**GROUP 4 PROJECT PROPOSAL**

Option 10: Multifaceted Data Analysis on a Unified Dataset

**Project Overview:** To understand the application of different database technologies: SQL (for structured data), MongoDB (document-oriented), and Neo4j (graph-based). The approach is to using a single, multifaceted dataset across these diverse database systems, analyzing their performance and flexibility in handling data beyond their typical use cases.

**Objective:**

* To explore and understand the capabilities of SQL, MongoDB, and Neo4j in handling a unified dataset that includes structured, semi-structured, and graph-oriented data
* Evaluate the performance of each database in managing data types and queries outside their standard domain

**Project Components:**

1. **Dataset Selection:**
   * Considering the given condition of using a single dataset to conduct the analysis, we plan to choose a complex dataset that encompasses elements suitable for all three database types: structured data for SQL, document-oriented data for MongoDB, and relationship-focused data for Neo4j
2. **Task Design:**
   * Conduct benchmark tests on primary functionalities like Create, Read, Update and Delete (CRUD) operations.
   * Develop a range of tasks applicable to all three databases, such as data retrieval, aggregation, and relationship analysis
3. **Hybrid Data Challenge:**
   * Challenge each database to handle data types and queries typically suited for the others. For instance, use SQL for document-oriented queries, MongoDB for relational-style queries, and Neo4j for structured data queries
   * Create a unified query set that can be executed on all three databases to retrieve comparable information, requiring creative adaptation of query structures
4. **Performance Metrics:**
   * Measure execution time, accuracy, and completeness of the results from each database
   * Monitor and compare resource utilization (CPU, memory, storage) during query execution
5. **Implementation:**
   * Set up each database in a similar environment for consistent testing conditions
   * Format and import the dataset into each database, ensuring a consistent data schema across them
6. **Comparative Analysis:**
   * Perform a performance comparison to assess each database's efficacy in handling non-native tasks
   * Analyze and document insights into the flexibility and limitations of each database system

**Expected Outcome**

* Practical experience working with multi-model databases and in-depth insights on strengths and weaknesses of each database type.
* Gain a comprehensive understanding of the strengths and limitations of different database models in processing various data types
* Develop hands-on experience in optimizing and adapting queries for different database systems.
* Enhance problem-solving skills by creatively tackling the challenge of making each database perform tasks outside its standard use-case.