Database Breakdown – Unoptimized Data Model & Critical Size

Context:

Your team manages a high-traffic e-commerce platform that relies on a central relational database. Over time, exponential data growth and a lack of proactive data management have caused the database to reach a critical size. As a result, performance issues have surfaced, including slow query responses, frequent timeouts, and connection failures—all negatively impacting customer experience and revenue.

Details:

- The database suffers from structural inefficiencies such as poor normalization, redundant data, and missing indexes.
- Slow queries and high resource usage are common, especially during peak periods.
- A lack of an archival strategy means historical data continues to burden the live system, exacerbating scalability issues.
- Implementing immediate fixes without a thorough plan risks compromising data integrity or prolonging downtime.

Challenges:

- Diagnose the root causes of performance degradation linked to the unoptimized data model and excessive database size.
- Develop solutions that offer immediate performance relief while planning for a sustainable, long-term data model refactoring.
- Communicate a clear recovery plan to both technical teams and non-technical stakeholders, ensuring minimal downtime and preserving data integrity.

Tasks:

- **Root cause analysis:** Determine which KPIs will effectively reflect database performance and scalability. Justify the selection of these metrics by explaining how they relate to the identified issues..
- **Evaluate the Data Model:** Evaluate the current data model to discover the actual sources of the performance degradation—such as poor schema design, missing indexes, or inefficient queries—that lead to the critical database size.
- **Propose Immediate (Short-Term) Fixes:**

  Suggest interventions that provide rapid relief, such as temporary indexing, query optimization, and selective data archival. Explain why these fixes will reduce current performance bottlenecks.
- **Develop a Medium-Term Strategy:**

  Outline a comprehensive plan for refactoring the data model, implementing partitioning, and establishing an archival strategy. Include a phased timeline that minimizes downtime while ensuring long-term sustainability.
- **Draft a Communication Plan:**

  Prepare concise messaging that explains the diagnosis, the impact on the project's future, and the detailed short- and medium-term actions. Ensure every point can be justified, supported, and defended during discussions with the evaluator(s).

Organization example

Day 1: Kick-Off & Initial Analysis

- **Morning:**
  - Introduce the workshop objectives and the scenario, emphasizing the lack of existing logs.
  - Discuss the importance of defining meaningful performance metrics and collaboratively select KPIs (e.g., query response time, transaction throughput, resource utilization).

  Afternoon:

- Plan how to collect these metrics using available monitoring tools or test queries.
- Document the chosen metrics and set clear objectives for data collection and analysis.

Day 2: Deep Dive Analysis & Brainstorming

- **Morning:**
  - With the defined metrics in place, simulate or gather initial performance data.
  - Analyze the existing schema to understand structural inefficiencies and potential redundancy issues.
- **Afternoon:**
  - Conduct a brainstorming session to propose immediate (short-term) technical fixes.
  - Document potential solutions, discussing the expected improvements based on the defined metrics.

Day 3: Strategy Development & Roadmap Formulation

Morning:

- Consolidate the findings from Days 1 and 2 into a unified recovery plan.
- Clearly identify the actual source(s) of the performance issues using the collected data and schema analysis.

Afternoon:

- Develop a detailed roadmap that outlines:
  - **Short-Term Actions:** Immediate fixes (e.g., indexing, query optimization, selective data archival) and their expected impact.
  - **Medium-Term Actions:** A phased strategy for data model refactoring, partitioning, and comprehensive archival.
- Quantify the impact of these issues on the future of the project, supporting your findings with the defined metrics.

Day 4: Refinement & Presentation Preparation

- **Morning:**
  - Finalize the recovery and optimization plan, ensuring all technical and communication strategies are clearly documented.
  - Perform a review to verify that proposed changes maintain data integrity and system continuity.
- **Afternoon:**
  - Rehearse the final 5-minute presentation, ensuring that the narrative is clear, concise, and well-coordinated across the team.
    - The actual source(s) of the problem.
    - A measurement of the impact on the future of the project.
    - Clearly defined short-term and medium-term actions.
  - Make final adjustments to the presentation materials to ensure every point is justified, supported, and defensible.

Day 5: Final 5-Minute Presentation

Each group delivers a concise 5-minute presentation that must include:

- **Identification of the Actual Source(s) of the Problem:**

  Clearly articulate which aspects of the data model (e.g., poor normalization, missing indexes) are causing performance degradation, supported by the defined metrics.

- **Measurement of the Impact on the Future of the Project:**

  Quantify the potential revenue loss, increased operational costs, or customer churn resulting from the performance issues.

- **Short-Term Actions:**

  Identify and define immediate interventions (e.g., adding essential indexes, optimizing queries) that will quickly relieve the current pressure on the system.

- **Medium-Term Actions:**

Detail a strategic roadmap for a complete data model refactoring, including partitioning strategies and a phased archival process to manage future growth.

- **Justification and Defensibility:**

  Ensure that every point presented is well-justified, supported by the established metrics, and can be defended in discussions with the evaluator(s).