

Task 9 – Star Schema SQL Code

```
-- Dimension Tables
CREATE TABLE dim_customer (
    customer_id SERIAL PRIMARY KEY,
    customer_name VARCHAR(50)
);

CREATE TABLE dim_product (
    product_id SERIAL PRIMARY KEY,
    product VARCHAR(50),
    category VARCHAR(50)
);

CREATE TABLE dim_region (
    region_id SERIAL PRIMARY KEY,
    region VARCHAR(20)
);

CREATE TABLE dim_date (
    date_id SERIAL PRIMARY KEY,
    order_date DATE
);

-- Fact Table
CREATE TABLE fact_sales (
    sales_id SERIAL PRIMARY KEY,
    order_id INT,
    customer_id INT,
    product_id INT,
    region_id INT,
    date_id INT,
    quantity INT,
    price INT,
    total_sales INT,
    FOREIGN KEY (customer_id) REFERENCES dim_customer(customer_id),
    FOREIGN KEY (product_id) REFERENCES dim_product(product_id),
    FOREIGN KEY (region_id) REFERENCES dim_region(region_id),
    FOREIGN KEY (date_id) REFERENCES dim_date(date_id)
);

-- Insert into Dimensions
INSERT INTO dim_customer(customer_name)
SELECT DISTINCT customer_name FROM retail_sales_raw;

INSERT INTO dim_product(product, category)
SELECT DISTINCT product, category FROM retail_sales_raw;

INSERT INTO dim_region(region)
SELECT DISTINCT region FROM retail_sales_raw;

INSERT INTO dim_date(order_date)
SELECT DISTINCT order_date FROM retail_sales_raw;

-- Load Fact Table
INSERT INTO fact_sales(order_id, customer_id, product_id, region_id, date_id, quantity, price, total_sales)
SELECT r.order_id, c.customer_id, p.product_id, g.region_id, d.date_id,
r.quantity, r.price, r.total_sales
FROM retail_sales_raw r
JOIN dim_customer c ON r.customer_name = c.customer_name
JOIN dim_product p ON r.product = p.product
JOIN dim_region g ON r.region = g.region
JOIN dim_date d ON r.order_date = d.order_date;

-- Indexes
CREATE INDEX idx_customer ON fact_sales(customer_id);
CREATE INDEX idx_product ON fact_sales(product_id);
CREATE INDEX idx_region ON fact_sales(region_id);
CREATE INDEX idx_date ON fact_sales(date_id);

-- Analytics Queries
SELECT g.region, SUM(f.total_sales)
FROM fact_sales f
JOIN dim_region g ON f.region_id = g.region_id
GROUP BY g.region;
```

```
SELECT p.product, SUM(f.total_sales)
FROM fact_sales f
JOIN dim_product p ON f.product_id = p.product_id
GROUP BY p.product;
```

```
SELECT d.order_date, SUM(f.total_sales)
FROM fact_sales f
JOIN dim_date d ON f.date_id = d.date_id
GROUP BY d.order_date;
```