.2 or greater. The historical imprint of the midwest and northeast regions as concentrated locations of manufacturing industries is revealed in this map. A portion of the southern region is also specialized in this cluster. The manufacturing supercluster comprises six sub-clusters containing mostly manufacturing and supporting industries.

Figure 5: Specialized locations of the manufacturing supercluster



ArcGIS *Spatial Analyst* extension was used to create density maps from the LQ point symbols (Figure 6). These Kernel density maps are weighted by the values of the LQ. Hence not only specialized locations are identified but also the degree of specialization is viewed by using this mapping technique. This map refines the snapshot of specialized locations and shows that portions of the midwest (northeastern Indiana and southwestern Ohio) are highly specialized in the manufacturing cluster due to a greater number of specialized locations as well as a higher degree of specialization.

Figure 6: Kernel density mapping of manufacturing supercluster

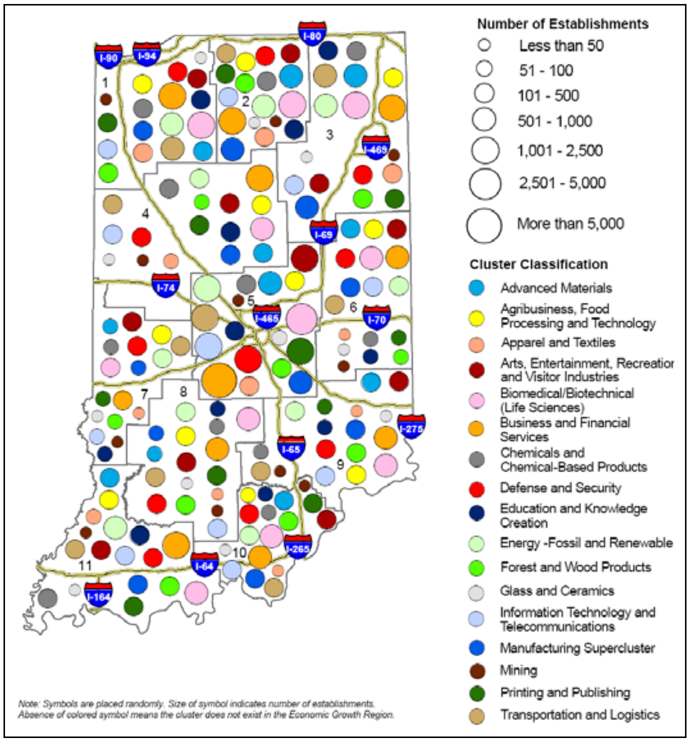


Graduated-symbol Mapping

Graduated symbols incorporate different ranges of the variable. Hence, the researcher can get an idea about the magnitude of the variable. Unlike proportional symbols, which represent precise values, graduated symbols represent a range the same as the graduated color mapping. In Figure 7, graduated symbols are used to show the ranges of establishments. The colors show different types of clusters present in the region. These symbols do not show specific geographic locations, however, they are good tools for providing an economic overview of the region. For example, Figure 7 shows different types of clusters existing in each of the 11 Economic Growth Regions (EGR) in Indiana.

A similar county level graduated-symbol map for the EGR 8 region in Indiana was used for economic development planning for the region. The employment and establishment cluster maps were useful in numerous public meetings and focus groups organized in the EGR 8 region.

Figure 7: Cluster establishment distribution by Economic Growth Regions in Indiana

