

Quiz Grade: 80.0% (A)

Quiz Submission

Data Science and Big Data Analytics

Student Name: Emmanuel Akarikumutima
Student ID: UM96010AC105232
Study Program: Accounting
Quiz ID: 68895fa2-999c-41fc-bb70-9d640a969a8d
Submission Date: July 07, 2026 at 4:44 AM
Total Questions: 10

Course Information

Course Title: Data Science and Big Data Analytics
Course Code: SCDESC 427
Credit Hours: 3

Quiz Questions, Student Answers, and Correct Answers

Question 1 of 10

Multiple Choice Question

Which of the following is a key characteristic of Big Data?

Available Options:

- (A) Variety
- (B) Velocity
- (C) Volume

(D) All of the above ← STUDENT SELECTED ← CORRECT ANSWER

Student's Answer: Option D: All of the above

Correct Answer: Option D: All of the above

Question 2 of 10

Multiple Choice Question

In data science, what is the primary purpose of data visualization?

Available Options:

- (A) To enhance data quality
- (B) To make data more accessible ← STUDENT SELECTED ← CORRECT ANSWER**
- (C) To reduce data volume
- (D) To increase data velocity

Student's Answer: Option B: To make data more accessible

Correct Answer: Option B: To make data more accessible

Question 3 of 10

Multiple Choice Question

Which of the following algorithms is commonly used for clustering in Big Data Analytics?

Available Options:

(A) Linear Regression

(B) K-Means ← STUDENT SELECTED ← CORRECT ANSWER

(C) Decision Trees

(D) Random Forest

Student's Answer: Option B: K-Means

Correct Answer: Option B: K-Means

Question 4 of 10

Multiple Choice Question

What is a primary benefit of using distributed computing in Big Data?

Available Options:

(A) Reducing data variety

(B) Improving data visualization

(C) Handling large volumes of data efficiently ← STUDENT SELECTED ← CORRECT ANSWER

(D) Enhancing data security

Student's Answer: Option C: Handling large volumes of data efficiently

Correct Answer: Option C: Handling large volumes of data efficiently

Question 5 of 10

Multiple Choice Question

Which of the following best describes supervised learning?

Available Options:

(A) Learning without labeled data ← STUDENT SELECTED

(B) Learning with labeled data ← CORRECT ANSWER

(C) Learning through reinforcement

(D) Learning with unlabeled data

Student's Answer: Option A: Learning without labeled data

Correct Answer: Option B: Learning with labeled data

Question 6 of 10

Multiple Choice Question

Which tool is commonly used for Big Data processing?

Available Options:

- (A) Excel
- (B) TensorFlow

(C) Hadoop ← STUDENT SELECTED ← CORRECT ANSWER

(D) SAS

Student's Answer: Option C: Hadoop

Correct Answer: Option C: Hadoop

Question 7 of 10

Multiple Choice Question

Which of the following is NOT a type of data analytics?

Available Options:

- (A) Descriptive Analytics
- (B) Predictive Analytics
- (C) Prescriptive Analytics

(D) Reactive Analytics ← CORRECT ANSWER

Student's Answer: No answer provided

Correct Answer: Option D: Reactive Analytics

Question 8 of 10

Text Answer Question

Explain the role of machine learning in data science and how it helps in making predictions.

Student's Answer:

Machine learning is a fundamental component of data science that enables computers to learn from data and make decisions or predictions without being explicitly programmed. Instead of relying on fixed rules, machine learning algorithms identify patterns and relationships within large datasets, allowing them to improve their performance over time as more data becomes available. In the data science process, machine learning is used after data has been collected, cleaned, and prepared. Data scientists train machine learning models using historical data, where the model learns the relationship between input variables (features) and the desired output (target). Once trained, the model can analyze new data and generate accurate predictions or classifications. Machine learning helps in making predictions by identifying trends and patterns that may not be obvious through traditional statistical methods. For example, businesses use machine learning to forecast sales, banks detect fraudulent transactions, healthcare providers predict disease risks, and e-commerce companies recommend products based on customer behavior. These predictive capabilities enable organizations to make informed decisions, improve efficiency, reduce costs, and enhance customer experiences. Overall, machine learning plays a crucial role in data science by transforming raw data into actionable insights and reliable predictions. Its ability to learn from data, adapt to new information, and automate decision-making makes it one of the most powerful tools in modern data analytics.

