

Date
25/08/2023



No. of Questions	23
No. of Pages	4

GENERAL SIR JOHN KOTELAWALA DEFENCE UNIVERSITY
BSc in Applied Data Science Communication
Supplementary Examination – Intake 38
2nd Year 2nd Semester
Advanced SQL and Cloud Databases [LB2224]

There are two sections in this paper.

Part 1 - 20 MCQ Questions

Part 2 - 3 Essay Type Questions

Answer all questions

Duration: 3 hours

Part 1 - (40 Marks)

1. Which of the following is an example of unstructured data?
 - a. Relational database tables
 - b. CSV files
 - c. JSON documents
 - d. Social media posts

2. Which one of the following tasks is the responsibility of an Azure database administrator?
 - a. Writing application code
 - b. Designing user interfaces
 - c. Managing database backups
 - d. Creating marketing campaigns

3. Which single service would you use to implement data pipelines, SQL analytics, and Spark analytics?
 - a. Azure Data Lake
 - b. Azure Databricks
 - c. Azure Stream Analytics
 - d. Azure SQL Data Warehouse

4. Which Azure services can you use to create a pipeline for data ingestion and processing?
 - a. Azure Logic Apps
 - b. Azure Functions
 - c. Azure Data Factory
 - d. Azure App Service

5. Which definition of stream processing is correct?
 - a. Analyzing large volumes of data in real-time batches
 - b. Processing data only when the system is idle
 - c. Handling data at rest in a data warehouse
 - d. Processing and analyzing data as it arrives

6. Which kind of visualization should you use to analyze pass rates for multiple exams over time?
 - a. Pie chart
 - b. Bar chart
 - c. Line chart
 - d. Scatter plot

7. Which service would you use to continually capture data from an IoT Hub, aggregate it over temporal periods, and store results in Azure SQL Database?
 - a. Azure Stream Analytics
 - b. Azure Event Grid
 - c. Azure Logic Apps
 - d. Azure Functions

8. What is an index?
 - a. A backup copy of a database
 - b. A list of reserved keywords in SQL
 - c. A schema for NoSQL databases
 - d. A data structure to optimize data retrieval

9. Which deployment option offers the best compatibility when migrating an existing SQL Server on-premises solution to the cloud?
 - a. Azure SQL Managed Instance
 - b. Azure SQL Database
 - c. Azure Cosmos DB
 - d. Azure Data Warehouse

10. What would be the place to store media content for sharing?
 - a. Azure SQL Database
 - b. Azure Blob Storage
 - c. Azure Cosmos DB
 - d. Azure Data Lake

11. When designing a cloud-based database solution, which factor should you consider for ensuring high availability?
 - a. Redundant on-premises backups
 - b. Limited network connectivity
 - c. Data locality
 - d. Geographic distribution

12. Which service would you use to perform complex transformations and data manipulations in real time on streaming data?
 - a. Azure Logic Apps
 - b. Azure Functions
 - c. Azure Stream Analytics
 - d. Azure Data Factory

13. What does OLAP stand for in the context of databases?
 - a. Online Load Analysis Process
 - b. Online Learning and Processing
 - c. Online Analytical Processing
 - d. Offline Archiving and Logging

14. Which term refers to the process of extracting, transforming, and loading data from various sources into a data warehouse?
 - a. ELT (Extract, Load, Transform)
 - b. ETL (Extract, Transform, Load)
 - c. ETLT (Extract, Transform, Load, Transform)
 - d. ETLA (Extract, Transform, Load, Analyze)

15. In a cloud database, what is horizontal scaling?
 - a. Adding more resources to a single server
 - b. Increasing the capacity of a single machine
 - c. Distributing the data and workload across multiple servers
 - d. Upgrading the software version of the database

16. Which Azure service would you use to analyze large volumes of data using Transact-SQL queries?
- Azure SQL Database
 - Azure Synapse Analytics (formerly SQL Data Warehouse)
 - Azure Cosmos DB
 - Azure Analysis Services
17. What is a data lake in the context of cloud databases?
- A storage solution for structured data only
 - A specialized database for time-series data
 - A storage repository that can store vast amounts of raw data in its native format
 - A high-performance, in-memory database
18. Which Azure service provides a fully managed platform for building, deploying, and scaling web apps?
- Azure App Service
 - Azure Kubernetes Service
 - Azure Functions
 - Azure Service Fabric
19. What is a data mart?
- A secure area of a database accessible only by administrators
 - A database optimized for operational data
 - A subset of a data warehouse focused on a specific business area
 - A schema-less NoSQL database
20. Which cloud deployment model provides dedicated infrastructure for a single organization?
- Public Cloud
 - Private Cloud
 - Hybrid Cloud
 - Community Cloud

Part 2 - (60 Marks)

1. Scenario: Imagine you are working as a data analyst for an e-commerce company. The company wants to analyze its sales data and create interactive reports to track sales performance and trends.
- Identify the data sources that are relevant for tracking sales data. *(2 marks)*
 - Explain the process of connecting Power BI to these data sources. *(4 marks)*
 - Describe the steps you would take to clean and transform the raw sales data. *(4 marks)*
 - Choose two types of visualizations that would be effective for displaying monthly sales trends and top-selling products. Justify your choices. *(4 marks)*
 - Walk through the process of creating these visualizations using Power BI. *(6 marks)*
2. Scenario: You are a database administrator for a healthcare organization. You are tasked with optimizing the organization's patient data management system using advanced T-SQL techniques.

- a. Identify two common performance bottlenecks that can occur in database queries. *(6 marks)*
 - b. Provide an example of creating an index on a specific column and explain the impact on query performance? *(6 marks)*
 - c. Explain the concept of "database normalization" and its significance in relational database design. *(8 marks)*
3. Scenario: You are an IT consultant tasked with helping a small business migrate its data management system to the cloud. The business currently uses an on-premises database to manage customer information, orders, and inventory. They are interested in the benefits of cloud computing and want to explore cloud database options.
- a. Explain three advantages of using cloud computing and cloud databases for this business. *(6 marks)*
 - b. Identify any potential challenges or risks associated with migrating to the cloud and how they can be mitigated. *(6 marks)*
 - c. Outline the steps involved in migrating the existing on-premises database to the chosen cloud provider. *(8 marks)*

-End of Paper-