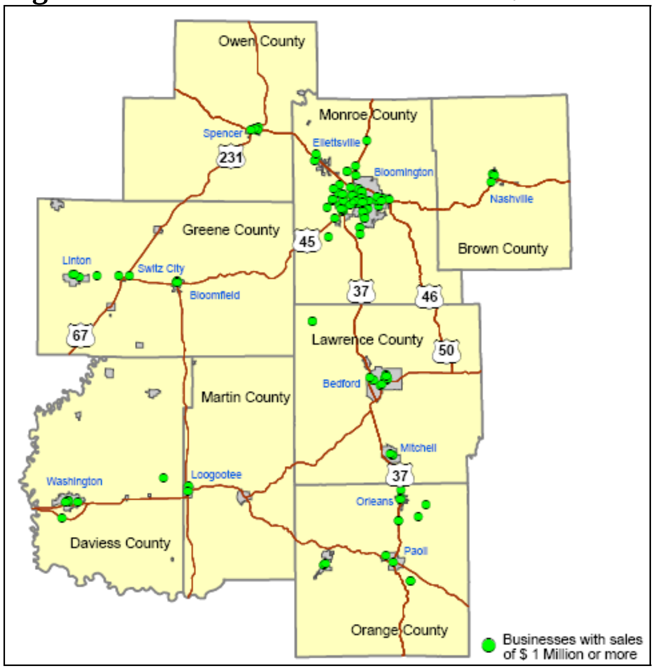


Geocoding

Geocoding is a process of locating actual latitudes and longitudes based on street addresses. It can be used to create maps showing actual locations of businesses and industries and used for economic development planning purposes. For this research, ESRI *StreetMap USA* was used to geocode businesses that have an annual sale of $ 1 million or more (Figure 8). The distribution shows that the maximum number of $ 1 million or more businesses is concentrated in Bloomington, the largest city in EGR 8.

These maps are a few examples of cluster mapping using GIS. In addition to quantitative and qualitative research, maps bring forth a unique geographical perspective in research. They are helpful in communicating strengths and opportunities in the region and engaging the community. This sheds light on the usefulness of the mapping process for planning purposes.

Figure 8: Businesses with more than $ 1 million sales by county, Indiana



Spatial Analysis:
Nearest Neighbor Analysis
Nearest neighbor is a spatial analysis technique that looks into distances to the nearest neighbors. This analysis helps determine if a pattern is a random distribution or clustered. For our research, nearest neighbor distances were included in studying interrelationships between rurality, metropolitan accessibility, economic performance, and clusters as required under the grant agreement. The research team needed distances to the five nearest neighbors for each of the 3,108 counties in each of the 17 clusters. A combination of ArcGIS processes and *Hawth's Analysis Tools* were used to find the five nearest distances between counties with clusters having LQ > 1.2 and all counties in the U.S.

Average Nearest Neighbor Distance is an index that compares average distances between the nearest neighbors to distances from a random distribution. This is a useful indicator to determine if a pattern is geographically clustered or a random distribution. By definition, business and industry clusters should have some degree of geographic co-location. However, given the advancements in communication and technology, geographic distances are no longer such a barrier to exchange knowledge and information. This is one of the measures which can reveal whether specialized clusters also co-locate. In other words, proximity might have some role in the development of business and industry clusters and further specializations.

Table 2: Average nearest neighbor distances of business and industry clusters, 2004

Business and Industry Cluster	Average Nearest Neighbor Distance		Business and Industry Cluster	Average Nearest Neighbor Distance
Advanced materials	0.52		Energy	0.72

Business and Industry Cluster	Average Nearest Neighbor Distance		Business and Industry Cluster	Average Nearest Neighbor Distance
Agribusiness, Food processing & technology	0.73		Forest and wood products	0.65
Apparel and Textiles	0.49		Glass and ceramics	0.61
Art, entertainment, recreation and visitor industries	0.64		Information technology and telecommunications	0.41
Biomedical/biotechnical	0.72		Manufacturing supercluster	0.55
Business and financial services	0.51		Mining	0.72
Chemicals and chemical-based	0.65		Printing and publishing	0.45
Defense and security	0.62		Transportation and logistics	0.70
Education and knowledge creation	0.78			

Note: Average nearest neighbor distances were calculated for specialized locations (LQ > 1.2)

The average nearest neighbor distance varies between 0 and 1. A smaller value shows more geographic clustering. The information technology and telecommunications cluster shows the most clustered pattern of specialized locations. These specialized locations are clustered or co-located primarily in the urbanized areas to the northeast and west of the US. The specialized locations of education and knowledge creation cluster have the least clustered pattern. Education establishments are required everywhere regardless of rural or urban areas.

