
Labor Market Data & Analysis for Site Selection



emsi

Economic Modeling Specialists Inc. (EMSI) is a professional services firm that offers integrated regional data, web-based analysis tools, data-driven reports, and custom consulting services. EMSI has served thousands of policy professionals and consultants in the U.S., Canada, and the United Kingdom, and the company's web-based **Analyst** research and analysis suite has over 2,500 users. For more information, call (866) 999-3674 or visit www.economicmodeling.com.

For more information and to schedule a live demonstration, contact:

Rob Sentz
rob@economicmodeling.com
(866) 999-3674

INTRODUCTION	3
LABOR MARKET ANALYSIS: A KEY FACTOR IN SITE SELECTION	3
Challenges of Labor Market Research	4
Public Data Sources.....	4
Proprietary Data Sources.....	4
Solution: Integrated, Accessible Labor Market Information	5
EMSI'S SOLUTION: ANALYST	5
Basic Industry and Occupation Analysis.....	5
Analyzing Industry Data with Analyst.....	7
Analyzing Occupation Data with Analyst.....	8
Advanced Labor Pool & Human Capital Analysis.....	8
Evaluating College Degree Output.....	10
Other Data and Indicators	11
Using Raw Data and Customized Analysis.....	11
CONCLUSIONS	12
APPENDIX: PUBLIC SOURCES FOR KEY LABOR MARKET DATA	13

Introduction

Businesses and economic development groups have increasingly recognized that innovation and human capital are the new economic drivers. Moreover, as retiring baby boomers are expected to create significant workforce gaps, businesses of all kinds—from international manufacturing firms to national retail chains to small businesses—are focusing on finding the right workers with the right skills and the right wage expectations.

According to a survey by *Site Selection* magazine, 74% of executives said that “finding available, skilled workers was their biggest challenge in today's economy.”¹ While that survey is from 2001, the same point is becoming even more relevant than ever. According to the Bureau of Labor Statistics, nearly 80% of the job openings in the next ten years will be in middle- to high-skill jobs, and that means that qualified labor is an indispensable factor for more site-selection decisions.

So labor market analysis is critical to the site selection process. But how to find the right data? The ideal solution, of course, is to get out in the community(ies), conduct surveys, meet the workers, and talk to the local employers. Unfortunately, in most cases such primary research is far too expensive and time-consuming, especially if there is a large pool of potential sites. A much more efficient solution is to use available secondary data sources to narrow the search to a handful of areas, and then follow up with targeted primary research. In this paper, we review the types of available data, common challenges, and EMSI's solutions for data collection and analysis.

Labor Market Analysis: A Key Factor in Site Selection

The main questions that workforce research seeks to answer are:

- What is the area's overall unemployment rate? A tight labor market will drive labor costs up.
- What is the area's industry mix? How has it been changing over time?
- How many area workers are in key occupations, or are in occupations that have knowledge/skill sets that match those of key occupations?
- What are the wage ranges for these kinds of workers?
- Have these kinds of jobs seen significant job growth or decline recently? Growth in specific industries and occupations indicates strong competition for labor, while decline indicates a readily available labor pool.
- Are local workforce investment officials and training providers—universities, community/technical colleges, and career schools—committed to supporting area employers? Do they currently offer relevant courses/programs, or would they offer them in the future?

Companies need detailed labor market and education data to perform an initial review of the areas surrounding potential sites. Labor market data at the metro, county, or even ZIP code level provides a quick insight into the local economy's unemployment rate, labor force, industry/occupation mix, job growth trends and forecasts, wages, relative industry sector sizes, students graduating in certain

¹ <http://www.siteselection.com/issues/2001/jul/p466/>.

fields, and more.

Challenges of Labor Market Research

Public Data Sources

Labor market information can be obtained from a number of public sources which, taken together, can provide a wealth of information to the determined researcher. For a list of major public sources, see the Appendix at the end of this paper.

