

Emsi Certification

UNDERSTANDING TRADITIONAL & REAL TIME
LABOR MARKET INFORMATION



Emsi Data Training

OUTLINE

- I. Tradition LMI
- II. Real-Time LMI
- III. Traditional and Real-Time LMI combined



Traditional LMI

OUTLINE

- I. Characteristics of Traditional LMI
 - I. Core Data Points
 - II. Coding Systems
- II. Specific Traditional LMI Datasets
- III. Data Processing



Traditional LMI

- Collected from Public, Government Sources
- Describes Trends & Projections



Core Data Points

- Job Counts: Number of computer programmers working in a region
- Earnings: Amount of money those computer programmers are making
- Job Change: Growth/Decline of computer programmers in the region
- Program Completions: Number of graduates to fill future openings



What is a “job”?

- Annualized
- Full and part-time
- Counted where they work, not where they live
- Employees & self-employed
- One person can hold multiple jobs—number of jobs don’t equal number of people.



Coding Systems

STANDARDIZE & TRACK DATA



Businesses



NAICS

(North American Industry Classification System)

- 1001 Codes
- 5 Levels
- Example:

31-Manufacturing

312-Beverage & Tobacco Product mfg.

3121-Beverage mfg.

31211-Soft Drink & Ice mfg.

312112-Ice mfg.



Occupation Codes

SOURCES USE EITHER SOC OR O*NET CLASSIFICATIONS



Occupations

SOC

(Standard Occupation Classification)

- 800 Codes
- 4 Levels
- Example:

15-0000 – Computer Occupations

15-1100 – Computer Occupations

15-1130 – Software Developers

15-1134 – Web Developer

O*NET SOC

(from Department of Labor)

- >900 Codes
- 5 Levels
- Example:

15-1143 Computer Network Architects

15-1143.00 – Computer Network
Architects

15-1143.01 – Telecommunications
Engineering Specialist



Programs



CIP

(Classification of Instructional Programs)

- 1700 Codes
- 3 Levels
- Example:

26 – Biological & Biomedical Sciences

26.01 – Biology, General

26.0102 – Biomedical Sciences



Traditional LMI

OUTLINE

I. Characteristics of Traditional LMI ✓

I. Core Data Points

II. Coding Systems

II. Specific Traditional LMI Datasets

III. Data Processing



Industry (Business) Sources

THESE DATSETS PROVIDE INFORMATION ON BUSINESSES AND THE
EMPLOYEES THEY HIRE



QCEW

(Quarterly Census of Employment and Wages)

- Crucial dataset
- County level data
- Includes:
 - Job Counts
 - Earnings
 - # of Establishments

Strength:

- Basis of our employee industry data due to job counts and annual earnings data

Weaknesses:

- Lots of suppressions
- Doesn't cover the self-employed



CBP & ZBP

(County and ZIP Business Patterns)

- CBP publishes employment ranges for industries in counties
- ZBP publishes employment ranges for industries in ZIPs

Strength:

- Ranges provide a starting point in dealing with suppressions in other data

Weaknesses:

- Not entirely consistent with QCEW
- Numbers are ranges



QWI

(Quarterly Workforce Indicators)

- Publishes industry:
 - Demographics
 - Hires
 - Separations

Strengths:

- Very accurate count of hiring & separations by industry
- Demographic breakouts of industries

Weaknesses:

- Significant lag time
- Suppressions



Industry Demographic Examples

Industry Gender Breakdown ?



Gender	2016 Jobs	2016 Percent
Males	1,354	75.4%
Females	442	24.6%

Industry Age Breakdown ?



Age	2016 Jobs	2016 Percent
14-18	11	0.6%
19-24	76	4.2%
25-34	364	20.2%
35-44	586	32.6%
45-54	466	25.9%
55-64	239	13.3%
65+	55	3.1%

Industry Race/Ethnicity Breakdown ?



Race/Ethnicity	2016 Jobs	2016 Percent
White	1,630	90.7%
Hispanic or Latino	69	3.9%
Asian	61	3.4%
Black or African American	18	1.0%
Two or More Races	13	0.7%
American Indian or Alaska Native	5	0.3%
Native Hawaiian or Other Pacific Islander	1	0.0%



ACS

(American Community Survey)

- Huge data set with many interesting data points

Strength:

- Publishes industry and occupation demographic proportions

Weakness:

- Sample-size survey
 - Not a complete picture of the nation



Worker Demographic Information

Occupation Gender Breakdown ?



Gender	2016 Jobs	2016 Percent
Males	724	72.3%
Females	278	27.7%

Occupation Age Breakdown ?



Age	2016 Jobs	2016 Percent
14-18	12	1.2%
19-24	72	7.1%
25-34	340	33.9%
35-44	315	31.5%
45-54	179	17.9%
55-64	68	6.8%
65+	16	1.6%

Occupation Race/Ethnicity Breakdown ?



Race/Ethnicity	2016 Jobs	2016 Percent
White	907	90.5%
Hispanic or Latino	37	3.7%
Asian	31	3.1%
Two or More Races	14	1.4%
Black or African American	11	1.1%
American Indian or Alaska Native	1	0.1%
Native Hawaiian or Other Pacific Islander	1	0.1%

