ATTENTION:
The E20-007 Data Science and Big Data Analytics Exam, that supports the Associate – Data Science Version 1.0 certification, will retire on December 28, 2018.

How will this impact me?
If you plan to take the E20-007 Data Science and Big Data Analytics Exam and have taken the Data Science and Big Data Analytics course, we highly recommend that you take the exam prior to December 28, 2018

View the E20-007 exam description and practice test »

Once you have achieved Associate – Data Science Version 1.0 certification, your credential will not expire. As a certified EMC Proven Professional, you will continue to gain access to exclusive program benefits.

Data Science and Big Data Analytics v2 Exam »

The new DEA-7TT2 exam is available starting June 01, 2018. This exam aligns with the Data Science and Big Data Analytics v2 course.

Thank you
Dell EMC Proven Professional Program Team
Certification Overview
This certification enables the learner to immediately participate in big data and other analytics projects. The certification validates the practical foundation skills required by a Data Scientist.

Certification Requirements
To complete the requirements for this certification you must:

Pass the following Associate level exam:
- E20-007 Data Science and Big Data Analytics Exam

This certification will qualify towards the Specialist level certification in the Dell EMC Proven Professional Data Scientist (DECA-DS) track including the following:
- Specialist - Data Scientist, Advanced Analytics

Note: These details reflect certification requirements as of 3/30/12.

The Proven Professional Program periodically updates certification requirements. *Please check the Proven Professional CertTracker website regularly for the latest information and for other options to meet the Associate level requirement.
E20-007 Data Science and Big Data Analytics Exam

Exam Description

Overview
This exam is a qualifying exam for the Associate - Data Science (DECA-DS) track.

This exam focuses on the practice of data analytics, the role of the Data Scientist, the main phases of the Data Analytics Lifecycle, analyzing and exploring data with R, statistics for model building and evaluation, the theory and methods of advanced analytics and statistical modeling, the technology and tools that can be used for advanced analytics, operationalizing an analytics project, and data visualization techniques. Successful candidates will achieve the Dell EMC Proven Professional – Data Science Associate credential.

Dell EMC provides free practice tests to assess your knowledge in preparation for the exam. Practice tests allow you to become familiar with the topics and question types you will find on the proctored exam. Your results on a practice test offer one indication of how prepared you are for the proctored exam and can highlight topics on which you need to study and train further. A passing score on the practice test does not guarantee a passing score on the certification exam.

Exam Topics
Topics likely to be covered on this exam include:

Big Data Analytics and the Data Scientist Role (7%)
- The characteristics of Big Data
- The practice of analytics
- The role and required skills of a Data Scientist

Data Analytics Lifecycle (9%)
- Discovery
- Data preparation
- Model planning and building
- Communicating results
- Operationalizing a data analytics project

Initial Analysis of the Data (15%)
- Using basic R commands to analyze data
- Using statistical measures and visualization to understand data
- The theory, process, and analysis of results to evaluate a model

Advanced Analytics for Big Data – Theory and Methods (40%)
- K-means clustering
- Association rules
- Linear regression
- Logistic Regression
- Naïve Bayesian classifiers
- Decision trees
- Time Series Analysis
- Text Analytics

Dell EMC
Hopkinton
Massachusetts
01748-9103
1-508-435-1000
In North America
1-866-464-7381
Advanced Analytics for Big Data – Technology and Tools (20%)

- MapReduce
- Hadoop Ecosystem
- SQL OLAP extensions, Windows functions, user defined functions, and aggregates
- MADlib

Operationalizing an Analytics Project and Data Visualization Techniques (9%)

- Best practices for operationalizing an analytics project
- Best practices for planning and creating effective data visualizations

The percentages after each topic above reflects the approximate distribution of the total question set across the exam.

Recommended Training

The following curriculum is recommended for candidates preparing to take this exam.

Please complete one of the following courses

<table>
<thead>
<tr>
<th>Course Title</th>
<th>Course Number</th>
<th>Mode</th>
<th>Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Science and Big Data Analytics</td>
<td>MR-1CP-DSBDA</td>
<td>Instructor-Led</td>
<td>3/30/12</td>
</tr>
<tr>
<td>Data Science and Big Data Analytics – Video IJT</td>
<td>MR-1TP-DSBDA</td>
<td>Video IJT-Stream</td>
<td>3/30/12</td>
</tr>
</tbody>
</table>

The course material is supplemented by the textbook. Text book is optional

<table>
<thead>
<tr>
<th>Course Title</th>
<th>Mode</th>
<th>Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data</td>
<td>Book</td>
<td>1/27/15</td>
</tr>
</tbody>
</table>


Note: These exam description details reflect contents as of **March 30, 2012**.

The Proven Professional Program periodically updates exams to reflect technical currency and relevance. Please check the Proven Professional website regularly for the latest information.