

A new client named ABC Ltd approached your company to do a small data analytics project.

Scope:

ABC Ltd is a food and beverage distributor. His business model is to sell to his clients through the agents who get paid mainly by commission.

Every agent has a commission percentage based on the product category. This means that a specific Agent will have different commissions based on the product category and every product category will have a commission assigned to it based on the agent selling the category.

Currently the commission rate for this matrix is being worked out and applied manually. In the next phase of the project, based on the information gathered by this phase (Phase 1), a recommendation engine needs to be implemented to make sure that the client is giving the agents the optimal rate per product category.

The main aim of phase one is to deliver a POC whereby a flat table with all the consolidated data is returned and sample reports using a querying language is returned.

Sample Reports to be Created:

- Commission per Agent
- Commission per Product Category
- Number of Sales per Agent
- Number of Sales per Product
- A matrix of the number of sales per agent per product category. (Agent at the row level, product category at the column level.)
- A matrix of the commission value sales per agent per product category. (Agent at the row level, product category at the column level.)

Following the business analysis, the following requirements were gathered:

- Data is currently in one excel file which is split into 3 sheets.
- Client is providing a file named Test Data.xlsx with a sample dataset of the source.
- Client is expecting to receive big volumes of data. This means that the queries for the data cleaning and data transformation must be optimized so that no rework for optimization is done.
- Additional Products, Product Categories and Agents can be added in the future. No changes in the algorithm are to be done when the client wants to add additional Products, Product Categories and Agents.
- Client is technical and will analyse the solution in detail.
- Client is not interested in seeing the transformed data stored in a table. He just wants to assess the transformation skills and the querying skills to create the reports mentioned to trust you with bigger projects. This means that no data should be written to a table/file.