IST722: Class Exercise 4

**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: | Jaya Varshini Prabakar |
| Your Email: | jprabaka@syr.edu |

# Instructions (Refer Units 4 & 5)

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] Summarize (SCD) Slowly Changing Dimensions and (RCD) Rapidly Changing Dimensions. Give examples.

**Slowly Changing Dimensions (SCD):** Slowly Changing Dimensions are the attributes in Data warehouse that experience slow changes. These are required to maintain historical data.

There are three types of SCD:

**Type1: Overwrite the existing attribute.**

**Type2: Add a new dimension row.**

**Type3: Add a new dimension attribute.**

**Example: Customer changing address.** In this case we can use Type2 SCD where we add inactive row for old address and an active row for new address.

**Rapidly Changing Dimensions:** RCD refers to attributes that experience frequent and often change with consistency.

**Example: Age in years** it changes every year with consistency. Another example **is Product Price.**

[2] Explain a) conformed dimensions b) role-play dimensions c) degenerate dimensions d) fact less facts. Give examples.

**a) Conformed Dimensions: Conformed Dimensions** are the master or common reference dimensions. These are shared across business processes. It also helps in reducing time in developing star schema. This is a key part of enterprise bus technical architecture.

**Types: Identical dimensions, subset.**

**Example: Dimension Product can be shared between Fact Sales and Fact Inventory snapshot.**

**b) Role-play Dimensions:** In role-play dimension the same physical dimension plays more than one logical dimension role. This is common among date dimensions. This is stored in the same physical table and aliased as a view. And this can be implemented as multiple foreign keys in fact table to the same dimension table.

**Example: Date** Can be implemented as **Order date, shipping date and delivery date.**

**c) Degenerate Dimensions:** Degenerate dimensions are dimension attributes we store in fact table that are not connected to s separate dimension table. They are used in scenarios where there is low-cardinality information, or where attributes change too frequently to remain in their own dimension or there is too many of them for their dimension.

**Example:** Order number, flight number, customer age.

**Fact less Facts:** These comes into picture when an event doesn’t have a quantifiable fact that the event occurred.

**Example: Student attendance** may have dimensions like date, student and class but would not contain any measures like attendance percentage or hours attended. It is a business process that doesn’t generate any quantifiable measures.

[3] What is the best choice for PK in a Dimension table? What is the best choice for PK in a Fact table?

The best choice for PK in a **Fact** table is composite key or it should be made up of FKs and degenerate dimensions.

For **Dimension** table it is a better practice to use surrogate key as PK.

[4] What are Database Schemas? How are they useful? Give examples.

Database Schema is a logical structure that defines the relationship of data within a database. It is a blueprint for how a database is designed and helps to maintain data integrity, security and consistency.

[5] What are Conceptual, Logical and Physical Models in the data warehouse context? Give examples of each.

**Conceptual Models:** It is the highest level and focus on representing the overall business requirements and concepts of data warehouse.

Example: For a retail company, Conceptual Model might include entities like Customer, Product, Order and Sales

**Logical Model:** It adds more detail to Conceptual Model by specifying logical structure of the data warehouse. It defines the business semantics of data.

Example: For Customer, CustomerID, Name, Email, Address.

**Physical Model:** At the most granular level of detail, the Physical Model represents the concrete implementation of the data warehouse on a specific database platform, outlining the storage, retrieval, and performance optimization strategies for the data.

Example: It provides details about database schema, indexes, partitions, data types and storage options for each table.

WORKS CITED: Professor Michael A Fudge. Dimensional Modeling [Video file].