While these sources are published for free, the data often require a good deal of time and effort before they can be called useful information. Some of the shared problems of public data sources are:

- **Lack of Data Integration:** Much of the data is housed on state labor agencies' web sites or various federal sites such as the Census Bureau, Bureau of Labor Statistics, and Bureau of Economic Analysis—all of which have multiple data programs with separate databases. Simply collecting all the raw data necessary to provide a complete picture of a local economy is surprisingly time-consuming.
- **Problems with Coverage and Scope:** Various sources have different strengths and weaknesses. Some are good at totaling all types of workers but don't give much detail about how they are distributed among various industries or occupations; others offer a high degree of sectoral detail but limited geographic detail; some are survey-based and others are not; some are published annually and others quarterly; and so on. Analysts must laboriously stitch all these views together to arrive at something approaching a full picture.
- **Non-Disclosed Data:** Due to government privacy policies, public data sources cannot publish detailed local data that could possibly be tied to one or two specific business establishments. Government statistical agencies have to "suppress" these elements, and usually multiple *other* pieces of data that could be used to "fill in" or calculate the original undisclosed data. This is a necessary limitation of public data that frustrates intensive labor market research.
- **Data Currency:** Most Census-based local data is anywhere from 5 to 10 years old, drastically reducing its usefulness. Other local (county/metro-level) economic data from the Bureau of Labor Statistics and BEA are usually more timely, but still 6 months to 2 years old. Unfortunately, the only way to overcome this issue (apart from primary, survey-based, on-the-ground research) is a method that estimates current and near-future data using past trends and published projections. However, public data sources rarely offer detailed and useful *annual* projections.
- **Overall Usability:** Data collection and integration is one task; actually analyzing the data is another. Most of the information collected from public sources is available only in various plain text, database, or spreadsheet formats, leaving the researcher with many hours of re-formatting and analysis required to produce "presentation-ready" information.

Proprietary Data Sources

There are also a number of proprietary data sets that offer some kind of labor market information. InfoUSA and Dun & Bradstreet specialize in business-level data (addresses, phone numbers, CEO names, number of employees, approximate sales), which are incredibly valuable for prospecting purposes. However, these data sources are based on periodic, voluntary phone surveys, making the

employee and/or sales numbers subject to significant error. They also provide no data on employment by occupation, or job type.

In contrast, the primary public labor market data source (the Quarterly Census of Employment and Wages), although it lacks the business-level detail, is a true census of employment based on legally-required filings for the government unemployment insurance program. Other programs use tax return data, Census returns, and large-scale surveys. Understandably, economists and labor market researchers who need accurate time-series employment data consider the government data more accurate, despite problems with its usability and lack of business-level detail.

Researchers often find that business-level data is a perfect “next step” complement to a rigorous initial analysis using broader, industry-level labor market information. The first step offers a macro view, allowing the micro view in the second step to be much more focused and manageable.

Solution: Integrated, Accessible Labor Market Information

By now the problem of quickly obtaining good labor market data should be clear. With a wide array of sources—all with different characteristics, coverage, ease of use, and overall strengths/weaknesses—the researcher is forced to settle for an incomplete picture of local labor markets. A better solution would be a system that intelligently stitches together these various pieces of information, in an attempt to yield the most accurate and comprehensive picture possible. Moreover, such a system would provide fast and easy access to data with time-saving features like custom-defined geographic areas, group analysis of multiple industries/occupations, easy time-series reports (past and projected), local-to-state/national comparisons, and ready-made reports that answer the researcher’s “frequently asked questions” about any area’s economy and workforce.

One such solution has been created by Economic Modeling Specialists, Inc. (EMSI), which has developed a web-based system and automated reports that integrate and harmonize over 80 state and federal databases, allowing quick and easy analytics for any county, ZIP, metro area, or custom-defined area.

EMSI’s Solution: Analyst

Below is an overview of the labor market data accessible in EMSI’s **Analyst**, how it provides a comprehensive understanding a region’s workforce, and examples of how it can be applied.

Basic Industry and Occupation Analysis

To capture the most complete possible picture of area employment, EMSI collects and combines data from the Bureau of Labor Statistics, the Bureau of Economic Analysis, the Census Bureau, and individual state agencies. This “Complete Employment” dataset covers all types of workers. (EMSI also produces a “Covered Employment” data set that is nearly identical to QCEW, except that it provides educated estimates for non-disclosed numbers and locally redistributes jobs that are reported only at the state level.) Drawing from dozens of sources and processed using sophisticated algorithms, it represents a significant advance in data integration, coverage, detail, and usability.

