March 2021

Labor Market

Analysis

Data Science

Prepared by the Central Valley/Mother Lode Center of Excellence

Table of Contents

COVID-19 Statement: This report includes employment projection data by Emsi. Emsi's projections are modeled on recorded (historical) employment figures and incorporate several underlying assumptions, including the assumption that the economy during the projection period will be at approximately full employment or potential output. To the extent that a recession or labor shock, such as the economic effects of COVID-19, can cause long-term structural change, they may impact the projections. At this time, it is not possible to quantify the impact of COVID-19 on projections of industry and occupational employment. Other measures such as unemployment rates and monthly industry employment estimates will reflect the most recent information on employment and jobs in the state and, in combination with input from local employers, may help validate current and future employment needs as depicted here.

If for any reason this document is not accessible or if you have specific needs for readability, please contact us and we will do our utmost to accommodate you with a modified version. To make a request, contact Nora Seronello by phone at (209) 575-6894 or by email seronellon@mjc.edu.

Summary

Please note the COVID-19 statement on page 2 when considering this report's findings.

This study conducted by the Central Valley/Mother Lode Center of Excellence examines labor market demand, wages, skills, and postsecondary supply for data science. Four occupations related to data science were identified for Fresno City College:

- 15-1211, Computer Systems Analysts
- 15-1221, Computer and Information Research Scientists
- 15-2041, Statisticians
- 15-2098, Data Scientists and Mathematical Science Occupations, All Other

Key findings:

- **Occupational demand —** Nearly 1,650 workers were employed in jobs related to data science in 2019 in the South Central Valley/Southern Mother Lode (SCV/SML) subregion. The largest occupation is computer systems analysts with 1,165 workers in 2019, a projected growth rate of 8% over the next five years, and 93 annual openings.

- **Wages —** Statisticians earn the highest entry-level wage, $38.66/hour in the subregion and $27.91/hour in the region.

- **Employers —** Employers with the most job postings in the subregion are Anthem Blue Cross, Bayer Corporation, and Jazz Pharma.

- **Occupational titles —** The most common occupational title in job postings in the subregion is computer systems analysts. The most common job title is business systems analyst.

- **Skills and certifications —** The top baseline skill is communication, the top specialized skill is SQL, and the top software skill is SQL. The most in-demand certification is a driver's license.

- **Education —** A master's degree is typically required for computer and information research scientists, and statisticians while a bachelor's degree is typically required for computer systems analysts and data scientists and mathematical science occupations (all other).

- **Supply —** Analysis of postsecondary completions in the region shows that on average 85 awards were conferred in the Central Valley/Mother Lode region each year.

Based on a comparison of occupational demand and supply, there is an undersupply of 95 trained workers in the subregion and 124 workers in the region. The COE recommends expanding existing computer systems analysis programs to meet the regional need for more workers in this field. To be clear, community college program completers will likely need to transfer to a four-year college or university to obtain a bachelor's degree before entering employment. Colleges considering this program should partner with four-year colleges and universities and local employers to ensure their training provides the skills and certifications needed to transfer to a four-year postsecondary institution and obtain a job in this field.

Introduction

The Central Valley/Mother Lode Center of Excellence was asked by Fresno City College to provide labor market information for data science. The geographical focus for this report is the South Central Valley/Southern Mother Lode (SCV/SML) subregion, but regional demand and supply data has been included for broader applicability and use. The average living wage for a single adult in the SCV/SML subregion is $10.30/hour.[1] Analysis of the program and occupational data related to data science resulted in the identification of applicable occupations. The Standard Occupational Classification (SOC) System codes and titles used in this report are:

- 15-1211, Computer Systems Analysts
- 15-1221, Computer and Information Research Scientists
- 15-2041, Statisticians
- 15-2098, Data Scientists and Mathematical Science Occupations, All Other

The occupational titles, job descriptions, sample job titles, and knowledge and skills from the Bureau of Labor Statistics and O*NET OnLine are shown below. O*NET data was not available for Data Scientists and Mathematical Science Occupations, All Other.

