Project Description (deadline 22 Jan 2023)

This is a group project with the number of students in a group cannot more than three (3). You are freely to pick the member in the group.

You are required to develop an appropriate relational database for a business data and performing queries that can answer some business-related analysis. Include the following in your report

1) Select your data

Find any business data from the internet. Examine the data to ensure you could possibly create queries to obtain some statistical metric from the data (i.e: the mean, median, average value). Size of data must be more that 10,000 rows (each row considered 1 record). (you may use more than 1 data set). State the source of the data.

2) Create a scenario for the data **(5 marks)**

Create a possible scenario that can reflect the data. Explain the attribute of the data or how you combine more than 1 set of data. For example, if the data is on e-commerce, first give some background to the company/enterprise; the business they are doing, the department, branches they have and etc. Explain what sort of data use in the business (i.e: customer data like name, address etc; sales data, item sale, item in stock, supplier data etc). Identify some insight/metrics for examples: -

i) Average sale of item in certain month, year
ii) Department/Item that have maximum sale
iii) Supplier performance
iv) The slow-moving item

any many more. **(10 marks)**

3) Define possible relations

Based on the scenario given in (2) identify entities (relation) and the relationship between relation. Identify the primary and foreign keys. The minimum number of relations in the database is five (5). **(10 marks)**

4) SQL queries

Perform queries to obtain the metric or data identified in (2). Include queries and the output from the queries in the report. The quires must involve multiple tables. Include index in certain data to improve the performance of the query. **(15 marks)**

The report must include all item (1), (2), (3) and (4).