# Data Science

*California*

## Introduction

This report aims to determine the demand for data science jobs in the Inland Empire/Desert Region. Data science is a multidisciplinary approach to the study of data, which analyzes large amounts of data to extract meaningful insights for organizations (AWS, 2023). Businesses use data science to discover patterns within business operations, innovate new products and solutions, and provide real-time business insights.

This report analyzes online job advertisements (ads) from employers seeking data science workers to understand the necessary knowledge, skills, and abilities (KSAs) to enter this field. The United States Bureau of Labor Statistics (BLS) does not identify any occupation that exclusively utilizes data science skills in its Standard Occupational Classification (SOC) system; therefore, this report analyzes all occupations requiring data science skills according to online job ads over the previous year.

The Taxonomy of Programs (TOP), developed by the California Community Colleges Chancellor's Office, does not contain a TOP code that directly provides data science training. According to the California Community Colleges Chancellor's Office Curriculum Inventory (COCI), two community colleges in the region currently offer programs that provide training related to data science (COCI, 2023).

## Data Science Summary of Findings

Traditional labor market information is not a reliable indicator of demand for data science workers. For this reason, online employer job advertisements are the basis of this report's job demand assessment. Online job ads are helpful in understanding KSAs, education requirements, and other employer hiring preferences. While efforts were made to de-duplicate job ads, employers may list jobs they do not end up filling or use a single ad to hire multiple workers. Therefore, the number of online job ads is neither comparable nor indicative of the number of annual openings for any given occupation. Following is an overview of this report's key findings.

**Key Findings:**

- Over the past 12 months (April 2022 to March 2023), there were 5,676 data science job ads posted in California. Of these job postings, 62 data science jobs were located in the Inland Empire/Desert Region.
    - Approximately 66% of employer job ads for data science workers were concentrated in the occupation of data scientists.
    - Overall, online job ad salary information reveals that employers in the state are willing to pay these positions $132,870 annually, well above the $45,386 annual ($21.82 hourly) MIT living wage standard.

    - Within job ads, nearly 97% of employers sought candidates with a bachelor's degree, while less than 1% sought candidates with an associate degree, and nearly 3% sought candidates with a high school diploma or equivalent.
    - Approximately 93% of employer job ads with experience requirements sought candidates with more than one year of previous work experience, indicating that employers value data science workers with previous work experience.
- Two community colleges in the region currently offer programs related to data science, issuing a combined annual average of one award over the last three academic years.
    - Other postsecondary education institutions in the region do not currently offer data analytics, general CIP code (30.7101).

## Job Advertisements

An online job ad search for jobs in data science workers was conducted to reveal the employers seeking these workers, including the median job ad duration, earnings information, and in-demand skills. Over the last 12 months, from April 2022 through March 2023, 62 job ads were posted for data science workers in the Inland Empire/Desert Region. The job ad search was expanded to include all the data science job ads listed over the last 12 months in California to ensure that the job ad information included in this report is reliable and generalizable.

Exhibit 1 shows the occupations that employers frequently associate with data science work, as well as the number of job ads posted over the last 12 months in the state, data science job ads as a percentage of the total, and the median posting duration. Approximately 2.8% of job ads for the occupations in the data science occupational group were specific to data science jobs. Despite its relation to data science, only 20% of the job ads for the occupation data scientists required candidates with data science skills.

On average, employers kept online job ads for data science workers open for 24 days. The average statewide online job is open for 28 days, indicating employers likely experience fewer challenges filling data science positions than they do with other jobs. More than two-thirds of data science job ads were for data scientists.