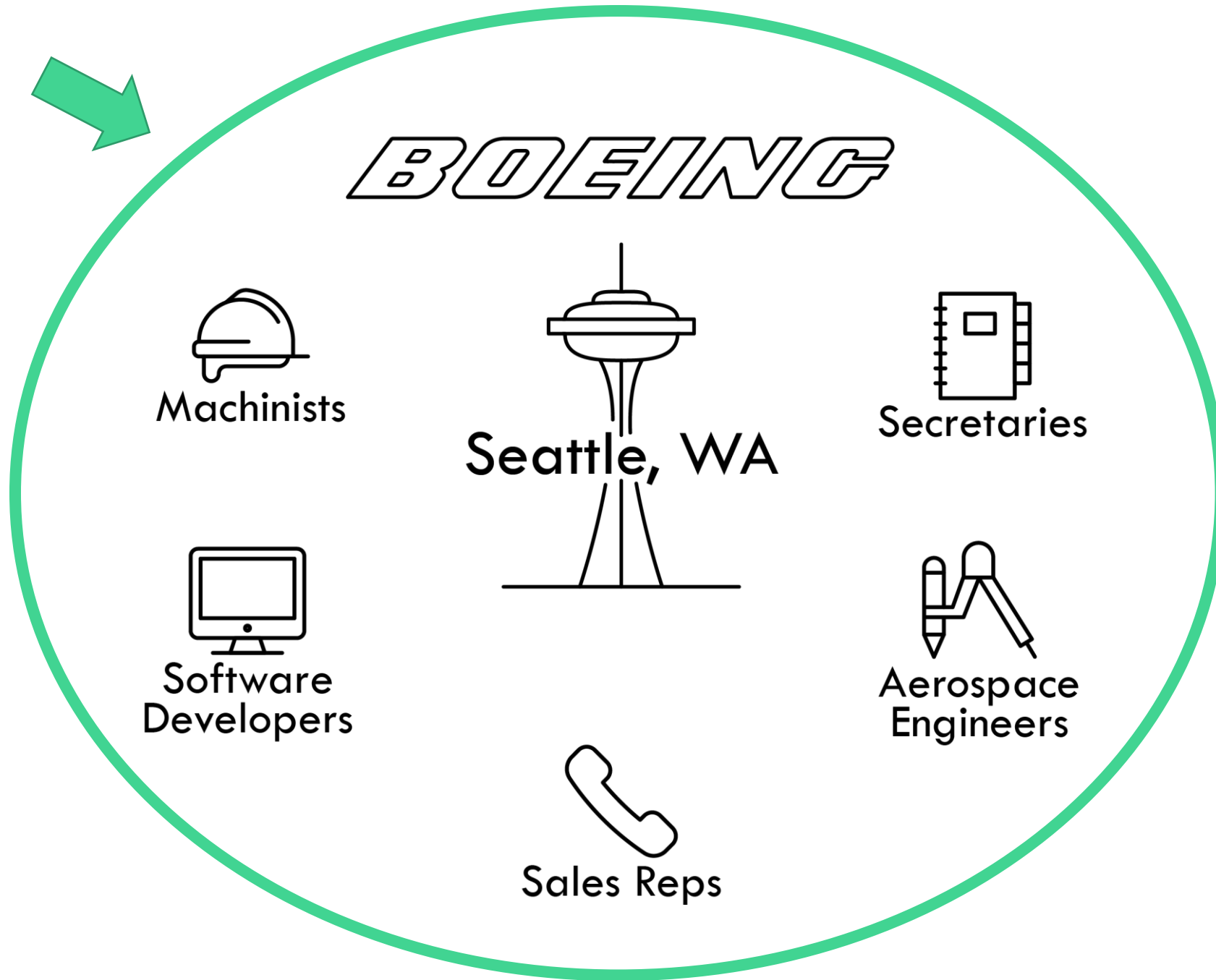


Staffing Pattern

CONNECTION BETWEEN BUSINESSES & WORKERS



Staffing Pattern



NIOEM

(National Industry-Occupation Employment Matrix)

- National table of what industries employ which occupations

Strength:

- Starting point for translating industry figures into occupations

Weakness:

- National level



Occupation Sources

PROVIDE ADDITIONAL INFORMATION ON WORKERS



OES

(Occupational Employment Statistics)

- Publishes estimates for occupations at the metropolitan level
 - Job counts
 - Hourly earnings figures

Strengths:

- Helps regionalize staffing patterns
- Publishes wages

Weaknesses:

- MSA level or Group of Counties
- Doesn't match industry employment
- Weak time-series



Staffing Pattern Example

Staffing Patterns

Computer Systems Design Services in Idaho

Export ▼

5-Digit ▼

Filter ▼

Keep

Hide

Jump To ▼

+ Create Group

SOC	Description	Employed in Industry (2010)	Employed in Industry (2016) ▼	Employed in Industry (2020)	Change (2010 - 2020)	% Change (2010 - 2020)	% of Total Jobs in Industry (2016)	Median Hourly Earnings	Typical Entry Level Education ?	Work Experience Required ?
15-1132	Software Developers, Applications	125	238	287	162	130%	13.3%	\$37.24	Bachelor's degree	None
15-1151	Computer User Support Specialists	71	137	164	93	131%	7.6%	\$19.23	Some college, no degree	None
15-1121	Computer Systems Analysts	76	134	162	86	113%	7.4%	\$37.52	Bachelor's degree	None
15-1133	Software Developers, Systems Software	58	112	135	77	133%	6.2%	\$39.57	Bachelor's degree	None
15-1131	Computer Programmers	71	100	103	32	45%	5.5%	\$32.49	Bachelor's degree	None



County Level Employment

- Modify regional OES estimates to bring them in line with our regional industry totals
- Transform the resulting OES estimates into a regional staffing pattern by converting the jobs to a percentage
- Adjusted staffing patterns are applied to Emsi's industry data to produce occupation data



County Level Earnings

- Emsi looks at which industries employ a given occupation within each county of an OES region
- Occupation wages are adjusted according to the difference between the county industry earnings & OES region industry earnings



O*NET

(Occupational Information Network)

- Publishes knowledge, skill and ability characteristics of occupations

Strengths:

- Allows us to find “similar” occupations.
- More detailed job titles helpful for using job posting analytics

Weakness:

