**CHALLENGE 12**

**REPORT**

**Introduction**:

This study is regarding data operations for data with flexible structure. The data is in JSON format rather than in a tabular form as in SQL. One of the most popular databases structured in this non-SQL fashion is called MongoDB.

In this study, data is retrieved from such a database, and analyzed using the library pyMongo.

**Data**:

**Part 1**

A database called ‘uk\_food’ includes a collection of documents named ‘establishments’.

The objective is to add data for a new establishment, update and modify data for further analysis.

Using the class ‘MongoClient’, a connection is established with the database uk\_food. A new restaurant called Penang Flavours in Greenwich, London is added to the database with the following information:

A screenshot of a computer code

Description automatically generated

Data from other similar establishments in the category "Restaurant/Cafe/Canteen" was used to determine the “Business Type ID” as 1.

Data from another county, ‘Dover’ is eliminated, as it is not relevant to the analysis of restaurants in Greenwich.

The ratings data, ‘Rating Value’ is changed from string to integers 1 through 5, or ‘None’, in case a number between 1 and 5 is not listed.

The longitude and latitude data, listed as string is converted to float.

The above manipulation of data is in preparation for analysis in Part 2.

**Part 2**

The database ‘uk\_food’ updated in Part 1 is used for analysis in Part 2.

In the area around the new restaurant, Penang Flavours, the collection ‘establishments’ is used to find other establishments related to food service, perhaps to assess the competition factor.

There are 145 establishments within ±0.01° of the location of Penang Flavours.

Number of establishments with ‘Hygiene score’ of 0 (55) is more than those with the score of 20 (41).

Under the ‘Local Authority Name’ that includes the string ‘London’, there are only 33 establishments with rating value ≥ 4.

All results are converted to pandas DataFrame for ease of readability in a tabular form.

**Conclusion**:

This study is to:

* Understand how to access and use MongoDB to add, delete, manipulate, or modify data to the user’s advantage.
* Gain insight into the flexible method of data-storage, in situations that do not require the rigid structure of SQL databases.