

# Experiment 1

**Name:** Danyl Fernandes (72)

**Class:** TE COMPS

**XIE ID:** 2020012004

**Date:** 24-07-2021

**Aim:** Write a detailed problem statement and design dimensional modelling (creation of star and snowflake schema)

**Problem statement:**

- All electronics may create a sales data warehouse in order to keep records of the store's sales with respect to the dimensions, time, item, branch and location.
- These dimensions allow the store to keep track of things like monthly sales of items and the branches and locations at which the items sold

**Theory:**

Surrogate key:

- A surrogate key is the unique key that identifies each row in the dimension or fact table.
- This key is mostly sequential since it is created for each new row added to the table.
- It is also deemed to be “meaningless” since it has no real significance apart from the fact that it identifies each row in a table uniquely.

Star Vs. Snowflake schema:

Star schema	Snowflake schema
The dimension tables are denormalized.	There is some degree of normalization in the dimension tables.
Queries take relatively less time to execute.	Queries take more time to execute.
Due to denormalization, redundancy is more.	Since the dimension tables are normalized, the level of redundancy is less.
More complex joins are required.	Less to no joins are required.
Good for speed and performance use-cases.	Good for data integrity and non-duplication use-cases.

## Star schema:

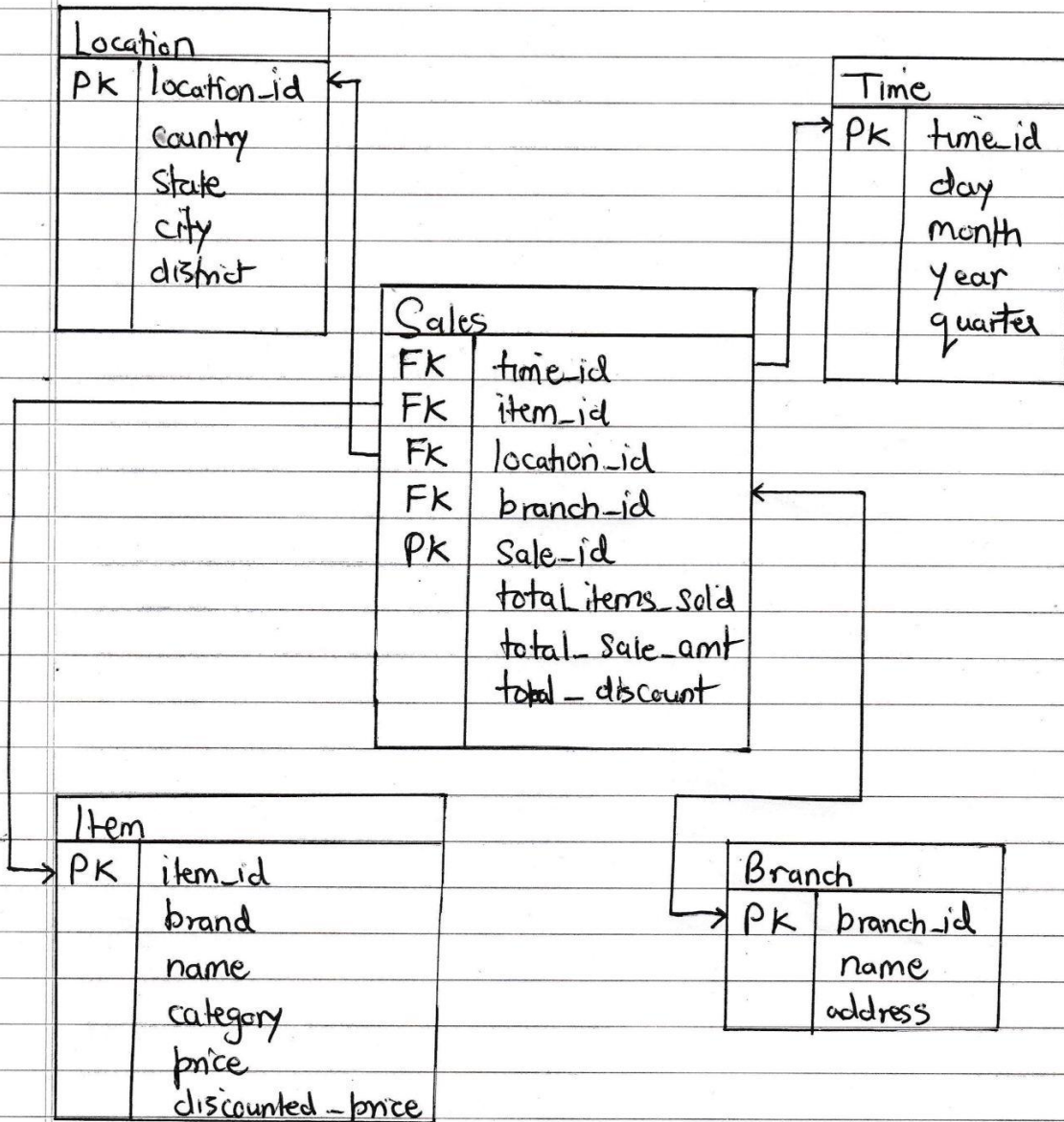
### Fact table:

**Sales** (  
    time\_id,  
    item\_id,  
    location\_id,  
    branch\_id,  
    sale\_id,  
    total\_items\_sold,  
    total\_sale\_amt,  
    total\_discount  
)

### Dimension tables:

1. **Time** (time\_id, day, month, year, quarter)
2. **Item** (item\_id, brand, name, category, price, discounted\_price)
3. **Location** (location\_id, country, state, city, district)
4. **Branch** (branch\_id, name, address)

# Danyl Fernandes (72) • Star Schema



Fact Table : Sales

Dimension Tables : Location, Time, Item, Branch

## Snowflake schema:

### Fact table:

**Sales** (  
    time\_id,  
    item\_id,  
    location\_id,  
    branch\_id,  
    sale\_id,  
    total\_items\_sold,  
    total\_sale\_amt,  
    total\_discount  
)

### Dimension tables:

1. **Time** (time\_id, day, month, year, quarter)
2. **Country** (country\_id, name, code)
3. **Item** (item\_id, brand, name, category, price, discounted\_price)
4. **Location** (location\_id, country, state, city, district)
5. **Branch** (branch\_id, name, address)
6. **Category** (category\_id, name, code)
7. **Brand** (brand\_id, name)

# Danyl Fernandes (72) • Snowflake Schema

