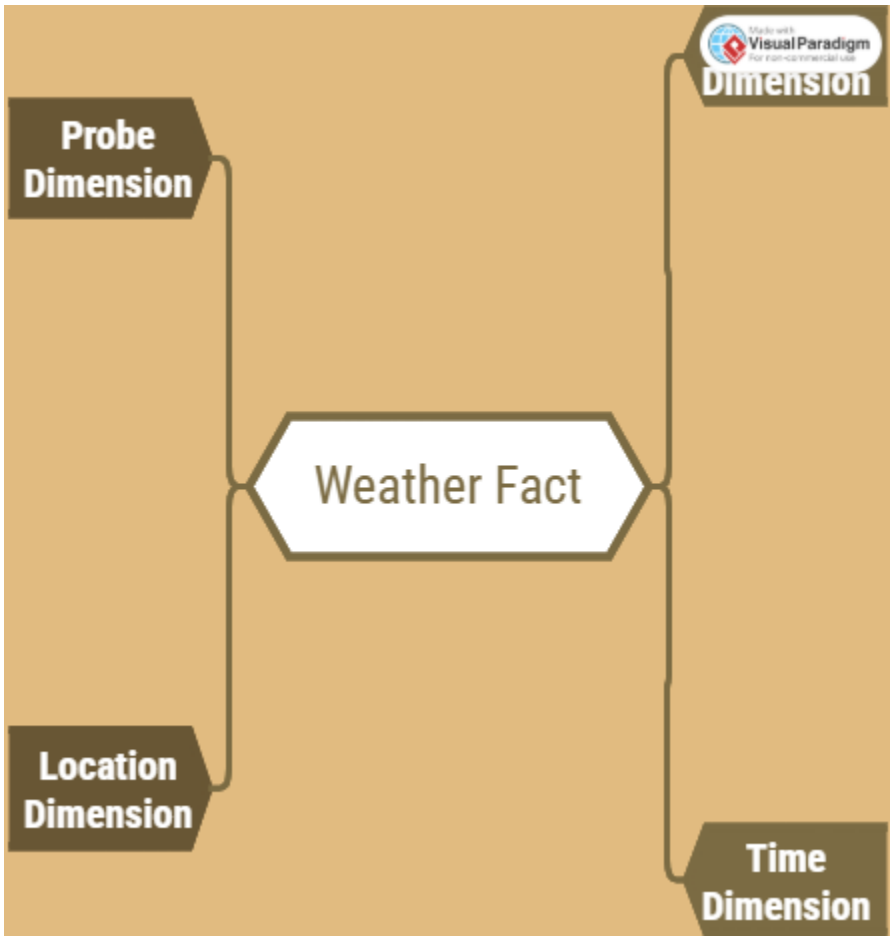
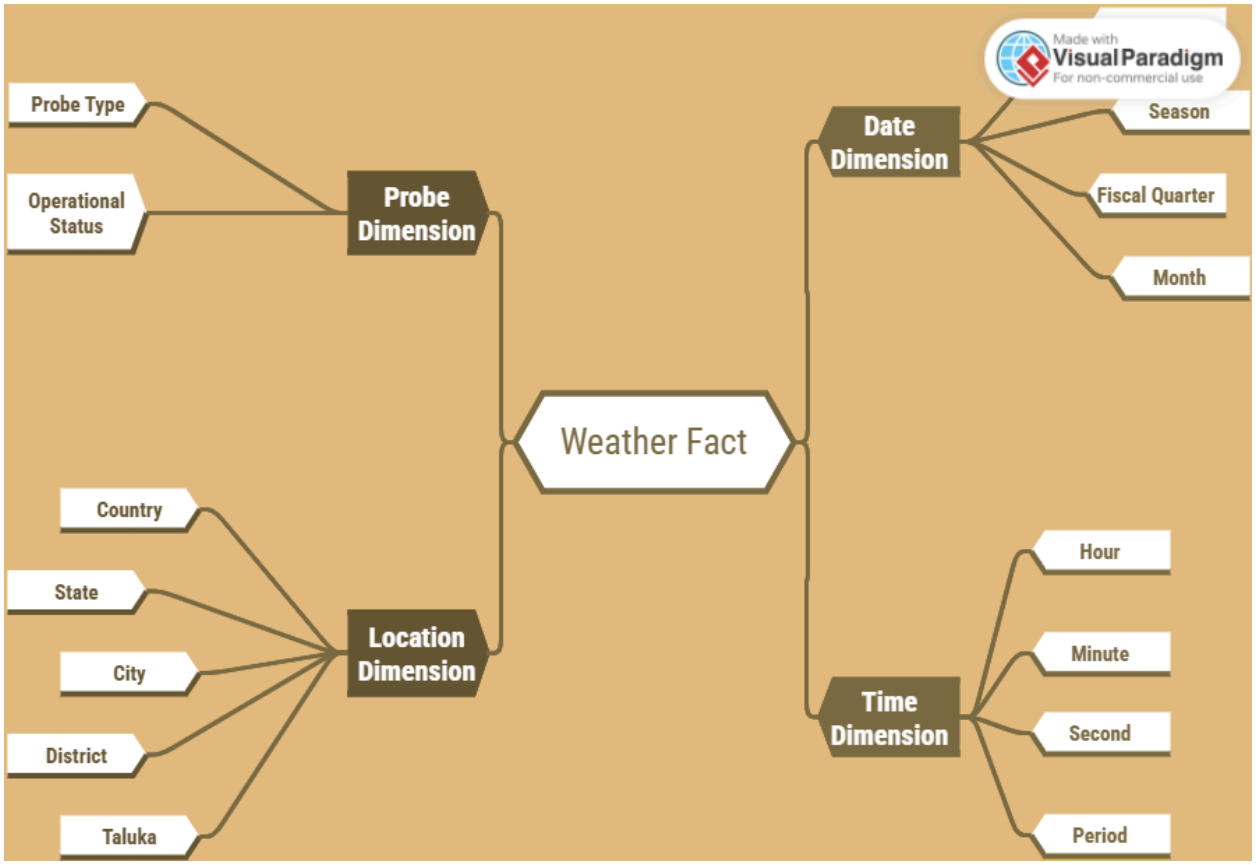
Name: Krish Nisar
Class: TY03 B Batch
Roll No. : 41
Subject: Data Warehouse and Mining

