# **Logical Model for Data Warehouse**

## **Business Requirements**

The business comes under health sector domain. The requirement of the business is given below:

* To be able to know if an employee was working on a particular day or not.
* If they worked, their star and left time, hours worked and if they were on charge should be known.
* If they didn’t work, were they on call should be known.
* There are two shifts in the company. Morning shift starts between 5:00 AM – 11:00 PM and evening starts after 12 PM. It should be known that which employee worked on which shift.
* To be able to know if the employees are working regularly on a weekend.
* To analyze the data on a biweekly basis starting from 2021-01-01.
* To analyze if any employee has to cover for other team members regularly.
* To analyze the data based on the employee role.
* To analyze the salary distribution by department.

There are two datasets provided for which we have to prepare a logical model which are employee and timesheet. The employee sheet contains data like employee name, their department, manager, salary, birth date etc. whereas timesheet dataset contains work day details like punch apply date, punch in time, punch out time etc.

According to the detail study on dataset and requirements provided by the client, the fact table and dimension table identified are as follows:

1. Fact Table

* fact\_timesheet – This table contains the data related to employee timesheet
* fact\_employee – This table contains data related to employee details like their department.

1. Dimension Table

* shift – Dimension table of fact\_timesheet
* department – Dimension table for both fact\_timesheet and fact\_employee
* role – Dimension table for fact\_employee
* time period – Dimension table for fact\_timesheet

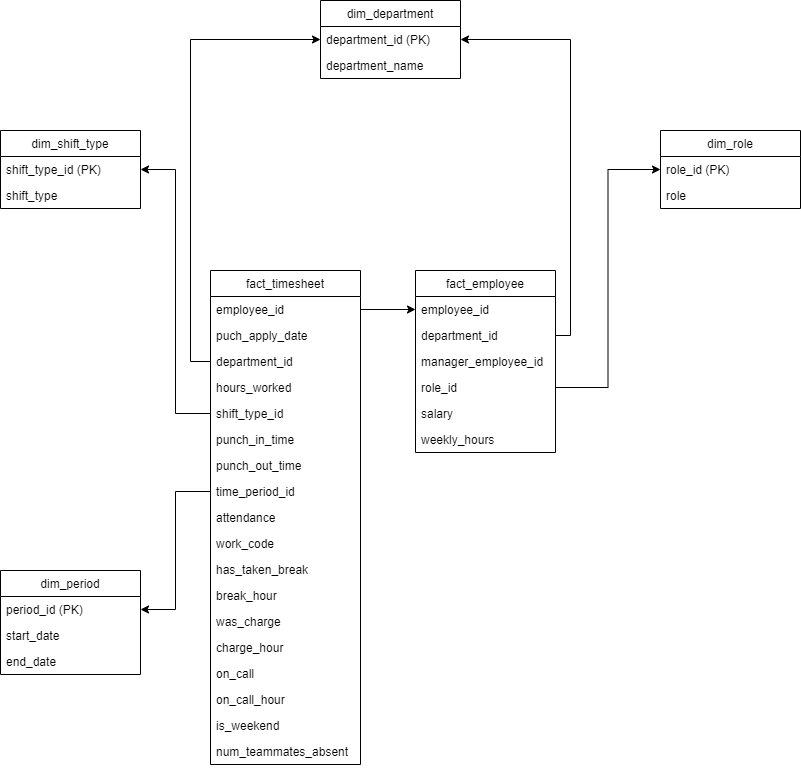


Figure 1: Logical Model for Data Warehouse

From above logical model, data related to the business requirement of the client can be made. Looking at punch\_apply\_date and attendance the client can find out whether the employee was working on a particular day or not. If they had attendance that day, to know the start and leave time punch\_in\_time and punch\_out\_time is there. Also, hours\_worked to find out how many hours they worked. The dimension table shift\_type helps to find out which shift the employee worked on. Aslo, the period table can be used to find the data on biweekly basis. The num\_teammates\_absent and hours\_worked will help to overlook any employee has to cover for other employees regularly. Similarly, department and salary of employee can also be known.

# **Physical Model**

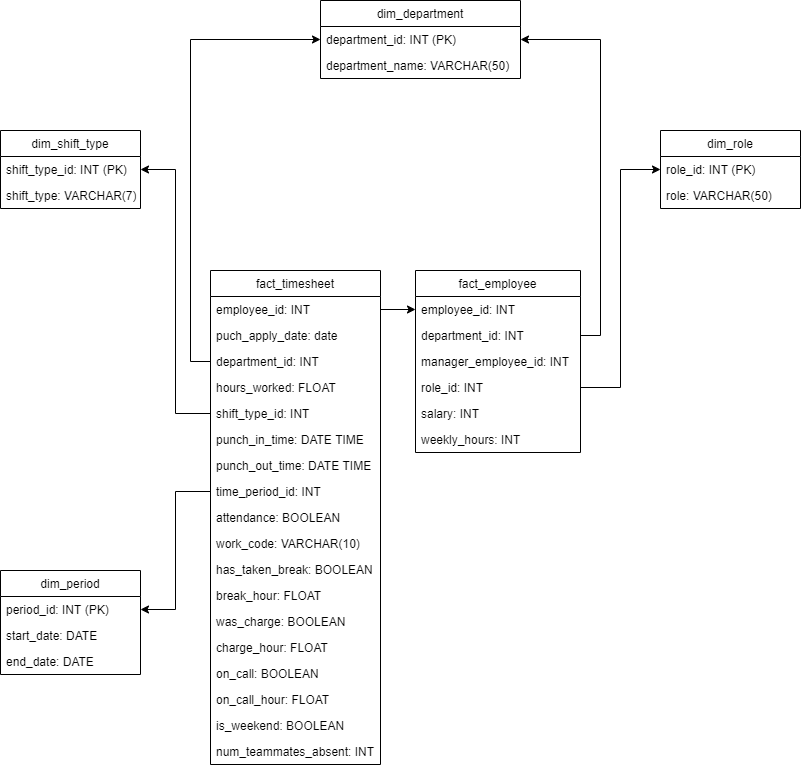


Figure 2: Physical Model

The above datatypes were chosen on the basis of type of data provided in the timesheet and employee dataset. Like, punch\_in\_time data had date and timestamp so datetime datatype was chosen whereas punch\_apply\_date contains only date.