IoT Fundamental: Big Data and Analysis

Started on	Saturday, 17 September 2022, 2:23 PM
State	Finished
Completed on	Saturday, 17 September 2022, 2:33 PM
Time taken	10 mins 26 secs
Marks	44.00/44.00
Grade	**100.00** out of 100.00

Question **1**

Correct

Mark 2.00 out of 2.00

What is an example of data in motion?

Select one:

- recording road traffic volumes and patterns for future highway planning
- medical information being transmitted from an ambulance to emergency department staff as a critically ill patient is being transported to the hospital ✔
- hourly weather information being collected in preparation for the next day weather forecast for a specific location
- collecting sales and transaction records in preparation for a monthly sales report from sales consultants as they travel between customers

Refer to curriculum topic: 1.1.1
Data in motion is dynamic data that requires real-time processing before the data becomes obsolete. It represents the continuous interactions between people, processes, data and things. In this example the real-time medical information enables the emergency staff to be appropriately prepared before the patient arrives at the hospital.

The correct answer is: medical information being transmitted from an ambulance to emergency department staff as a critically ill patient is being transported to the hospital

Question **2**

Correct

Mark 2.00 out of 2.00

What is a purpose of applying data anonymization process to a data sets?

Select one:

- to compress the data sets
- to reduce the size of the data sets
- to remove identifiable personal information ✔
- to adjust the value length of certain data fields

Refer to curriculum topic: 1.2.2
Data anonymization is a process of either encrypting or removing identifiable personal information from data sets to achieve privacy protection.

The correct answer is: to remove identifiable personal information

Question **3**

Correct

Mark 2.00 out of 2.00

What is Hadoop?

Select one:

- a method of preventing loops when analyzing Big Data
- a groundbreaking method of moving large amounts of data through micro loops
- a framework that allows distributed processing of data across clusters of computers ✔
- a method of sharing data across multiple companies using computing resources housed within each respective company

Refer to curriculum topic: 1.3.2
Data management and analysis today are characterized by the use of flat file databases, relational database management system (RDBMS), and the Hadoop framework that allows distributed processing of data across clusters of computers using simple programming models.

The correct answer is: a framework that allows distributed processing of data across clusters of computers

Question **4**

Correct

Mark 2.00 out of 2.00

Which statement describes SQLite?

Select one:

- It is an example of flat file database.
- It is an embedded SQL database engine. ✔
- It is a free version of RDBMS suitable for enterprises.
- It is a fully functional RDBMS for distributed data processing.

Refer to curriculum topic: 1.3.2
SQLite is an embedded SQL database engine in that it does not follow the traditional client/server model like SQL RDBMS (relational database management system). SQLite reads and writes directly to ordinary disk files.

The correct answer is: It is an embedded SQL database engine.

Question **5**

Correct

Mark 2.00 out of 2.00

A multi-campus school wants to perform analytics on classes held during the past 5 years. The school wants to know which classes filled up the quickest across all campuses and which classes filled up the quickest at each campus. The school also wants to know if there is a relationship between the number of passing students and the speed in which a class taught by a particular teacher fills. If the school could only choose one type of database to store the data on one server, which type would be best suited for this task?

Select one:

- flat
- local
- Hadoop
- relational ✔

Refer to curriculum topic: 1.3.2
A relational database, even though it has multiple, connected tables, can reside on one server and would be best for this type of data. A local database is typically used to collect and store local data, for example, a database of all movies and music for a particular family. A flat database would most likely not be used in a multi-location school to store student data such as this. Hadoop is best to use when distributing processing power across server clusters.

The correct answer is: relational

Question **6**

Correct

Mark 2.00 out of 2.00

What are two examples of unstructured data? (Choose two.)

Select one or more:

- video content ✔
- user account data
- SQL queries
- blog entry ✔
- customer account spreadsheet

Refer to curriculum topic: 1.2.3
Unstructured data is raw data, data that is not organized in a predefined way. Examples of unstructured data would be contents of photos, audio, video, web pages, blogs, books, journals, and white papers.

The correct answers are: video content, blog entry

Question 7

Correct

Mark 2.00 out of 2.00

What is a characteristic of structured data?

Select one:

- Structured data is subject to intellectual property restrictions.
- It has a predefined organization. ✔
- It is raw data.
- It generates new knowledge.

Refer to curriculum topic: 1.2.3
Structured data is data that is structured and can be entered, classified, and queried by a computer. Data that is found in databases and spreadsheets is an example of structured data.

The correct answer is: It has a predefined organization.

