

Exploring the Skills Needed by the Data Scientist

Kirk Paul Lafler; @sasNerd

Abstract

As 2.5 quintillion bytes (1 with 18 zeros) of new data are created each and every day, the age of big data has taken on new meaning. More and more organizations across industries are embracing Data Science / Computer Research Scientist skills resulting in an emerging demand for qualified and experienced talent. According to the Bureau of Labor Statistics (BLS) the number of data science jobs is expected to grow 19 percent over the next two decades – nearly three times as fast as the average growth rate for all jobs. Energized by this employment outlook, students and professionals across job functions are preparing for tomorrow's growing data science / analytic demands by acquiring a comprehensive skill set. To prepare for this growing demand, many colleges, junior colleges, Universities, and vocational training organizations offer comprehensive degrees and certificate programs to fulfill the increasing demand for analytical skills. This paper and presentation explores the skills needed by the Data Scientist / Analytics professional including non-technical skills such as critical thinking; business acumen and verbal/written communications; and technical skills such as data access; data wrangling; statistics; use of statistical programming languages like Python, R and SAS®; Structured Query Language (SQL); Microsoft Excel; and data visualization.

Keywords

SAS, SQL, Python, R, Data Science, Data Scientist, Analytics, skills, data science skills, employment websites, job postings, employment outlook, data science projections, data science degrees, data science certifications, Lafler, sasNerd.

Introduction

Wikipedia describes data science, “as a multi-disciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data.” Wikipedia further describes data science, “as the profession to unify statistics, data analysis, mathematics, computer programming, operations research, machine learning and related methods in order to understand and analyze actual phenomena with data and enhance decision making.”

Professionals with skills in data science and analytics are a “hot” commodity in the employment marketplace. The Bureau of Labor Statistics (BLS) (accessed on April 9th, 2021) projects that Employment of computer and information research scientists is projected to grow 15 percent from 2019 to 2029, much faster than the average for all occupations. One thing is certain, with the rapid growth in data collection by organizations everywhere will lead to an increased need for data-mining services.

Brian Holak, a site editor at Search Business Analytics, (January 31st, 2019) predicts, “The demand for data scientists continues to grow sharply with more employers than ever looking to hire data scientists. Data Scientists will experience a 29% increase in demand year over year, a 344% increase since 2013, with the supply of “qualified” data scientists lagging demand.” Merrimack College projects, “The Federal Government will experience a 19% growth rate in the field of computer and information research scientists by 2026.” Also many employment / career websites (2021), professionals possessing data science and analytics skills are highly sought after skills and are experiencing an exponential rate of growth. But, while demand is huge the supply of skilled applicants is growing at a much slower pace.

Data Science / Analytics Annual Base Salaries

The average annual base salaries for Data Scientists / Analytics professionals by city are displayed in the following table.

