**ASSIGNMENT 9 (16-09-25)**

**Data Modelling**

**Normalization: It is a process of splitting the database into multiple tables so that we can reduce redundancy and improve data integrity.**

**In Power BI, Fact and Dimension tables are two types of tables used to build a data model.**

**Fact: Fact tables will have measurable or quantitative data.**

**The dimension table will contain unique categorical data.**

**Model Creation**

**Cardinality is the relationship between two tables in a data model.**

**It is defined by 4 types.**

1. **One to One**
2. **One to Many**
3. **Many to One**
4. **Many to Many**

* **Why is data modelling necessary?**
* **Remove unnecessary columns and rows**
* **Group by and summarize**
* **Optimize column data types**
* **Use custom columns.**

**A schema is the structure of a data model that defines how data is connected and organized.**

1. **Star schema**

* **One fact table will connect all the dimension tables.**
* **The dimension table in a star schema aligns with the fact table, giving it a star shape.**

**FACT**

1. **Snowflake schema**

**The dimension table has a sub-dimension table.**

**Snowflake schema uses sub-dimensions to represent additional joins in queries.**

**SUB DIM 2**

**SUB DIM 1**

**SUB DIM 4**

**SUB DIM 3**

FACT

1. **Galaxy Schema**

**This schema has more than one fact table linked to dimension tables.**

**SUB DIM 2**

**SUB DIM 1**

**FACT 2**

**FACT 1**