Read the following case and answer the questions:

Kizoka Building and Construction Company (KBCC) requests you (Data mining expert) to design and implement a data warehouse that will enable a company to analyse business data with respect to items, manufacturer, time and location. The company is based in Tanzania and dealing with selling items. The company has three stores located in three regions (Dar es Salaam, Arusha and Mwanza). Moreover, the company need to analyse.

- Its items with respect to their names, type, size and category.
- Manufactures with respect to their names, items and location.

The company is interested in analysing and learning at least the quantity and income of its sales.

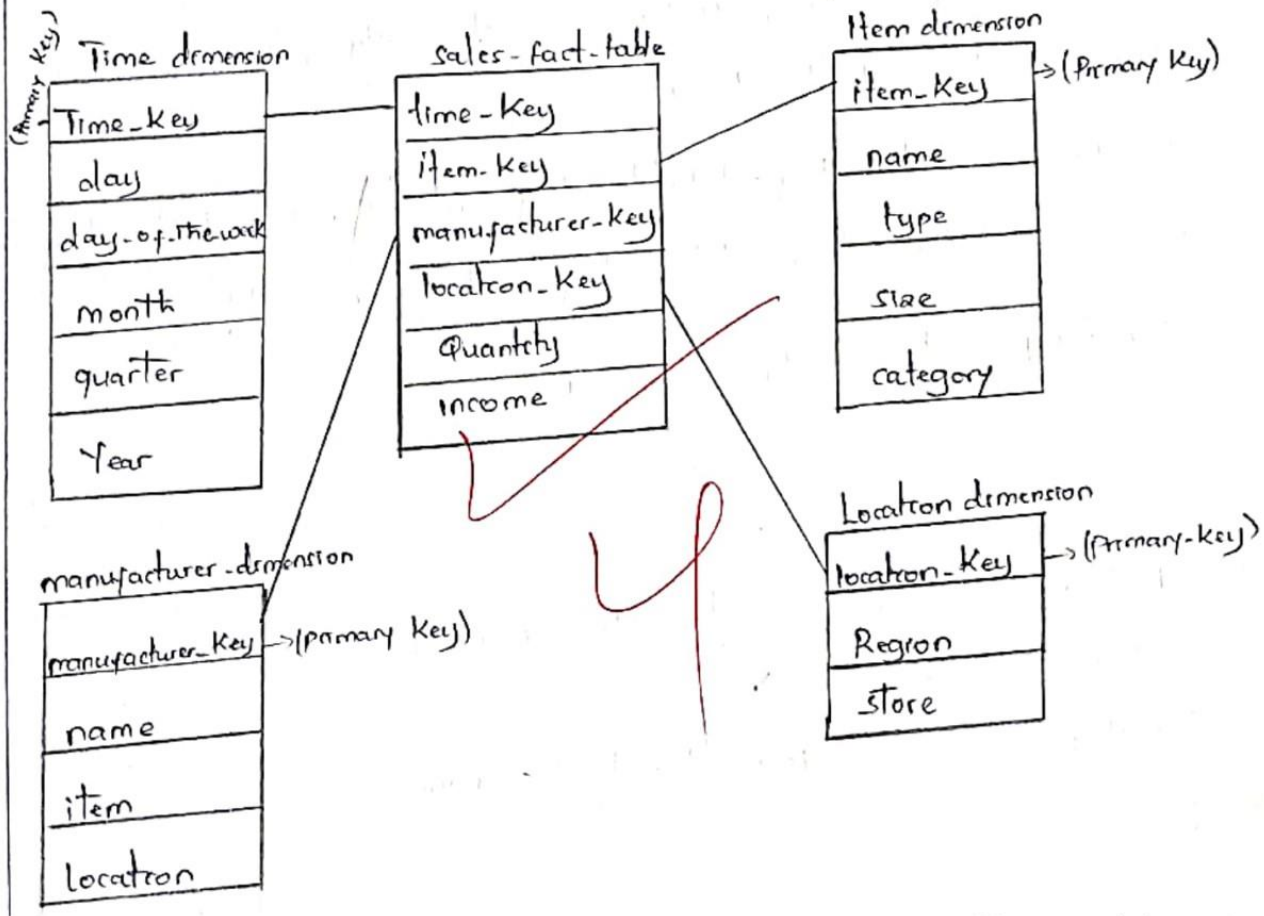
Questions:

1) List steps for Data warehouse design. **(2 marks)**
2) Draw appropriate schema for this data warehouse. State any assumptions you have made. **(5 marks)**
3) Define a schema in (2) using DMQL **(3 marks)**
4) Translate DMQL in (3) to SQL. **(3 marks)**
5) Write the SQL query that finds the average quantity and total income in 2019 with respect to each store, item type and year. **(3 marks)**
6) *OLAP operations:*
   a) What specific *OLAP operation(s)* should be performed in order to list the total income for each item in 2013? **(2 marks)**
   b) What specific *OLAP operation(s)* should be performed in order to list the number of sold Twiga cements in July 2018 in Arusha. **(1 mark)**
   c) What specific *OLAP operation(s)* should be performed in order to list the number of sold doors and bolts in Dar es Salaam and Mwanza. **(1 mark)**

Qn 1. The steps for datawarehouse design are:-

i) Understand the business process
ii) Identify the grain
iii) Identify the facts
iv) Identify the dimensions.

Qn 2. Assume the schema used is a star-schema.



Qn 3. define cube sales_star [Time, Item, manufacturer, location]: Quantity = count (*)
income = sum (income)

define dimension time as (time_key, day, day-of-the-week, month, quarter, year)
define dimension Item as (item_key, name, type, size, category)
define dimension manufacturer as (manufacturer_key, name, item, location)
define dimension location as (location_key, region, store)

Qn4. SELECT s.time_key, s.item_key, s.manufacturer_key, s.location_key, sum (income), count (quantity)

FROM time t, item i, manufacturer m, location l

WHERE s.time_key = t.time_key and s.item_key = i.item_key and s.manufacture m.manufacturer_key and s.location_key = l.location_key

GROUP BY s.time_key, s.item_key, s.manufacturer_key, s.location_key

Qn.5. SELECT Avg (quantity), Sum (income), store, type, year

From sales s, location l, item i, time t

WHERE s.location_key = l.location_key and s.item_key = i.item_key and s.time_key = t.time_key and year = "2019"

GROUP BY store, type, year

Qn 6 a) - Roll up for time from day to year
- slice for year = "2013"
- Roll up for item from item_key to name.

b) - Roll up for time from day to year
- slice for year = "2018"
- Roll up for item from item_key to type.

c) Dice for type = "doors" or "bolts" and Region = "Dar es salaam" or "Mwanza"