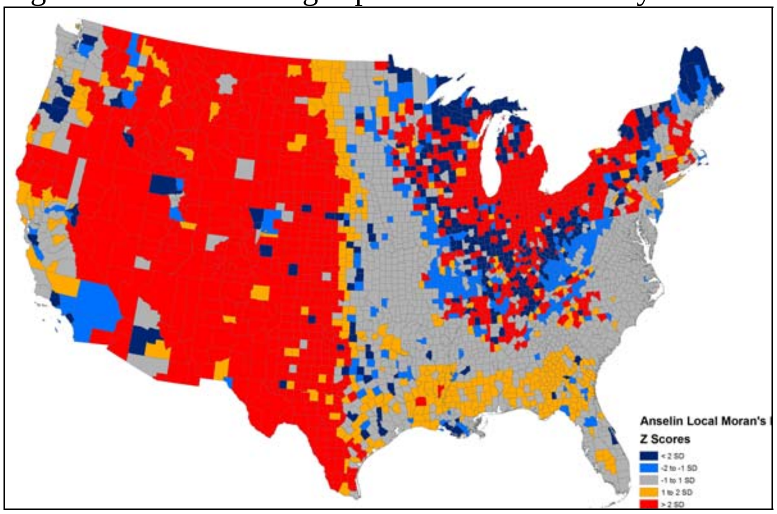
A temporal study of the most and least clustered patterns form 2001 to 2004 does not indicate any major change in the degree of geographic clustering. Information technology and telecommunications showed an increase in clustering from 2001 (0.43) to 2004 (0.41) as the average nearest neighbor distance decreased. However, the education and knowledge creation cluster showed a decrease in clustering from 2001 (0.77) to 2004 (0.78) as the average nearest neighbor distance increased.

Getis-Ord General G Statistic

Getis-Ord is an index that measures the characteristics of the clustering. It shows if higher or lower values are clustering, or that a distinct pattern exists or not. In this case, the location quotients (LQ) of specialized locations are used to find if higher LQ values are clustering or vice versa. If higher LQ values do not cluster for a particular business and industry cluster, we can infer that co-location might not be a criterion to gain specialization in that particular business and industry cluster.

For example, the General G Statistic shows that high LQ values in the Manufacturing supercluster are clustering geographically or co-locating over space. In other words, the business and industry cluster is becoming specialized in a group of closely located counties. We can infer that proximity might have some role in achieving specialization in manufacturing activities. A distinct pattern of clustering either in high or low LQ values does not emerge in the information technology and telecommunications cluster. This is one cluster that makes extensive use of internet-based and computer technologies. Hence, physical distance is not a barrier in exchanging resources.

Figure 9: Manufacturing supercluster-outlier analysis



Note: SD refers to standard deviation

Spatial Association

An interesting study within this research looks into spatial associations of counties. If a business and industry cluster is specialized in a county, what happens to the surrounding counties? Are they specialized in the same business and industry cluster, showing a uniform distribution of LQs? Figure 9 shows “z score” values of *Anselin’s Local Moran’s I* by counties. LQ values for the manufacturing supercluster for all U.S. counties were used in this analysis. A value greater than two standard deviation infers that counties are clustering either with high or low LQs. A large region in the midwest, west, and central regions emerges with a uniform distribution of LQ. This means that if a county is specialized in a cluster, the neighboring counties are also specialized and vice versa. The blue-colored distribution shows non uniform distribution of LQs, which means a county specialized in a cluster might not have neighbors specialized in that cluster.