- Survey data

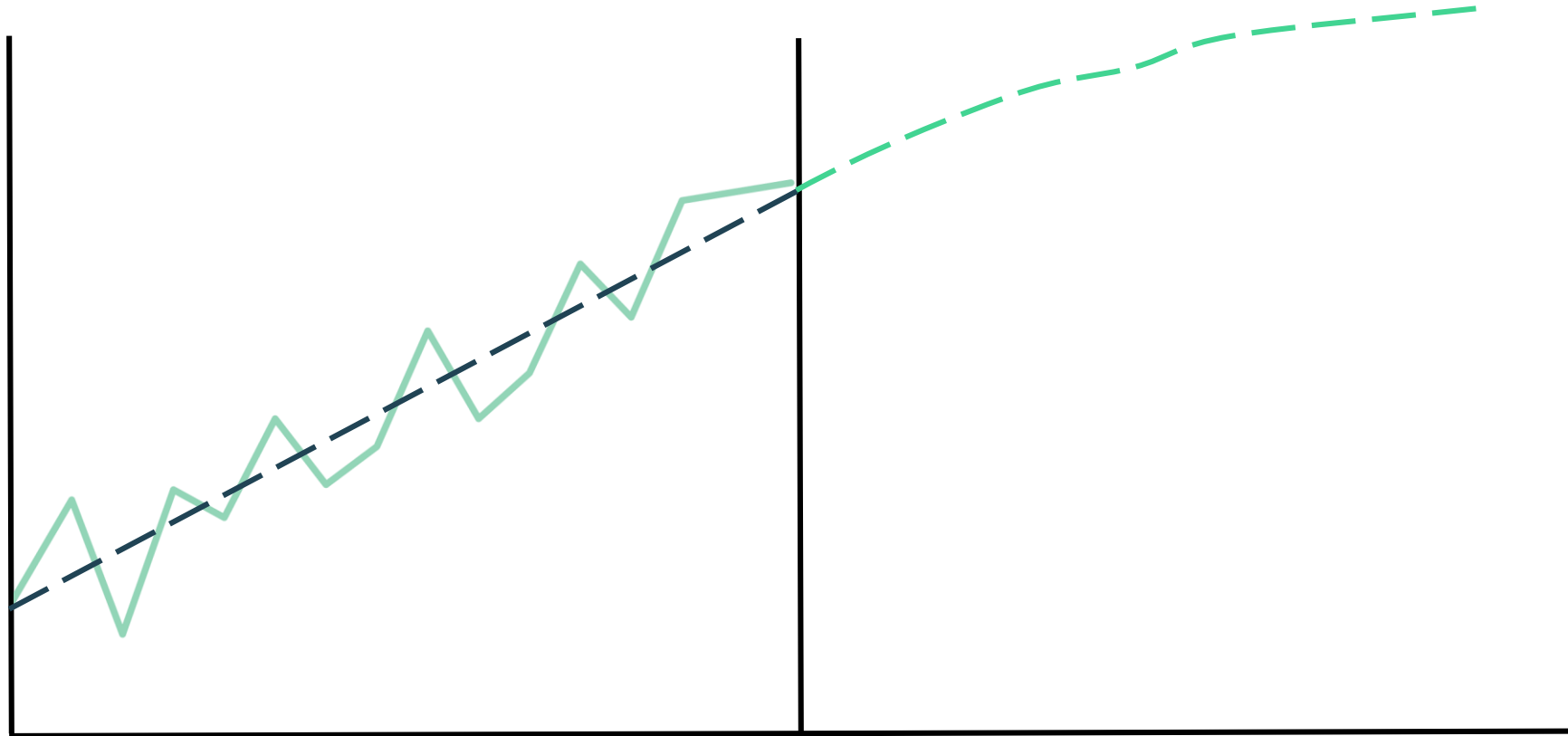


Projections



Historic

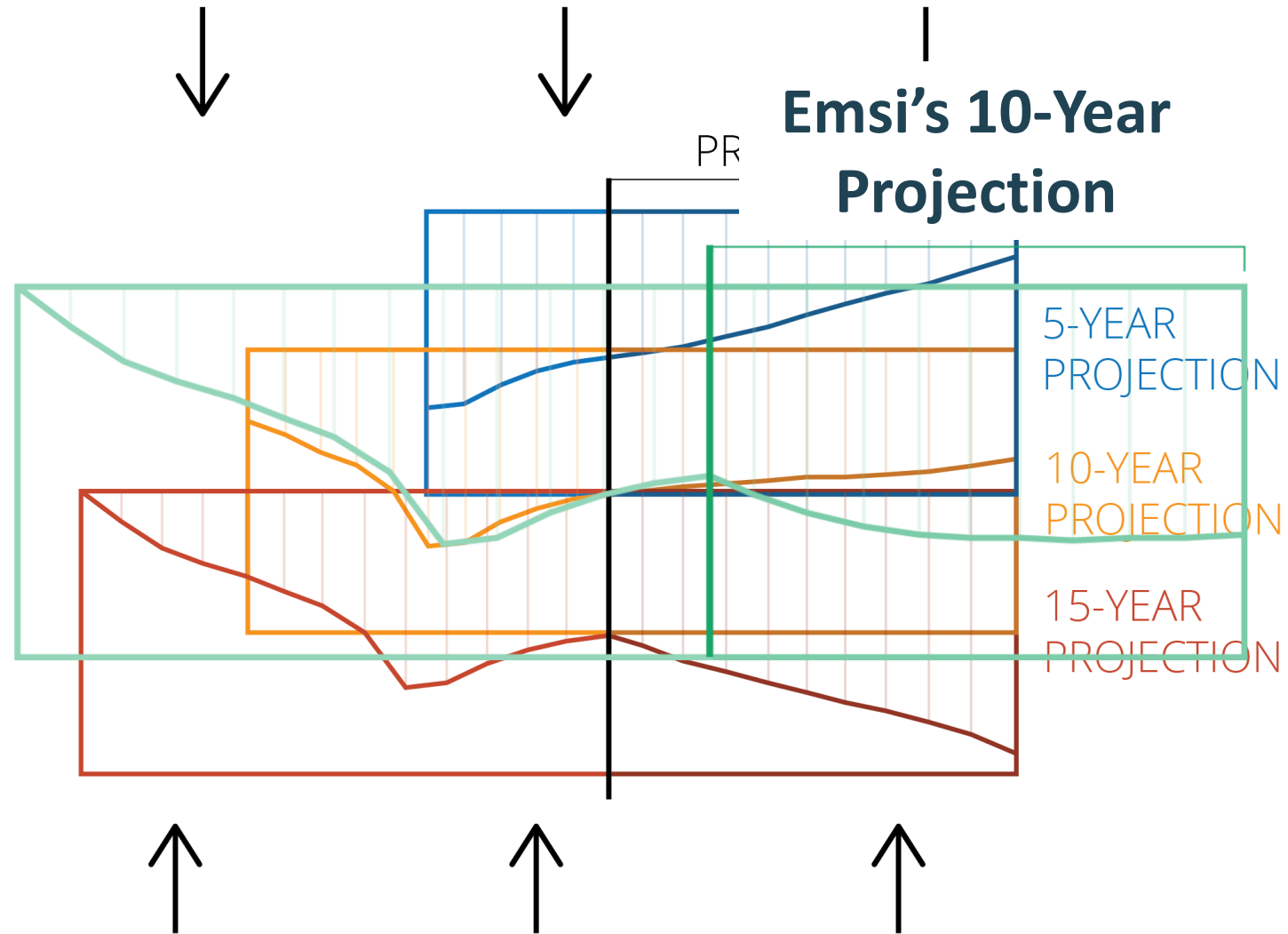
Projected



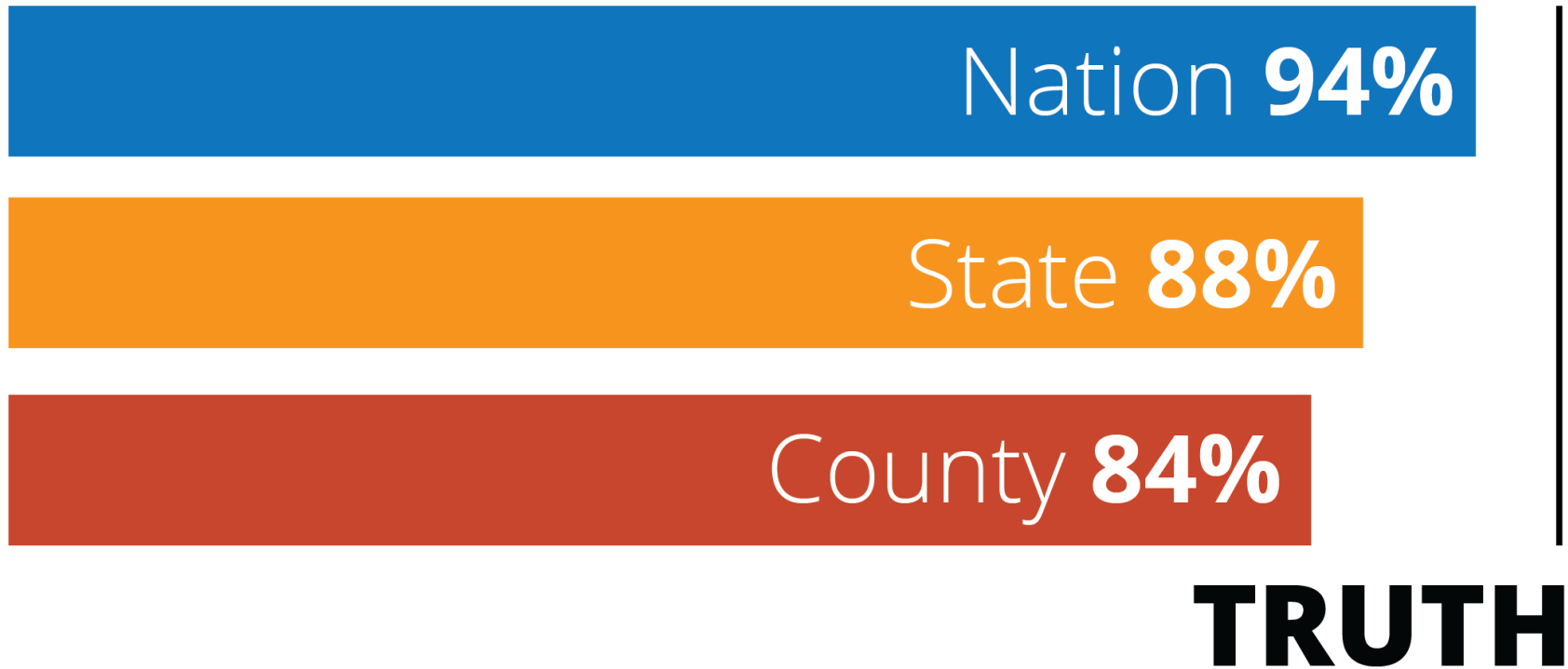
Projections don't take unexpected things into account – because they're unexpected