*Exhibit 1: Job ads and posting duration, California, April 2022 – March 2023*

| Occupations | Data Science Job Ads | All Job Ads | Percentage of Data Science Job Ads | Median Posting Duration (Days) |
|---|---|---|---|---|
| Data Scientists | 3,796 | 19,058 | 19.9% | 24 |
| Database Administrators | 1,072 | 9,631 | 11.1% | 24 |
| Database Architects | 334 | 5,682 | 5.9% | 22 |

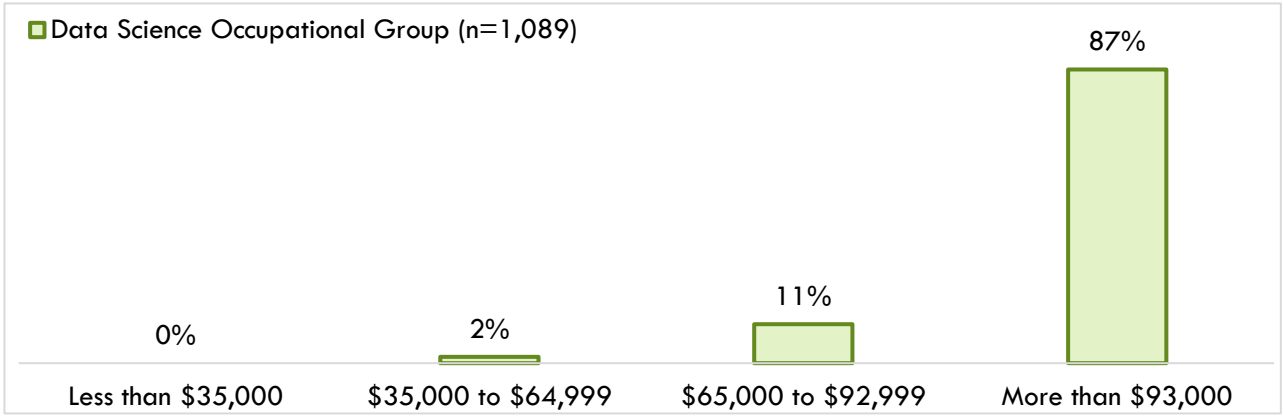
| Occupations | Data Science Job Ads | All Job Ads | Percentage of Data Science Job Ads | Median Posting Duration (Days) |
|---|---|---|---|---|
| Software Developers | 327 | 105,193 | 0.3% | 26 |
| Computer Systems Analysts | 88 | 17,365 | 0.5% | 23 |
| Computer Network Architects | 24 | 6,743 | 0.4% | 25 |
| Information Security Analysts | 16 | 11,828 | 0.1% | 24 |
| Software Quality Assurance Analysts and Testers | 11 | 12,107 | 0.1% | 21 |
| Computer Programmers | 7 | 5,706 | 0.1% | 28 |
| Network and Computer Systems Administrators | 1 | 8,314 | 0.01% | - |
| **Total** | **5,676** | **201,627** | **2.8%** | **24** |

Source: Lightcast 2023.1

## Advertised Salary from Online Job Ads

Exhibit 2 displays the online advertised salaries for the data science occupational group over the last 12 months. Online job ad salary information reveals that employers in the state are willing to pay data science workers an annual salary of $132,870 or $63.88 per hour, well above the $21.82 hourly living wage standard. Consider the salary information with caution since only 19% (1,089 of 5,676) of online job ads for this occupational group provided salary information.

*Exhibit 2. Online advertised salaries for the data science occupational group, California, April 2022 – March 2023*




Source: Lightcast 2023.1

## Job Titles, Employers, Skills, Education, and Work Experience

Exhibit 3 displays the job titles most frequently requested for data science worker jobs over the last 12 months in the state. Job titles may provide insight into the types of positions held by data science workers. The job title most frequently associated with data science workers in California over the last 12 months was data scientist, accounting for nearly 27% of job ads.

*Exhibit 3: Job titles most frequently requested for data science workers, California, April 2022 – March 2023*

| Job Titles | Job Ads |
|---|---|
| Data Scientists | 1,519 |
| Data Engineers | 725 |
| Data Analysts | 526 |
| Data Science Managers | 346 |
| Data Analytics Scientists | 167 |
| Data Software Engineers | 153 |
| Directors of Data Science | 142 |
| Principal Data Scientists | 117 |
| Data Science and Analytics Managers | 110 |
| *All other job titles* | *1,871* |
| **Total** | **5,676** |

Source: Lightcast 2023.1

Exhibit 4 displays the employers that posted the most job ads for data science workers during the last 12 months. Showing employer names provides insight into where students may find employment after completing a program. Walmart posted the most job ads for data science workers over the last 12 months in the state, primarily seeking data science managers to increase the effectiveness of the organization's supply chain and logistics operations.

