

1. What is a characteristic of streaming data? (0.5 points)
 - A. Data is finite in size and size determines the time and space of processing the data.
 - B. The data is unbounded in size and the size determines the time and space of processing the data.
 - C. Data is unbounded in size but requires only finite time and space to process it.
 - D. The data is finite and requires only finite time and space to process the data.

2. What type of algorithm is required for analyzing streaming data? (0.5 points)
 - A. Fast and Complex
 - B. Accurate and Memory Efficient
 - C. Fast and Simple
 - D. Accurate and Consistent

3. What is lambda architecture? (0.5 points)
 - A. A specific method for processing streaming data using special real time processes.
 - B. A specific hardware architecture for a server made specifically for processing real time data.
 - C. A method to process streaming data by utilizing batch processing and real time processing.

4. Of the following, which best represents the challenge regarding the size and frequency of data? (0.5 points)
 - A. The size and frequency of the streaming data may be sporadic.
 - B. There may not be data to produce the notion of size and frequency.
 - C. The size and frequency of the streaming data may be too small.

5. What is the difference between data lakes and data warehouses? (0.5 points)
 - A. Data lakes contain only files while data warehouses contain only databases.
 - B. Data lakes utilize hierarchical systems while data warehouses use object storage.
 - C. Data lakes house raw data while data warehouses contain pre-formatted data.

6. What is schema-on-read? (0.5 points)
 - A. The process where formatted data is given structure when read.
 - B. The process where data is pre-formatted prior to being read but the schema is loaded on read.
 - C. Another name for data lakes.
 - D. Data is stored as raw data until it is read by an application where the application assigns structure.

7. Which definition best describes a data lake? (0.5 points)
- A. A data warehouse storing data in files or folders
 - B. An information depository supplying concentrated marketing data
 - C. A hierarchical storage unit caching sensitive data.
 - D. A storage repository holding raw data in its native format
8. The desired characteristics of a BDMS include (select four choices): (0.5 points)
- A. Support for ACID
 - B. Continuous data ingestion
 - C. A full query language
 - D. Narrow range of query sizes
 - E. A flexible semi-structured data model
 - F. Support for common "Big Data" data types
9. Fill in the blank with the best answer: CAP theorem states that _____ all at once within a distributed computer system? (0.5 points)
- A. it is necessary to have consistency, availability, and partition tolerance
 - B. it is impossible to have consistency, accuracy, and partial tolerance
 - C. it is necessary to have consistency, accuracy, and partial tolerance
 - D. it is impossible to have consistency, availability, and partition tolerance
10. What is the purpose of the acronym BASE? (0.5 points)
- A. To overcome CAP theorem.
 - B. To impose properties on a BDMS in order to guarantee certain results.
 - C. The same as ACID.
 - D. Enables stricter enforcement of ACID type design.
11. Read Chapter 24.2 in the book "*Fundamentals of Database Systems (7th Edition)*". Describe the difference between the use of the word "**Consistency**" in CAP theorem and its use in ACID in relational databases. (1 points)
12. Define and explain the differences between: Data Warehouse, Data Mart, Data Lake. (2 points)