** Data Science Occupations**

**Labor Market Information Report**

**Foothill College**

Prepared by the San Francisco Bay Center of Excellence

for Labor Market Research

February 2019

# Recommendation

Based on all available data, it is difficult to determine if there is an unmet need for Data Science workers that community college students in the Bay region are qualified for, since this is an emerging occupation and field. A traditional supply vs. demand “gap analysis” is difficult to perform. The demand is not completely clear (as is outlined in the Introduction section below) and the supply data from the TOP code selected by Foothill College – TOP 0708.00 - Computer Infrastructure and Support, is a program of study that would be preparing students for not only Data Science Occupations, but a cluster of other occupations more traditionally aligned with Computer Infrastructure and Support. So it is impossible to determine what portion of the supply/students from this TOP code would be seeking employment in a Data Science occupation (i.e. Computer and Information Research Scientist).

This report also provides student outcomes data on employment and earnings for programs on TOP 0708.00 - Computer Infrastructure and Support in the state and region. It is recommended that these data be reviewed to better understand how outcomes for students taking courses on this TOP code compare to potentially similar programs at colleges in the state and region, as well as to outcomes across all CTE programs at Foothill College and in the region.

# Introduction

This report profiles Data Science Occupations in the 12 county Bay region and in the Silicon Valley sub-region for a proposed new Data Science program at Foothill College. The best available occupation aligned with Data Science is "Computer and Information Research Scientists" (SOC 15-1111).  The issue is that the education level required for employment is very high - with 90% of workers who are currently employed in this occupation having an education level of Bachelors degree or higher: Bachelors degree (30%), Masters degree (32%), Ph.D. (28%).

Only about 10% of those employed currently in this occupation have less than a Bachelors degree - so many community college students who completed a 12-18 unit certificate, would not be qualified for this occupation, unless they already had a Bachelor’s degree and were looking to add new skills to increase their employability in the labor market. In fact, Foothill College reports that 20% - 30% of their student population already have Bachelors degrees, hence the proposed Data Science Certificate could serve this student population who already have a four-year degree.

Job Postings data (from Burning Glass) supports strong demand for this occupation with over 6,700 job postings annually in the region (2018) and 2,700 job postings annually in the Silicon Valley sub-region for Computer and Information Research Scientists, at all education levels. However, the majority of employers posting job ads desire candidates with a Bachelor’s degree or higher.

Of the total 6,732 job postings in the Bay region, there are annually 2,138 job postings listed as requiring a Bachelor’s degree or less. Of the total 2,732 job postings in the sub-region, there are annually 879 job postings listed as requiring a Bachelor’s degree or less. This seems to indicate demand for Computer and Information Research Scientists at this education level, using job postings data.

|  |
| --- |
| * **Computer and Information Research Scientists (SOC 15-1111):** 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.
 |
| *Entry-Level Educational Requirement: Master's degree* |
| *Training Requirement: None* |
| *Percentage of Community College Award Holders or Some Postsecondary Coursework: 2%* |

# Occupational Demand

**Table 1. Employment Outlook for Data Science Occupations in Bay Region**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Occupation  | 2017 Jobs | 2022 Jobs | 5-Yr Change | 5-Yr % Change | 5-Yr Open-ings | Annual Open-ings | 10% Hourly Wage | Median Hourly Wage |
| Computer and Information Research Scientists | 2,729 | 3,168 | 439  | 16% | 1,403 | 281 | $32.12  | $66.58  |

*Source: EMSI 2019.1*

**Bay Region** includes Alameda, Contra Costa, Marin, Monterey, Napa, San Benito, San Francisco, San Mateo, Santa Clara, Santa Cruz, Solano and Sonoma Counties

**Table 2. Employment Outlook for Data Science Occupations in Silicon Valley Sub-Region**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Occupation  | 2017 Jobs | 2022 Jobs | 5-Yr Change | 5-Yr % Change | 5-Yr Open-ings | Annual Open-ings | 10% Hourly Wage | Median Hourly Wage |
| Computer and Information Research Scientists | 1,273 | 1,418 | 145  | 11% | 587 | 117 | $31.16  | $53.58  |

*Source: EMSI 2019.1*

**Silicon Valley Sub-Region** includes Santa Clara County

### Job Postings in Bay Region and Silicon Valley Sub-Region

**Table 3. Number of Job Postings by Occupation for latest 12 months (Feb 2018 - Jan 2018)**

| Occupation | Bay Region | Silicon Valley |
| --- | --- | --- |
| Computer and Information Research Scientists | 6,732 | 2,732 |