Computer Systems Analysts

Job Description: Analyze science, engineering, business, and other data processing problems to develop and implement solutions to complex applications problems, system administration issues, or network concerns. Perform systems management and integration functions, improve existing computer systems, and review computer system capabilities, workflow, and schedule limitations. May analyze or recommend commercially available software.
Knowledge: Computers and Electronics, English Language, Customer and Personal Service, Mathematics, Administration and Management
Skills: Active Listening, Critical Thinking, Reading Comprehension, Speaking, Systems Analysis

Computer and Information Research Scientists

Job Description: Conduct research into fundamental computer and information science as theorists, designers, or inventors. Develop solutions to problems in the field of computer hardware and software.
Knowledge: Computers and Electronics, Mathematics, Engineering and Technology, English Language, Administration and Management
Skills: Complex Problem Solving, Critical Thinking, Judgement and Decision Making, Active Listening, Reading Comprehension

Statisticians

Job Description: Develop or apply mathematical or statistical theory and methods to collect, organize, interpret, and summarize numerical data to provide usable information. May specialize in fields such as biostatistics, agricultural statistics, business statistics, or economic statistics. Includes mathematical and survey statisticians.
Knowledge: Mathematics, English Language, Computers and Electronics
Skills: Mathematics, Critical Thinking, Reading Comprehension, Active Listening, Complex Problem Solving

---

[1] The term "living wage" in Center of Excellence reports is calculated by averaging the self-sufficiency wages from the Insight Center's California Family Needs Calculator for each county in the subregion: https://insightcced.org/tools-metrics/self-sufficiency-standard-tool-for-california/.

Occupational Demand

The South Central Valley/Southern Mother Lode subregion employed 1,649 workers in data science occupations in 2019 (Exhibit 1). The largest occupation is computer systems analysts with 1,165 workers in 2019. This occupation is projected to grow by 8% over the next five years and has the greatest number of projected annual openings, 93.
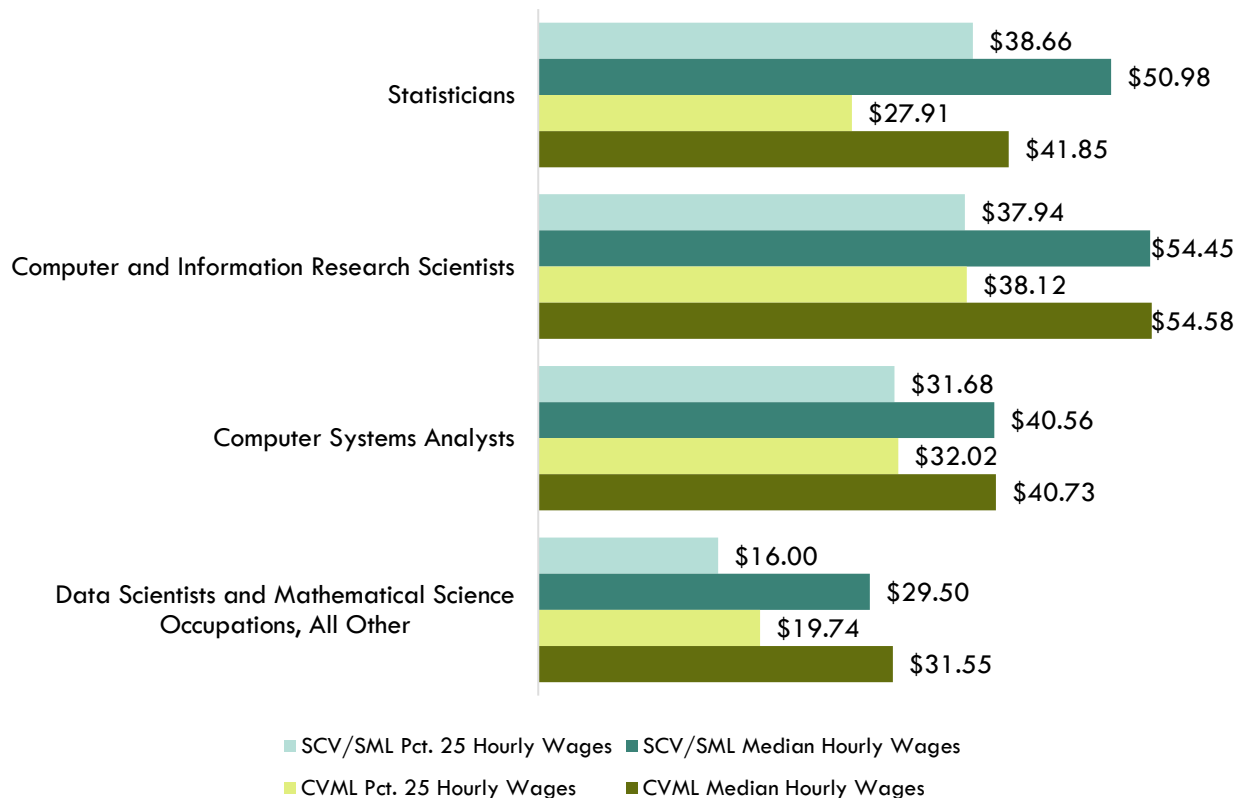
Exhibit 1. Data science employment and occupational projections in the SCV/SML subregion

