

**CMT 428 Module 4: Hadoop Ecosystem Tools: quiz 2**

Question	Options	Correct Answer
<b>1. What language does Hive use to query data?</b>	A) Pig Latin B) HiveQL C) SQL D) MapReduce	<b>B) HiveQL</b>
<b>2. Which Hadoop tool is primarily used for data warehousing?</b>	A) HBase B) Flume C) Hive D) Oozie	<b>C) Hive</b>
<b>3. What is Pig Latin in the Hadoop ecosystem?</b>	A) A data warehousing language B) A scripting language for parallel computation C) A language for data ingestion D) A query language like SQL	<b>B) A scripting language for parallel computation</b>
<b>4. Which tool is used for NoSQL data storage in Hadoop?</b>	A) Sqoop B) Hive C) Pig D) HBase	<b>D) HBase</b>
<b>5. Which of the following is a key feature of HBase?</b>	A) Row-based storage B) Column-oriented storage C) SQL support D) Data partitioning	<b>B) Column-oriented storage</b>
<b>6. Which of the following allows for efficient bulk data transfer between Hadoop and relational databases?</b>	A) Hive B) Oozie C) Sqoop D) Pig	<b>C) Sqoop</b>
<b>7. Which tool can be used for log data ingestion from multiple sources into Hadoop?</b>	A) Flume B) Hive C) Pig D) HBase	<b>A) Flume</b>
<b>8. Which component in Flume is responsible for transporting data between sources and sinks?</b>	A) Sources B) Channels C) Agents D) Sinks	<b>B) Channels</b>

Question	Options	Correct Answer
9. What is Oozie used for in the Hadoop ecosystem?	A) Querying large datasets B) Scheduling Hadoop jobs C) Collecting log data D) Data storage	<b>B) Scheduling Hadoop jobs</b>
10. What is a typical use case of Hive in Hadoop?	A) Performing real-time data updates B) Writing complex MapReduce programs C) Querying large datasets using SQL-like queries D) Managing workflows	<b>C) Querying large datasets using SQL-like queries</b>
11. What is a common use of Pig in Hadoop?	A) Real-time data analysis B) Loading, transforming, and analyzing large datasets C) Scheduling Hadoop jobs D) Managing NoSQL databases	<b>B) Loading, transforming, and analyzing large datasets</b>
12. What is an HBase “region”?	A) A set of rows stored in the same file B) A type of query used in Hive C) A collection of columns D) A node in a distributed Flume setup	<b>A) A set of rows stored in the same file</b>
13. Which tool can you use to import a MySQL table into Hadoop?	A) Hive B) Sqoop C) Flume D) HBase	<b>B) Sqoop</b>
14. What is a key feature of Flume?	A) It supports only structured data B) It's a NoSQL storage solution C) It's used for log data collection and aggregation D) It schedules MapReduce workflows	<b>C) It's used for log data collection and aggregation</b>

Question	Options	Correct Answer
15. Which of the following Oozie components is responsible for managing control flow?	A) Action Nodes B) Coordinators C) Control Flow Nodes D) Action Nodes	<b>C) Control Flow Nodes</b>
16. Which tool would you use to perform Extract, Transform, and Load (ETL) tasks in Hadoop?	A) Hive B) Pig C) HBase D) Oozie	<b>B) Pig</b>
17. What is the default storage format used in Hive?	A) JSON B) XML C) ORC D) TextFile	<b>D) TextFile</b>
18. Which of the following operations would Pig be best suited for?	A) Loading structured data from a database B) Joining multiple datasets and performing data transformations C) Scheduling workflows D) Querying data using SQL-like syntax	<b>B) Joining multiple datasets and performing data transformations</b>
19. What type of database is HBase?	A) Relational database B) NoSQL database C) In-memory database D) Object-oriented database	<b>B) NoSQL database</b>
20. What does Sqoop primarily connect to when transferring data?	A) NoSQL databases B) Flat files C) Relational databases D) HDFS	<b>C) Relational databases</b>