*Source: Burning Glass*

**Table 4. Top Job Titles for Data Science Occupations for latest 12 months (Feb 2018 - Jan 2018)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Common Title | Bay | Silicon Valley | Common Title | Bay | Silicon Valley |
| Data Scientist | 2,045 | 727 | Data Consultant | 33 | 21 |
| Senior Data Scientist | 687 | 232 | Manager, Data, Science | 33 | 15 |
| Machine Learning Scientist | 493 | 299 | Machine Learning Engineer | 28 | 16 |
| Computer Scientist | 207 | 109 | Data Scientist, Information And Technology Industry | 27 | 15 |
| Data Analyst | 125 | 61 | Data Scientist, Python | 26 | 12 |
| Staff Data Scientist | 112 | 54 | Junior Data Scientist | 25 | 4 |
| Lead Data Scientist | 108 | 34 | Senior Manager, Data, Science | 25 | 7 |
| Principal Data Scientist | 90 | 50 | Research Scientist | 21 | 15 |
| Director, Data, Science | 88 | 15 | Associate Data Scientist | 20 | 5 |
| Research Engineer | 81 | 54 | Data Scientist/Engineer | 20 | 4 |
| Data Science Manager | 77 | 14 | Engineering Manager | 19 | 6 |
| Data Scientist, Analytics | 62 | 8 | Chief Data Scientist | 18 | 1 |
| Natural Language Processing Scientist | 50 | 27 | Data Scientist, Learning | 18 | 7 |
| Staff Scientist | 47 | 5 | Senior Research Engineer | 14 | 12 |

*Source: Burning Glass*

# Industry Concentration

**Table 5. Industries hiring Data Science Workers in Bay Region**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Industry – 6 Digit NAICS (No. American Industry Classification) Codes | Jobs in Industry (2017) | Jobs in Industry (2022) | % Change (2017-22) | % in Industry (2017) |
| Custom Computer Programming Services (541511) | 448 | 462 | 21% | 16.4% |
| Research and Development in the Physical, Engineering, and Life Sciences (except Nanotechnology and Biotechnology) (541715) | 395 | 417 | 1% | 14.8% |
| Software Publishers (511210) | 296 | 316 | 29% | 11.2% |
| Computer Systems Design Services (541512) | 270 | 277 | 19% | 9.8% |
| Research and Development in Biotechnology (except Nanobiotechnology) (541714) | 216 | 230 | 47% | 8.2% |
| Federal Government, Civilian, Excluding Postal Service (901199) | 177 | 174 |  (2%) | 6.2% |
| Internet Publishing and Broadcasting and Web Search Portals (519130) | 157 | 166 | 38% | 5.9% |
| Electronic Computer Manufacturing (334111) | 112 | 116 | 10% | 4.1% |
| Colleges, Universities, and Professional Schools (902612) | 83 | 83 |  (2%) | 3.0% |
| Other Computer Related Services (541519) | 62 | 64 | 27% | 2.3% |
| Computer and Computer Peripheral Equipment and Software Merchant Wholesalers (423430) | 55 | 53 |  (5%) | 1.9% |
| Data Processing, Hosting, and Related Services (518210) | 44 | 46 | 20% | 1.6% |
| Engineering Services (541330) | 43 | 43 | 7% | 1.5% |

*Source: EMSI 2019.1*

**Table 6. Top Employers Posting Data Science Occupations in Bay Region and Silicon Valley Sub-Region (Feb 2018 - Jan 2018)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Employer | Bay | Employer | Bay | Employer | Silicon Valley |
| Apple Inc. | 189 | Intel Corporation | 25 | Apple Inc. | 186 |
| Facebook | 187 | Lawrence Berkeley National Laboratory | 25 | Walmart / Sam's | 60 |
| Capital Markets Placement | 135 | Sandia Corporation | 25 | Intuit | 59 |
| Walmart / Sam's | 96 | Target | 24 | Adobe Systems | 44 |
| Uber | 90 | Oracle | 22 | Amazon | 41 |
| Intuit | 73 | Samsung America, Inc. | 22 | Google Inc. | 38 |
| Amazon | 61 | KLA-Tencor | 21 | Linkedin Limited | 31 |
| Adobe Systems | 58 | Wework | 21 | Cisco Systems Incorporated | 28 |
| Genentech | 56 | Electronic Arts Incorporated | 20 | SAP | 26 |
| Google Inc. | 51 | Stitch Fix | 20 | Intel Corporation | 25 |
| Accenture | 36 | Autodesk Incorporated | 19 | Target | 24 |
| IBM | 35 | Paypal | 19 | eBay | 23 |
| eBay | 35 | Allstate | 18 | KLA-Tencor | 21 |
| Linkedin Limited | 34 | Capital One | 18 | Samsung America, Inc. | 21 |
| Cisco Systems Incorporated | 32 | Hewlett-Packard | 18 | IBM | 19 |
| Airbnb | 31 | Salesforce | 18 | Hewlett-Packard | 18 |
| SAP | 31 | Social Finance | 18 | Paypal | 18 |
| General Electric Company | 26 | Lawrence Livermore National Laboratory | 17 | Microsoft Corporation | 17 |
| Microsoft Corporation | 26 | Slac National Accelerator Laboratory | 17 | Visa | 17 |
| Twitter | 26 | Square Incorporated | 16 | Wework | 16 |
| Visa | 26 | Harnham | 15 | Nvidia Corporation | 15 |
| Deloitte | 25 | Nvidia Corporation | 15 | Stanford University | 13 |

