VISHWAKARMA
UNIVERSITY
*Maximising Human Potential*

**T. Y. B. Tech Computer Engineering**
**2024-2025**
**Pursued in Department of Computer Engineering Faculty of Science & Technology**
**Vishwakarma University, Pune-411048**

| | |
|---|---|
| **NAME** | **Swanand Garge** |
| **YEAR** | **THIRD YEAR** |
| **DIV** | **D** |
| **BATCH** | **D2** |
| **ROLL NO** | **39** |
| **SRN NO** | **202201589** |
| **PRN NO** | **2280030433** |
| **COURSE NAME** | **DATAWAREHOUSE & DATA MINING LAB** |
| **COURSE CODE** | **BTECCE22509** |
| **COURSE TEACHER NAME** | **PROF. RAHUL PAPALKAR** |

# ASSIGNMENT NO: 5

## Problem Statement:

## Assignment 5

**Problem Statement**:Assignment 5 Gather Business Requirementsfor Employee Shift Analysis and design it using a Multi-dimensional data model, namely snowflake schema.

## THEORY:

- **Snowflake Schema:**
  - A type of database schema that organizes data into fact and dimension tables but normalizes the dimension tables.
  - The fact table remains central, but dimension tables are split into multiple related tables, forming a snowflake-like structure.
  - It reduces redundancy and can improve data integrity and query performance.
- **Fact Table:**
  - **Fact_EmployeeShift**:
    - Stores quantitative data related to employee shifts, such as overtime hours.
    - Each record contains foreign keys referencing normalized dimension tables.
- **Dimension Tables:**
  - **Dim_Employee**:
    - Contains employee details.
  - **Dim_Department**:
    - Contains department details.
  - **Dim_Job**:
    - Contains job details.
  - **Dim_Shift**:
    - Contains shift details.
  - **Dim_ShiftDay**:
    - Contains information about which day of the week each shift occurs.
  - **Dim_ShiftWeek**:
    - Contains information about which week of the year each shift occurs.
  - **Dim_ShiftMonth**:
    - Contains information about which month each shift occurs.

# Source Code:

```
CREATE DATABASE employee_shift_analysis;
USE employee_shift_analysis;

-- Fact Table
CREATE TABLE Fact_EmployeeShift (
    EmployeeId int NOT NULL,
    DepartmentId int NOT NULL,
    Job_Id int NOT NULL,
    Shift_Id int NOT NULL,
    Date date NOT NULL,
    Overtime int,
    PRIMARY KEY (EmployeeId, DepartmentId, Job_Id, Shift_Id, Date)
);

-- Employee Dimension
CREATE TABLE Dim_Employee (
    EmployeeId int NOT NULL,
    First_Name varchar(30) NOT NULL,
    Last_Name varchar(30) NOT NULL,
    Hire_date date NOT NULL,
    PRIMARY KEY(EmployeeId)
);

-- Department Dimension
CREATE TABLE Dim_Department (
    DepartmentId int NOT NULL,
    DepartmentName varchar(30) NOT NULL,
    PlantNo int NOT NULL,
    PRIMARY KEY(DepartmentId)
);

-- Job Dimension
CREATE TABLE Dim_Job (
    Job_Id int NOT NULL,
    Job_name varchar(30) NOT NULL,
    Job_type varchar(30) NOT NULL,
    PRIMARY KEY(Job_Id)
);
-- Shift Dimension
```

```
CREATE TABLE Dim_Shift (
    Shift_Id int NOT NULL,
    Shift_Name varchar(30) NOT NULL,
    StartTime time NOT NULL,
    EndTime time NOT NULL,
    PRIMARY KEY(Shift_Id)
);

-- Shift Day Dimension
CREATE TABLE Dim_ShiftDay (
    Shift_Id int NOT NULL,
    Day varchar(10) NOT NULL,
    PRIMARY KEY(Shift_Id, Day)
);

-- Shift Week Dimension
CREATE TABLE Dim_ShiftWeek (
    Shift_Id int NOT NULL,
    WeekId int NOT NULL,
    PRIMARY KEY(Shift_Id, WeekId)
);

-- Shift Month Dimension
CREATE TABLE Dim_ShiftMonth (
    Shift_Id int NOT NULL,
    Month varchar(10) NOT NULL,
    PRIMARY KEY(Shift_Id, Month)
);

-- Insert values into Fact_EmployeeShift
INSERT INTO Fact_EmployeeShift VALUES
(1, 2, 3, 1, '2024-09-01', 2), (2, 1, 4, 2, '2024-
09-01', 3), (3, 3, 5, 3, '2024-09-02', 1);


-- Insert values into Dim_Employee
INSERT INTO Dim_Employee VALUES
(1, 'John', 'Doe', '2020-01-15'),
(2, 'Jane', 'Smith', '2019-03-22'),
(3, 'Jim', 'Brown', '2018-07-30');
```

```
-- Insert values into Dim_Department
INSERT INTO Dim_Department VALUES
(1, 'HR', 101),
(2, 'Finance', 102),
(3, 'IT', 103);

-- Insert values into Dim_Job
INSERT INTO Dim_Job VALUES
(1, 'Manager', 'Full-time'),
(2, 'Analyst', 'Part-time'),
(3, 'Developer', 'Full-time');

-- Insert values into Dim_Shift
INSERT INTO Dim_Shift VALUES
(1, 'Morning', '08:00:00', '16:00:00'),
(2, 'Afternoon', '16:00:00', '00:00:00'),
(3, 'Night', '00:00:00', '08:00:00');

-- Insert values into Dim_ShiftDay
INSERT INTO Dim_ShiftDay VALUES
(1, 'Monday'),
(1, 'Tuesday'),
(2, 'Wednesday'),
(3, 'Thursday');

-- Insert values into Dim_ShiftWeek
INSERT INTO Dim_ShiftWeek VALUES
(1, 35),
(2, 36),
(3, 37);

-- Insert values into Dim_ShiftMonth
INSERT INTO Dim_ShiftMonth VALUES
(1, 'September'),
(2, 'September'),
(3, 'October');
```

