

IE301  
Analysis and Design of Data Systems

Lecture 16

General Definitions of 2NF & 3NF  
Boyce-Codd Normal Form

Aram Keryan

November 4, 2015

# General Definitions of 2NF & 3NF

So far definitions of 2NF and 3NF were based on Primary Keys, hence the normalization procedure was useful in situations for a given database when the primary keys were already been defined.

Now, let's give definitions of 2NF and 3NF that take all candidate keys into account.

General definition of **prime attribute**:

An attribute that is part of *any candidate key* will be considered as prime

**Definition:**

If a functional dependency  $X \rightarrow Y$  holds true where  $Y$  is not a subset of  $X$  then this dependency is called **non-trivial** Functional dependency.

# General Definition of 2NF

## Definition based on primary key:

A relation schema  $R$  is in 2NF if every nonprime attribute  $A$  in  $R$  is *fully functionally dependent* on the primary key of  $R$ .

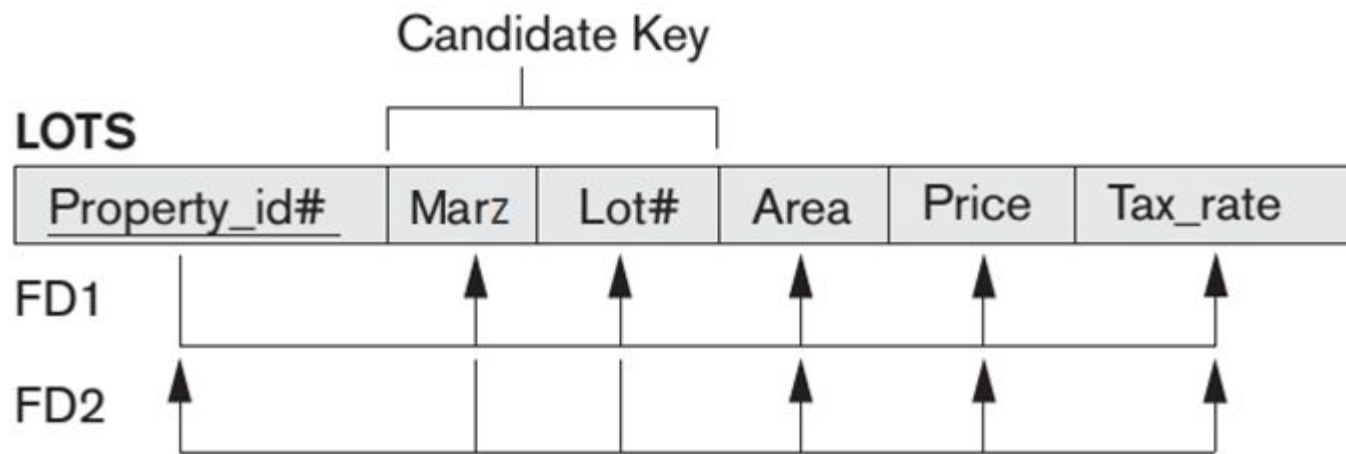
## General Definition:

A relation schema  $R$  is in **second normal form (2NF)** if every nonprime attribute  $A$  in  $R$  is *fully functionally dependent* on every key of  $R$ .

# Example

Relation LOTS describes pieces of land for sale in various Marzes of Armenia

- Lot numbers are unique only within each county
- Property\_id# numbers are unique across the country

