**Logical modelling**

The list of possible entities are:

1. department (Dimension table)
2. employee (Dimension table)
3. punch\_apply\_date (Dimension table)
4. paycode (dimension table)
5. attendence (Fact table)

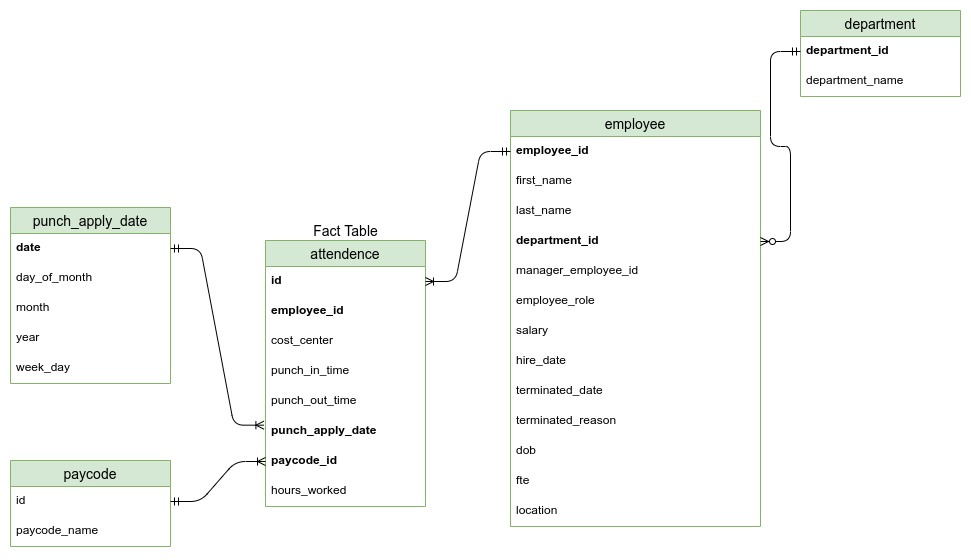
Now I will give the description of the entities , description of their attributes and the attributes domain.

| Entity | Description | Domain |
| --- | --- | --- |
| department | The list of departments present in the hospital. |  |
| **Attributes**:  department\_id  department\_name | Identifier of the entity department,PK  The name of the department | Fixed length character  Text |
| employee | The person working on the hospital |  |
| **Attributes:**  employee\_id  name  department\_id  manager\_employee\_id  employee\_role  salary  hire\_date  terminated\_date  dob  fte  location | Identifier of the entity employee,PK  First\_name+last\_name  ID referencing the entity department,FK  ID referencing the employee\_id of the same table,FK  The role of an employee  The salary of the employee  The hire date of employee  The terminated date of employee  Data of birth  Full time equivalent  The address of an employee | Integer  Text  Valid id from table department  Valid employee\_id from same table  Text  Money  Date  Date  Date  Float  Text |
| punch\_apply\_date | Everyday attendance sheet |  |
| **Attributes:**  date  day\_of\_month  month  Year  week\_day | Valid date,PK  The day of the date  The month of the date  The year of the date  The week day,  monday=0...sunday=6 | Date  CHAR(2)  CHAR(2)  CHAR(4)  CHAR |
| paycode | The paycode of the customer |  |
| **Attributes:**  Id  paycode\_name | Identifier of the entity paycode,PK  The name of the paycode | CHAR(1)  Text |
| attendence | The daily attendence of the employee |  |
| **Attributes:**  id  employee\_id  cost\_center  punch\_in\_time  punch\_out\_time  punch\_apply\_date  paycode\_id  hours\_worked | Identifier of the entity attendence,SK,FK  Id referencing the entity employee\_id,FK  The cost center code  The date and time of punch in.  The date and time of punch out.  The date of the punch apply.  ID referencing the paycode table.  The number of hours worked on a particular paycode. | Auto generated  Valid id from table employee  CHAR(5)  Time Stamp  Time Stamp  Date  CHAR  Float |

*FIG:Table showing the entities,attributes,attributes description and attributes domain of the entities.*

As per the above listed entities and looking at their relationship , I go for snowflake schema for dimensionality modelling.

The ER diagram is drawn below.



*FIG:Logical ER diagram of Hospital warehouse(SNOWFLAKE SCHEMA)*

**Physical modelling**

This is the part where I will physically implement the model into the database.For that purpose we are gonna use PostgreSQL relational database.It is an open source DBMS.

It uses the Standard Query Language(SQL) to create,update,alter ,delete the tables or any rows in the database.

Firstly , I made the dummy datas for each table with some tuples in it in the Excel sheet.

Then , I have used python and the driver of PostgreSQL which is Psycopg2 to push the

Data of .xlsx file into the Postgre DBMS.I have used Jupyter Notebook to write all of the code .

The physical implementation code of the hospital warehouse is presented below in my github repository link:

VISIT THE LINK PLEASE FOR PHYSICAL IMPLEMENTATION + REQUIREMENT IMPLEMENTATION:

LINK:

<https://github.com/callingsandesh/Leapfrog-Assignment/blob/main/Data/3rd%20Week(OLAP%20Design)/Day1/hospital%20datawarehouse.ipynb>

**Requirements**:

A

i)

The clients can know what time the employee started and left from the fact table by querying the punch\_in\_time and punch\_out\_time from the Fact table.

The number of hours worked is also in the dimension table .So it can be easily fetched , and to get the total number of hours worked we can use the GROUP BY statement.

We can use the paycode\_id to know the paycode , and fetch the respective paycode,

ii)

So, the employee can be on paycode “CALL” if they have “CALL” on the column paycode.

B. To solve this issue we can make a separate column in the Fact table which has columns like morning\_shift,day\_shift,evening\_shift.

C. Their is a attribute called week\_day which denotes the day of the week .So,

We can extract the weekend .

D. So, to know this we have see the attendance table , and see the paycode\_type is “CHARGE”

E. So, to analyze the data on a biweekly basis , we can simple get data skipping 13 days from the punch\_apply\_date.

F.We have the attributes employee\_role in the employee table(dimention\_table) from which we can fetch the information.

G.We have the department\_id in the employee table , so we can GROUP BY department of employee table and SUM up the salary, to get the total salary distribution of the respective departments.

**References:**

1. *Draw.io (*[*https://app.diagrams.net/*](https://app.diagrams.net/) *)*
2. PostgreSQL ( <https://www.postgresql.org/> )
3. Psycopg2 ( <https://www.psycopg.org/docs/> )
4. Python Programming language (<https://www.python.org/> )
5. Jupyter Notebook (<https://jupyter.org/> )
6. VS Code
7. Leapfrog slides