## OUTPUT:

Dim_ShiftMonth

| Shift_Id | Month |
|---|---|
| 1 | September |
| 2 | September |
| 3 | October |
| NULL | NULL |

Dim_ShiftWeek;

| Shift_Id | WeekId |
|---|---|
| 1 | 35 |
| 2 | 36 |
| 3 | 37 |
| NULL | NULL |

Dim_ShiftDay;

| Shift_Id | Day |
|---|---|
| 1 | Monday |
| 1 | Tuesday |
| 2 | Wednesday |
| 3 | Thursday |

Dim_Shift

| Shift_Id | Shift_Name | StartTime | EndTime |
|---|---|---|---|
| 1 | Morning | 08:00:00 | 16:00:00 |
| 2 | Afternoon | 16:00:00 | 00:00:00 |
| 3 | Night | 00:00:00 | 08:00:00 |
| NULL | NULL | NULL | NULL |

Dim_Job;

| Job_Id | Job_name | Job_type |
|---|---|---|
| 1 | Manager | Full-time |
| 2 | Analyst | Part-time |
| 3 | Developer | Full-time |
| NULL | NULL | NULL |

Dim_Department;

| DepartmentId | DepartmentName | PlantNo |
|---|---|---|
| 1 | HR | 101 |
| 2 | Finance | 102 |
| 3 | IT | 103 |
| NULL | NULL | NULL |

Dim_Employee;

| EmployeeId | First_Name | Last_Name | Hire_date |
|---|---|---|---|
| 1 | John | Doe | 2020-01-15 |
| 2 | Jane | Smith | 2019-03-22 |
| 3 | Jim | Brown | 2018-07-30 |
| NULL | NULL | NULL | NULL |

Fact_EmployeeShift;

| EmployeeId | DepartmentId | Job_Id | Shift_Id | Date | Overtime |
|---|---|---|---|---|---|
| 1 | 2 | 1 | 1 | 2024-09-01 | 2 |
| 2 | 1 | 2 | 2 | 2024-09-01 | 3 |
| 3 | 3 | 3 | 3 | 2024-09-02 | 1 |
| NULL | NULL | NULL | NULL | NULL | NULL |

Snowflake Schema