ASSIGNMENT 4 – Normalization

Normalization

Database normalization is the process of organizing the attributes and tables of a relational database to minimize data redundancy.

Normalization involves refactoring a table into smaller (and less redundant) tables but without losing information. The objective is to isolate data so that additions, deletions, and modifications of an attribute can be made in just one table and then propagated through the rest of the database using the defined foreign keys.

First normal form (1NF)

If a relation contains a composite or multi-valued attribute, it violates the first normal form or a relation is in first normal form if it does not contain any composite or multi-valued attribute. A relation is in first normal form if every attribute in that relation is a single valued attribute.

Before Normalization Doctors table Schema:

```
CREATE TABLE Doctors (
    Doctor_Id INT NOT NULL,
    License_Number TEXT,
    Doctor_Name VARCHAR(150) NOT NULL,
    Degree TEXT,
    Practice_Specialities TEXT,
    Hospital_Id BIGINT NOT NULL,
    Gender VARCHAR(20) NOT NULL,
    PRIMARY KEY (Hospital_Id),
    FOREIGN KEY (Doctor_Id)
        REFERENCES Diseases (Disease_Id)
);
```

Post Normalization Doctors table,Degree table, Practice_Specialities Table Schema:

Doctors Table

```
CREATE TABLE Doctors (
Doctor_Id INT NOT NULL,
License_Number TEXT,
Doctor_Name VARCHAR(150) NOT NULL,
Hospital_Id INT NOT NULL,
Gender VARCHAR(20) NOT NULL,
PRIMARY KEY (Doctor_Id)
);
```

Degree Table

```
CREATE TABLE DEGREE(
Doctor_Id INT NOT NULL,
Doctor_Degree TEXT,
PRIMARY KEY (Doctor_Id),
FOREIGN KEY (Doctor_Id) References Doctors(Doctor_Id)
);
```

Practice_Specialities Table

```
CREATE TABLE Practice_Specialities (
Doctor_Id INT NOT NULL,
Practice_Specialities VARCHAR(300),
PRIMARY KEY (Doctor_Id,Practice_Specialities),
FOREIGN KEY (Doctor_Id) REFERENCES Doctors(Doctor_Id))
);
```

3.We also have split Hospitals table into Hospital_Affiliations table to meet 1st NF

Before Normalization Hospitals table Schema:

```
CREATE TABLE Hospitals (
    Hospital_Id BIGINT NOT NULL,
    Hospital_Affliations TEXT,
    City VARCHAR(150) NOT NULL,
    State VARCHAR(20),
    Zip_Code BIGINT NOT NULL,
    PRIMARY KEY (Zip_Code),
    FOREIGN KEY (Hospital_Id)
        REFERENCES Doctors (Hospital_Id)
);
```

Post Normalization Hospitals table and Hospital_Affiliations Table Schema:

Hospitals Table

```
CREATE TABLE Hospitals (
Hospital_Id INT NOT NULL,
Zip_Code INT NOT NULL,
PRIMARY KEY (Hospital_Id),
FOREIGN KEY (Hospital_Id)
REFERENCES Doctors (Doctor_Id)
);
```

Hospital_Affiliations Table

```
CREATE TABLE Hospital_Affiliations(
Hospital_Id INT NOT NULL,
Hospital_Affliations VARCHAR(300),
PRIMARY KEY (Hospital_Id,Hospital_Affliations),
FOREIGN KEY (Hospital_Id)
```

REFERENCES Hospitals (Hospital_Id)

);

Justifications:

1.Every table of our database has primary key with minimal set of attributes which can uniquely identify a record

2. The values in each column of a table are atomic and there are no multivalued attributes. We have split Doctors table into Degree table and Practice_Specialities which initially had multivalued attributes to meet the 1st NF.