Correct Answer:

Machine learning allows computers to learn from data without being explicitly programmed. It is used in data science to build models that can identify patterns and make predictions based on historical data.

Question 9 of 10

Text Answer Question

Describe the relationship between data science and big data analytics.

Student's Answer:

Data science and big data analytics are closely related fields that work together to extract valuable insights from data. Data science is a broad discipline that combines statistics, mathematics, programming, and machine learning to collect, process, analyze, and interpret data for decision-making. Big data analytics is a specialized area that focuses on analyzing extremely large, complex, and rapidly growing datasets that cannot be processed efficiently using traditional methods. Big data provides the vast amounts of structured, semi-structured, and unstructured data that data scientists use to build predictive models and generate insights. Big data analytics uses distributed computing technologies, such as Hadoop and Spark, to store, process, and analyze these massive datasets efficiently. Data science then applies analytical techniques, statistical methods, and machine learning algorithms to discover patterns, identify trends, and make accurate predictions based on the analyzed data. Together, data science and big data analytics enable organizations to make data-driven decisions, improve operational efficiency, detect fraud, personalize customer experiences, optimize business processes, and forecast future outcomes. While big data analytics focuses on managing and processing large-scale data, data science focuses on extracting meaningful knowledge and creating predictive models from that data. In summary, big data analytics provides the infrastructure and techniques for handling massive datasets, while data science uses those datasets to generate insights, solve complex problems, and support strategic decision-making.

Correct Answer:

Data science involves extracting insights from data, while big data analytics focuses on analyzing large volumes of data to uncover hidden patterns, correlations, and other insights. Big data analytics is a subset of data science that deals specifically with the challenges and opportunities of processing large datasets.

Question 10 of 10

Text Answer Question

What challenges do data scientists face when dealing with big data, and how can they be addressed?

Student's Answer:

Data scientists face several challenges when working with big data due to its size, complexity, and speed of generation. One of the biggest challenges is managing the enormous volume of data, which often exceeds the capacity of traditional databases and computing systems. This challenge can be addressed by using distributed storage and processing technologies such as Hadoop and Spark, which allow data to be stored and analyzed across multiple computers. Another challenge is ensuring data quality. Big data often contains missing values, duplicate records, inconsistent formats, and inaccurate information. Data scientists address this issue through data cleaning, validation, preprocessing, and quality control techniques before analysis begins. The variety of data is also a significant challenge. Big data includes structured, semi-structured, and unstructured data such as text, images, videos, and social media content. To manage this diversity, data scientists use flexible storage systems such as NoSQL databases and specialized tools capable of processing different data formats. High data velocity presents another challenge because data is generated continuously and must often be processed in real time. Technologies such as Apache Spark Streaming, Apache Kafka, and real-time analytics platforms help organizations process and analyze streaming data quickly. Data privacy and security are also major concerns, especially when handling sensitive personal or financial information. These risks can be reduced by implementing strong encryption, access controls, authentication, data governance policies, and compliance with data protection regulations. Finally, data scientists require advanced technical skills in statistics, programming, machine learning, and distributed computing. Organizations can address this challenge by providing continuous training, encouraging collaboration among multidisciplinary teams, and adopting modern analytical tools that simplify complex tasks. In conclusion, although big data presents challenges related to volume, variety, velocity, data quality, security, and technical complexity, these issues can be effectively addressed through modern technologies, strong data management practices, and continuous skill development. These solutions enable organizations to gain valuable insights and make informed, data-driven decisions.

Correct Answer:

Challenges include handling the volume, variety, and velocity of data, ensuring data quality, and maintaining privacy. These challenges can be addressed through the use of scalable data processing frameworks, data cleaning and preprocessing techniques, and implementing robust security measures.