Experiment 1: Design a data warehouse for a regional weather bureau. The weather bureau has about 100 probs, which are scattered throughout various land and ocean locations in the region to collect basic weather data, including air pressure, temperature and precipitation at each hour. All data are sent to the central station, which has collected such data for more than 10 years. Design Star schema and Snowflake schema such that it should facilitate efficient querying and online analytical processing and derive general weather patterns in multidimensional space. Explain all aspects of the diagram. Design Star and Snowflake schema for above case.

Star Schema:



Snowflake Schema:



Review Questions:

Q1) In a star schema, how is the fact table typically related to the dimension tables?

Ans: Fact Table Relationship in Star Schema:

1. The fact table is centrally located and contains measurable data (e.g., sales, revenue, quantity)
2. It is directly connected to multiple dimension tables through foreign keys.
3. Each dimension table contains descriptive attributes (e.g., product details, customer info, time periods).
4. This structure forms a star-like shape when visualized.

Q2) What is the main difference between a star schema and a snowflake schema?

Ans:

Feature	Star Schema	Snowflake Schema
Structure	Central fact table with directly connected dimension tables	Central fact table with normalized dimension tables (split into sub-tables)
Normalization	Denormalized (data redundancy)	Normalized (less redundancy, more joins)
Query Performance	Faster due to fewer joins	Slower due to multiple joins
Storage Efficiency	Requires more storage (redundant data)	Requires less storage (eliminates redundancy)
Ease of Use	Simple to understand and query	More complex due to normalization
Use Case	Best for **OLAP** systems needing quick analytics	Useful when storage optimization is a priority