Concepts from Power BI Session 03

Primary Key vs Foreign Key:

Primary Key: A Primary Key is a unique identifier for each record in a table, ensuring no duplicates and no null values. It's typically a single column, but can be a composite of multiple columns, and doesn't change over time.

Foreign Key: A Foreign Key: is a field in one table that links to a primary key in another table, establishing a relationship. It ensures referential integrity, can contain duplicate values, and can be null, reflecting optional relationships.

Fact Table vs Dimension Table:

Fact Table: Contains quantitative data (like sales, costs) for analysis; usually large with many records. Fact tables contain the core numerical data, often involving measurements or metrics of business processes, and they usually have foreign keys that link to other tables.

Dimension Table: Stores descriptive attributes (like customer details, product info); smaller, provides context to facts. Dimension tables typically contain descriptive attributes to categorize and describe business entities. They often have primary keys that are referenced by fact tables.

Star Schema vs Snowflake Schema:

Star Schema: A database schema with a central fact table surrounded by dimension tables, forming a star-like pattern, optimized for querying large data sets.

Snowflake Schema: A more complex version of the star schema where dimension tables are normalized into multiple related tables, forming a snowflake-like pattern, which reduces data redundancy.

Relationships:

One-to-Many (1:∞): A single record in one table corresponds to multiple entries in another, enabling hierarchical data structures and clear aggregation paths.

Many-to-One (∞:1): Multiple records in one table associated with a single record in another, typically used for detailed reporting on a summary field.

One-to-One (1:1): Each record in one table matches a single record in another, often redundant, suggesting a possible schema consolidation.

Many-to-Many (∞:∞): Records in one table relate to multiple records in another, can create complex and ambiguous relationships, leading to potential inaccuracies and performance issues.

Reading material:
https://www.linkedin.com/pulse/data-modeling-power-bi-introduction-chester-king/
https://www.datacamp.com/tutorial/data-modeling-in-power-bi-tutorial

Youtube Video:
https://youtu.be/LjTXGFpXhoI?si=ZFZnTgIF8us5PBC1

Datasets:

We used these 6 datasets in Power BI session 03**:**
1- Sales 2020
2- Product
3- Territories
4- Returns
5- Product categories
6- Product subcategories