This integrated data set reveals the basic composition of a region’s industries and occupations, including:

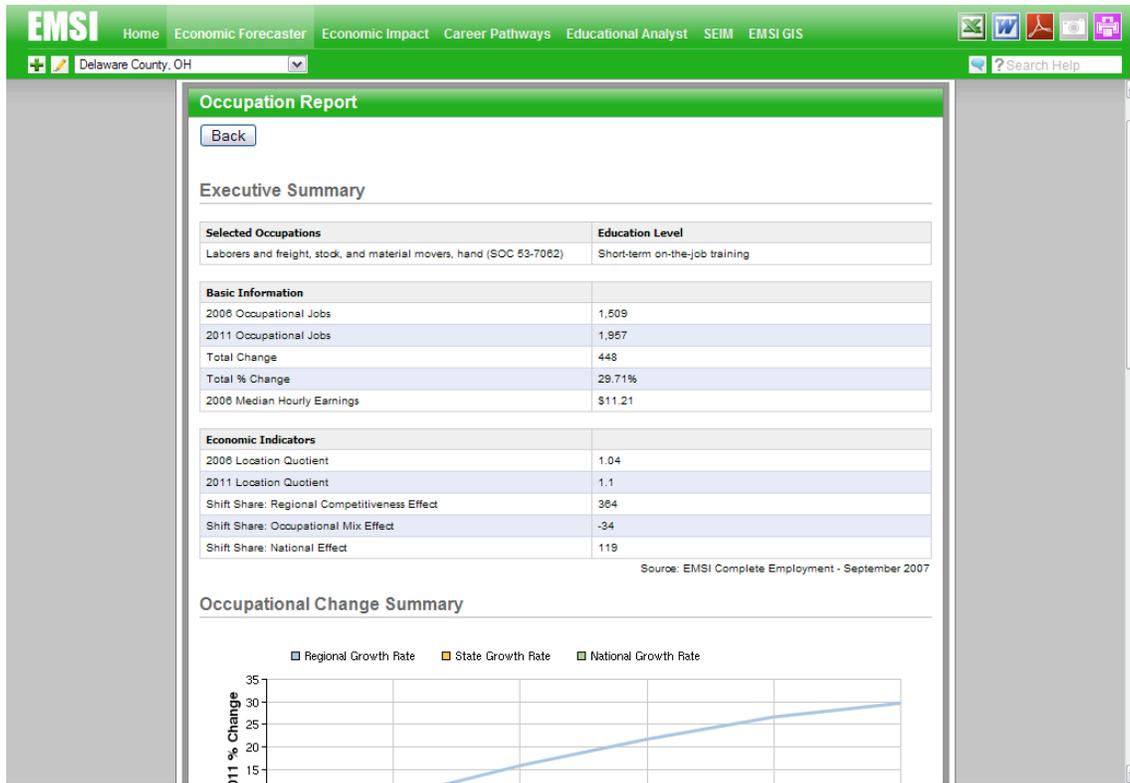
- Total jobs and past/projected job trends for detailed industries and occupations
- Number of establishments by industry
- Earnings by industry and wages by occupation, including percentile wage levels for occupations
- Relative local industry/occupation employment size and growth, compared to state(s) or U.S.
- Overall labor force and unemployment rates by county
- Unemployment by broad industry and occupational sector
- Staffing patterns showing the types of workers employed within a certain industries.

Figures 1 and 2 are examples of the interface and analysis features of EMSI’s web-based system, called *Analyst*. The first shows a ranking of detailed local industries; the second provides information on a specific occupation in a local area. The system can rank industries and occupations using multiple variables, and supports any custom-defined geographic area and custom-defined industry or occupation aggregations. In addition, it provides an interactive table of all the raw data which can be sorted and filtered instantly.

Figure 1



Figure 2



Analyzing Industry Data with Analyst

Site selectors can use industry data in two basic ways to better serve their clients:

1. Profile a local area's industry mix and detect structural changes that might work in your client's favor. For example, economically distressed or stagnant areas may be more willing to offer incentive packages. Or, the decline of an industry similar to your client's business signals an oversupply of labor with the right skills they need.
2. Analyze a region's relevant market dynamics:
 - The strength of the client's industry in the area (similar businesses / competitors)
 - Potential local/regional market (area businesses that might buy your client's services)
 - Potential local/regional supplier base (are businesses that produce inputs that your client needs)

Let's walk through an example. Your client wants to set up more local distribution centers for its winery supply products to supplement its existing center in Napa, California. In particular, the client is interested in the nearby, emerging wine-making regions in Oregon and Washington State.

