CSCI 5408

DATA MANAGEMENT AND WAREHOUSING

GROUP PROJECT
SPRINT -1

GitLab Link: https://git.cs.dal.ca/patel38/csci_5408_s24_3

DATA STRUCUTURES:

In our code, we use a combination of a LinkedHashMap and a List for data storage. The reason for using LinkedHashMap is to ensure that the data is stored in a linear order and maintains the appending order when retrieving the data.

Our main data structure is Map<String, LinkedHashMap<String, List<String>>>, where:

- The first key represents the table name.
- The LinkedHashMap stores the columns, with the inner key representing the column name and the associated List<String> holding the column values.

For the table schema, we use List<Map<String, String>>, where:

- The key represents the column name.
- The value represents the schema of the column.

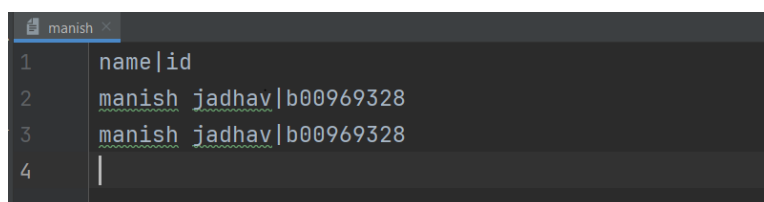This structure is intended for temporary storage.

For persistent storage, we use a file directory system:

- The main folder is named "TinyDB".
- Each database created will be a subfolder within "TinyDB".
- Each table will be stored as a text file within its respective database subfolder, along with a schema file.

CUSTOM FILE FORMAT:

For storing the information of the databases we use .txt files. We use the pipe symbol (|) to differentiate between the table's columns. The reason for choosing this operator is very rarely used in terms of inserting data.

The same operator is used to differentiate between column names as well as the data of those columns in rows.

Figure 1: Custom file format for data in a table.