

Lecture 5 Notes - Wednesday 09/14/16

Reading Quiz

Question 1: Ans = B

Question 2: Ans = A

Question 3: Ans = C

Question 4: Ans = C

Question 5: Ans = B

Notes

We use the idea of a *functional dependency* (see last lecture's slides) to define 2nd and 3rd Normal Form

1NF to 2NF slide: note that we're decomposing the table into two new tables, so that we don't have any columns that are determined by only *part* of the primary key

2NF to 3NF slide: we're decomposing the table into two new tables so that no column is determined by any other except for its primary key column

The point of these decompositions is to reduce the redundancy of our data, so that we aren't storing two columns in a table if we could know both column's values by knowing only one:

For example, if knowing Alice's *major* tells us her *college*, then we shouldn't store both these in the table -- because if UT decides to change the name of the *Natural Sciences* college to *Data and Physical Sciences*, we would have to change every single row in the table -- but if we're in 3NF, we only have to change one row in the *Major_College* table. Not only is this easier, it keeps us from accidentally leaving some rows as *Natural Sciences* somehow when we are updating.

CQ1: B Since *drug_description* is dependent on a non-key column, *drug_name*, this table is no longer in 3NF. So at most it is 2NF. But nothing has changed with respect to 1NF or 2NF conditions, so we're in 2NF.

Note in the revised version of this table (back to 3NF) queries become more difficult because we have to join the new tables to match a drug description to the drug's number, but database people think this is worth it to get rid of the redundancy.