Using the GIS module of Analyst, we rank all the counties in those states, or the entire US, by 2002-2007 percent job growth in the "Wineries" industry, further refining our results with the requirement that each county must have a winery industry big enough to support at least 150 jobs, and (to reduce inflated percentages) had at least 10 winery jobs in 2002. For this analysis, we chose to use "covered" or payroll employment numbers only.

As a result, we find that the biggest growth areas are actually in California counties, but outside the Napa Valley region:

1. Stanislaus County (includes Modesto) – about 2,000 new jobs or 390% growth.
2. Riverside County (from Riverside east to Palm Springs and beyond) – about 650 new jobs or 270% growth.

In Oregon, we see over 600 new winery jobs in Yamhill, Lane, Polk, Washington, and Marion counties, for a combined growth of 80%. Washington State shows smaller, but still very promising numbers, with a total of about 420 new jobs in Benton, Walla Walla, Yakima, and King counties, for a combined growth of about 55% (70% excluding King county).

With this information, we can estimate the sales potential for these areas and make recommendations accordingly—such as shifting the focus of the Napa center to the nearby Modesto area, and possibly opening new centers in Riverside, CA and Portland, OR.

Analyzing Occupation Data with Analyst

So far the data we've seen has an industry focus. But what about businesses that need a skilled, affordable workforce? Can they evaluate a region's workforce by occupation to factor specific labor availability and costs into their site selection decision? The answer is yes, and Analyst makes such analysis efficient.

The Bureau of Labor Statistics has occupational wage levels for metropolitan areas as well as the "leftover" non-metro areas. EMSI combines this information with a model based on its detailed industry data and staffing patterns to produce well-founded estimates of average and 10th, 25th, 50th, 75th, and 90th percentile occupational wages at the county and even ZIP code levels.

Suppose your client wants to open a computer support call center somewhere in the United States, due to concerns about rising costs overseas and negative customer attitudes toward off-shoring. However, they still need to keep costs down, especially labor costs. So what U.S. regions have the most affordable wage rates for computer support specialists?

An initial analysis, which filters out counties having fewer than 150 estimate computer support jobs, points to Hays, TX (midway between San Antonio and Austin) and Fayette, PA (south of Pittsburgh) as the top two candidates, each with estimated wage rates hovering just below \$9/hour, but less than 200 estimated computer support workers.

If a larger labor pool was desired, one interesting result of our analysis was that three New York City boroughs—Kings, Queens, and Bronx—were ranked high on the list, with median wages of only \$10-\$11 per hour. This was surprising since the median wage in the entire New York/New Jersey/Long Island/SE Pennsylvania metropolitan area is significantly higher at about \$23.50. However, these high wages are concentrated in New York proper and neighboring New Jersey—information that could not have been obtained from a less detailed data source.

Advanced Labor Pool & Human Capital Analysis

EMSI has also created a unique tool called Career Pathways that leverages O*NET's occupational competency profiles in combination with local area occupational jobs data. This powerful combination allows local analysis of extended labor pools and embedded human capital.