Question 8

Correct

Mark 2.00 out of 2.00

Which method does openPDS use to protect user privacy of GPS records on a mobile device?

Select one:

- requiring authentication to be completed first
- encrypting the communication from data to the app
- providing answers to specific queries instead of raw data ✔
- removing identifiable personal information before sending data to the app

Refer to curriculum topic: 1.2.2
Using the SafeAnswers framework, openPDS provides only answers to specific queries and no raw data is sent. The calculation for the answer is done within the personal data store (PDS) of the user.

The correct answer is: providing answers to specific queries instead of raw data

Question 9

Correct

Mark 2.00 out of 2.00

When is data considered to be information?

Select one:

- when it is stored
- when it is recorded
- when it is processed and analyzed ✔
- when it is generated

Refer to curriculum topic: 1.1.1
Data that has been processed, organized, analyzed, or presented in a meaningful way becomes information.

The correct answer is: when it is processed and analyzed

Question **10**

Correct

Mark 2.00 out of 2.00

What are two key components in creating data analysis tools from scratch? (Choose two.)

Select one or more:

- coding ✔
- modeling ✔
- data sets
- performance
- program length

Refer to curriculum topic: 1.3.2
Modeling and coding are the two key components in the process of creating data analysis tools from scratch. Modeling consists of deciding what to do with the data to achieve the desired results and conclusions. A well-developed model can be used to handle multiple types of data sets. The code is the program that implements the model and processes the data according to the model already developed. The length and performance are factors and features of a program.

The correct answers are: coding, modeling

Question **11**

Correct

Mark 2.00 out of 2.00

What has contributed to the exponential growth in data generation?

Select one:

- the increasing number of mobile devices ✔
- the increasing number of standalone devices
- the increasing number of isolated software applications
- the increasing number of physical installations for protecting environment facilities

Refer to curriculum topic: 1.1.2
An increased number of sensors and other end devices as well as mobile devices are contributing to an exponential growth in data generation.

The correct answer is: the increasing number of mobile devices

Question 12

Correct

Mark 2.00 out of 2.00

In the data analysis process, which sequence depicts the work flow suitable for data at rest?

Select one:

- act > analyze > store > notify
- analyze > notify > act > store
- notify > store > act > analyze
- store > analyze > notify > act ✔

Refer to curriculum topic: 1.2.4
Data at rest is static data that is stored in a database first and then analyzed and interpreted. Data at rest follows the traditional analysis flow of store > analyze > notify > act. Once the data is analyzed, decision makers are notified and determine whether action is needed.

The correct answer is: store > analyze > notify > act

Question 13

Correct

Mark 2.00 out of 2.00

Which statement describes the paradigm that is promoted in the Cisco Fog Computing Model?

Select one:

- All data analysis and decision making should take place near the data source.
- Some data analysis should take place at the edge of infrastructure rather than at a central location. ✔
- Data generated by edge devices should be sent to the nearest regional data analysis center for data aggregation.
- Data collected at the edge of the infrastructure should be stored in a central data center for security and backup operations.

Refer to curriculum topic: 1.2.4
The Cisco Fog Computing Model states that some of the analysis work should take place at the network edge instead of at a centralized location. Sensors and controllers at the edge can make smart decisions based on the data collected locally, factors that facilitate faster response and action.

The correct answer is: Some data analysis should take place at the edge of infrastructure rather than at a central location.

Question 14

Correct

Mark 2.00 out of 2.00

Which term describes the growth rate of data in the IoT?

Select one:

- linear
- cyclical
- uniform
- exponential ✔

Refer to curriculum topic: 1.1.2
Current trends and forecasts all indicate that today data grows exponentially.

The correct answer is: exponential

Question **15**

Correct

Mark 2.00 out of 2.00

Which type of information is captured and stored as events happen?

Select one:

- critical
- analytical
- comparative
- transactional ✔

Refer to curriculum topic: 1.2.1
The two primary types of business information useful to a company are transactional information and analytical information. Transactional information is captured and stored as events happen. Transactional information can be used to analyze daily sales reports and production schedules to determine how much inventory to carry. Analytical information supports managerial analysis tasks like determining whether the organization should build a new manufacturing plant or hire additional sales personnel.

The correct answer is: transactional

Question **16**

Correct

Mark 2.00 out of 2.00

What are three examples of a NoSQL database? (Choose three.)