How do we calculate the Emsi 10 year Projection?



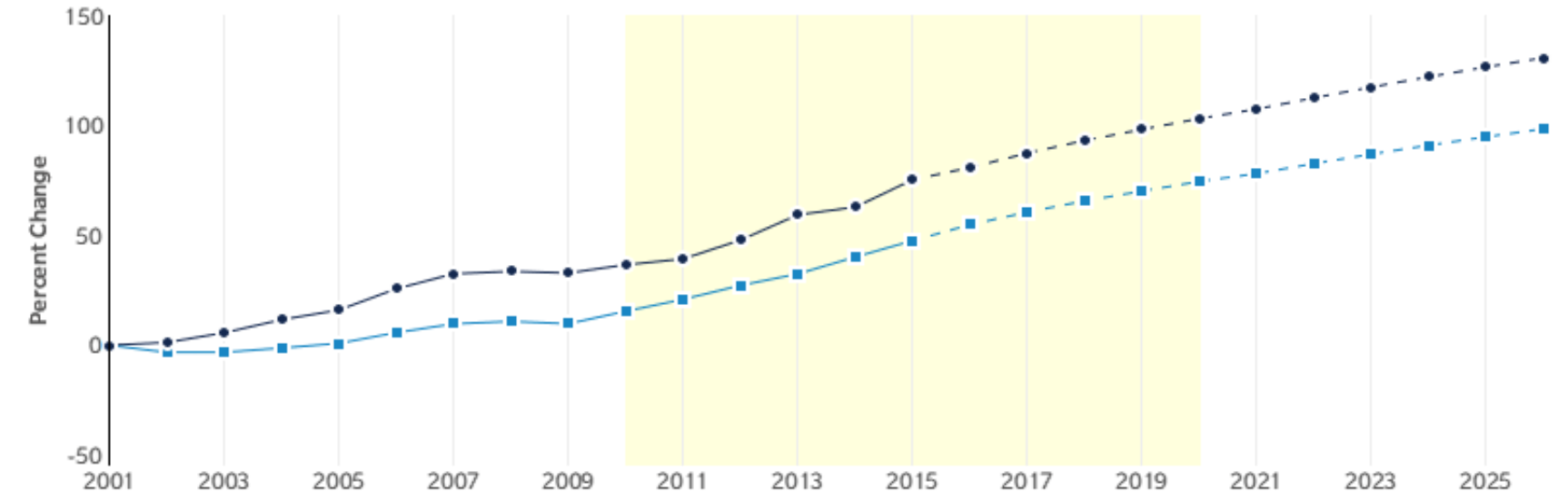
Emsi's Projection Accuracy



Example of Projection

Regional Trends

[Jump to Regional Comparison by Occupation](#)



Region	2010 Jobs	2020 Jobs	Change	% Change
Region	756	1,125	369	48.8%
Nation	129,917	196,468	66,551	51.2%



Program Sources

PROVIDE ADDITIONAL INFORMATION ON STUDENTS



NCES / IPEDS

(National Center for Education Statistics /
Integrated Postsecondary Education Data System)

- Annual counts of completions by program for each college
- Crosswalk of programs and the occupations they typically train for
- Full-time equivalent enrollment
- Demographic data of enrolled students

Strengths:

- Only source for graduates by program
- Great starting point for connecting programs to occupations

Weaknesses:

- Self-reported by institution
- Schools not participating in federal financial aid program not required to report
- Some mappings are too general



CIP to SOC Crosswalk



Programs
(CIP)



Occupations
(SOC)

Emsi Data Experts have updated the NCES matrix and added
~1000 new links to improve mapping!



General Program Connections



Crosswalk Example

Target Occupations

[edit](#)[Jump to Occupation Overview](#)

7,730

Jobs (2016) ?

28% **below** National average

+22.7%

⊕ % Change (2010-2020)

Nation: +27.4%

\$34.44/hr

⊕ Median Hourly Earnings ?

Nation: \$44.07/hr

Occupation	2010 Jobs	Annual Openings	Median Hourly Earnings	Growth (2010 - 2020)	Location Quotient (2010)
Network and Computer Systems Administrators	1,392	44	\$29.74/hr	+16.52%	0.83
Computer Systems Analysts	1,336	61	\$37.52/hr	+29.19%	0.57
Computer Occupations, All Other	1,073	32	\$36.02/hr	+12.86%	0.98
Computer and Information Systems Managers	1,005	34	\$49.34/hr	+20.20%	0.68
Web Developers	756	51	\$19.36/hr	+48.81%	1.21



Traditional LMI

OUTLINE

I. Characteristics of Traditional LMI ✓

I. Core Data Points

II. Coding Systems

II. Specific Traditional LMI Datasets ✓

III. Data Processing



Suppressions

- Data not disclosed by the government
 - When one business in a region makes up the majority of employment in that industry
 - When there are not many businesses in an industry
- How does Emsi account for these?
 - CBP & ZBP ranges create a starting point
 - Data Processing to fill in the rest



Is our method reliable?