*Source: Burning Glass*

# Educational Supply

There are six community colleges in the Bay Region issuing 62 awards on average annually (last 3 years) on TOP 0708.00 - Computer Infrastructure and Support. Gavilan College is the only college in the Silicon Valley Sub-Region issuing an award on this TOP code, issuing one award on average annually (last 3 years). It is important to note that TOP 0708.00 - Computer Infrastructure and Support, is a program of study that would be preparing students for not only Data Science Occupations, but a cluster of other occupations more traditionally aligned with Computer Infrastructure and Support.

**Table 7. Awards on TOP 0708.00 - Computer Infrastructure and Support in the Bay Region**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| College | Sub-Region | Headcount | Associates | Certificates | Total |
| **Cabrillo** | Santa Cruz - Monterey | 335 |   |   |   |
| **Contra Costa** | East Bay |   |   | 1 | 1 |
| **De Anza** | Silicon Valley | 312 |   |   |   |
| **Diablo Valley** | East Bay | 125 |   |   |   |
| **Gavilan** | Silicon Valley | 95 | 1 |   | 1 |
| **Las Positas** | East Bay | 184 |   |   |   |
| **Los Medanos** | East Bay |   | 4 | 4 | 8 |
| **Mission** | Silicon Valley | 43 |   |   |   |
| **Ohlone** | East Bay | 64 |   | 1 | 1 |
| **San Francisco** | Mid-Peninsula | 345 | 39 | 9 | 48 |
| **San Mateo** | Mid-Peninsula |   | 1 | 2 | 3 |
| **Santa Rosa** | North Bay | 205 |   |   |   |
| **Total Bay Region** |  **1,708**  |  **45**  |  **17**  |  **62**  |
| **Total Silicon Valley Sub-Region** |  **450**  |  **1**  |  **0**  |  **1**  |

# *Source: IPEDS, Data Mart and Launchboard*

NOTE: Headcount of students who took one or more courses is for 2016-17. The annual average for awards is 2014-17 unless there are only awards in 2016-17. The annual average for other postsecondary is for 2013-16.

# Gap Analysis

Based on all available data, it is difficult to determine if there is an unmet need for Data Science workers that community college students in the Bay region are qualified for, since this is an emerging occupation and field. A traditional supply vs. demand “gap analysis” is difficult to perform. The demand is not completely clear (as is outlined in the Introduction section above) and the supply data from the TOP code selected by Foothill College – TOP 0708.00 - Computer Infrastructure and Support, is a program of study that would be preparing students for not only Data Science Occupations, but a cluster of other occupations more traditionally aligned with Computer Infrastructure and Support. So it is impossible to determine what portion of the supply/students from this TOP code would be seeking employment as a Computer and Information Research Scientist.

# Student Outcomes

**Table 8. Four Employment Outcomes Metrics for Students Who Took Courses on TOP 0708.00 - Computer Infrastructure and Support**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 2015-16 | Bay (All CTE Programs) | Foothill College (All CTE Programs) | State (0708.00) | Bay (0708.00) | Silicon Valley (0708.00) | Foothill College (0708.00) |
| % Employed Four Quarters After Exit | 74% | 77% | 66% | 72% | 76% | n/a |
| Median Quarterly Earnings Two Quarters After Exit | $23,396 | $55,862 | $38,862 | $47,642 | $53,243 | n/a |
| Median % Change in Earnings | 46% | 82% | 47% | 47% | 73% | n/a |
| % of Students Earning a Living Wage | 63% | 76% | 62% | 68% | 76% | n/a |