*Exhibit 4: Employers posting the most job ads for data science jobs, California, April 2022 – March 2023*

| Top Employers | Jobs Ads |
|---|---|
| Walmart | 136 |
| Amazon | 108 |
| UnitedHealth Group | 75 |
| Meta | 68 |
| Apple | 65 |
| Deloitte | 63 |
| Shopify | 52 |
| Cognizant Technology Solutions | 52 |
| FocusKPI | 50 |
| *All other employers* | *5,007* |
| **Total** | **5,676** |

Source: Lightcast 2023.1

Exhibit 5 lists a sample of specialized, employability, and software and programming skills employers seek when looking for data science workers. Specialized skills are occupation-specific skills that employers request

for industry or job competency. Employability skills are foundational skills that transcend industries and occupations; this category is often referred to as "soft skills." The skills requested in job ads may be utilized to guide curriculum development. The programming language Python was included in approximately 74% of employer job ads, indicating that this skill is important for data science workers.

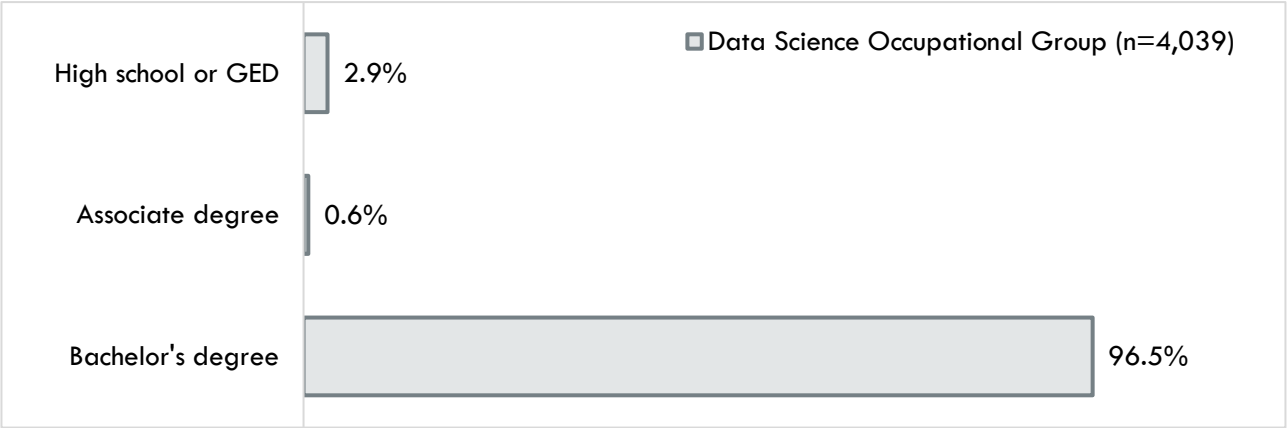
*Exhibit 5: Sample of in-demand skills from employer job ads for data science workers, California, April 2022 – March 2023, (n=5,676)*

| Specialized Skills | Employability Skills | Software and Programming Skills |
|---|---|---|
| • Data Science | • Communication Skills | • Python |
| • Machine Learning | • Mathematics | • SQL |
| • Computer Science | • Leadership Skills | • R |
| • Data Analysis | • Research | • Apache Spark |
| • Statistics | • Problem-Solving | • Tableau |
| • Data Modeling | • Operations | • Amazon Web Services |
| • Data Visualization | • Management | • Java |

Source: Lightcast 2023.1

Exhibit 6 displays the minimum advertised education requirements for data science workers. Approximately 29% of employer job ads did not include minimum education requirements in job ads for data science workers. The vast majority of employer job ads (96.5%) for data science workers sought candidates with a bachelor's degree.