Occupation	2019 Jobs	2024 Jobs	5-Year Change	5-Year % Change	Annual Openings
Computer Systems Analysts	1,165	1,256	91	8%	93
Computer and Information Research Scientists	319	334	14	5%	26
Statisticians	107	133	26	24%	13
Data Scientists and Mathematical Science Occupations, All Other	57	67	10	18%	7
TOTAL	**1,649**	**1,789**	**141**	**9%**	**139**

Wages

Exhibit 2 compares the entry-level and experienced wages of the data science occupations. Statisticians earn the highest entry-level wage, $38.66/hour in the subregion and $27.91/hour in the region.

Exhibit 2. Entry-level and experienced wage comparison in the SCV/SML subregion and region

Statisticians
- $38.66
- $50.98
- $27.91
- $41.85

Computer and Information Research Scientists
- $37.94
- $54.45
- $38.12
- $54.58

Computer Systems Analysts
- $31.68
- $40.56
- $32.02
- $40.73

Data Scientists and Mathematical Science Occupations, All Other
- $16.00
- $29.50
- $19.74
- $31.55

Legend: SCV/SML Pct. 25 Hourly Wages ■ SCV/SML Median Hourly Wages ■ CVML Pct. 25 Hourly Wages ■ CVML Median Hourly Wages

Job Postings

There were 281 job postings for the four occupations in the SCV/SML subregion from August 2020 to January 2021.[2] The employers with the most job postings are listed in Exhibit 3.

Exhibit 3. Top employers of data science occupations by number of job postings

Employer	Job Postings	% Job Postings
Anthem Blue Cross	37	17%
Bayer Corporation	6	3%
Jazz Pharma	6	3%
Bixly Inc	5	2%
Amentum	4	2%
General Dynamics	4	2%
Jt4 LLC	4	2%
Northrop Grumman	4	2%
Onepoint Hcm	4	2%
Advanced Network Solutions	3	1%

Exhibit 4 shows how job postings for the targeted occupations in the SCV/SML subregion are distributed across six O*NET OnLine occupations. The occupational title computer systems analysts is listed in 166 job postings. Note how this occupational title dominates the job posting results. Common job titles in postings include business systems analyst in 14 job postings, data scientist in seven job postings, and data scientist co-op in five job postings.

Exhibit 4. Top occupational titles in job postings for data science

Occupational Title	Job Postings	% of Job Postings
Computer Systems Analysts	166	59%
Clinical Data Managers	51	18%
Computer and Information Research Scientists	42	15%
Statisticians	12	4%
Biostatisticians	9	3%
Informatics Nurse Specialists	1	0%

Salaries

Exhibit 5 shows the "Market Salaries" for data science occupations that are calculated by Burning Glass which uses a machine learning model built off of millions of job postings every year, and accounts for adjustments based on locations, industry, skills, experience, education requirements, among other variables.

