# DWM PARTICAL NO 1

## Assignment No 1: Build Data Warehouse and Explore WEKA

Ware House (Star Schema) – MySql

Weka - https://www.cs.waikato.ac.nz/ml/weka/

**Code:**

```
CREATE DATABASE DATA_WAREHOUSE;
USE DATA_WAREHOUSE;

CREATE TABLE date_dim (
    date_id INT PRIMARY KEY,
    full_date DATE,
    day_name VARCHAR(10),
    month_name VARCHAR(10),
    year INT
);

CREATE TABLE product_dim (
    product_id INT PRIMARY KEY,
    product_name VARCHAR(100),
    category VARCHAR(50),
    price DECIMAL(10,2)
);

CREATE TABLE customer_dim (
    customer_id INT PRIMARY KEY,
    customer_name VARCHAR(100),
```

```
    city VARCHAR(50),
    gender VARCHAR(10)
);

CREATE TABLE store_dim (
    store_id INT PRIMARY KEY,
    store_name VARCHAR(100),
    location VARCHAR(100)
);

CREATE TABLE sales_fact (
    sale_id INT PRIMARY KEY,
    date_id INT,
    product_id INT,
    customer_id INT,
    store_id INT,
    sales_amount DECIMAL(10,2),
    FOREIGN KEY (date_id) REFERENCES date_dim(date_id),
    FOREIGN KEY (product_id) REFERENCES product_dim(product_id),
    FOREIGN KEY (customer_id) REFERENCES customer_dim(customer_id),
    FOREIGN KEY (store_id) REFERENCES store_dim(store_id)
);

INSERT INTO product_dim VALUES
(1, 'Laptop', 'Electronics', 80000.00),
(2, 'Smartphone', 'Electronics', 30000.00),
(3, 'Shoes', 'Apparel', 2000.00),
(4, 'Tablet', 'Electronics', 25000.00),
(5, 'Headphones', 'Electronics', 1500.00),
```

```
(6, 'T-shirt', 'Apparel', 800.00);

INSERT INTO date_dim VALUES
(101, '2025-04-10', 'Thursday', 'April', 2025),
(102, '2025-04-11', 'Friday', 'April', 2025),
(103, '2025-04-12', 'Saturday', 'April', 2025),
(104, '2025-04-13', 'Sunday', 'April', 2025),
(105, '2025-04-14', 'Monday', 'April', 2025),
(106, '2025-04-15', 'Tuesday', 'April', 2025);

INSERT INTO customer_dim VALUES
(201, 'Tony Stark', 'Mumbai', 'Female'),
(202, 'Steve Rogers', 'Delhi', 'Male'),
(203, 'Bruce Banner', 'Bangalore', 'Male'),
(204, 'Natasha', 'Chennai', 'Female'),
(205, 'Clint', 'Pune', 'Male');

INSERT INTO store_dim VALUES
(301, 'Store A', 'Mumbai'),
(302, 'Store B', 'Delhi'),
(303, 'Store C', 'Bangalore'),
(304, 'Store D', 'Chennai');

INSERT INTO sales_fact VALUES
(1, 101, 1, 201, 301, 80000.00),
(2, 102, 2, 202, 302, 30000.00),
(3, 102, 3, 201, 301, 2000.00),
(4, 103, 4, 203, 303, 25000.00),
(5, 104, 5, 204, 304, 1500.00),
```

```
(6, 105, 6, 205, 301, 800.00),
(7, 104, 2, 203, 302, 30000.00),
(8, 106, 1, 204, 303, 80000.00),
(9, 105, 5, 201, 304, 1500.00),
(10, 102, 6, 202, 303, 800.00),
(11, 103, 3, 205, 302, 2000.00),
(12, 101, 2, 201, 304, 30000.00),
(13, 106, 1, 203, 301, 80000.00);

-- Total sales by category
SELECT p.category, SUM(f.sales_amount) AS total_sales
FROM sales_fact f
JOIN product_dim p ON f.product_id = p.product_id
GROUP BY p.category;

-- Sales by city
SELECT c.city, SUM(f.sales_amount) AS total_sales
FROM sales_fact f
JOIN customer_dim c ON f.customer_id = c.customer_id
GROUP BY c.city;

CREATE TABLE sales_by_category AS
SELECT p.category, SUM(f.sales_amount) AS total_sales
FROM sales_fact f
JOIN product_dim p ON f.product_id = p.product_id
GROUP BY p.category;

CREATE TABLE sales_by_city AS
SELECT c.city, SUM(f.sales_amount) AS total_sales
```

```
FROM sales_fact f
JOIN customer_dim c ON f.customer_id = c.customer_id
GROUP BY c.city;
SELECT * FROM sales_by_category;
SELECT * FROM sales_by_city;
```

**Schema (Diagram):**