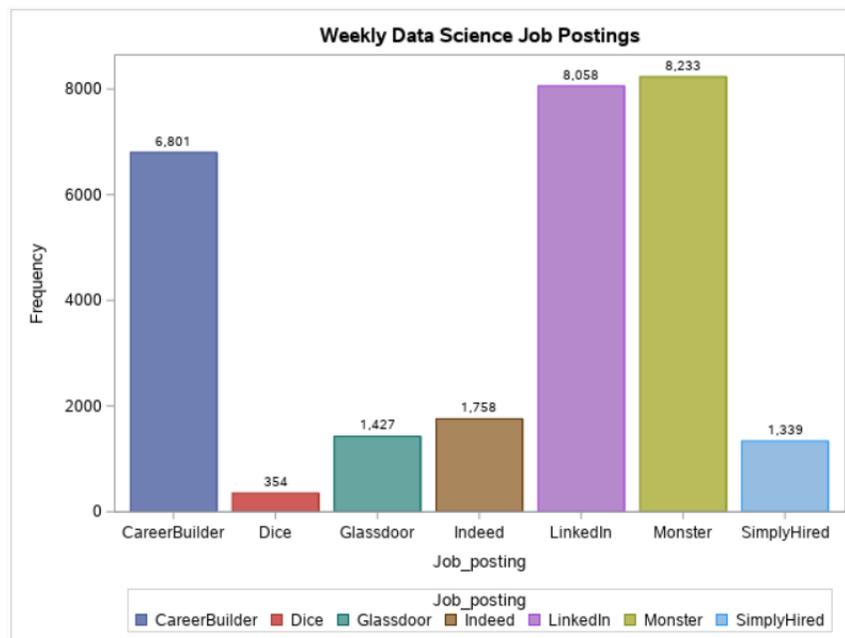
Average Annual Base Salaries by City

San Francisco	\$142,338 / year	Austin	\$102,470 / year
Seattle	\$129,332 / year	Dallas	\$101,208 / year
Los Angeles	\$120,189 / year	Miami	\$99,167 / year
New York	\$115,815 / year	Atlanta	\$98,202 / year
San Diego	\$114,637 / year	Raleigh	\$98,091 / year
Boston	\$112,059 / year	Charlotte	\$97,935 / year
Houston	\$110,284 / year	Indianapolis	\$96,778 / year
Chicago	\$106,135 / year	Jacksonville	\$96,764 / year
Denver	\$106,025 / year	Cincinnati	\$96,229 / year
Washington, DC	\$105,975 / year	Tampa	\$95,752 / year
Phoenix	\$104,502 / year	San Antonio	\$95,255 / year
Philadelphia	\$103,995 / year	Orlando	\$95,092 / year

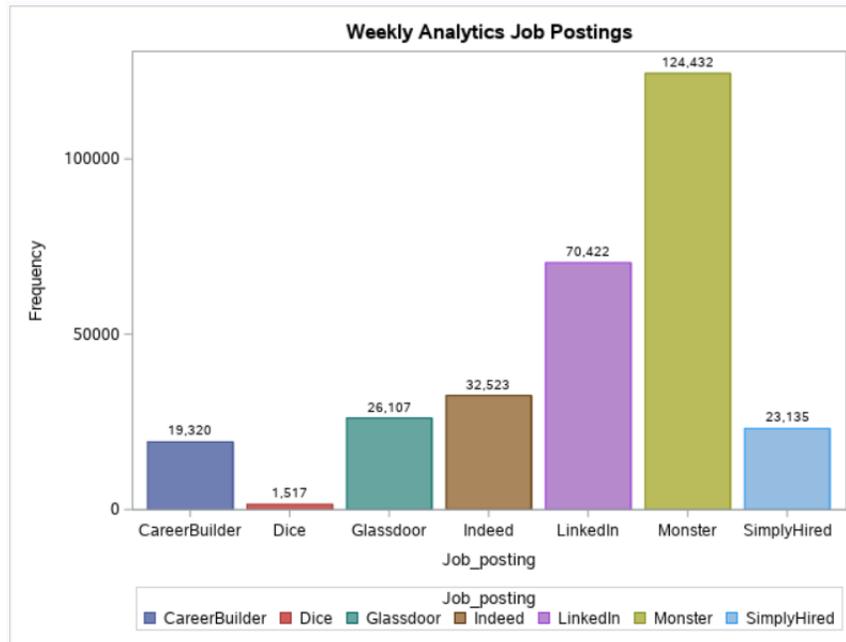
Salary Information from www.Glassdoor.com (September 2019)

Data Science / Analytics Employment / Career Opportunities

When searching job listing websites using the keyword, “data scientist”, there appears to be considerable demand for “data scientists.” Although cross-postings do exist, an exact match search for data scientist jobs was performed to ensure the results were relevant and the search keyword was processed similarly. The following VBAR chart shows the number of data scientist job postings by job listing website for the week of May 8th, 2021.



Using an exact match search, results using a VBAR chart for “analytics” job postings show the number of analytics job postings by job listing websites for the week of May 8th, 2021.



Weekly "Data Science" job postings on a number of employment / career websites from March 5th, 2021 are displayed, below.

6,801 Data Science Jobs

Sort by: Relevancy | Date

Data Scientist
NWN Corporation | Full-Time
Easy Apply

Data Scientist
AARATECH INC. | NJ - Jersey City | Contractor
\$60k - \$70k/year
Easy Apply

Data Scientist
Emerson | Sidney, OH, US | Not Specified

Data Scientist
NWN Corporation | Work From Home | Full-Time
Apply Now

Job Details | Company Overview

NWN Corporation is an established leader in technology-enabled services, transforming the customer experience for commercial, enterprise and public sector organizations through its solution-as-a-service platform. With over 1,300 customers throughout the U.S., NWN provides a unified customer experience with integrated offerings for unified communications, security, contact center, DaaS (device-as-a-service), connectivity and advanced technology solutions.

NWN is looking for a Data Scientist to join our team. This is a full time position with full benefits.

This is a REMOTE opportunity!!