---

[2] Other than occupation titles and job titles, the categories below can be counted one or multiple times per job posting, and across several areas in a single posting. For example, a skill can be counted in two different skill types, and an employer can indicate more than one education level.

Exhibit 5. Salaries for data science occupations

Market Salary Percentile	Salary Amount
10th Percentile	$43,645
25th Percentile	$61,827
50th Percentile	$82,978
75th Percentile	$101,182
90th Percentile	$118,873

Education

Of the 281 job postings, 200 listed an education level preferred for the positions being filled. Of those, 77% requested a bachelor's degree, 28% requested a master's degree, and 18% requested a doctoral degree (Exhibit 6). A job posting can indicate more than one education level. Hence, the percentages shown in the chart below total more than 100%.

Exhibit 6. Education levels requested in job postings for data science

Education Level	Job Postings	% of Job Postings
Bachelor's degree	153	77%
Master's degree	55	28%
Doctoral degree	35	18%
High school or vocational training	15	8%
Associate degree	10	5%

Baseline and Specialized Skills

Exhibit 7 depicts the top baseline and specialized skills for the targeted occupations. The three most important baseline skills are communication, 25% of job postings, troubleshooting, 25%, and teamwork/collaboration, 20%. The top three specialized skills are SQL, 21% of job postings, project management, 18%, and technical support, 17%.
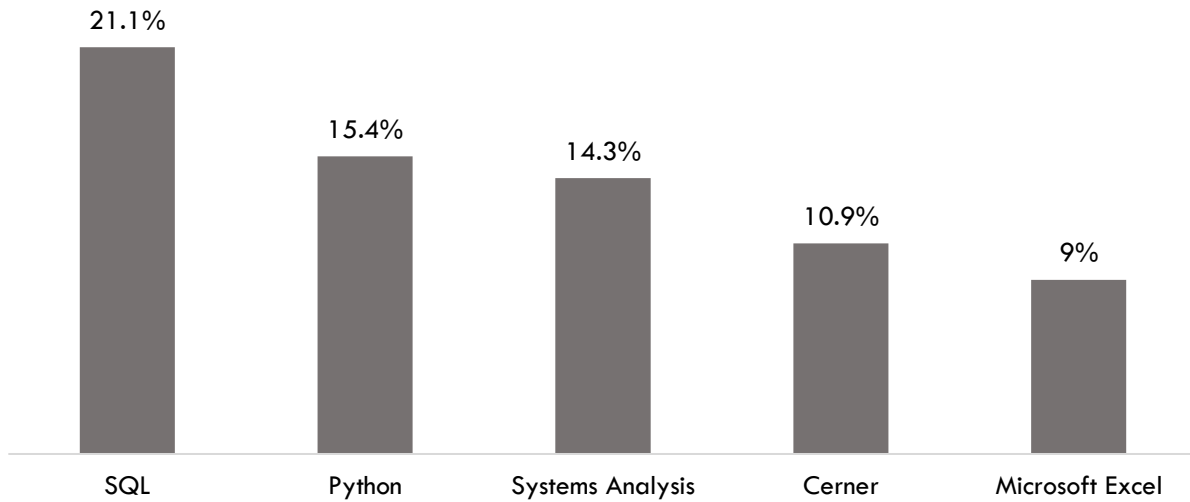
Exhibit 7. In-demand data science baseline and specialized skills
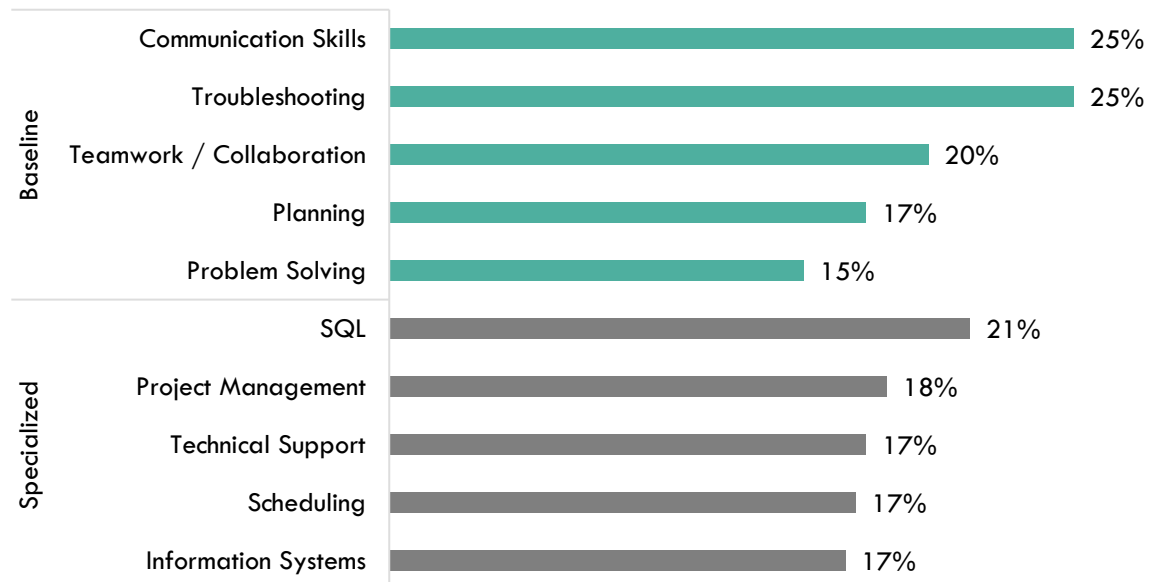
Software Skills

