



---

# E20-007

---

**Dell EMC Data Science and Big Data Analytics**  
Exam Summary – Syllabus – Questions



---

## Table of Contents

<b>Introduction to E20-007 Exam on Dell EMC Data Science and Big Data Analytics</b> .....	<b>2</b>
<b>Dell EMC E20-007 Certification Details:</b> .....	<b>2</b>
<b>Dell EMC E20-007 Exam Syllabus:</b> .....	<b>2</b>
<b>E20-007 Sample Questions:</b> .....	<b>3</b>
<b>Answers to E20-007 Exam Questions:</b> .....	<b>6</b>

# Introduction to E20-007 Exam on Dell EMC Data Science and Big Data Analytics

This page is a one-stop solution for any information you may require for Dell EMC Data Science and Big Data Analytics (E20-007) Certification exam. The Dell EMC E20-007 Exam Summary, Syllabus Topics and Sample Questions provide the base for the actual Dell EMC Certified Associate - Data Science (DECA-DS) exam preparation, we have designed these resources to help you get ready to take your dream exam.

The Dell EMC Data Science and Big Data Analytics credential is globally recognized for validating Dell EMC Data Science Associate knowledge. With the Dell EMC Certified Associate - Data Science (DECA-DS) Certification credential, you stand out in a crowd and prove that you have the Dell EMC Data Science Associate knowledge to make a difference within your organization. The Dell EMC Data Science and Big Data Analytics Certification (E20-007) exam will test the candidate's knowledge on following areas.

## Dell EMC E20-007 Certification Details:

Exam Name	Dell EMC Data Science and Big Data Analytics
Exam Code	E20-007
Exam Duration	90 minutes
Exam Questions	60
Passing Score	60%
Exam Price	\$165 (USD)
Exam Registration	<a href="#">Pearson VUE</a>
Sample Questions	<a href="#">Dell EMC Data Science Associate Certification Sample Question</a>
Practice Exam	<a href="#">Dell EMC Data Science Associate Certification Practice Exam</a>

## Dell EMC E20-007 Exam Syllabus:

Objective	Details
Big Data Analytics and the Data Scientist Role (7%)	<ul style="list-style-type: none"> <li>- The characteristics of Big Data</li> <li>- The practice of analytics</li> <li>- The role and required skills of a Data Scientist</li> </ul>
Data Analytics Lifecycle (9%)	<ul style="list-style-type: none"> <li>- Discovery</li> <li>- Data preparation</li> <li>- Model planning and building</li> </ul>

Objective	Details
	<ul style="list-style-type: none"> <li>- Communicating results</li> <li>- Operationalizing a data analytics project</li> </ul>
Initial Analysis of the Data (15%)	<ul style="list-style-type: none"> <li>- Using basic R commands to analyze data</li> <li>- Using statistical measures and visualization to understand data</li> <li>- The theory, process, and analysis of results to evaluate a model</li> </ul>
Advanced Analytics for Big Data – Theory and Methods (40%)	<ul style="list-style-type: none"> <li>- K-means clustering</li> <li>- Association rules</li> <li>- Linear regression</li> <li>- Logistic Regression</li> <li>- Naïve Bayesian classifiers</li> <li>- Decision trees</li> <li>- Time Series Analysis</li> <li>- Text Analytics</li> </ul>
Advanced Analytics for Big Data – Technology and Tools (20%)	<ul style="list-style-type: none"> <li>- MapReduce</li> <li>- Hadoop Ecosystem</li> <li>- SQL OLAP extensions, Windows functions, user defined functions, and aggregates</li> <li>- MADlib</li> </ul>
Operationalizing an Analytics Project and Data Visualization Techniques (9%)	<ul style="list-style-type: none"> <li>- Best practices for operationalizing an analytics project</li> <li>- Best practices for planning and creating effective data visualizations</li> </ul>

## E20-007 Sample Questions:

**Q1: Your organization has a website where visitors randomly receive one of two coupons. It is also possible that visitors to the website will not receive a coupon.**

**You have been asked to determine if offering a coupon to visitors to your website has any impact on their purchase decision. Which analysis method should you use?**

