Methodology

Using both qualitative and quantitative data, this project seeks to analyze New Jersey's industry clusters and the infrastructure necessary to support industry cluster development. Research on basic industry cluster theory, cluster development, and examples from other states was completed to establish a conceptual foundation for this work. The analysis section of this report is supported by further research on the composition of the eight specific clusters, and a review of the various types of infrastructure necessary for industry development in New Jersey. In addition to basic research, the analysis is also informed by expert opinions from industry leaders participating in the project's Advisory Committee. Finally, GIS mapping technology was used to map the location of New Jersey's physical infrastructure, including: hard infrastructure such as roads, port, airports, freight rail, public transit systems; utilities such as electricity lines and sewer service areas; and soft infrastructure or institutions such as hospitals, places of education and public amenities like parks and recreational areas. Together, the research, expert opinions from local leaders and mapping analysis offer an extensive analysis of the relationship between industry and infrastructure in New Jersey. The following is an overview of the data inputs and quantitative methods used in this project.

Put simply, the definition, measurement, and depiction of industry clusters in a spatial environment is a complicated endeavor. Methods can range from simply showing point locations of firms to mapping location quotient values, statistical cluster analysis (Moran's I, Getis Ords Gi*), and other density analyses. Some studies have also examined the distribution of patent citations to identify clusters of innovation. Hofe & Chen (2006) address the confusion over the different methods used to identify clusters. They assert that the varying concepts of agglomeration economies and industrial clusters lead to a "tremendous number of methods" to identify clusters. Quantitative methods used in this project include an analysis of statewide Location Quotient Values and the use of the Nearest Neighbor Index and the Kernel Density function in ArcGIS.

Location Quotient

Location Quotient Value was obtained from the U.S. Department of Labor Bureau of Labor Statistics Location Quotient Calculator for each of our regional industry clusters.

According to EMSI Economic Modeling Specialists Intl. we can use Location Quotient to quantify how "concentrated an industry is in a region compared to a larger geographic area, such as the state or nation."

Industry LQs are calculated by comparing the industry's share of regional employment with its share of national employment. Our LQ calculation is augmented by two other pieces of information: size of industry/cluster/occupation in terms of jobs, and percent change in LQ over a given time period (Economic Modeling).

Interpreting Location Quotient Values (Bureau of Labor Statistics)	
LQ = 1	Industry has the same share employment in New Jersey as it does in the USA.
LQ > 1	Indicates an industry with a greater share of employment in New Jersey than the USA. This industry makes up the "economic base" and exports services and products out of state.
LQ < 1	Indicates an industry has less share of the employment in New Jersey than the USA.

As promoted by Harvard Professor Michael Porter, the calculations of the LQ as well as looking at the historical change over time can help to target efforts of economic development. Keeping in mind the goal of investing developing policies to support New Jersey's Regional Innovation Clusters, there are four potential combinations (Location Quotients)

- If the industry has a LQ of less than one and is declining over time, this industry is considered to be "weak and declining" and generally should not be considered a potential cluster.
- If the industry has a LQ of less than one but is increasing over time, the industry is considered "weak and growing" and may be a potential industry to focus economic policy efforts.

- If the industry has a LQ of greater than one but is declining over time, it is considered "strong and declining" In this case, policy makers should determine if the decline of this industry presents a potential risk to the state economy. If so, focused policies to revitalize the industry may be appropriate.
- If the LQ is greater than one and growing over time, it is considered "strong and growing" These are potential clusters for economic growth and development should be a focus of policy efforts. Additionally, these industries have a competitive advantage over other regions and may have further growth potential.

Identification of Clusters	
Weak and Growing Location Quotient < 1 Location Quotient Increases Over Time	***Strong and Growing*** Location Quotient > 1 Location Quotient Increases Over Time (Potential Cluster)
Weak and Declining Location Quotient < 1 Location Quotient Decreases Over Time	***Strong and Declining*** Location Quotient > 1 Location Quotient Decreases Over Time

Location Quotient values will be used to reinforce and highlight the industries that have been chosen for the study. First, we calculated the Location Quotient using the BLS calculator. Then, in an attempt to show how each of the industry clusters faired through the recession of 2007 the time series starts in 2005. The time series ends in 2011, which represented the most up to date Bureau of Labor Statistics employment data sets at the time of retrieval. Then, we charted the change of the LQ values from 2005 to 2011 to show change over time. We also included a trend line that overlays the data points. The trend line's R-squared value explains how well the trend line matches the original data points based on a zero to one scale. The line represents the ability of the model to project data points forward based on the data already in the model.