Analysis also included the software skills most in demand by employers. SQL and Python were the top two software skills identified in job postings (Exhibit 8).

Exhibit 8. In-demand data science software skills

Certifications

Of the 281 job postings, 62 contained certification data. Among those, 48% indicated a need for a driver's license. The next top certifications are security clearance and IT Infrastructure Library (Exhibit 9). (Due to the low number of job postings with certifications listed, the chart below may not be representative of the full sample.)

Exhibit 9. Top data science certifications requested in job postings
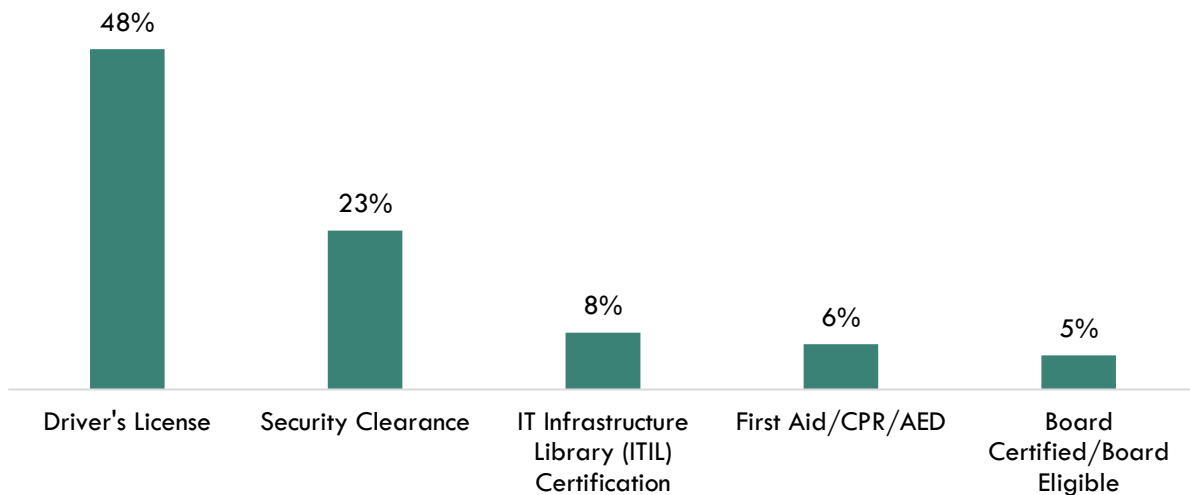
Education, Work Experience & Training

A master's degree is typically required for computer and information research scientists, and statisticians while a bachelor's degree is typically required for computer systems analysts and data scientists and mathematical science occupations (all other) (Exhibit 10).