Role Overview

The Data Scientist role will create customer facing analytics solutions, leveraging your mathematics and statistical experiences in an Agile software development environment to

354 Data Science Jobs

Search: Data Scientist

Remote Options: Remote Only, Exclude Remote, Work From Home Available

Posted Date: Any Date, Today, Last 3 Days, Last 7 Days

Employment Type: Full-time (254), Part-time (1), Contract (100), Third Party (50)

Distance: [Filter]

Data Scientist Co-op
Healthcare Financial, Inc. | Quincy, MA, USA
Full-time | 11 days ago
JOB SUMMARY The Data Scientist Co-op is responsible for processing structured and unstructured data, analyzing large amounts of raw information to find patterns that will help improve the company by building

Data Scientist
Emerson Electric Co | Sidney, OH, USA
Full-time | 15 days ago
At Emerson, we are innovators and problem-solvers, focused on a common purpose: leaving our world in a better place than we found it. Each and every day, our foundational values integrity, safety and quality.

Data Scientist
Emerson Climate Technologies, Inc. | Sidney, OH, USA
Full-time | 24 days ago
At Emerson, we are innovators and problem-solvers, focused on a common purpose: leaving our world in a better place than we found it. Each and every day, our foundational values integrity, safety and quality.

1,427 Data Science Jobs

Walmart 3.4 ★ Data Scientist Sunnyvale, CA New 1d

ImageTrend Inc 3.1 ★ Data Scientist Lakeville, MN New 3d

Rare Collectibles TV Data Scientist Torrance, CA Easy Apply 24h

Provident Research Inc Data Scientist- Biotech Remote \$115K - \$130K (Employer est.) Easy Apply 3d

Walmart 3.4 ★ Senior Data Scientist Sunnyvale, CA New 3d

Utility Associates 3.7 ★ Data Scientist Decatur, GA \$110K - \$135K (Employer est.) Easy Apply 2d

BASF 4.2 ★ Data Scientist - Weather and Climate Research Triangle Park, NC New 2d

Cox Automotive Data Scientist II

Walmart 3.4 ★ Data Scientist Sunnyvale, CA Apply Now Save

Rating Highlights
 Compensation & Benefits: 3.4 ★
 Culture & Values: 3.1 ★
 Career Opportunities: 3.3 ★
 Work/Life Balance: 3.0 ★

Job & Company Insights
 Job Type: N/A
 Job Function: Data Scientist
 Industry: Retail
 Size: 10000+ Employees

Position Summary...
What you'll do...
 Analytical Modeling - Feature relevance and selection, Exploratory data analysis methods and techniques, Advanced statistical methods and best- practice advanced modelling techniques (e.g., graphical models, Bayesian inference, basic level of NLP, Vision, neural networks, SVM, Random Forest etc.), Programming languages like R/Python.
 Model Assessment & Validation - Model fit testing, tuning, and validation techniques (e.g., Chi square, ROC

1,758 Data Science Jobs

What Data Scientist Where City, state, zip code, or "remote" Find jobs

Technology > Data & Database Occupations > Data Scientists & Statisticians

Last 7 days Remote Salary Estimate Job Type Location Company Experience Level

Data Scientist Knowsis 4.3 ★ 11 reviews Coronado, CA Apply Now

Resume insights
 Here's how your resume aligns with the job description

Experience & Skills
 ✓ R ✓ Statistical software ✓ SPSS show more

Education & Certificates
 ✓ Bachelor's degree

Ensure your resume details are correct before applying Update Resume

Job details
Job Type
 Full-time

8,058 Data Science Jobs

Cambridge, MA 3 alumni work here Promoted 1 applicant

Data Science Lead Vivante Health United States Remote Promoted 3 applicants

Sr. Data Scientist (Clinical Data Management) - Remote Premier Research Seattle, WA Remote Promoted 19 connections work here Promoted 4 applicants

SAS Analyst-Clinical data Iris Software Inc, Lawrenceville, GA Actively recruiting Promoted 20 applicants Easy Apply

Principal Data Scientist - Telecommute Optum Eden Prairie, MN Remote Vision, Dental, +1 benefit Promoted 37 connections work here 4 days ago 5 applicants

Principal Clinical Data Scientist Seagen Field, MN Medical, Vision, Dental, 401(k) Promoted 20 connections work here

Sr. Data Scientist IQVIA - Cambridge, MA Posted 5 days ago - 26 views Apply Save

See recent hiring trends for this company Retry Premium Free

Job
 • 1 applicant
 • Full-time

Company
 • 10001+ employees
 • Hospital & Health Care

Connections
 310 connections
 9 company alumni

Job Overview
 Responsible for data extract, transformation and loading (ETL) between different systems and databases, and on a number of key large-scale projects. The data is primarily in the areas of medical, pharmaceutical and life sciences, and as such, requires a strong understanding of related security considerations and practices.

Essential Functions
 • Responsible for data extract, transformation and loading (ETL) between different systems and databases, and on a number of key large-scale projects.