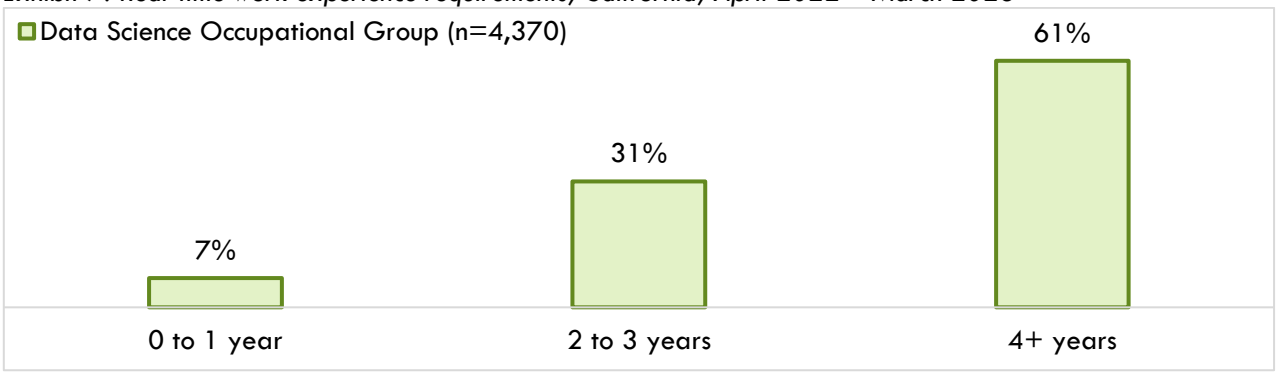
*Exhibit 6: Minimum advertised education requirements for data science workers, California, April 2022 – March 2023*




Source: Lightcast 2023.1

Exhibit 7 displays the real-time work experience requirements from employer job ads. Approximately 23% of employer job ads did not include work experience requirements in job ads for data science workers. Approximately 93% of employer job ads with experience requirements sought candidates with more than one year of previous work experience, indicating that employers value data science workers with previous experience.

*Exhibit 7: Real-time work experience requirements, California, April 2022 – March 2023*




Source: Lightcast 2023.1

## Training Programs

Two community colleges in the region currently offer programs related to data science. Regional community colleges use two program codes when coding their data science programs. Combined, regional community college data analytics programs have issued one award annually over the last three academic years, 2019-2022. Exhibit 8 displays each regional data science program and award types students earn upon program completion.

*Exhibit 8: Data science-related programs, Inland Empire/Desert Region, 2022-23 academic year*

| College | TOP Program (TOP Code) | Local Program Title | Award |
|---|---|---|---|
| Moreno Valley College | Computer Systems Analysis (0707.30) | Data Science | A.S. Degree |
| | | Data Analytics | Certificate |
| | | Information Assurance Auditing | Certificate |
| Riverside City College | Database Design and Administration (0707.20) | Data Quick Start | Certificate |

Source: COCI, 2022-23 Community College Catalogs

Exhibits 9 and 10 display student completions for computer systems analysis (TOP 0707.30) and database design and administration (0707.20) programs related to data science over the last three academic years, 2019-2022. Over the last three academic years, Moreno Valley College has not issued any known awards in computer systems analysis programs. Riverside City College's database design and administration program last issued two awards in the 2019-2020 academic year. Program completion and student outcome methodologies can be found in the appendix.

*Exhibit 9: Annual average community college awards for computer systems analysis programs related to data science, Inland Empire/Desert Region, Academic Years 2019-2022*

| **TOP 0707.30 – Computer Systems Analysis** (Local Program Title) | **Academic Year 2019-20** | **Academic Year 2020-21** | **Academic Year 2021-22** | **Total CC Annual Average Awards, Academic Years 2019-22** |
|---|---|---|---|---|
| **Moreno Valley** (Data Analytics/Data Science/Information Assurance Auditing) | | | | **0** |
| Associate Degree | 0 | 0 | 0 | 0 |
| Certificate 30 < 60 semester units | 0 | 0 | 0 | 0 |
| Certificate 16 < 30 semester units | 0 | 0 | 0 | 0 |
| **Total** | **0** | **0** | **0** | **0** |

Source: MIS Data Mart, COCI

*Exhibit 10: Annual average community college awards for database design and administration programs related to data science, Inland Empire/Desert Region, Academic Years 2019-2022*

| **TOP 0707.20 – Database Design and Administration** (Local Program Title) | **Academic Year 2019-20** | **Academic Year 2020-21** | **Academic Year 2021-22** | **Total CC Annual Average Awards, Academic Years 2019-22** |
|---|---|---|---|---|
| **Riverside City** (Data Quick Start) | | | | **1** |
| Certificate 6 < 18 semester units | 2 | 0 | 0 | 1 |
| **Total** | **2** | **0** | **0** | **1** |

Source: MIS Data Mart, COCI

California program outcome data may provide useful insight into the likelihood of success for the proposed program. Community college student outcome information based on the selected TOP code and region is provided in Exhibits 11 and 12.