Second normal form (2NF)

To be in second normal form, a relation must be in first normal form and relation must not contain any partial dependency. A relation is in 2NF if it has No Partial Dependency, i.e., no non-prime attribute (attributes which are not part of any candidate key) is dependent on any proper subset of any candidate key of the table.

Justifications:

1. All the above tables fulfill the requirements of 1st NF
2. No partial dependencies- Since all our tables had a Candidate key with single valued attribute : It conforms to the property Rule : "If a table candidate key is a single valued attribute then that table is in 2NF form"
3. None of our tables have calculated data

Third normal form (3NF)

A relation is in third normal form, if its in 2NF and there is **no transitive dependency**.

Justifications:

1. All the above tables fulfill the requirements of 2nd NF
2. Transitive dependencies in Hospitals Table : Zip_Code-> State & City .Hence we have a new table called **PINCODES_STATE_CITY Table** which consists of City, State, Zip_Code.

Post Normalization

PINCODES_STATE_CITY Table

CREATE TABLE PINCODES_STATE_CITY(

City VARCHAR(150) NOT NULL,

```
State VARCHAR(20),
Zip_Code INT NOT NULL,
Hospital_Id INT NOT NULL,
PRIMARY KEY (Hospital_Id),
FOREIGN KEY (Hospital_Id)
References Hospitals(Hospital_Id)
);
```

NORMALIZED PHYSICAL MODEL (SQL)

1. Diseases Table

```
CREATE TABLE Diseases (
    Disease_Id INT ,
    Disease_name VARCHAR(100) NOT NULL,
    Signs_and_symptoms TEXT,
    Diagnosis TEXT,
    Treatment TEXT,
    Disease_Category_Id INT,
    Doctor_Id INT,
    PRIMARY KEY (Disease_Id),
    FOREIGN KEY (Doctor_Id)
    REFERENCES Doctors (Doctor_Id),
    FOREIGN KEY (Disease_Category_Id)
    REFERENCES Disease Category(Disease_Category_Id)
```

2. Disease_Category Table

```
CREATE TABLE DiseaseCategory (
```

```
    Disease_Category_Id INT NOT NULL,
    Disease_Category_Name VARCHAR(150) NOT NULL UNIQUE,
    PRIMARY KEY (Disease_Category_Id)
   );
```

3. Degree Table

```
CREATE TABLE DEGREE(
Doctor_Id INT NOT NULL,
Doctor_Degree TEXT,
PRIMARY KEY (Doctor_Id),
FOREIGN KEY (Doctor_Id) References Doctors(Doctor_Id)
);
```

4. Practice_Specialities Table

```
CREATE TABLE Practice_Specialities (
Doctor_Id INT NOT NULL,
Practice_Specialities VARCHAR(300),
PRIMARY KEY (Doctor_Id,Practice_Specialities),
FOREIGN KEY (Doctor_Id) REFERENCES Doctors(Doctor_Id))
);
```

5. Doctors Table

```
CREATE TABLE Doctors (
Doctor_Id INT NOT NULL,
License_Number TEXT,
Doctor_Name VARCHAR(150) NOT NULL,
Hospital_Id INT NOT NULL,
Gender VARCHAR(20) NOT NULL,
```

```
PRIMARY KEY (Doctor_Id)
);
```

6.Hospitals Table

```
CREATE TABLE Hospitals (
Hospital_Id INT NOT NULL,
Zip_Code INT NOT NULL,
PRIMARY KEY (Hospital_Id),
FOREIGN KEY (Hospital_Id)
REFERENCES Doctors (Doctor_Id)
);
```

7. PINCODES_STATE_CITY Table

```
CREATE TABLE PINCODES_STATE_CITY(
City VARCHAR(150) NOT NULL,
State VARCHAR(20),
Zip_Code INT NOT NULL,
Hospital_Id INT NOT NULL,
PRIMARY KEY (Hospital_Id),
FOREIGN KEY (Hospital_Id)
References Hospitals(Hospital_Id)
);
```