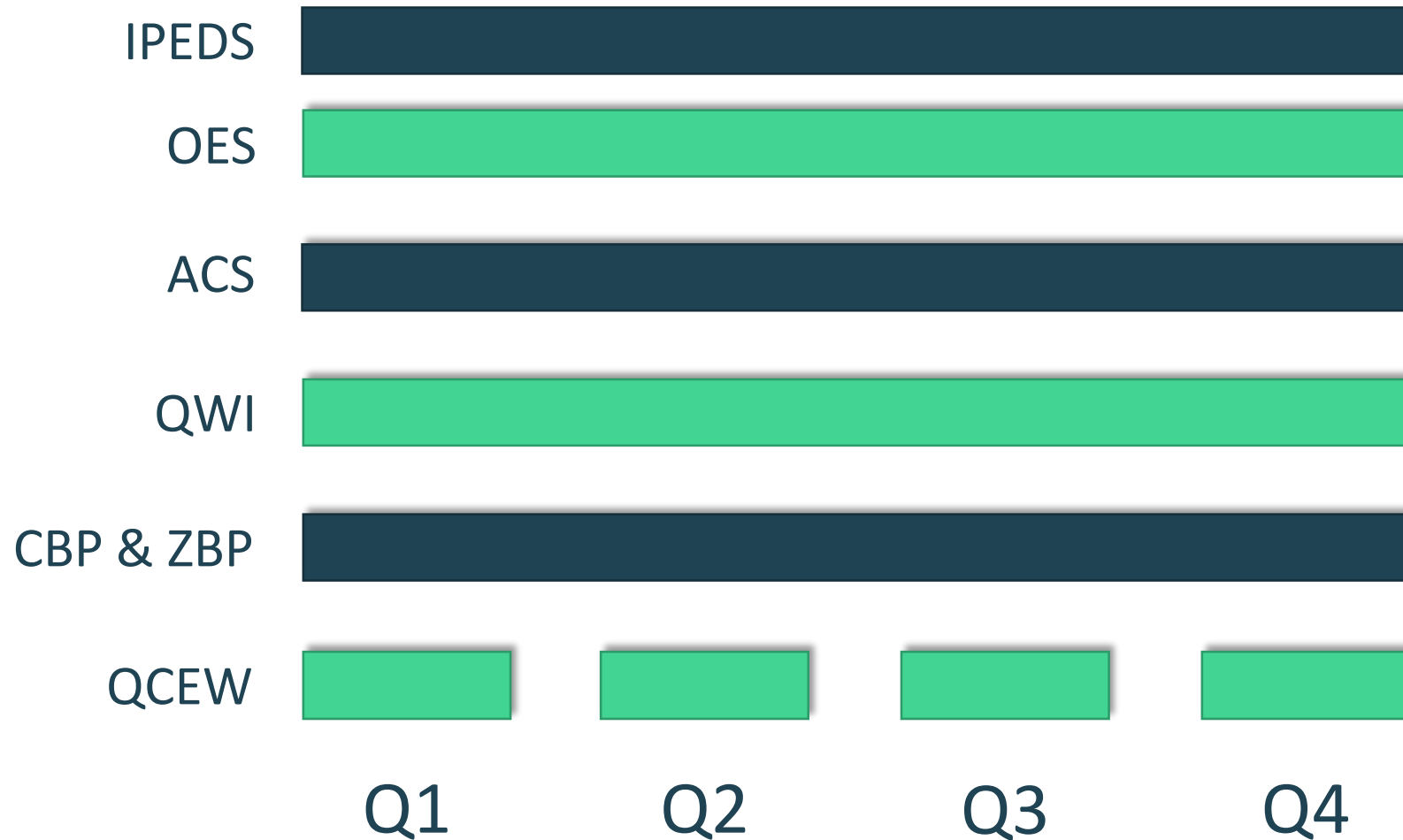


Compiling the Sources

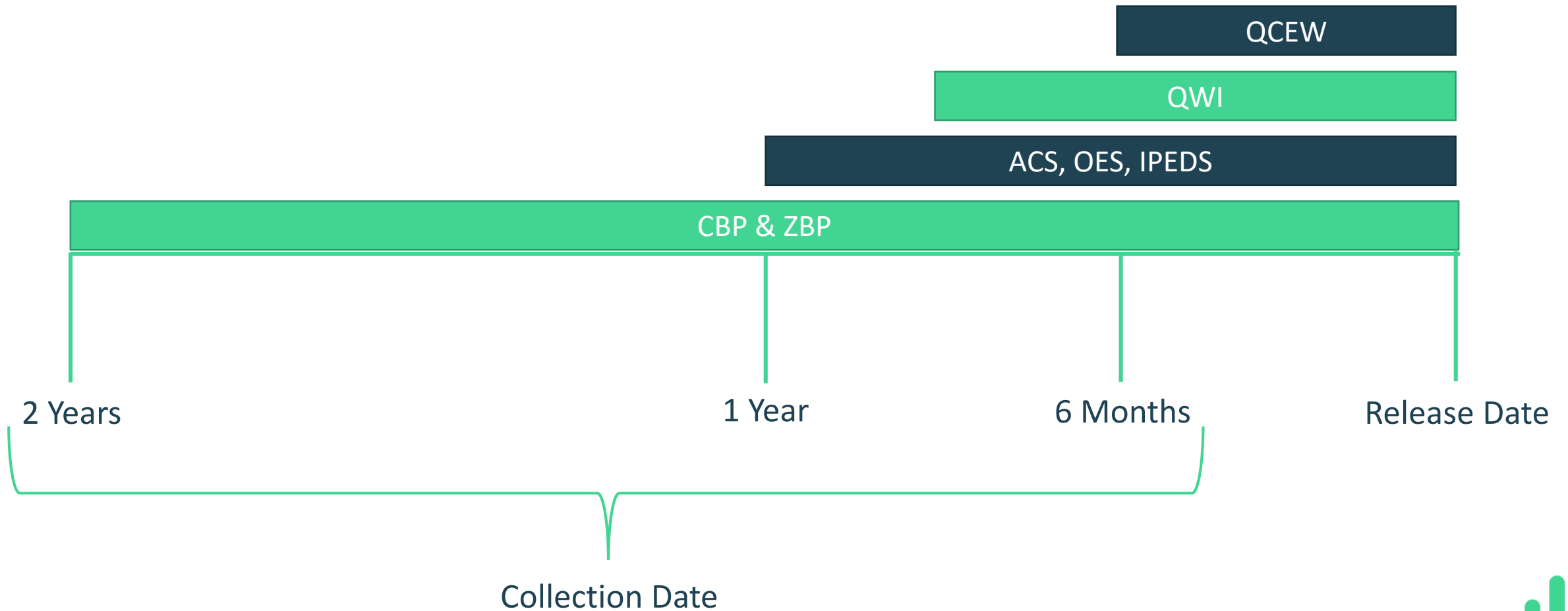
QCEW	ACS	QCEW	QCEW	CES	QCEW	NES
QCEW	QCEW	QCEW	CES	QCEW	QCEW	QCEW
QCEW	CBP & ZBP	QCEW	QCEW	QCEW	QCEW	LAPI
QCEW	QCEW	QCEW	QCEW	CBP & ZBP	QWI	QCEW



Release Schedules: Traditional LMI



Lag Time: Traditional LMI



Emsi Traditional LMI Data Updates

To account for release differences, we update Emsi Traditional LMI Data quarterly!



Traditional LMI

OUTLINE

I. Characteristics of Traditional LMI ✓

I. Core Data Points

II. Coding Systems

II. Specific Traditional LMI Datasets ✓

III. Data Processing ✓



Emsi Data Training

OUTLINE

- I. Tradition LMI 
- II. Real-Time LMI
- III. Traditional and Real-Time LMI combined



Real-Time LMI OUTLINE

- I. Characteristics of Real-Time LMI
 - I. Core Data Points
- II. Specific Real-Time LMI Datasets
- III. Data Processing



Real-Time LMI

- Collected from Private, On-line Sources
 - Job & Resume Posting sites
 - Profiles
- Detailed Look at Current Outlook



Core Data Points

- Skills sought after by potential employers
- Employers who are advertising for the position
- Specific job titles requested by the employer
- Total Monthly Active Postings: Count of all postings
- Unique Monthly Active Postings: Unduplicated count of job postings



What is a “Skill”?

