PDF 8: Azure Synapse End-to-End Design

Subjective Case Study Questions and Answers

Q1. Designing a Data Warehouse
Scenario: A retail company wants to consolidate sales, inventory, and customer data.
Answer: Use **Dedicated SQL Pool** for structured historical data. Design **fact and dimension tables** using Star or Snowflake schema. Partition large fact tables for performance and enable **columnstore indexes**.

Q2. Integrating External Data Sources
Scenario: Data exists in on-prem SQL Server, Azure Blob, and REST APIs.
Answer: Use **Linked Services** in Synapse to connect external sources. Use **Copy Activity** to ingest data, transform with **Mapping Data Flow**, and load into dedicated SQL pool. Ensure security with Managed Identity and encryption.

Q3. Optimizing Query Performance
Scenario: Queries on consolidated sales data are slow.
Answer: Apply **materialized views** for common aggregations, **partition fact tables**, update **statistics**, and consider **result set caching** in Synapse SQL.

Q4. Security and Compliance
Scenario: Sensitive customer data must be protected.
Answer: Enable **TDE**, **column-level encryption**, and **Row-Level Security (RLS)**. Use **Azure AD authentication** and audit access logs for compliance.

MCQs

1. Best schema for analytical queries?
   Answer: B. Star Schema — simple joins and fast aggregations.
2. Partitioning improves?
   Answer: C. Query performance — reduces scanned data.
3. Precomputed aggregations?
   Answer: A. Materialized views — faster analytics.
4. Secure external connections?
   Answer: B. Managed Identity + encryption — authentication and data protection.
5. Protecting sensitive columns?
   Answer: C. Always Encrypted + RLS — ensures privacy.
6. Recommended for large historical data?
   Answer: A. Dedicated SQL Pool — high performance for analytical queries.
7. Optimizing ad-hoc queries?
   Answer: B. Result set caching — reduces repeated computation.
8. Monitoring user activity?
   Answer: C. Audit logs in Synapse — track access and compliance.