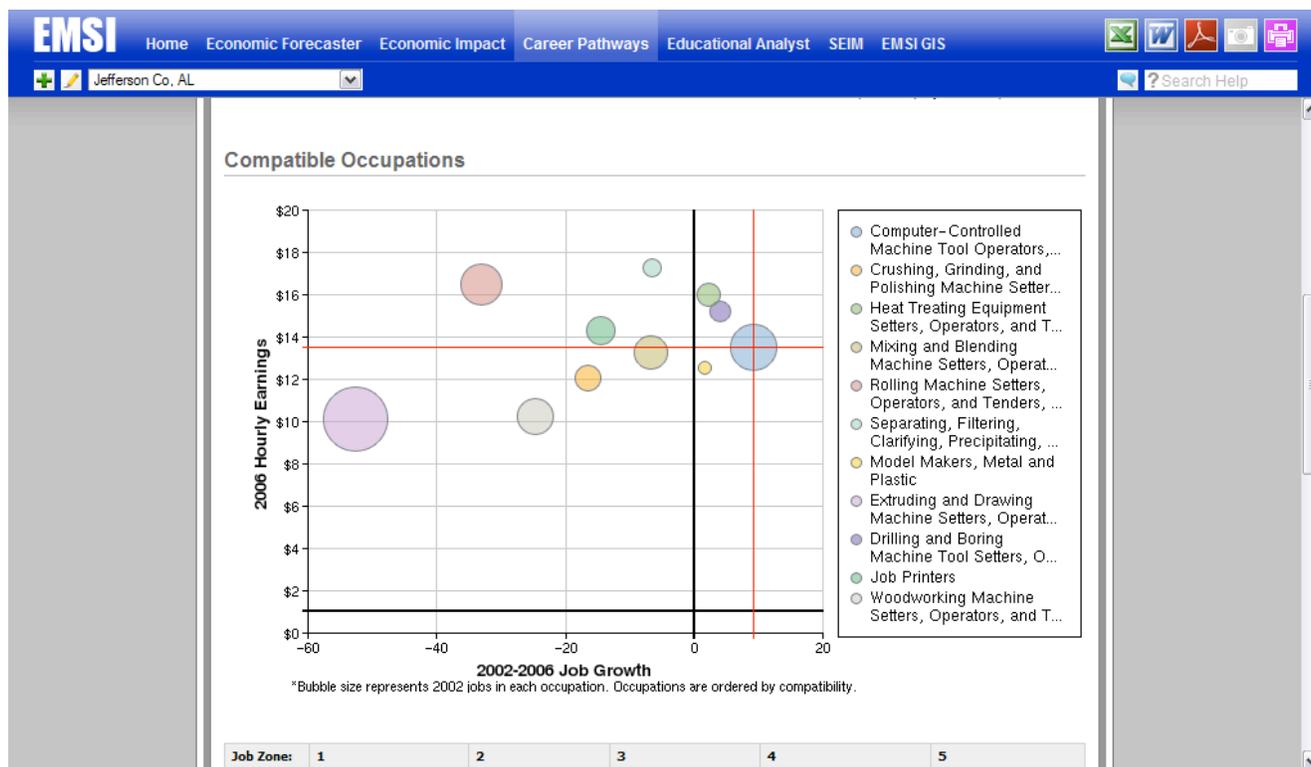
Suppose that a company is evaluating sites for a new facility, and is interested in the local labor market for *operators of computer-controlled machine tools*. The data may show some 350 current jobs, with moderate job growth trends. Based merely on these numbers, the labor market looks tight. However, there are other jobs with similar skill sets. If workers could be persuaded to switch jobs, and the local college offered relevant training, the labor market picture would look very different.

To enable this kind of research, EMSI's Career Pathways analyzes the O*NET competency profile for operators of computer-controlled machine tools, comparing it to the profile of hundreds of other occupations. It would find similar occupations that include:

- Crushing, grinding, and polishing machine setters, operators, and tenders
- Rolling machine setters, operators, and tenders
- Extruding and drawing machine setters, operators, and tenders

An easy-to-read graph (**Figure 3**) might show that several compatible occupations are not only declining in the area, but also offer lower pay than the target occupation. This indicates a much looser labor market, since there is likely a significant pool of workers who would be very willing to upgrade their skills to get a higher-paying and more secure job as operators of computer-controlled machine tools. This promising trend can be confirmed and further investigated using on-the-ground research.

Figure 3



Another feature of the Career Pathways tool is the linkage between occupations and local postsecondary training programs. The researcher can select an occupation of interest and scan regional universities, community/technical colleges, and many private career colleges for related

programs.² The results even show how many recent graduates the programs produced.