Areas for Further Research:

The examples mentioned in this study are a few ways of exploring business and industry clusters using spatial analysis techniques. An in-depth study might reveal interesting spatial patterns in the distribution of business and industry clusters. Each of the 17 clusters has unique as well as overlapping characteristics. A study of their spatial distributions as well as associations will help in policy decisions at the regional level tailored to each of the clusters. It should be noted that clusters exist regionally, hence, policies for cluster activation, expansion, or retention should be at the regional level.

Conclusion:
Business and industry clusters are considered as economic engines that drive the regional economy. Various studies are looking into different aspects of business and industry clusters such as methods of cluster identification, the role of clusters in a regional economy, etc. GIS mapping and spatial analysis tools help explore the geographic aspects of business and industry clusters bringing new insights and information to decision makers and researchers. Spatial analysis is a vast field and there are numerous opportunities to explore business and industry clusters through the lens of spatial statistics and other tools.

Glossary:
1) Location quotient (LQ) – The formula for LQ is

$$LQ = \frac{(R1/R2)}{(N1/N2)}$$

Where: R1 = Regional Employment in Industry X
R2 = Total Regional Employment
N1 = National Employment in Industry X
N2 = Total National Employment

If LQ < 1, region is less specialized in industry X, and needs to import goods to satisfy local demand; if LQ = 1, region produces just enough in industry X to satisfy local demand; and if LQ > 1, region is more specialized in industry X and exports the industry's output to other regions.

2) Nearest-neighbor analysis- This analysis pertains to study of distances between a feature and its nearest neighbor of similar characteristics. It is often used in spatial statistics.

3) Getis-Ord General G Statistic- A multiplicative measure of overall association of values that fall within a given distance of each other. [4] According to Jared and Getis the Local Getis and Ord G statistics evaluate the extent to which a location is surrounded by a cluster of high or low values.

4) Local Anselin's Moran's I- Moran's I measures global spatial autocorrelation. The Local Anselin's Moran's I is a local autocorrelation statistic based on Moran's I.[5]

References:
1) Arlinghaus, Sandra Lauch. The Evidence of Cartographic Example: Hierarchy and Centrality. http://www-personal.umich.edu/~copyrght/image/solstice/sum05/dotdensitymap/overviewI.html.

[4] Aldstadt, Jared and Arthur Getis. Point Pattern Analysis in an ArcGIS Environment. http://www.dpi.inpe.br/gilberto/csiss/papers/aldstadt.pdf . San Diego State University. Accessed on 06/25/2006.
[5] Same as Footnote 5.

2) ArcGIS 9.1 Spatial Statistics Toolbox.

3) Aldstadt, Jared and Arthur Getis. Point Pattern Analysis in an ArcGIS Environment. Published by San Diego State University.

4) Boston Consulting Group; http://www.netmba.com/strategy/matrix/bcg/.

5) GeoDa™ 0.9 User's Guide. Center for Spatially Integrated Social Science. Spatial Analysis laboratory, University of Illinois Urbana Champaign, 2003.

6) Greater Cincinnati's Target Industries, Center for Economic Education, University of Cincinnati, 1996.

7) Hawth's Analysis Tools for ArcGIS; www.spatialecology.com.

8) Industry Clusters: An Economic Development Strategy for Minnesota Preliminary Report. http://www.hhh.umn.edu/centers/slp/economic_development/econdev_strategy_industry_cluster.pdf.

9) Point Pattern Mapping and Analysis. http://mailer.fsu.edu/~jastallins/dir/course/home/Geo3140/Lectures/Point.pdf.

Authors:

1) Christine Nolan, Senior Planner and Associate, Purdue Center for Regional Development, Purdue University, 1201 W. State Street, West Lafayette, IN 47907-2057; Tel: 765-494-9262; Fax: 765-494-9870; email: cenolan@purdue.edu

2) Indraneel Kumar, Spatial and GIS Analyst, Purdue Center for Regional Development, Purdue University, 1201 W. State Street, West Lafayette, IN 47907-2057; Tel: 765-494-9485; Fax: 765-494-9870; email: ikumar@purdue.edu