Exhibit 10. Education, work experience, training, and Current Population Survey results for data science occupations[3]

Occupation	Typical Entry-level Education	Work Experience Required	Typical On-The-Job Training	CPS
Computer Systems Analysts	Bachelor's degree	None	None	20.9%
Computer and Information Research Scientists	Master's degree	None	None	1.8%
Statisticians	Master's degree	None	None	6.9%
Data Scientists and Mathematical Science Occupations, All Other	Bachelor's degree	None	None	6.9%

---

[3] "Labor Force Statistics from the Current Population Survey," Bureau of Labor Statistics, https://www.bls.gov/cps/.

Supply

Analysis of program data from the California Community Colleges Chancellor's Office Data Mart included the TOP code and title: 070600 - Computer Science (Transfer). Analysis of the last three years of data shows that, on average, 85 awards were conferred in the Central Valley/Mother Lode region each year (Exhibit 11).

Please note: There were no community-college-level computer-data-science-specific programs found in the region. There are zero completions for 0707.30 Computer Systems Analysts and 0799.00 Other Information Technology.

Exhibit 11. Postsecondary supply for data science occupations in the region

TOP Code - Title	Colleges	Associate Degree	Associate for Transfer Degree	Subtotal
070600 - Computer Science (Transfer)	Bakersfield	0	14	14
	Cerro Coso	1		1
	Clovis	20		20
	Merced		13	13
	Modesto	24	1	26
	Reedley College	9		9
	San Joaquin Delta		2	2
TOTAL		**55**	**30**	**85**

There is an undersupply of 95 data science workers in the SCV/SML subregion and 124 workers in the region (Exhibit 12).

Exhibit 12. Data science workforce annual demand and supply in the SCV/SML subregion and region

	Demand	Gap	Supply
SCV/SML	139	95	44
CVML	209	124	85
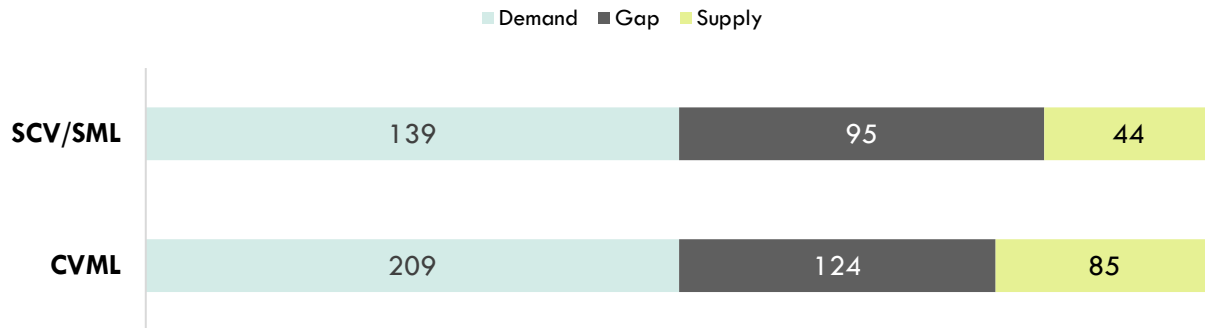
Student Outcomes

Exhibit 13 summarizes employment and wage outcomes from the California Community College Chancellor's Cal-PASS Plus LaunchBoard for the TOP code related to data science. Among students who completed a computer systems analysis program, 22% reported a median change in earnings and 75% attained a living wage.

Exhibit 13. Regional metrics for the TOP code related to data science

Metric	Computer Systems Analysis 070730
Students Who Got a Degree or Certificate or Attained Apprenticeship Journey Status	*
Number of Students Who Transferred	*
Job Closely Related to Field of Study	*
Median Change in Earnings	22%
Attained a Living Wage	75%
* denotes data not available.	