- Common Skills: Broad, basic competencies
 - Character traits: Ethics, assertiveness, etc.
 - General cognitive or physical abilities: Critical thinking, Creativity, etc.
 - Basic interpersonal skills: Leadership, Cooperation, etc.
- Hard Skills: Technical, subject matter specific
 - Sales techniques, online marketing, foreign languages, etc.
- Qualifications: Credentials or licenses granted
 - Registered Nurse, Certified Public Accountant, Oracle Certified Associate Database Administrator, etc.



Posting Intensity

$$\text{Posting Intensity} = \frac{\text{Total Monthly Active Postings}}{\text{Unique Monthly Active Postings}}$$

- A proxy for the hiring effort that is going in to filling a particular opening in a given region
- A high value could indicate that an employer is intensely looking to fill that opening

Job Postings Overview

[Jump to Job Posting Analytics](#)



4,162

Unique Postings (Apr 2011 - Aug 2016) ⓘ
17,837 Total Postings

4 : 1

Posting Intensity (Apr 2011 - Aug 2016) ⓘ

Regional Average: 6 : 1



I. Characteristics of Real-Time LMI

I. Core Data Points

II. Specific Real-Time LMI Datasets

III. Data Processing

Real-Time LMI

OUTLINE



Real-Time LMI Sources



Job Postings

- Used to get information about businesses and the potential employees
- Comprised of 300 million+ unique job postings that are collected and curated.
 - Job boards, recruiting temp agencies & individual company's career pages



Job Postings

- 6-8 million unique active posts/month
- ~400,000 unique businesses over the last 5 years, with >90,000 unique companies/month



Job Posting Strengths

- Data on skills & qualifications
- Companies hiring
- Posting Intensity
- Gives context to Traditional LMI

Job Posting Weaknesses

- Voluntary
- Not all occupations covered
 - Professional, high education types greater representation
 - E.g.: RNs = >80,000 postings
Welders = 901 postings



Profiles Data

- 108 Million Alumni Profiles in Alumni Insights
- Data points:
 - Alma Mater, Occupation, Location, Employer, Job Title, Skills
- Uses:
 - Tracking Outcomes, Building Partnerships, Engage Students



I. Characteristics of Real-Time LMI



I. Core Data Points

II. Specific Real-Time LMI Datasets



III. Data Processing

Real-Time LMI OUTLINE



JPA Processing Steps

- Normalization: Company & Geography
- Job Title & Occupation Classification Tags
- Deduplication
- Skills & Qualifications Tagging
- Geography Reallocation
- Taxonomy Refinements
- Skills Filtering



Company Normalization

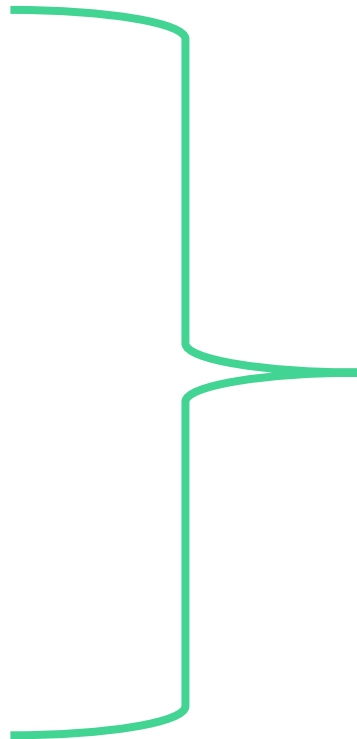
HomeDepot, 10



Home Depot, 49



The Home Depot, 100



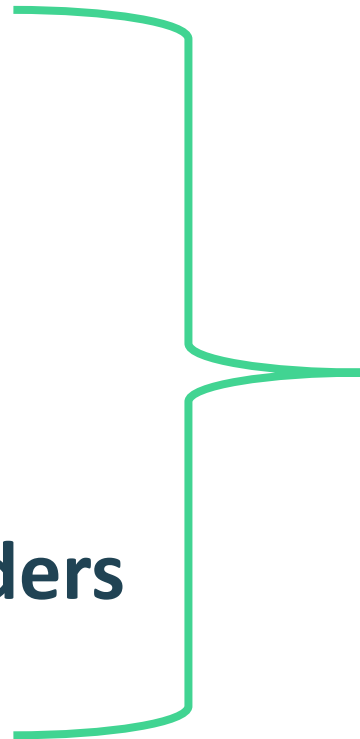
The Home Depot
159



Company Normalization

**Delta Airlines,
50 Flight Attendants**

**Delta Healthcare Providers
10 Physicians**



**Delta Airlines
50 Flight Attendants
&
10 Physicians**



Geography Normalization

Program looks through text to find mention of a location



If Successful

Latitude & Longitude

Any Administrative Location Information (country, state, county, etc.)



Job Title & Occupation Classification

- Tags posts with standardized job title &
- Tags posts with O*NET SOC code
- Necessary to group by occupations &
- Link to Traditional LMI



Deduplication

- One of the most critical processes
 - Unique Job Postings
 - April 2016: 90% of postings were duplicates
- 2 Step process:
 - Exact Matching Step
 - Fuzzy Matching Step



Deduplication: Exact Matching

- Location: Moscow, Idaho → ■ Location: Moscow, Idaho
- Job Title: Sales Representative → ■ Job Title: Sales Representative
- Company: Emsi → ■ Company: Emsi
- Job Description: We are looking for highly-motivated Sales Representatives to join our team. In this position, you will play a key role providing clients with the personal attention and professional, knowledgeable customer service. → ■ Job Description: We are looking for highly-motivated Sales Representatives to join our team. In this position, you will play a key role providing clients with the personal attention and professional, knowledgeable customer service.



Deduplication: Fuzzy Matching

- Location: Moscow, Idaho → ■ Location: Moscow, ID
- Job Title: Sales Representative → ■ Job Title: Sales Representative
- Company: Emsi → ■ Company: EMSI
- Job Description: We are looking for highly-motivated Sales Representatives to join our team. In this position, you will play a key role providing clients with the personal attention and professional, knowledgeable customer service. → ■ Job Description: Play a key role providing clients with the personal attention and professional, knowledgeable customer service and join our Sales team.



Skills & Qualifications Tagging

- Necessary to get skills and certifications associated with occupations
- Library of ~25,000 items
- Postings are tagged with these standardized skills
 - Improvements to tagging
- Skills & qualifications derived by counting their occurrence in all postings available for that occupation



Processing Cont.