**Options:**

- A. One-way ANOVA
- B. K-means clustering
- C. Association rules
- D. Student T-test

**Q2: Your customer provided you with 2, 000 unlabeled records and asked you to separate them into three groups. What is the correct analytical method to use?**

**Options:**

- A. K-means clustering

- B. Naive Bayesian classification
- C. Linear regression
- D. Logistic regression

**Q3: Your company has 3 different sales teams. Each team's sales manager has developed incentive offers to increase the size of each sales transaction.**

**Any sales manager whose incentive program can be shown to increase the size of the average sales transaction will receive a bonus.**

**Data are available for the number and average sale amount for transactions offering one of the incentives as well as transactions offering no incentive.**

**The VP of Sales has asked you to determine analytically if any of the incentive programs has resulted in a demonstrable increase in the average sale amount.**

**Which analytical technique would be appropriate in this situation?**

**Options:**

- A. Wilcoxon Rank Sum Test
- B. Student's t-test
- C. One-way ANOVA
- D. Multi-way ANOVA

**Q4: Your colleague, who is new to Hadoop, approaches you with a question. They want to know how best to access their data. This colleague has a strong background in data flow languages and programming. Which query interface would you recommend?**

**Options:**

- A. HSQL
- B. Pig
- C. Hive
- D. HBase

**Q5: Your colleague, who is new to Hadoop, approaches you with a question. They want to know how best to access their data.**

**This colleague has previously worked extensively with SQL and databases. Which query interface would you recommend?**

**Options:**

- A. Howl
- B. Hive
- C. HBase
- D. Pig

**Q6: You submit a MapReduce job to a Hadoop cluster and notice that although the job was successfully submitted, it is not completing. What should you do?**

**Options:**

- A. Ensure that the TaskTracker is running.
- B. Ensure that the JobTracker is running
- C. Ensure that the NameNode is running
- D. Ensure that a DataNode is running

**Q7: You submit a Map Reduce job to a Hadoop cluster. However, you notice that although the job was successfully submitted, it is not completing. What should be done to identify the issue?**

**Options:**

- A. Ensure NameNode is running
- B. Ensure DataNode is running
- C. Ensure TaskTracker is running
- D. Ensure JobTracker is running

**Q8: You have used k-means clustering to classify behavior of 100, 000 customers for a retail store. You decide to use household income, age, gender and yearly purchase amount as measures.**

**You have chosen to use 8 clusters and notice that 2 clusters only have 3 customers assigned. What should you do?**

**Options:**

- A. Decrease the number of measures used
- B. Decrease the number of clusters
- C. Increase the number of clusters
- D. Identify additional measures to add to the analysis

**Q9: You have two tables of customers in your database. Customers in cust\_table\_1 were sent an e-mail promotion last year, and customers in cust\_table\_2 received a newsletter last year.**

**Customers can only be entered in once per table. You want to create a table that includes all customers, and any of the communications they received last year. Which type of join would you use for this table?**

**Options:**

- A. Full outer join
- B. Left outer join
- C. Inner join
- D. Cross join

**Q10 : You have run the association rules algorithm on your data set, and the two rules {banana, apple} => {grape} and {apple, orange}=> {grape} have been found to be relevant. What else must be true?**

**Options:**

- A. {banana, apple, grape, orange} must be a frequent itemset.
- B. {banana, apple} => {orange} must be a relevant rule.
- C. {grape} => {banana, apple} must be a relevant rule.
- D. {grape, apple, orange} must be a frequent itemset.

**Answers to E20-007 Exam Questions:**

Question: 1	Answer: A	Question: 2	Answer: A
Question: 3	Answer: C	Question: 4	Answer: B
Question: 5	Answer: B	Question: 6	Answer: A
Question: 7	Answer: C	Question: 8	Answer: B
Question: 9	Answer: A	Question: 10	Answer: D

Note: If you find any typo or data entry error in these sample questions, we request you to update us by commenting on this page or write an email on [feedback@analyticsexam.com](mailto:feedback@analyticsexam.com)