Conclusion

The entry-level wages of the four occupations exceed the SCV/SML subregion's average living wage. There were 281 job postings in the past six months for occupations related to data science in the subregion. Analysis of skills and certification requirements in job postings indicates:

- The top baseline skill is communication, and the top specialized skill is SQL.
- The top software skill is SQL.
- The top certification is a driver's license.

There is an undersupply of trained workers, a shortage of 95 in the SCV/SML subregion and 124 in the region.

Recommendation

Based on these findings, it is recommended that Fresno City College work with the regional director, the college's advisory board, and local industry in the development of programs to address the shortage of data science in the region.

The COE recommends expanding existing computer systems analysis programs to meet the regional need for more workers in this field. To be clear, community college program completers will likely need to transfer to a four-year college or university to obtain a bachelor's degree before entering employment.

Colleges considering this program should partner with four-year colleges and universities and local employers to ensure their training provides the skills and certifications needed to transfer to a four-year postsecondary institution and obtain a job in this field.

Appendix A: Methodology & Data Sources

Data Sources

Labor market and educational supply data compiled in this report derive from a variety of sources. Data were drawn from external sources, including the Economic Modeling Specialists, Inc., the California Community Colleges Chancellor's Office Management Information Systems Data Mart and the National Center for Educational Statistics (NCES) Integrated Postsecondary Education Data System (IPEDS). Below is the summary of the data sources found in this study.

Data Type	Source
Labor Market Information/Population Estimates and Projections/Educational Attainment	Economic Modeling Specialists, Intl. (EMSI). EMSI occupational employment data are based on final EMSI industry data and final EMSI staffing patterns. Wage estimates are based on Occupational Employment Statistics (QCEW and Non-QCEW Employees classes of worker) and the American Community Survey (Self-Employed and Extended Proprietors). Occupational wage estimates also affected by county-level EMSI earnings by industry: economicmodeling.com.
Typical Education Level and On-the-job Training	Bureau of Labor Statistics (BLS) uses a system to assign categories for entry-level education and typical on-the-job training to each occupation for which BLS publishes projections data: https://www.bls.gov/emp/tables/educational-attainment.htm.
Labor Force, Employment and Unemployment Estimates	California Employment Development Department, Labor Market Information Division: labormarketinfo.edd.ca.gov.
Job Posting and Skills Data	Burning Glass: burning-glass.com/.
Additional Education Requirements/ Employer Preferences	The O*NET Job Zone database includes over 900 occupations as well as information on skills, abilities, knowledge, work activities and interests associated with specific occupations: onetonline.org.

Key Terms and Concepts

Annual Job Openings: Annual openings are calculated by dividing the number of years in the projection period by total job openings.

Education Attainment Level: The highest education attainment level of workers age 25 years or older.

Employment Estimate: The total number of workers currently employed.

Employment Projections: Projections of employment are calculated by a proprietary Economic Modeling Specialists, Intl. (EMSI) formula that includes historical employment and economic indicators along with national, state, and local trends.

Living Wage: The cost of living in a specific community or region for one adult and no children. The cost increases with the addition of children.

Occupation: An occupation is a grouping of job titles that have a similar set of activities or tasks that employees perform.

Percent Change: Rate of growth or decline in the occupation for the projected period; this does not factor in replacement openings.

Replacements: Estimate of job openings resulting from workers retiring or otherwise permanently leaving an occupation. Workers entering an occupation often need training. These replacement needs, added to job openings due to growth, may be used to assess the minimum number of workers who will need to be trained for an occupation.

Total Job Openings (New + Replacements): Sum of projected growth (new jobs) and replacement needs. When an occupation is expected to lose jobs, or retain the current employment level, number of openings will equal replacements.

Typical Education Requirement: represents the typical education level most workers need to enter an occupation.

Typical On-The-Job Training: indicates the typical on-the-job training needed to attain competency in the skills needed in the occupation.