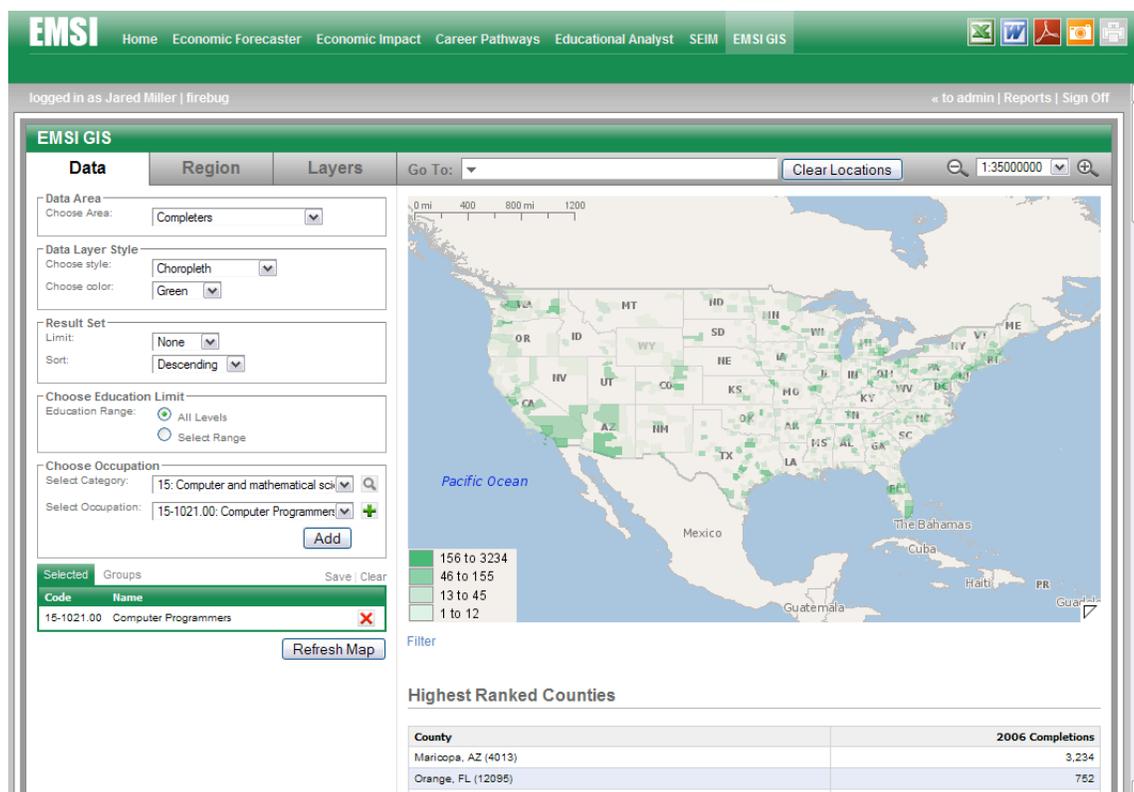
In summary, the power of the Career Pathways module is that it integrates industry/occupation employment data, staffing patterns, standard O*NET competency profiles, and local education/training program information into one user-friendly package.

Evaluating College Degree Output

Before they locate to a region, businesses need to know whether they'll be able to access a knowledgeable, skilled workforce for years to come—whether they need accountants, computer programmers, or manufacturing workers with customized skill sets. A good way to gauge this is the region's output of graduates with postsecondary degrees and awards. EMSI gathers this information from the U.S. Department of Education and makes it easy to use by aggregating it for custom-defined areas and connecting it to specific occupations.

For example, the following map (**Figure 4**) shows recent college completions in programs linked to the occupation “computer programmers.” A ranking table under the map reveals that Maricopa County, Arizona (which includes Phoenix) leads the nation in producing graduates in this field—more than four times that of the second-ranked county.

Figure 4



² Any institution that reports data to the U.S. Department of Education IPEDS program is included. Note that some occupations—especially those requiring less than a 2-year degree—are not linked to any standard postsecondary program.

For a single area of interest, the Career Pathways tool will also list all institutions or training providers in the area that offer programs linked to a given occupation. If none are found, the researcher might search using related occupations to find potential institutions that could be persuaded to offer the right program(s).

Other Data and Indicators

Analyst excels in regional workforce analysis, but it also includes other useful information about any U.S. region, including:

- Demographics: population and projections by race, age, and gender; educational attainment
- Various indicators such as annual county-to-county migration flows, housing costs, poverty rates, per-capital personal income, and patents granted
- Customizable regional economic models that show economic base, economic ripple effects of job loss or creation, and regional inter-industry supply chain gaps

