Q76. Design anomaly detection.

I would use statistical thresholds to detect deviations.

Q77. Design error dashboards.

I would visualize failures and trends.

Q78. Metrics vs logs vs traces.

I use metrics for health, logs for debugging, traces for flow.

Q79. Debug production data issues.

I would trace lineage and replay data.

Q80. Design self-healing pipelines.

I would automate retries and recovery.

Q81. Design RBAC for data.

I would assign role-based permissions.

Q82. Design column-level security.

I would restrict sensitive columns.

Q83. Design encryption strategy.

I would encrypt data at rest and in transit.

Q84. Secure data ingestion.

I would authenticate producers and validate payloads.

Q85. Cross-region access design.

I would replicate data and control access.

Q86. Secret management.

I would store secrets in a secure vault.

Q87. Data masking system.

I would mask sensitive fields dynamically.

Q88. Audit logging.

I would log all access and changes.

Q89. Secure API access.

I would use authentication and rate limiting.

Q90. Tenant isolation.

I would isolate data logically and enforce quotas.

Q91. Metadata-driven platform.

I would drive pipelines via configuration and metadata.

Q92. Self-service analytics platform.

I would provide governed access and tools.

Q93. Feature store design.

I would centralize feature definitions with consistency.

Q94. Real-time recommendation pipeline.

I would combine streaming features with online serving.

Q95. CDC pipeline design.

I would capture changes and stream them reliably.

Q96. Multi-region replication.

I would replicate asynchronously for resilience.

Q97. Event replay without duplication.

I would use idempotent processing.

Q98. Stream-batch unification.

I would reuse streaming pipelines for batch via replay.

Q99. ML data pipeline design.

I would ensure reproducibility and freshness.

Q100. Company-wide data platform.

I would standardize tooling, governance, and best practices.