- Geography reallocation: Assign counties to those without a county
- Taxonomy refinements: Adjust O*NET & SOC codes to match Emsi's
- Skills filtering:
 - Discard those skills and certifications not meeting the 0.70 confidence interval Emsi set
 - Discard generic skills – to remove noise and make posting data more useful



Job Postings: Updates

- Updated Monthly
- Lag time of <2 weeks following the end of the month



I. Characteristics of Real-Time LMI

I. Core Data Points

II. Specific Real-Time LMI Datasets

III. Data Processing

Real-Time LMI OUTLINE



Summary of Real-Time LMI

- Provides: Skills, Employer information, Job Titles, Total & Unique Postings
- To create our Real-Time LMI dataset:
 - Emsi collects and processes Job Posting data
- Updated monthly



I. Characteristics of Real-Time LMI ✓

I. Core Data Points

II. Specific Real-Time LMI Datasets ✓

III. Data Processing ✓

IV. Miscellaneous LMI Sources ✓

Real-Time LMI OUTLINE



Emsi Data Training

OUTLINE

- I. Tradition LMI ✓
- II. Real-Time LMI ✓
- III. Traditional and Real-Time LMI combined



Two Types of LMI

TRADITIONAL LMI



Government Sources

REAL-TIME LMI



Private On-line Sources



Traditional LMI

Strengths

- Highly Structured
- Projections & Trends
- Total Coverage

Traditional LMI

Weaknesses

- Lack of Detail
- No Localized Skills Information
- Outdated (maybe)
- No Employer Data



Real-Time LMI

Strengths

- Provides Details on Job Descriptions
- Connects Jobs to Specific Businesses
- Reflects Current Outlook & Demand

Real-Time LMI

Weaknesses

- Smaller Portion of Economy Represented
- Non-Standardized Structure
- Unable to Produce Projections



Traditional LMI

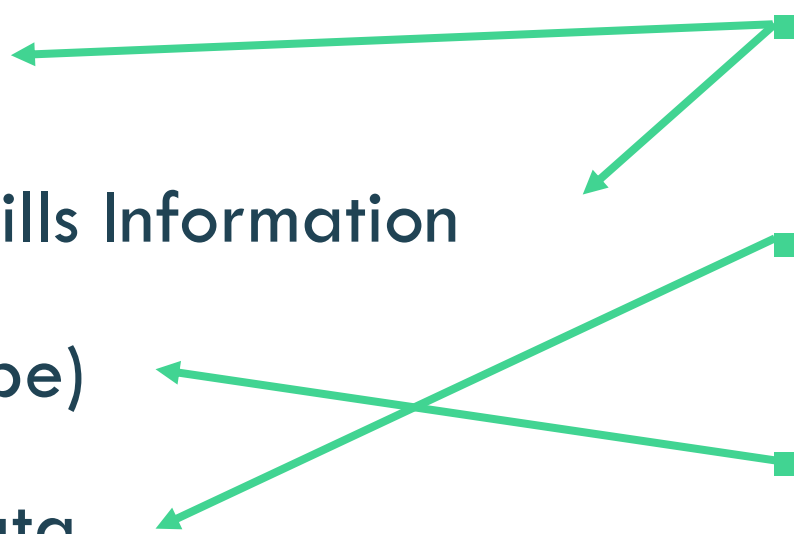
Weaknesses

- Lack of Detail
- No Localized Skills Information
- Outdated (maybe)
- No Employer Data

Real-Time LMI

Strengths

- Provides Details on Job Descriptions
- Connects Jobs to Specific Businesses
- Reflects Current Outlook & Demand



- I. Why Data? ✓
- II. What is LMI? ✓
- III. Lenses to View LMI
- IV. Regions Covered
- V. Summary

LMI

Introduction

OUTLINE



Lenses to View LMI Data



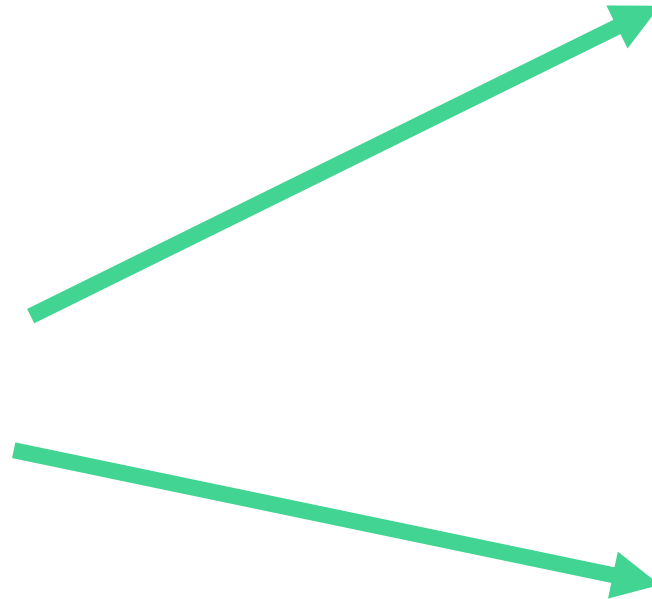
People



Businesses



People



Students



Workers



Businesses



- I. Why Data? ✓
- II. What is LMI? ✓
- III. Lenses to View LMI ✓
- IV. Regions Covered
- V. Summary