8.**Hospital_Affiliations Table**

```
CREATE TABLE Hospital_Affiliations(
Hospital_Id INT NOT NULL,
Hospital_Affliations VARCHAR(300),
PRIMARY KEY (Hospital_Id,Hospital_Affliations),
FOREIGN KEY (Hospital_Id)
      REFERENCES Hospitals (Hospital_Id)
);
```

TABLE SCREENSHOTS :

Diseases Table

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

Disease_Id	Disease_name	Signs_and_symptoms	Diagnosis	Treatment	Disease_Cat	Doctor_Id
1	Acinetobacter infections	Blood infection: Fever chills vomiting c...	Culture	Supportive care	18	23
2	Actinomycosis	Painful abscesses	Histologic findings	Penicillin doxycycline and sulf...	18	34
3	Adenovirus infection	high fever that lasts 4 to 6 daysphary...	Antigen detection polymerase ...	Most infections are mild and r...	19	40
4	African sleeping sickness (African ...	Hemolymphatic phase: Fever lymphad...	Identification of trypanosome...	Fexinidazole by mouth or pen...	19	59
5	AIDS (acquired immunodeficiency ...	Opportunistic infections	Antibody test p24 antigen tes...	Treatment is typically a non-n...	18	67
6	Amoebiasis	Most are asymptomatic.	Microscopy	Those with symptoms require ...	19	77
7	Anaplasmosis	Fever leukopenia abnormally elevated ...	indirect immunofluorescence a...	Tetracycline drugs (including t...	19	7
8	Angiostrongyliasis	Abdominal pain nausea vomiting and w...	Lumbar puncture brain imagin...	Albendazole	18	36

Degree Table

Doctor_Degree	Doctor_Id
M.D	1
M.D	2
M.D	3
M.D	4
D.O.	5
M.D	6
M.D	7
M.D	8

Practice_Specialities Table

Practice_Specialities	Doctor_Id
Child and Adolescent Psychiatry	1
Psychiatry	1
Forensic Psychiatry	1
Geriatric Psychiatry	1
Child and Adolescent Psychiatry	2
Psychiatry	2
Forensic Psychiatry	2
Forensic Psychiatry	3

Disease Category Table

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Disease_Category_Id	Disease_Category_Name
1	Child and Adolescent Psychiatry
2	Forensic Psychiatry
3	Medical Microbiology Pathology
4	Anesthesiology
5	Addiction Medicine
6	Consultation-Liaison Psychiatry/Psychosomatic ...
7	Adult Reconstructive Orthopaedic Surgery
8	Adult Cardiac Anesthesiology

Doctors Table

Doctor_Id	License_Number	Doctor_Name	Hospital_Id	Gender
1	54713.0	Ward Bein	1	Male
2	58952.0	Donald Condie	2	Male
3	51517.0	Jeffrey Friedman	3	Female
4	255819.0	Mihae Platt	4	Transgender
5	288074.0	Allison Nussbaum	5	Male

Hospitals Table

Hospital_Id	Zip_Code
1	1720
2	2141
3	2139
4	1752
5	2139
6	2139
7	1803
8	1805

PINCODES_STATE_CITY Table

City	State	Zip_Code	Hospital_Id
Acton	MA	1720	1
Cambridge	MA	2141	2
Cambridge	MA	2139	3
Marlborough	MA	1752	4
cambridge	MA	2139	5
Cambridge	MA	2139	6
Burlington	MA	1803	7
Burlington	MA	1805	8

Hospital_Affiliations Table

Hospital_Id	Hospital_Affliations
1	Private Practice
1	Tewksbury Hospital
2	People Care Clinic
2	Vinfen Corporation
2	Private Office
3	Beth Israel Deaconess Medical Center
3	Brigham & Women's Hospital
3	Private Practice

Contributors:

1. Rohit Panicker : 002791446
2. Rohit Paul : 002908933
3. Akshatha Patil : 00265158