After that, we have provided an analysis of where each industry cluster fits into the "Identification of Clusters" graphic. For each industry we have provided a close examination of location quotient. From that categorization and examination we can then recommend what levels of economic development policy are appropriate for that industry cluster. The evaluative criteria used to determine if an industry was growing or declining was based on the change in that Industry's Location Quotient Value from 2010 to 2011. The evaluative Criteria used to determine if an industry is strong or weak is based on if that industry's Location Quotient value is more or less than 1.

Nearest Neighbor and Kernel Density

The Nearest Neighbor and Kernel Density surface calculations are based on firm-level data obtained from ReferenceUSA, a continuously updated database of U.S. businesses. Data for Advanced Manufacturing, Finance, Healthcare, Life Sciences, TLD, and Technology industries was downloaded in March 2012. Data for the Defense and Tourism industries was downloaded in March 2013. To establish whether or not such clusters were statistically significant, a nearest neighbor ratio was computed for each of the eight industry clusters. This index measures spatial clustering by calculating the average distance from one point to that point's closest neighbor. The nearest neighbor ratio evaluates the existence of clustering on a scale between less than one, indicating clustering, and greater than one, implying dispersion. The nearest neighbor ratio for all eight industry types measured less than one and supported the visual observation that firms in the same sector locate close to one another.

The ReferenceUSA database, which is used by the NJ Department of Labor and Workforce Development, allowed us to display and analyze individual firms with 2 or more employees within each industry cluster using the latitude and longitude of each firm. The decision to use Kernel Density as a means of depicting clusters was based on a review of the available methods and the type of data inputs used in this project. Examples of other studies that implement Kernel Density include *Spatial Analysis of Knowledge-based Occupation Clusters* (Kumar & Nolan, 2010), *Geography of Opportunity: Poverty, Place, and Educational Outcomes* (Tate, 2008), and

Transportation infrastructure impacts on firm location: the effect of a metro line in the suburbs of Madrid (Mejia-Dorantes, Paez & Vassallo, 2011).

After downloading descriptive data about firms, the data points were transformed using the Kernel Density function to create a smooth density map that depicts clusters using the density of firms and their respective employment values. It is important to note that we used the location size (actual number of employees) for the population field to weigh each data point. Therefore, those firms with large amounts of employees represent more jobs within the cluster and result in great density. Furthermore, we utilized the default search radius generated by ArcGIS and the raster cell size to produce a detailed map of RICs. The search radius can be made smaller or larger, though changing the radius does not significantly change the density values. Larger search radius values produce a smoother, more generalized density raster, while smaller values produce a raster that shows more detail (ESRI, 2011). The kernel density application produced a raster surface showing locations of industry clusters, with each cell displaying a value of employees per square mile (ESRI, 2012). These density maps were then juxtaposed with vector data showing key infrastructure obtained from the NJ Geographic Information Network, NJ Transit, NJTPA Geoportal, PASDA, and NJ OGIS.

Limitations

The methodology of this project faces a number of limitations in both data and procedure. Firstly, many of the desired infrastructure data points were unavailable or outdated. Future analysis could include more detailed data regarding capacity, as well as power facilities, updated sewer service areas, housing, and other desirable infrastructure such as corporate parks and distribution centers. Secondly, there is also a great deal of subjective judgment involved in defining clusters in a GIS environment. There are many available methods and lack of consensus among professionals on which are most appropriate. Finally, the kernel density function uses certain parameters, including weights, search radius and cell size, which are user-defined and depend on individual needs. For example, the kernel density could be calculated based on purely the clustering of firms, or it can be weighted with employment numbers. There is a lack on consensus in the literature as to which method (establishments or employment) best indicates clustering of economic activity (Sweeney & Feser, 2004). Future work using kernel density could examine the use of employment as an indicator or change the search radius to reflect the characteristics of each cluster and/or specific policy goals.