**44-560 Advanced Topics in Database Systems**

**Assignment-02: Normalization**

Answer the following questions (Answers should be short and precise, preferably one line. Place the answer below the question and highlight your answer in yellow like This is the Answer):

1. **Provide three reasons why we do normalization?**

**1. To reduce the number of duplicates in database.**

**2. To minimize the redundancy of data in the table.**

**3. To easier mapping of object to data.**

**4. It reduces the chances for the anomalies to occur in a database.**

**5. It is helpful to divide the large database table into smaller tables and link them using relationship**

1. **Condition(s) for a relation to be in First Normal Form?**

**1. if every attribute in the relation is single valued attribute.**

**2. Attribute domain doesn’t change in the relation.**

**3. The order in which data is stored does not matter.**

**4. There is unique name for every column in the normal form.**

1. **Condition(s) for a relation to be in Second Normal Form?**

**1. The relation must be in First normal form.**

**2. Relation must not contain any partial dependencies.**

**3. It does not have any non-prime attribute that is functionally dependent on any proper subset of any candidate key of the relation.**

1. **What is Partial Dependency?**

**1. Partial Dependency occurs when a non-prime attribute is functionally dependent on part of a candidate key.**

**2. The second normal form eliminates the Partial dependencies.**

1. **What is Transitive Dependency?**

**1. A functional dependency is said to be transitive if it is indirectly formed by two functional dependencies.**

**2. A transitive dependency can only occur in a relation of three of more attributes.**

1. **Condition(s) for a relation to be in Third Normal Form?**

**1. A relation will be in 3NF if it is in 2NF and not contain any transitive partial dependency.**

**2. 3NF is used to reduce the data duplication.**

**3. If there is no transitive dependency for non-prime attributes, then the relation must be in third normal form.**

Consider a clinic that gives health services on different health-related issues. New patients are enrolled. Every doctor and patient is provided with ID number. The hospital has maintained a single spreadsheet to store all details so far. As the result, it has become very difficult for the hospital to keep track of the information being accumulated.

**Normalization:**

Here, the data is not atomic, and it consists of partial dependencies and transitive dependencies. As per database developer perspective, it is required to normalize the data into 3NF. Apply the necessary normalizations on the given data and transform the data into the third normal form.

The excel sheet has a list of data, consisting of doctor details, Patient details, Clinic code, Room Number, Room Type and Bed Type and Registration details.

Download the excel sheet from assignment page.

**To submit your assignment, you are responsible for the following.**

* Rename this given file with your Lastname\_Assignment02.
* Include 3NF normalized excel sheet. Consider one tab in the excel sheet as a table. Highlight the primary key attribute cell in yellow color \_\_\_\_ and foreign keys in \_\_\_\_.
* Design and draw the ER diagram with all normalized entities and relationships and submit it via Screenshot from Lucid chart, ERD plus, or other diagramming tools. Clearly show key attributes, including primary keys and foreign keys, and the relationships between the entities.

Note:

* Do not introduce new surrogate keys as primary keys in the entities in your solution.
* Do not include any composite keys in your solution.

Attach the excel sheet and diagram screenshot separately in your submission. For the diagram, make sure to add enough of your desktop to show the screen is running in your laptop (for example, a command prompt window with your SID in command line).