LMI

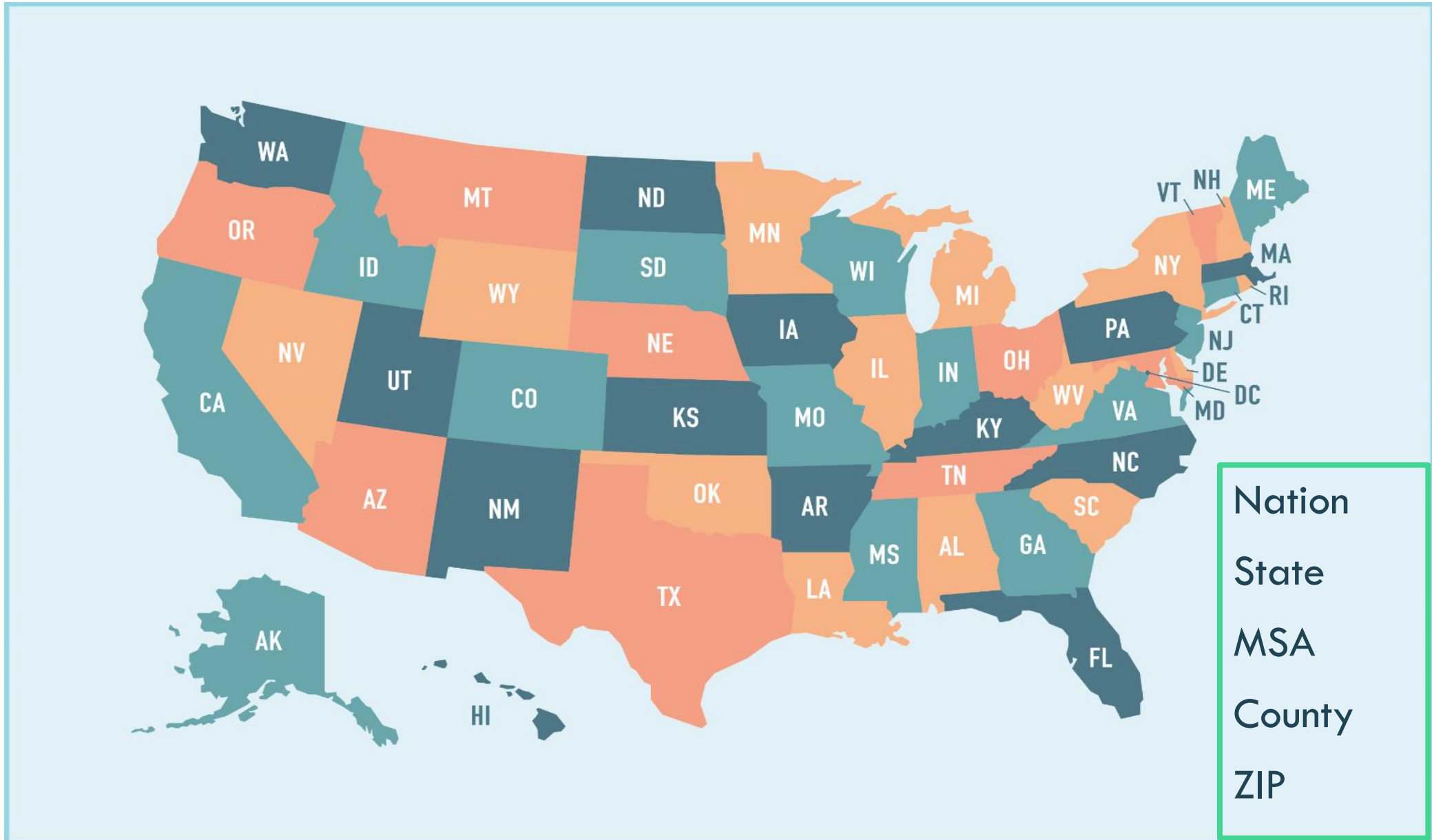
Introduction

OUTLINE



Regions

WE CREATE NATIONAL LMI



MSA

Metropolitan Statistical Area and Micropolitan Statistical Area

MSAs

- Population 50K+
- 374 MSAs

μSAs

- Population 10-50K+
- 581 μSAs

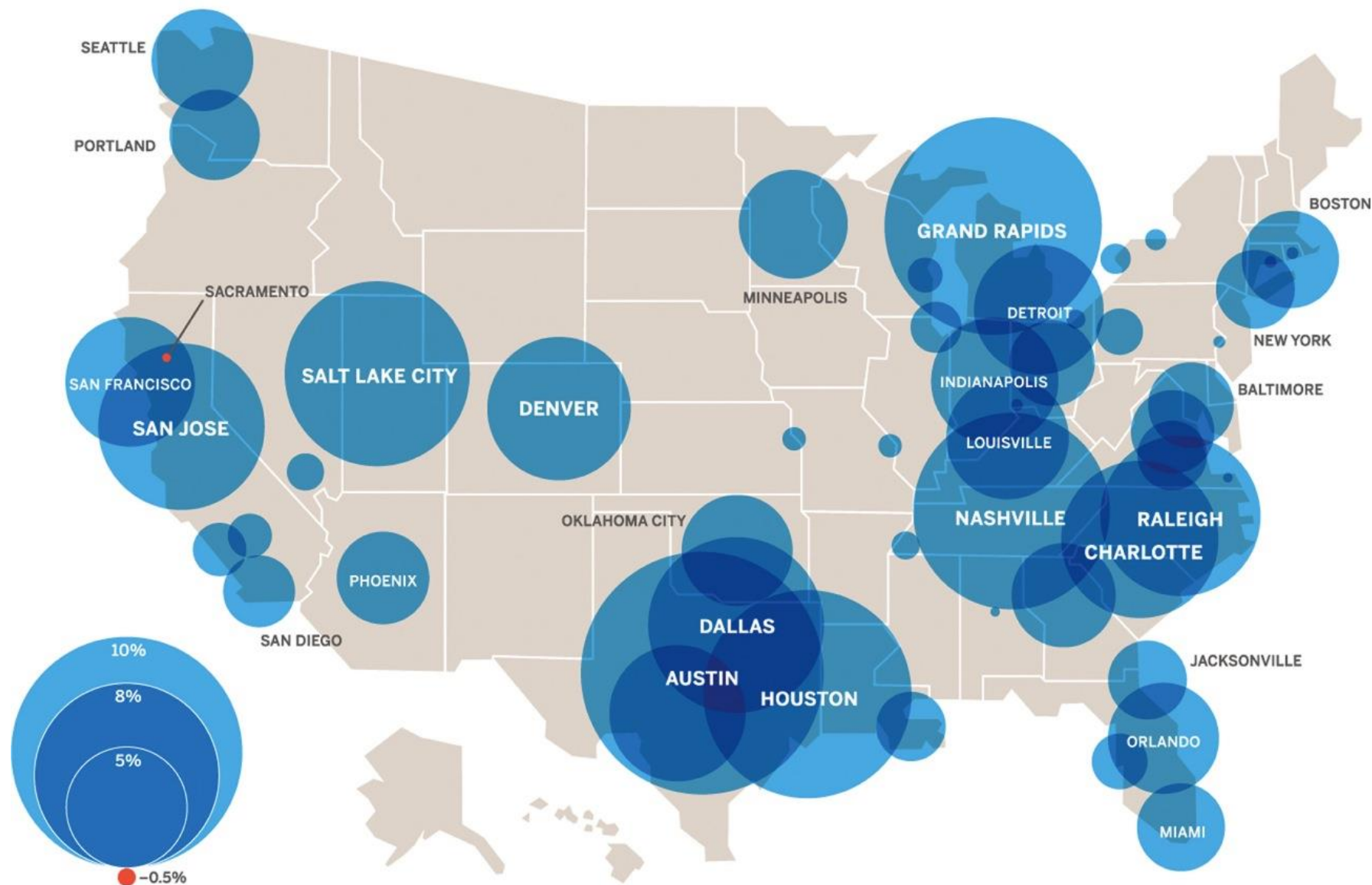


MSAs

- Larger, multi-county regions
- Describe major economies well.
- Shouldn't be combined to create larger regions.
- Shouldn't be broken up into individual counties.
- Powerful for national comparisons.



MSAs



Counties

- Can simply be a single county.
 - Many users focus on one or two counties.
- Easy to break out into component counties and analyze.
- Combine to build larger regions



ZIPs

- Regions within counties
- Can be combined to build larger regions
- Best when combined
- Weak when used one at a time



Summary

- Describes businesses, workers and students
- Provides: Job counts, earnings, change and program completions.
- To create our Traditional LMI dataset:
 - Gather dozens of sources
 - Compile them together
- Updated quarterly



- I. Why Data? ✓
- II. What is LMI? ✓
- III. Lenses to View LMI ✓
- IV. Regions Covered ✓
- V. Summary

LMI

Introduction

OUTLINE



Summary

- Connect people and businesses to inform
 - Students on career and program options
 - Colleges in program alignment and development
- Both Traditional LMI & Real-time LMI are in Emsi's LMI Data



Certification Summary

- Best use case of Labor Market Information is to use Traditional & Real-Time LMI together!
 - Traditional LMI can give perspective on how employment has been changing for occupations filled by program graduates.
 - Real-Time LMI can give perspective on content relevance of a given program.



Certification Summary Continued

- What could happen if you only use Real-Time and don't incorporate Traditional LMI?
 - Nursing program is being created because there was a plethora of postings.
 - Already ~1000 nurses being graduated in your region
 - Waste of resources to create program that won't fill gaps in your region
- Real-Time & Traditional LMI provide complete picture to help you make the best decisions!



Thank You

For more information, please visit our website:
economicmodeling.com

