PDF 5: Data Integration with Power BI

Subjective Case Study Questions and Answers

Q1. Connecting Power BI to Synapse
Scenario: Analysts want real-time dashboards on Synapse data.
Answer: Use DirectQuery connection for near real-time data without moving data. Ensure Synapse performance optimization to handle queries efficiently. Use aggregations in Power BI for faster reporting.

Q2. Row-Level Security in Power BI
Scenario: Different regional managers should see only their region's data.
Answer: Implement RLS in Power BI dataset using roles. Map users to roles via Azure AD groups. Test with multiple user accounts to validate security.

Q3. Refreshing Large Datasets
Scenario: Daily refresh of 100M+ rows causes timeout.
Answer: Incremental refresh in Power BI datasets, partition tables in Synapse, use query folding, and reduce unnecessary columns. Schedule refresh during off-peak hours.

Q4. Integrating Multiple Sources
Scenario: Dashboard combines SQL, Excel, and REST API data.
Answer: Use Power Query for ETL, merge datasets using appropriate joins, transform and clean data before loading into data model. Optimize model size by removing unused columns and summarizing data.

MCQs

1. DirectQuery vs Import mode in Power BI?
   Answer: A. DirectQuery — real-time queries; Import — cached data.
2. Implementing RLS requires?
   Answer: B. Roles and Azure AD user mapping.
3. Handling large dataset refresh?
   Answer: C. Incremental refresh — improves efficiency.
4. Power BI data source combining method?
   Answer: D. Power Query — merge and transform.
5. Aggregations in Power BI improve?
   Answer: A. Query performance — precomputed summaries.
6. Testing RLS roles?
   Answer: B. View as role — verify access.
7. Reducing dataset size?
   Answer: C. Remove unused columns + pre-aggregate data.
8. Schedule refresh for large datasets?
   Answer: B. Off-peak hours — reduces resource contention.