Select one or more:

- Ceph
- HDFS
- Redis ✔
- MongoDB ✔
- GlusterFS
- Apache Cassandra ✔

Refer to curriculum topic: 1.3.2
MongoDB, Apache Cassandra, and Redis are examples of a NoSQL database. The Hadoop Distributed File System (HDFS), Ceph, and GlusterFS are examples of distributed file systems (DFS).

The correct answers are: Redis, MongoDB, Apache Cassandra

Question 17

Correct

Mark 2.00 out of 2.00

How do sensors relate to Big Data?

Select one:

- They are types of multimedia applications that are sources for Big Data.
- They are devices that collectively generate large amounts of data. ✔
- They are devices that can only be used with static data.
- They produce structured data.

Refer to curriculum topic: 1.2.1
The use of sensors in IoT systems is growing exponentially. Each sensor has a multiplicative effect on the amount of data generated. Sensors are quickly becoming the greatest contributors toward Big Data.

The correct answer is: They are devices that collectively generate large amounts of data.

Question 18

Correct

Mark 2.00 out of 2.00

What is a characteristic of open data?

Select one:

- data that lacks intellectual property restrictions ✔
- data that lacks predefined organization
- data that does not need to be stored
- data that does not generate new knowledge

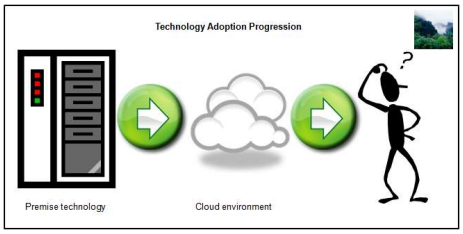
Refer to curriculum topic: 1.2.2
Open data is not protected by intellectual property restrictions and can be used and redistributed without legal, technical, or social restrictions.

The correct answer is: data that lacks intellectual property restrictions

Question **19**

Correct

Mark 2.00 out of 2.00



Refer to the exhibit. To remain competitive, a company has progressed from on-premise technology to the cloud environment. What technology environment would a manager need to consider to accommodate long-term storage and immediate analysis of data in motion.

Select one:

- an analytic model
- a hybrid model ✔
- the fog model
- a cloud model will accompish both requirements
- on-premise clouds

Refer to curriculum topic: 1.3.1
A manager should consider a hybrid option that includes cloud computing for long-term storage of data and fog computing for immediate access to streaming data. The immediate access to the data at the company edge would allow for rapid analysis for time-sensitive applications.

The correct answer is: a hybrid model

Question **20**

Correct

Mark 2.00 out of 2.00

What is true of Big Data in comparison to traditional data?

Select one:

- Traditional data is represented through binary strings, whereas Big Data is represented through hexadecimal strings.
- Both types of data require the same hardware for processing and storage.
- Big Data requires a different approach to analysis, computing, and storage mechanisms. ✔
- Big Data means that the data sets are being sent through the network in larger packets than the sets that contain legacy data.

Refer to curriculum topic: 1.2.1
Scale defines the difference between Big Data and the data that existed before the term Big Data existed. Based on the increased volume and type of data, big data requires a different approach to data analysis, computing, and storage. Different hardware and applications are required to handle the quantity of data produced. Both types, however, still involve binary strings. There is no difference between Big Data packets and packets that are not Big Data.

The correct answer is: Big Data requires a different approach to analysis, computing, and storage mechanisms.

Question **21**

Correct

Mark 2.00 out of 2.00

Which characteristic of big data describes different types of datasets that include both structured and unstructured data?

Select one:

- velocity
- volume
- variety ✔
- veracity

Refer to curriculum topic: 1.2.1
The characteristics of big data can be described in four Vs:

- **Volume** - the amount of data being transported and stored
- **Velocity** - the rate at which this data is generated
- **Variety** - the different types of data both structured and unstructured: video, audio, text
- **Veracity** - the process of preventing inaccurate data from spoiling the data sets

The correct answer is: variety

Question **22**

Correct

Mark 2.00 out of 2.00

Which two statements describe characteristics of data in motion? (Choose two.)

Select one or more:

- Its value changes over time. ✔
- It is stored at a central data center.
- It requires real-time processing close to the source. ✔
- It is the data in RAM during a data analysis process.
- It is stored in removable devices for easy transportation.

Refer to curriculum topic: 1.2.4
Data in motion describes the status of data to be distributed among different locations, the need of data to be analyzed close to the source, and how its value changes dynamically over time.

The correct answers are: Its value changes over time., It requires real-time processing close to the source.

◄ Chapter 1 Terms and Concepts Practice

Jump to...

Read Chapter 2: Fundamentals of Data Analysis ►