*Exhibit 11: 0707.30 – Computer systems analyst strong workforce program outcomes, Inland Empire/Desert Region, Academic Year 2019-2020 (Unless Noted)*

| Strong Workforce Program Metrics: 0707.30 – Computer System Analyst Academic Year 2019-20, unless noted otherwise | Inland Empire/Desert Region | California |
|---|---|---|
| Unduplicated count of enrolled students (2020-21) | 302 | 2,156 |
| Completed 9+ career education units in one year (2020-21) | 49% | 39% |
| Students who completed a noncredit CTE or workforce preparation course (2020-21) | - | 96% |
| Students who earned a degree, certificate, or attained apprenticeship (2020-21) | - | 149 |
| Transferred to a four-year institution (transfers) | 17 | 213 |
| Job closely related to the field of study (2018-19) | - | 62% |
| Median annual earnings (all exiters) | $36,016 | $37,580 |
| Median change in earnings (all exiters) | -21% | 17% |
| Attained a living wage (completers and skills-builders) | 58% | 53% |

Sources: LaunchBoard Community College Pipeline and Strong Workforce Program Metrics

*Exhibit 12: 0707.20 – Database design and administration strong workforce program outcomes, Inland Empire/Desert Region, Academic Year 2019-2020 (Unless Noted)*

| Strong Workforce Program Metrics: 0707.20 – Database Design and Administration Academic Year 2019-20, unless noted otherwise | Inland Empire/Desert Region | California |
|---|---|---|
| Unduplicated count of enrolled students (2020-21) | 158 | 2,473 |
| Completed 9+ career education units in one year (2020-21) | 50% | 48% |
| Students who earned a degree, certificate, or attained apprenticeship (2020-21) | - | 60 |
| Transferred to a four-year institution (transfers) | 19 | 150 |
| Job closely related to the field of study (2018-19) | 30% | 70% |
| Median annual earnings (all exiters) | $46,190 | $67,400 |
| Median change in earnings (all exiters) | 23% | 11% |
| Attained a living wage (completers and skills-builders) | 71% | 71% |

Sources: LaunchBoard Community College Pipeline and Strong Workforce Program Metrics

Other postsecondary institutions may utilize the data science, general CIP code (30.7001) for their data analytics programs. However, over the last three academic years, no other regional postsecondary education institution issued awards using this program code. The following is the program description for data science, general CIP programs.

> *A program that focuses on the analysis of large scale data sources from the interdisciplinary perspectives of applied statistics, computer science, data storage, data representation, data modeling, mathematics, and statistics. Includes instruction in computer algorithms, computer programming, data management, data mining, information policy, information retrieval, mathematical modeling, quantitative analysis, statistics, trend spotting, and visual analytics (IPEDS, 2022).*

**Contact**

Michael Goss
Paul Vaccher
Centers of Excellence, Inland Empire/Desert Region
michael.goss@chaffey.edu
May 2023

## References

Amazon Web Services (AWS). (2023). *What is Data Science?* Retrieved from https://aws.amazon.com/what-is/data-science/

California Community Colleges Chancellor's Office. LaunchBoard. (2023). *California Community Colleges LaunchBoard*. Retrieved from https://www.calpassplus.org/Launchboard/Home.aspx

California Community Colleges Chancellor's Office. LaunchBoard. (2023a). *Strong Workforce Program Metrics Data Element Dictionary*. Pg. 3. Retrieved from https://www.calpassplus.org/MediaLibrary/calpassplus/launchboard/Documents/SWP-DED.PDF

