Project Statement for Milestone 4

You team name
(Team members' names)

Overview:

At the end of Milestone 4 of the project, teams should have developed a prototype of an end-to-end application with a graphical user interface. Teams should have integrated the graphical user interface with a NoSQL database and Hadoop/Spark framework. In cases where teams are focused on a research topic, they should have validated their findings and compared them against the existing methods.

Teams should have updated their database schema and algorithms for robustness and scalability, and should have validated the scalability of their solution using a large dataset on distributed instances of a NoSQL database and Hadoop/Spark.

All team members are expected to make a significant contribution to the project milestone tasks.

Project Report Topics:

The report should cover the following subtopics and answer the questions listed:
1. User Interface and Data Visualization:
    a. Describe the user interface and data visualization components of the software. The user interface should be interactive, taking inputs from a user and providing outputs, i.e., results, to the user. The user interface should include data visualization.

2. User Queries and Results:
    a. Provide all user queries and their results to describe the overall functionality of the software.

3. Scalability:
    a. Describe how data is stored in a cluster configuration of a NoSQL database. Provide performance benchmarks for data ingestion/query in a cluster deployment. How do the results compare to the performance benchmarks on the reduced dataset or non-cluster deployment?
    b. Describe how the algorithms are run on a cluster configuration of Hadoop/Spark. Provide performance benchmarks for algorithm in a cluster deployment. How do the results compare to the performance benchmarks on the reduced dataset or non-cluster deployment?
    c. What hardware (computers with CPU, memory and disk storage) did the team use to test the scalability of the solution?

4. Source Code:

a. Provide the source code of your application prototype, including user interface, data ingestion, data query, and analytics algorithms, in a ZIP file. Make sure to **exclude** the dataset.

Peer Evaluation:

Each team members should complete the CATME Peer Evaluation survey.

Grading:

- 50 pts: The team has successfully completed all the tasks described in this milestone template document. The team has successfully implemented an end-to-end application (or a research solution) and has demonstrated the scalability of the solution. The team has provided all relevant information, including descriptions, pseudo-code / code snippets, and diagrams / images.