Katie Bartolotta

CMPT 308N 114

Professor Labouseur

9/10/16

Lab 2

2. A candidate keys consists of individual columns in a table that has all unique rows. This is the most minimal subset of fields that uniquely identifies a tuple. For example, if there is a candidate key in a column called “user\_id”, there will never be more than one tuple with the same user\_id. Primary keys are the columns that maintain a unique state in a table. This key is the one that a user tells the database to work on. A super key occurs when any other column or attribute is added to a primary key. This means that it is a set of fields that contains a key. So, if you used the example for candidate keys which was user\_id and added a user\_name, then it would classify as a super key.

3. Data types identify different types of data which determine the possible values that type could have. A table could be created that contains customer data. The fields in the customer data table include customer ID, first name, last name, address, phone number, and email. While the customer ID and phone number will be numeric strings, the other fields will be alphanumeric strings. The only fields in this table that can be null are the phone number and email. Sometimes people would rather a company doesn’t have this data, but the other fields are necessary to complete transactions and figure out who the customer is. Without data types, data can’t be defined and used.

4a. In relational databases the first normal form rule is the basis for database normalization. The structure relates to a single table that is within a relational database system. The first normal form says that every column in a table has to be unique, so separate tables are created for related data so this rule is not broken. Also, no rows or columns can be duplicated which builds off of the previous rule that every column is unique. Another parameter is that no intersections of rows or columns can have either a null value or multiple values. Following these guidelines is important because searching records that contains duplicate entries can be very difficult and prone to mistakes. An example of an issue relating to the first normal form rule is when in a data table that includes customer information, a customer could have multiple phone numbers. This violates the 1NF rule, but can be fixed by creating a separate table for telephone numbers that relates back to the main table.

b. Another relational database rule is the access rows by content only rule. When accessing data from a table, the user is not allowed to access it by where it is located in a table. For example, a person can’t say they want the data in the second column and third row. The column names and content that is in a row must be referenced. This way, mistakes with finding data occur less often as there may be an issue when counting the rows and columns. So, data can be accessed by what it is in a table, not where it is.

c. A very important relational rule is that all rows must be unique. There can be similarities in the data, but not exact matches. If there are two rows that contain the same data, it will be almost impossible to refer to them without it being confusing. Duplicate data has a chance of being incorrect. When creating a table, the creator has to be smart about the primary keys. For example, a student ID number would be a good choice, but using a student’s first or last name would not be a good idea because more than one student may have the same name. If this is the case, then when referring to the data, the wrong student could be chosen.