# Task 1 (Explanation):

Keeping in mind the requirements provided by the client, a start schema is created. This start schema includes one fact table and five dimension-tables. The fact table is named as Sales whereas the dimension tables are Clerk, Sales Branch, Products, Supplier and Time. The granularity of the fact table data is kept being one record of one sales date. Similarly, one branch from one supplier by one clerk. The time dimension will have years, month, week, and day.

There were other alternatives for fulfilling the same requirements as put forward by the client. A major one of them would be to simply use the ETL process to collect the data as it is and put it in physical tables using aggregation. In case of this approach the process would have been become very simple. But the main issue with this approach is that, according to the given conditions and requirements the business under consideration is a fairly large-scale business and offers a variety of items to its customers. So, it is safe to the say that the data volume will be very high which will make the final data warehouse very large in terms of storage space. In this way, the redundancy will increase significantly due to which the requirement for the storage space will also increase and in turn will also require a lot of cost to fulfill that need of storage space.

So, to protect the customer from these unnecessary costs, the granularity of the data is kept low. The aggregation of the data will be done according to the requirements of the client depending on the frequency of refreshing of data that is required. Views or Materialized views will be used for this purpose.

# Task 2 (Star Schema):

