TODAY'S TOP STORIES

26 big data certifications that will pay off

Data scientists, data analysts and data engineers are in high demand. Here are the certifications that will give your career an edge.

By Thor Olavsrud

Senior Writer, CIO

SEP 20, 2018 2:11 PM PT

Data and big data analytics are the lifeblood of any successful business. Getting the technology right can be challenging but building the right team with the right skills to undertake big data initiatives can be even harder. Not surprisingly, that challenge is reflected in the rising demand for big data skills and certifications.

If you're looking for a way to get an edge, big data certification is a great option. Certifications measure your knowledge and skills against industry- and vendor-specific benchmarks to prove to employers that you have the right skillset. The number of big data certs is expanding rapidly.

[Learn the secrets of highly successful data analytics teams. | Beware the 12 myths of data analytics and the sure-fire ways organizations fail at data analytics. | Get ahead with the top certs for project management, agile, data science, IT management and the cloud. | Get weekly career tips by signing up for our CIO Leader newsletter.]

Below is our guide to the most sought after big data certifications to help you decide which cert is right for you.

If you would like to submit a big data certification to this directory, please email us.

Is big data certification worth pursuing?

Organizations are on the hunt for data scientists and analysts with expertise in dealing with big data. They also need big data architects to translate requirements into systems, data engineers to build data pipelines, developers who know their way around Hadoop and other technologies, and systems administrators and managers to tie everything together.

These skills are in high demand and are relatively rare. Individuals with the right mix of experience and skills can demand high salaries. The right certifications can help.

According to recent research by tech analyst firm Foote Partners, the explosion of interest in the internet of things (IoT) is creating large staffing deficits. IDC is predicting a 30 percent CAGR over the next five years, while McKinsey is expecting IoT to have a $4 trillion to $11 trillion global economic impact by 2025 as businesses look to IoT technologies to provide more insight.

[Looking to upgrade your career in tech? This comprehensive online course teaches you how.]

Foote Partners believes advanced data analytics will be a major force driving the IoT. Foote notes that key advanced analytics skills in the IoT area include:

- Apache Hadoop and related modules (HDFS, HBase, Flume, Oozie, Hive, Pig, YARN)
- NoSQL and NewSQL
- Apache Spark
- Machine learning and data mining

"Advanced data analytics capabilities are just too critical for staying competitive," says David Foote, co-founder, chief analyst and chief research officer of Foote Partners. "They've expanded in popularity from a few industries to nearly every industry and market."

The top 26 big data and data analytics certifications

- Amazon Web Services (AWS) Certified Big Data - Specialty
- Analytics: Optimizing Big Data Certificate
- Certification of Professional Achievement in Data Sciences
- Certified Analytics Professional
- Cloudera Certified Associate (CCA) Administrator
- Cloudera Certified Associate (CCA) Data Analyst

- Cloudera Certified Associate (CCA) Spark and Hadoop Developer
- Cloudera Certified Professional (CCP): Data Engineer
- EMC Proven Professional Data Scientist Associate (EMCDSA)
- HDP Apache Spark Developer
- HDP Certified Developer Big Data Hadoop
- Hortonworks Certified Associate (HCA)
- IBM Certified Data Architect – Big Data
- IBM Certified Data Engineer – Big Data
- MapR Certified Data Analyst 1.9
- MapR Certified Hadoop Developer 1.0
- MapR Certified Spark Developer 2.1
- Microsoft Certified Solutions Expert (MCSE): Data Management and Analytics
- Mining Massive Data Sets Graduate Certificate
- MongoDB Certified DBA Associate
- MongoDB Certified Developer Associate
- Oracle Business Intelligence Foundation Suite 11 Certified Implementation Specialist
- Post Graduate Program (PGP) in Big Data Analytics
- SAS Certified Big Data Professional
- SAS Certified Data Scientist Using SAS 9
- Stanford Data Mining and Applications Graduate Certificate

Amazon Web Services (AWS) Certified Big Data – Specialty

The AWS Certified Big Data – Specialty certification validates technical skills and experience in designing and implementing AWS services to derive value from data. It is intended to validate the ability to:

- Implement core AWS big data services according to basic architecture practices
- Design and maintain big data
- Leverage tools to automate data analysis

Organization: Amazon Web Services

Price: $300 registration fee for exam

How to prepare: Eligible candidates should have the AWS Certified Cloud Practitioner or current Associate-level certification (AWS Certified Solutions Architect – Associate, AWS Certified Developer – Associate, or AWS Certified SysOps Administrator – Associate). In addition, candidates should have a minimum of five years hands-on experience in a data analytics field, a background in defining and architecting AWS Big Data services, and experience in designing a scalable, cost-effective architecture to process data.

Analytics: Optimizing Big Data Certificate

The Analytics: Optimizing Big Data Certificate is an undergraduate-level program intended for business, marketing and operations managers, data analyst and professionals, financial industry professionals, and small business owners. The program brings together statistics, analysis, and written and oral communications skills. It introduces students to the tools needed to analyze large datasets, covering topics including importing data into an analytics software package, exploratory graphical and data analysis, building analytics models, finding the best model to explore correlation among variables and more.

Organization: University of Delaware

Price: $2,895 course fee

How to prepare: A basic background in statistics and some prior college coursework is recommended.

Certification of Professional Achievement in Data Sciences

The Certification of Professional Achievement in Data Sciences is a non-degree program intended to develop facility with foundational data science skills. The program consists of four courses: Algorithms for Data Science, Probability & Statistics, Machine Learning for Data Science, and Exploratory Data Analysis and Visualization.