Qualifications
 • Bachelor's Degree
 • Exposure to healthcare \ life sciences \ clinical trials \ pharmaceutical data is highly preferred
 • Experience of Master & Reference Data Management
 • Understanding of Event-driven Messaging and SOA approaches
 • Extensive and recent experience of data Extraction, Transformation and Loading with at least one of:
 • Modern open-source ETL tools e.g. Talend, Pentaho, Mulesoft
 • Healthcare integration engines e.g. Mirth, Rhapsody, Ensemble
 • Enterprise ETL tools e.g. Informatica, PowerCenter, Oracle Data Integrator
 • Proven experience with a scripting language such as JavaScript, Perl, Python
 • Comfortable working with XML data and XSD Schema
 • At least one RDBMS (e.g. Oracle PL/SQL, SQL Server T-SQL)

8,233 Data Science Jobs

Data Scientist
Brilliant Infotech Inc.
Houston, TX

Role: Data Scientist – Focus specifically on experience with Deep learning algorithms and experience with time series data.
Houston, TX, Atlanta, GA & Albany, NY
Longterm contract
Data Scientist/ ML Scientist / Engineer:
Expertise in various predictive modeling
Expertise in supervised learning regression and classification algo...

Data Scientist
Whitney, Bradley & Brow...
Arlington, VA

JOB TITLE Data Scientist
PROJECT OVERVIEW: Provide support to the Deputy Assistant Secretary of Defense for Force Readiness (DASDR) in advising DoD leadership on key military readiness issues. Provide assessment of military readiness issues with the objective of ensuring ready and responsive military forces.
JOB DESCRIPTION AND R...

Machine Learning Data ...
Odyssey Systems Cons...
Lexington, MA

Position Summary: Our client develops novel technologies to detect and characterize threats for Homeland Security applications, and to provide decision support to operators and analysts. The client is in need of a data scientist and machine learning specialist to aid

Data Scientist
Brilliant Infotech Inc.
Houston, TX 77001

Job Type: full-time
Posted: 12 days ago
Location: Houston, TX

Description
Role: Data Scientist – Focus specifically on experience with Deep learning algorithms and experience with time series data.
Houston, TX, Atlanta, GA & Albany, NY
Longterm contract
Data Scientist/ ML Scientist / Engineer:
Expertise in various predictive modeling
Expertise in supervised learning regression and classification algorithms: Linear regression, Logistics regression, Decision Tree, SVM, KNN
Expertise in unsupervised learning such as PCA/SVD/ICA/LDA, K-means, t-SNE, Hierarchical clustering
Expertise in ensemble modeling such as Random forest and Gradient Boosting
Experience building deep learning models: DNN, CNN, RNN, LSTM
Deep Knowledge of ML/AI software and packages such as python: scikit-learn, Tensorflow and R: CARET
Expertise in model validation, hyperparameter tuning, and model selection techniques such as cross validation

1,339 Data Science Jobs

SimplyHired Job Title, Skills or Company: data scientist City, State, ZIP or "Remote": Search Jobs

Technology > Data & Database > Data Scientists & Statisticians

Relevance Date Job Type Minimum Salary 7 days Clear All

Data Scientist
Knowesis - Coronado, CA 4.3★
Skilled in using statistics and data visualization techniques to provide an accessible way to see and understand trends, outliers, and patterns in data.
Estimated: \$57,000 - \$81,000 a year Quick Apply 3d

Data Scientist
Walmart - Sunnyvale, CA 3.5★
Sound knowledge of machine learning and deep learning concepts. 5 years' experience in data science, machine learning, optimization models, or related field. 1d

Sr Data Scientist
Cox Communications Inc - Atlanta, GA 3.8★
The Sr. Data Scientist is required to have strong skills in advanced analytics algorithms, tools, and technologies to drive solution development at scale. Today
Estimated: \$93,000 - \$140,000 a year

Data Scientist II
Cox Automotive - Irvine, CA 3.4★
Our data scientists are responsible for leading the development of advanced analytics models to solve our customers problems through innovative products and... 2d
Estimated: \$80,000 - \$110,000 a year

Data Scientist
Knowesis - 4.3★
Coronado, CA Quick Apply

Job Details
Full-time Estimated: \$57,000 - \$81,000 a year
3 days ago

Qualifications
Analysis skills Statistical software SPSS Microsoft Excel SAS
Bachelor's degree

Full Job Description
Knowesis is looking for personnel to support the United States Special Operations Command Preservation of the Force and Family (USSOCOM POTFF) program in identifying and implementing innovative, valuable solutions across the Special Operations Forces (SOF) enterprise aimed at improving the short and long-term well-being of SOF warriors and their families. This work will be performed at the client's location on NAB, Coronado (San Diego, CA).
U.S. Citizenship Required (applicants without proof of US citizenship will not be considered due to the position's security clearance requirement).