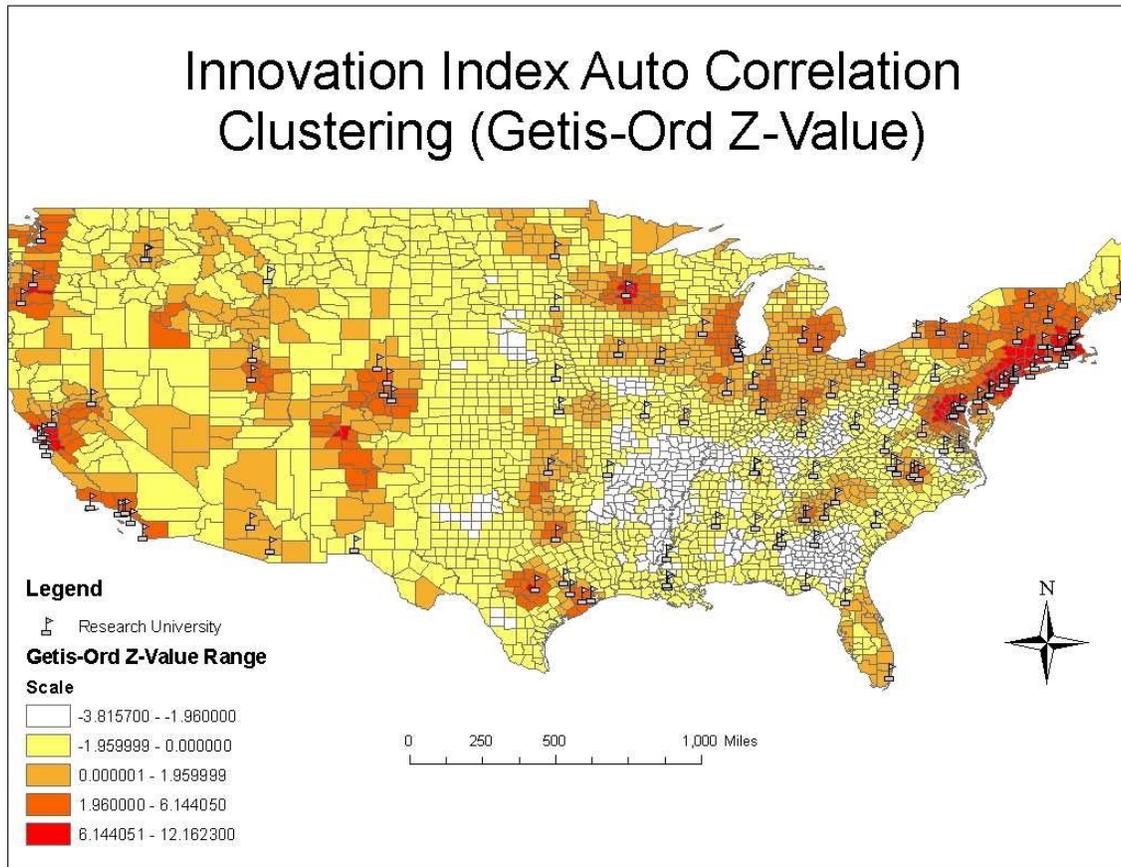
These indicators enhance the tool's specialization in labor market data and analysis.

Using Raw Data and Customized Analysis

For some specialized clients, EMSI can offer access to its raw data in a relational database format. This dramatically increases the possibilities for analysis since it can be combined with proprietary data or any public data source not already included.

An example of customized analysis using raw EMSI data is a recent conference paper given by an EMSI economist at the Pacific Northwest Regional Economic Conference. The paper, titled "Identification of Knowledge and Innovation Clusters: A GIS Application of Concentration, Co-Existence, and Correlation," uses a combination of data (high-tech and high-knowledge employment, patent generation, educational attainment, density of postsecondary institutions, and other factors) with specialized analysis in order to create a regional "innovation index" for areas across the U.S. The result was mapped in ESRI's ArcGIS and clearly shows regional innovation "hotspots."

Figure 5



This is just one application of custom analysis using EMSI's diverse data sets.

Conclusions

Site selection is a complex process that involves dozens of factors, including physical infrastructure, land parcels or buildings, quality workforce, tax incentives, and more. However, the availability of a quality workforce is quickly becoming one of the most important factors for many companies.

Traditionally, researching local labor markets required laboriously collecting, integrating, and analyzing publicly available data from multiple sources. Today, the rise of private-sector data integration platforms—led by EMSI's Analyst—has greatly accelerated this process, providing a host of valuable data along with value-added processing and presentation.

Appendix: Public Sources for Key Labor Market Data

In this appendix, we evaluate a few of the most important sources of public labor market data. EMSI's Analyst uses an integrated database created using all these sources (except Local Employment Dynamics, which is currently being incorporated).