California Community Colleges Chancellor's Office, Curriculum and Instructional Unit, Academic Affairs Division. (2012). *Taxonomy of Programs, 6th Edition, Corrected Version*. Retrieved from https://www.cccco.edu/-/media/CCCCO-Website/About-Us/Divisions/Digital-Innovation-and-Infrastructure/Research/Files/TOPmanual6200909corrected12513.ashx?la=en&hash=94C709CA83C0380828415579395A5F536736C7C1

California Community Colleges Chancellor's Office Management Information Systems (MIS) Data Mart. (2023). *Data Mart*. Retrieved from https://datamart.cccco.edu/datamart.aspx

California Community Colleges Chancellor's Office. (2023). *Chancellor's Office Curriculum Inventory (COCI), version 3.0*. Retrieved from https://coci2.ccctechcenter.org/programs

Carnevale, A. P., Jayasundera, T., & Repnikov, D. (n.d.). Understanding Online Job Ads Data. Retrieved from https://cew.georgetown.edu/wp-content/uploads/2014/11/OCLM.Tech_.Web_.pdf

Glasmeier, A. K. (2023). Massachusetts Institute of Technology. Living Wage Calculator. *Living Wage Calculation for California*. Retrieved from https://livingwage.mit.edu/states/06

Lightcast. (2023). *Datarun 2023.1*. Retrieved from https://www.economicmodeling.com/

National Center for O*NET Development. (2023). *O*NET OnLine*. Retrieved from https://www.onetonline.org/

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. Integrated Postsecondary Education System (IPEDS). (2023). *Detail for CIP Code 30.7001 Data Science, General*. Retrieved from https://nces.ed.gov/ipeds/cipcode/cipdetail.aspx?y=56&cipid=92953

## Appendix: Methodology

Exhibits 9 and 10 display the average annual California Community College (CCC) awards conferred during the three academic years between 2019 and 2022 from the California Community Colleges Chancellor's Office Management Information Systems (MIS) Data Mart. Awards are the combined total of associate degrees and certificates issued during the timeframe, divided by three in this case to calculate an annual average. This is done to minimize the effect of atypical variations that might be present in a single year.

Community college student outcome information is from LaunchBoard and based on the selected TOP code and region. These metrics are based on records submitted to the California Community Colleges Chancellor's Office Management Information Systems (MIS) by community colleges, which come from self-reported student information from CCC Apply and the National Student Clearinghouse. Employment and earnings metrics are sourced from California's Employment Development Department's Unemployment Insurance database records. When available, outcomes for completers are reported to demonstrate the impact that earning a degree or certificate can have on employment and earnings. For more information on the types of students included for each metric, please see the web link for LaunchBoard's Strong Workforce Program Metrics Data Element Dictionary in the References section (LaunchBoard, 2023a). Finally, employment in a job closely related to the field of study comes from self-reported student responses on the CTE Employment Outcomes Survey (CTEOS), administered by Santa Rosa Junior College (LaunchBoard, 2023a).

Job ad data is limited to the information provided by employers and the ability of artificial intelligence search engines to identify this information. Additionally, preliminary calculations by Georgetown Center on Education and the Workforce found that "just 30 to 40 percent of openings for candidates with some college or an associate degree, and only 40 to 60 percent of openings for high school diploma holders appear online" (Carnevale et al., 2014). Online job ads often do not reveal employers' hiring intentions; it is unknown if employers plan to hire one or multiple workers from a single online job ad or collecting resumes for future hiring needs. A closed job ad may not be the result of a hired worker.

## Appendix: Job Advertisement Search Parameters

**Area Filter:**

- California
- Riverside-San Bernardino-Ontario MSA

**Occupation Filter:**

- Computer Systems Analysts (SOC 15-1211)
- Information Security Analysts (15-1212)
- Computer Network Architects (15-1241)
- Database Administrators (15-1242)
- Database Architects (15-1243)
- Network and Computer Systems Administrators (15-1244)
- Computer Programmers (15-1251)
- Software Developers (15-1252)
- Software Quality Assurance Analysts and Testers (15-1253)
- Data Scientists (15-2051)

**Education Filter:**

- High school or GED
- Associate degree
- Bachelor's degree
- Education Not Listed

**Skill Filter:**

- Data Skills

**Job Title Filter:**

- Data

**Keyword Filter:**

- Data Science
- Data Analysis
- Data Analyst