Data Science Graduate Degree Programs

The majority of Data Scientists hold an undergraduate degree in a quantitative field such as Mathematics, Statistics, Decision Sciences, Computer Science, Management Information Systems, and Economics to name a few. Many also hold graduate degrees and/or certificates in the field of data science. The next table represents a starting point for professionals seeking Colleges and Universities offering graduate degrees in Data Analytics, Big Data Analytics and Data Science. Although this table identifies many academic institutions, it may not provide a complete list of all the academic institutions offering these prestigious graduate degrees. Readers are encouraged to perform your own search using one or more search engine(s) for a more complete list of all the colleges and universities offering these popular degree programs.

**Graduate Degree Programs in Data Analytics, Big Data Analytics and Data Science
(alphabetical)**

Boston University	New York University	University of Massachusetts Amherst
Columbia University – New York	North Carolina State University	University of Michigan
Cornell University	Oklahoma State University	University of Minnesota
George Mason University	San Diego State University	University of Notre Dame
Georgetown University	Stanford University	University of San Francisco
Georgia Tech	Texas A&M University	University of Southern California
Indiana University Bloomington	University of California – Berkeley	University of Washington
Johns Hopkins University	University of California San Diego	University of Wisconsin Madison
Louisiana State University	University of Chicago	

Data Science / Analytics Skills

Technical Skills	Non-Technical Skills
Programming including Python, R and SAS	Critical Thinking
SQL	Intellectual Curiosity
Statistics	Business Acumen
Excel	Verbal / Written Communication
Hadoop	Ability to Work in a Team
Machine Learning (ML)	
Artificial Intelligence (AI)	
Knowledge with Structured and Unstructured Data	
Data Wrangling	
Knowledge of Analytical Functions	
Data Visualization	

References

- Bureau of Labor Statistics (BLS) (accessed on April 9th, 2021), [BLS Employment Projections for Computer and Information Research Scientists from 2019 to 2029](#).
- Holak, Brian (January 31st, 2019), "[Demand for Data Scientists is Booming and will only Increase.](#)"
- Lafler, Kirk Paul (2019), "*Exploring the Skills Needed by the Data Scientist*", Proceedings of the MidWest SAS Users Group (MWSUG) Conference – 2019.
- Lafler, Kirk Paul (2019), "*Exploring the Skills Needed by the Data Scientist*", Proceedings of the Western Users of SAS Software (WUSS) Conference – 2019.
- Lafler, Kirk Paul (2019). [PROC SQL: Beyond the Basics Using SAS, Third Edition](#), SAS Institute Inc., Cary, NC, USA.
- Lafler, Kirk Paul (2016), "*What's Hot – Skills for SAS Professionals*", Proceedings of the SouthEast SAS Users Group (SESUG) Conference – 2016.
- Lafler, Kirk Paul (2016), "*What's Hot – Skills for SAS Professionals*", Proceedings of the Pharmaceutical SAS Users Group (PharmaSUG) Conference – 2016.

Acknowledgments

The author thanks the 2021 SESUG Conference Committee, particularly Soft Skills aren't Soft Section Chairs Brian Varney and Jason Brinkley for accepting my abstract and paper; the 2021 SESUG Executive Committee for organizing and supporting this virtual conference event; SAS Institute Inc. for providing SAS users with wonderful software; and SAS users everywhere for being the nicest people anywhere!

Trademark Citations

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.

Author Information

Kirk Paul Lafler is a lecturer and adjunct professor at San Diego State University; an advisor and adjunct professor at the University of California San Diego Extension; and teaches SAS, SQL, Python, R and Excel courses, seminars, workshops, and webinars to students, professionals and users around the world. Kirk has been a SAS user since 1979 and is the author of several books including, PROC SQL: Beyond the Basics Using SAS, Third Edition (SAS Press. 2019) along with papers and articles on a variety of SAS topics. Kirk has also been selected as an Invited speaker, educator, keynote and section leader at SAS conferences; and is the recipient of 25 "Best" contributed paper, hands-on workshop (HOW), and poster awards.

Comments and suggestions can be sent to:

Kirk Paul Lafler

SAS® Consultant, Application Developer, Programmer, Data Analyst, Educator and Author

E-mail: KirkLafler@cs.com

LinkedIn: <https://www.linkedin.com/in/KirkPaulLafler/>

LinkedIn: <https://www.linkedin.com/in/Order-of-Magnitude-Analytics/>

Twitter: @sasNerd