Organization: Columbia University

Price: $1,936 per credit (a minimum of 12 credits, including the four courses, are required to complete the program). In addition, there is an $85 non-refundable application fee for the on-campus program and $150 for the online program. The online program also includes an additional

non-refundable technology fee of $395 per course.

How to prepare: An undergraduate degree and prior quantitative and introductory to computer programming coursework are required.

Certified Analytics Professional

The Certified Analytics Professional (CAP) credential is a general analytics certification that certifies end-to-end understanding of the analytics process, from framing business and analytic problems to acquiring data, methodology, model building, deployment and model lifecycle management. It requires completion of the CAP exam and adherence to the CAP Code of Ethics.

Organization: INFORMS

Price: $495 if you are an INFORMS member, or $695 if you're not. Team pricing is available for organizations.

How to prepare: A list of study courses and a series of webinars are available through registration.

Cloudera Certified Associate (CCA) Administrator

The CCA Administrator credential certifies an individual has demonstrated the core systems and cluster administrator skills required by organizations deploying Cloudera in the enterprise, including:

- An understanding of the installation process for Cloudera Manager, Cloudera Hadoop (CDH), and the ecosystem projects
- The ability to perform basic and advanced configuration needed to effectively administer a Hadoop cluster
- The ability to maintain and modify the cluster to support day-to-day operations in the enterprise
- An understanding of how to enable relevant services and configure the cluster to meet goals defined by security policy, as well as knowledge of basic security practices
- The ability to benchmark the cluster operational metrics, and test system configuration for operation and efficiency
- The ability to troubleshoot, including finding the root cause of a problem, optimize inefficient execution, and resolve resource contention scenarios

The credential requires passing the remote-proctored CCA Administrator Exam (CCA131), which consists of eight to 12 performance-based, hands-on tasks on a pre-configured Cloudera Enterprise cluster. Each question requires the candidate to solve a particular scenario. Some require making configuration and service changes via Cloudera Manager, while others demand knowledge of command line Hadoop utilities and basic competence with the Linux environment. Candidates have 120 minutes to complete the exam.

Organization: Cloudera

Price: $295

How to prepare: There are no prerequisites required, but Cloudera says the exam follows the same objectives as the Cloudera Administrator Training, making it excellent preparation for the exam.

Cloudera Certified Associate (CCA) Data Analyst

A SQL developer who earns the CCA Data Analyst certification demonstrates core analyst skills to load, transform and model Hadoop data to define relationships and extract meaningful results from the raw output. It requires passing the CCA Data Analyst Exam (CCA159), a remote-proctored set of eight to 12 performance-based, hands-on tasks on a CDH 5 cluster. Candidates have 120 minutes to implement a technical solution for each task. They must analyze the problem and arrive at an optimal approach in the time allowed.

Organization: Cloudera

Price: $295

How to prepare: Cloudera recommends candidates take the Cloudera Data Analyst Training course, which has the same objectives as the exam.

Cloudera Certified Associate (CCA) Spark and Hadoop Developer

The CCA Spark and Hadoop Developer credential certifies a professional has proven their core skills to ingest, transform and process data using Apache Spark and core Cloudera enterprise tools. It requires passing the remote-proctored CCA Spark and Hadoop Developer Exam (CCA175), which

consists of eight to 12 performance-based, hands-on tasks on a Cloudera Enterprise cluster. Each question requires the candidate to solve a particular scenario. Some cases may require a tool such as Impala or Hive, others may require coding. Candidates have 120 minutes to complete the exam.

Organization: Cloudera

Price: $295

How to prepare: There are no prerequisites required, but Cloudera says the exam follows the same objectives as the Cloudera Developer Training for Spark and Hadoop course, making it excellent preparation for the exam.

Cloudera Certified Professional (CCP): Data Engineer

The CCP: Data Engineer credential certifies the ability to perform core competencies required to ingest, transform, store and analyze data in Cloudera's CDH environment. It requires passing the remote-proctored CCP: Data Engineer Exam (DE575), a hands-on, practical exam in which each user is given five to eight customer problems each with a unique, large data set, a CDH cluster and four hours. For each problem, the candidate must implement a technical solution with a high degree of precision that meets all the requirements.

Organization: Cloudera

Price: $400

How to prepare: Cloudera suggests professionals seeking this certification have hands-on experience in the field and take the Cloudera Developer Training for Spark and Hadoop course.

EMC Proven Professional Data Scientist Associate (EMCDSA)

The EMCDSA certification demonstrates an individual's ability to participate and contribute as a data science team member on big data projects. It includes deploying the data analytics lifecycle, reframing a business challenge as an analytics challenge, applying analytic techniques and tools to analyze big data and create statistical models, selecting the appropriate data visualizations and more.

Organization: Dell EMC Education Services

Price: $600 for video-ILT streaming; $5,000 for instructor-led

How to prepare: EMC offers a training course, available as either a video or an instructor-led course.

HDP Apache Spark Developer

The HDP Apache Spark Developer certification is intended to validate an individual's understanding of Spark Core and Spark SQL applications in Scala or Python. The exam consists of a series of tasks that must be performed successfully on a live cluster.

Organization: Hortonworks

Price: $250 for exam

How to prepare: Hortonworks offers courses on its website with options that include live training, self-paced e-learning or a blended experience.

HDP Certified Developer Big Data Hadoop

The HDP Certified Developer Big Data Hadoop certification validates a developer's proficiency in Pig, Hive, Sqoop and Flume. The exam consists of a series of data ingestion, data transformation and data analysis tasks that must be performed on an HDP 2.4 cluster.

Organization: Hortonworks

Price: $250 for exam

How to prepare: Hortonworks offers courses on its website with options that include live training, self-paced e-learning or a blended experience.

💡 **NEW! Download the Fall 2018 digital issue of CIO**

IDC