*Source: Launchboard Pipeline (version available on 2/12/19)*

# Skills, Certifications and Education

**Table 9. Top Skills for Data Science Occupations in Bay Region (Feb 2018 - Jan 2018)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Skill | Postings | Skill | Postings | Skill | Postings |
| Data Science | 5,504 | C++ | 992 | Software Engineering | 554 |
| Python | 4,784 | Scala | 946 | MapReduce | 513 |
| Machine Learning | 4,254 | TensorFlow | 936 | Product Development | 499 |
| SQL | 2,969 | Physics | 896 | Pandas | 473 |
| Apache Hadoop | 1,894 | Predictive Models | 882 | Cluster Analysis | 468 |
| Java | 1,769 | Economics | 782 | Computer Vision | 465 |
| Big Data | 1,567 | Tableau | 777 | Regression Algorithms | 452 |
| Data Analysis | 1,407 | Statistics | 747 | Scikit-learn | 448 |
| Experiments | 1,361 | SAS | 736 | Big Data Analytics | 441 |
| Deep Learning | 1,202 | Statistical Analysis | 735 | NoSQL | 436 |
| Data Mining | 1,183 | MATLAB | 682 | PIG | 408 |
| R | 1,179 | Data Visualization | 632 | Predictive Analytics | 406 |
| Artificial Intelligence | 1,056 | Software Development | 612 | Linux | 391 |
| Apache Hive | 1,039 | Clustering | 608 | Classification Algorithms | 385 |
| Natural Language Processing | 1,029 | Neural Networks | 603 | Distributed Computing | 385 |

*Source: Burning Glass*

**Table 10. Certifications for Data Science Occupations in the Bay Region (Feb 2018 - Jan 2018)**

Note: 98% of records have been excluded because they do not include a certification. As a result, the chart below may not be representative of the full sample.

|  |  |  |  |
| --- | --- | --- | --- |
| Certification | Postings | Certification | Postings |
| Project Management Certification | 17 | CompTIA Server+ | 2 |
| Project Management Professional (PMP) | 12 | Driver's License | 2 |
| American Board for Engineering and Technology (ABET) Accredited | 11 | IT Infrastructure Library (ITIL)  | 2 |
| Security Clearance | 11 | Investment Advisor | 2 |
| Certified Scrum Trainer (CST) | 7 | Six Sigma Certification | 2 |
| Certified Novell Administrator | 6 | Six Sigma DFSS-Green Belt | 2 |
| Certified Scrum Professional (CSP) | 6 | Advanced Engineering Certificate | 1 |
| Certified Professional in Healthcare Quality | 5 | Certified Business Analysis Professional  | 1 |
| Capability Model Maturity Integration (CMMI) Certification | 4 | Certified Information Systems Security Professional (CISSP) | 1 |
| Certified Scrum Developer (CSD) | 4 | Certified ScrumMaster (CSM) | 1 |
| Clinical Laboratory Scientist (ClS) | 3 | Certified Teacher | 1 |
| Financial Accounting Standards Board  | 3 | Citrix Certified Advanced Administrator  | 1 |
| Six Sigma Green Belt Certification | 3 | Fellow of the Casualty Actuarial Society  | 1 |

*Source: Burning Glass*

**Table 11. Education Requirements for Data Science Occupations in Bay Region**

Note: 75% of records have been excluded because they do not include a degree level. As a result, the chart below may not be representative of the full sample.

|  |  |
| --- | --- |
| Education (minimum advertised) | Latest 12 Mos. Postings |
| High school or vocational training | 0 (0%) |
| Associate Degree | 0 (0%) |
| Bachelor’s Degree or Higher | 5,036 (100%) |

*Source: Burning Glass*

# Methodology

Occupations for this report were identified by use of skills listed in O\*Net descriptions and job descriptions in Burning Glass. Labor demand data is sourced from Economic Modeling Specialists International (EMSI) occupation data and Burning Glass job postings data. Educational supply and student outcomes data is retrieved from multiple sources, including CTE Launchboard and CCCCO Data Mart.

# Sources

O\*Net Online

Labor Insight/Jobs (Burning Glass)

Economic Modeling Specialists International (EMSI)

CTE LaunchBoard [www.calpassplus.org/Launchboard/](http://www.calpassplus.org/Launchboard/)

Statewide CTE Outcomes Survey

Employment Development Department Unemployment Insurance Dataset

Living Insight Center for Community Economic Development

Chancellor’s Office MIS system

# Contacts

For more information, please contact:

* Doreen O’Donovan, Data Research Analyst, for Bay Area Community College Consortium (BACCC) and Centers of Excellence (CoE), doreen@baccc.net or (831) 479-6481
* John Carrese, Director, San Francisco Bay Center of Excellence for Labor Market Research, jcarrese@ccsf.edu or (415) 267-6544