- **The Quarterly Census of Employment and Wages (QCEW),¹** Bureau of Labor Statistics.
 - *Pros:* Includes state, metro area, and county employment, establishments, and payroll numbers by industry. Its attempted geographic and industry detail coverage are impressive. It produces quarterly data with a 6-9 month lag for preliminary data and a 1-year lag for final data.
 - *Cons:* (1) This source only includes payroll (unemployment insurance covered) employment; other types of workers, such as the self-employed, small business owners with no paid employees, railroad workers, and others are left out. (2) It does not include data for occupations—only industries. (3) Much of the detailed local data is not disclosed due to government privacy policies. (4) The available data access tools are difficult to work with, especially for customized multi-industry and multi-area groupings.

- **Occupational Employment Statistics & Employment Projections,²** Bureau of Labor Statistics.
 - *Pros:* Includes state, metro, and non-metro region employment and wage numbers by *occupation* (job type) rather than by industry, although it also shows how occupations are distributed across industries and vice versa. You can also find 10-year job projections for each occupation. It generally has an 18-month lag time.
 - *Cons:* (1) Because it is a survey, there is possibility for error, especially for smaller areas. (2) No county-level data, which makes analysis of local and custom geographic areas difficult or impossible. (3) Data access is difficult, requiring the researcher to click through multiple pages to get a single piece of data for a single year.

- **Census 2000 and American Community Survey,** Census Bureau.
 - *Pros:* Comprehensive data for hundreds of demographic, social, and labor force variables, including labor force status, industry/occupation, wage, and journey to work. Has detailed geographic coverage.
 - *Cons:* (1) Census 2000 is long out of date for real-time corporate decision-making. (2) American Community Survey is susceptible to survey error, does not include as much data as the centennial census, and excludes smaller / more rural counties. (3) Respondents “self-classify” their industry and occupation, which is translated to standard codes by Census workers—introducing a good deal of error.

¹ <http://www.bls.gov/cew/>.

² <http://www.bls.gov/oes/>, <http://www.bls.gov/emp/home.htm>.

- **Local Employment Dynamics - Quarterly Workforce Indicators,**³ Census Bureau.
 - *Pros:* Provides payroll employment statistics by industry for local areas, including total jobs, new jobs, separations, turnover, and other variables by worker age and sex. It represents a step forward in government data integration, detail, and flexibility. Includes “On the Map” tool. Approximately 1 year data lag.
 - *Cons:* (1) Eight states don’t provide data, and a few have a longer lag time. (2) Geographic and industry aggregation is difficult. (3) No occupation data. (4) Interface is somewhat clumsy.

- **Nonemployer Statistics,**⁴ Census Bureau.
 - *Pros:* This unique dataset uses tax returns to show establishments and earnings by industry for self-employed workers and proprietors without paid employees, who do not appear in QCEW and OES. It is available for states, metros, and counties with mid-level industry detail.
 - *Cons:* Much detailed data is undisclosed, there is a 2-3 year time lag, and there can be self-classification errors (people misclassifying their industry).

- **Local Area Personal Income Reports,**⁵ Bureau of Economic Analysis.
 - *Pros:* This county-level source (also available for metros and states) specializes in comprehensive, high-level totals for personal income, population, employment by industry sector, and levels of total proprietor vs. payroll employment. Unlike QCEW, it includes all types of workers—covered and non-covered.
 - *Cons:* (1) Only provides totals for very broad industry sectors. (2) Data time lag of about 1 year for metro areas and 2 years for counties.

- **Occupational Information Network (O*NET),**⁶ Department of Labor / Employment & Training Administration.
 - *Pros:* Provides profiles of knowledge, skills, abilities, work interests, work activities, workplace tools, and other characteristics for hundreds of occupations. When connected to local labor market data, it can provide quantifiable estimates of available human capital.
 - *Cons:* Requires a good deal of analysis to make the raw data useful for constructing total labor pool estimates (jobs in all compatible occupations), especially since it uses a slightly different occupational classification system than occupational labor market data.

³ <http://lehd.did.census.gov/led/datatools/qwi-online.html>.

⁴ <http://www.census.gov/epcd/nonemployer/>.

⁵ <http://www.bea.gov/bea/regional/reis/>.

⁶ <http://www.onetcenter.org/>, <http://online.onetcenter.org/>.

- **State Labor and Employment Agencies:**⁷ Nearly all states have their own labor market information agencies. They reproduce data also available on federal sites (such as QCEW), but most states also produce their own state-level and sub-state regional projections.

⁷ See <http://www.bls.gov/bls/ofolist.htm>.