IST722: Unit 04 Participation Questions

This is an individual assignment.

Before you begin, please make sure you've read and understand 1) our class honor code, 2) course policies on late work and 3) participation policies as posted on the syllabus. "I didn't know" is not an excuse.

You should cite your sources in a standard format like MPA or APA and include a list of works cited.

Your Name:	Akshay Bhala
Your Email:	abhala@syr.edu

Instructions

Answer each of the following questions as concisely as possible. More is not necessarily better. Please justify your answer by citing your sources from the assigned readings from our textbooks, our class lectures, or online if directed to do so. Be sure to cite in text and include a list of works cited. Place your answer below each question. When you're finished, print out this document and bring it to class as part of your participation grade.

Questions

[1] Explain a) conformed dimensions b) role-play dimensions. Give examples.

a. **Conformed Dimensions (I am an octopus with eight links to different fact table):** A conformed dimension allow a consistent analysis of different business process metrics used across multiple dimensional models. A product dimension may join to the inventory, sales, procurement, and other fact tables. It is important to closely collaborate with data governance team to have a consistent definition for this dimension.

b. **Role-Play Dimensions:** A role playing dimension is a dimension that can be aliased multiple times to fulfill analysis in different context. The Time dimension is a role-playing dimension. Suppose the fact table has an order date, shipment date, and delivery date keys. We can create three aliases of the Time dimension (three roles) in the semantic layer of the reporting/OLAP tool and name them as Order time dimension, Shipment time dimension and delivery time dimension. The key is to remember that we do not create multiple physical copies of the time dimension table but only create aliases in the logical layer.

[2] Explain (SCD) Slowly Changing Dimensions. Give examples.

Slow Changing Dimensions (SCD): A slowly changing dimension as the name suggest, has attributes that do not change frequently but they do change once in a while. A slowly changing dimension according to business requirements can be tracked for historic changes. Notably, the

Type 1 (Overwrite/update the attribute), Type 2 (Insert the new record with current status) and Type 3 (Maintain the current value column) are the most common slowly changing dimension scenarios

[3] Explain (RCD) Rapidly Changing Dimensions. Give examples.

Rapidly Changing Dimensions (RCD): Rapidly Changing Dimensions are dimensions which change rapidly, and this rapid change can affect maintenance and performance. The changes are not consistent. Example- person wearing different shirts every day.

[4] Explain a) junk dimensions b) degenerate dimensions c) mini-dimensions d) factless facts. Give examples.

a. **Junk Dimensions:** There are some attributes that have really low cardinality (distinct possible values). Neither this attributes specifically belong to a dimension table nor they are important enough to create an individual dimension for them. A junk dimension creates combinations of all the distinct values of such attributes and stores it in a single junk dimension table. The surrogate keys for these rows are inserted in the fact table. So now the fact table instead of having multiple foreign keys for each of these attributes, has a single foreign key for each row in the junk dimension table. Example-: Status (pass, fail) Income category (lower, middle, high)

b. **Degenerate Dimensions:** A degenerate dimension is a special dimension like invoice number, check number that is an identifier for a transaction. However, it does not have attributes linked to it as all the important attributes are already part of their respective dimensions. So an invoice may have a customer name attribute but it's already a part of the customer dimension. A degenerate dimension is a part of the fact table but it's not a measure, it is still a dimension. Example- Order Number, Flight Number.

c. **Mini-Dimensions:** Suppose the customer dimension has a million rows. Every year when the age of the customer changes we add one more row to reflect this change and preserve history in accordance with the Type 2 SCD, this adds another 1 million records to the table. If we talk about the changes to the income attribute, then millions of more records will be added. We want to avoid such large dimension tables as they are an overhead to the query performance and filtering. A solution is to create a separate dimension called as Customer Demographics dimension and insert its surrogate key in the fact table. This customer demographic table is called as a mini dimension table.

d. **Factless Facts:** A factless fact table is fact table that does not contain fact. They contain only dimesional keys and it captures events that happen only at information level but not included in the calculations level. just an information about an event that happen over a period. A factless fact table captures the many-to-many relationships between dimensions, but contains no numeric or textual facts. They are often used to record events or coverage information. Common examples of factless fact tables include:

Identifying product promotion events (to determine promoted products that didn’t sell)

Tracking student attendance or registration events

WORKS CITED:

Classroom Discussions

https://www.linkedin.com/pulse/different-types-dimensions-data-warehouse-amit-deshmukh/

Professor Fudge Videos

Professor Humayun Explanations