

# Tutorial 3 — Normalization, Functional Dependencies

Richard Wong  
`rk2wong@edu.uwaterloo.ca`

Department of Electrical and Computer Engineering  
University of Waterloo

January 29, 2018

What are the non-trivial functional dependencies in the following table?

Also, what are the superkeys in the table? What about candidate keys?

id	name	address
1	Alice	123 Park Place
2	Alice	85 Seagram Drive
3	Bob	161 University Avenue W
4	Bob	85 Seagram Drive

$\text{id} \rightarrow \text{name}$

$\text{id} \rightarrow \text{address}$

$\text{id} \rightarrow \text{name, address}$

Since  $\text{id}$  can functionally determine all of the other attributes in the relation, it is a superkey. In fact, any set of attributes containing  $\text{id}$  is also a superkey. However, because  $\text{id}$  on its own is a minimal superkey, it is also a candidate key.

What is the highest normal form that the following table fits?

PersonID	Name	FavouriteColourID	ColourName
1	Alice	1	Green
2	Bob	1	Green
3	Eve	2	Blue

The table is in 1NF, because given reasonable business logic assumptions, the domains of each attribute are atomic.

The table is also in 2NF, because each non-prime attribute (not part of any candidate key) is dependent on the whole of every candidate key (there is just one, the PersonID). Since the candidate key consists of only one attribute, there can be no partial dependencies.

The table is NOT in 3NF, since ColourName, a non-prime attribute, is transitively dependent on PersonID. We know this because (we would assume)  $\text{PersonID} \rightarrow \text{FavouriteColourID}$  and  $\text{FavouriteColourID} \rightarrow \text{ColourName}$ .

So the highest normal form that the table is in is 2NF.

What is the highest normal form that the following relation  $R=ABCD$  fits?

$$F = \{$$
$$A \rightarrow B$$
$$A \rightarrow C$$
$$C \rightarrow D$$
$$\}$$

We assume the table is in 1NF, because that's a reasonable assumption.

There are no partial dependencies (dependencies on just part of any candidate key), so the relation is in 2NF.

There is a transitive dependency  $A \rightarrow D$ , where  $D$  is a non-prime attribute, so the relation is not in 3NF.

What is the highest normal form that the following relation  $R=ABCDE$  fits?

$F = \{$   
 $A \rightarrow BCDE$   
 $E \rightarrow ABCD$   
 $C \rightarrow D$   
 $\}$



A is a candidate key since it alone can determine the remainder of the attributes in the relation.

There are no partial dependencies, so the relation is in 2NF.

D, a non-prime attribute, is transitively dependent on the candidate key A, since A functionally determines C, which functionally determines D.

So the table is not in 3NF.

What is the highest normal form that the following relation  $R=ABCDEG$  fits?

$F = \{$   
 $AB \rightarrow CDEG$   
 $EG \rightarrow ABCD$   
 $C \rightarrow E$   
 $\}$

The candidate keys of R are AB and EF, so the non-prime attributes are C and D.

There are no partial dependencies, so the relation is in 2NF.

There are also no transitive dependencies to non-prime attributes, so the relation is in 3NF.

C is not a superkey of R, but it is on the left side of a functional dependency of R, so R is not in BCNF.

What is a canonical cover for the following set of FDs?

$$F = \{$$
$$A \rightarrow BC$$
$$CD \rightarrow E$$
$$B \rightarrow D$$
$$E \rightarrow A$$
$$\}$$

The canonical cover for that set of FDs is the same set. There are no extraneous attributes.

Recall that an FD is extraneous within a set if it can be inferred from the other FDs in that set.

What is a canonical cover for the following set of FDs?

$$F = \{$$
$$A \rightarrow BC$$
$$A \rightarrow B$$
$$B \rightarrow C$$
$$AB \rightarrow C$$
$$\}$$

First, combine all FDs with the same left side to make all FDs contain unique left sides. FDs 1 and 2 give us FD 1, so we can remove FD 2.

Then, remove all extraneous FDs. FD 4 is extraneous, because we can derive it from FD 1 via augmentation (add B to both sides) and decomposition (remove B from the right side).

So we are left with FDs 1 and 3.

Then, remove all extraneous attributes. FD 1 contains the extraneous attribute C. We know it is extraneous because if we remove it from FD 1, the closure of the set does not change, i.e. we can rederive FD 1 from the reduced set.

So the canonical cover is  $A \rightarrow B$  and  $B \rightarrow C$ .

Decompose the following relation  $R=ABCDE$  into BCNF.

$$F = \{ \\ A \rightarrow BC \\ C \rightarrow DE \\ \}$$



First, compute attribute closures. In doing so, we find that A is the only candidate key.

Check each FD to see if they satisfy BCNF.

$A \rightarrow BC$  satisfies BCNF because A is a candidate key.

$C \rightarrow DE$  does not satisfy BCNF because C isn't a candidate key, so we need to decompose.

Split ABCDE into CDE (everything in FD 2), and ABC (the original, minus the right side of FD 2).

Now we are left with ABC and CDE, and with this formulation, both FDs satisfy BCNF.

Decompose the following relation  $R=ABCD$  into BCNF.

$$F = \{$$
$$AB \rightarrow C$$
$$B \rightarrow D$$
$$C \rightarrow A$$
$$\}$$

First, compute attribute closures. We find that AB and BC are candidate keys.

Check each FD to see if they satisfy BCNF.

No violation in FD 1.

FD 2 violates BCNF, since B is not a candidate key.

Split ABCD into BD and ABC. (All attributes in FD 2 go in one relation; the right side of FD 2 is removed from the bigger relation to produce the other relation)

With this decomposition, FD 2 no longer violates BCNF.

FD 3 still violates BCNF, since C is not a candidate key.

Split ABC into CA and BC.

Now, all FDs either satisfy BCNF, or are no longer applicable (as in the case of FD 1; the three attributes it pertains to no longer exist in a single relation), so we are done.