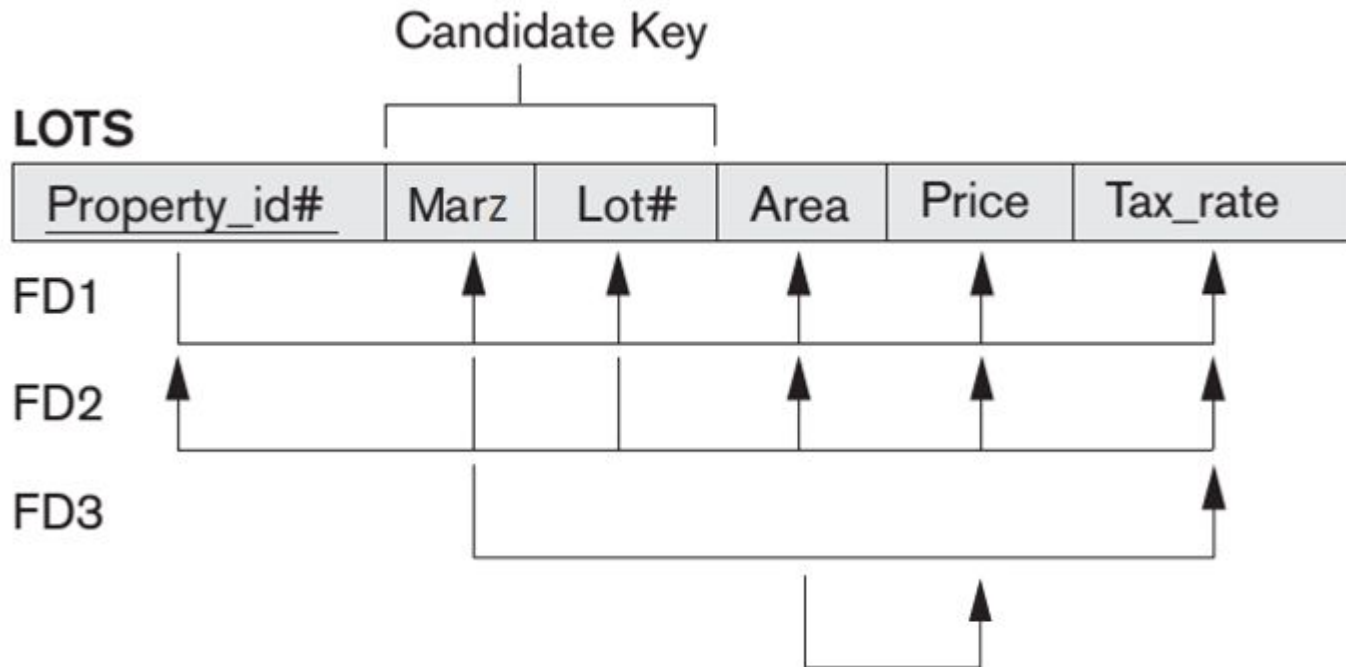


Since the primary key consists of only one attribute, it means that all the nonprime attributes are fully functionally dependent on the primary key

## Example (cont.)

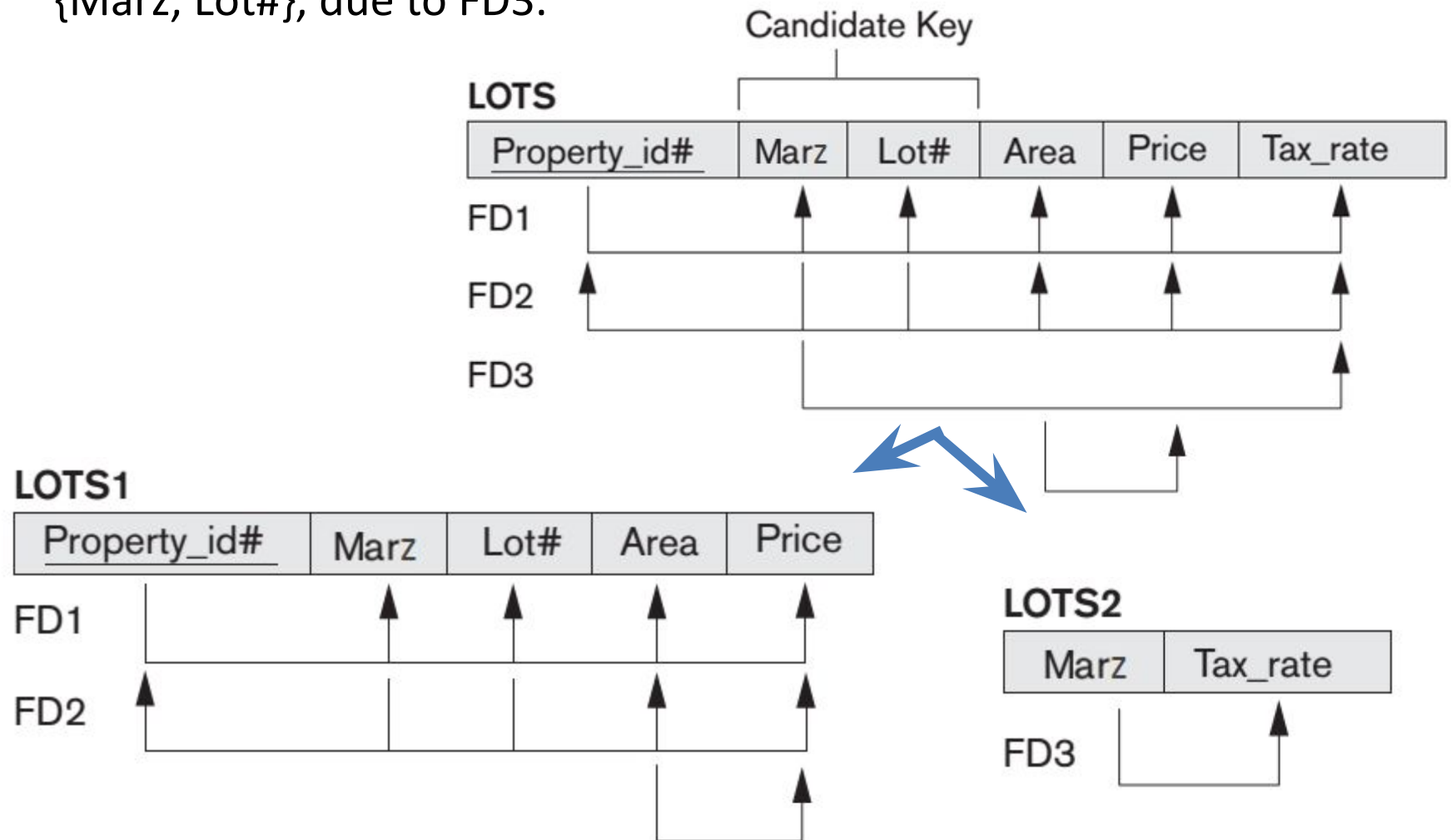
Suppose that the following two additional functional dependencies hold in LOTS: FD3:  $\text{Marz} \rightarrow \text{Tax\_rate}$ ,  $\text{Area} \rightarrow \text{Price}$

- ✓ FD3 says that the tax rate is fixed for a given Marz (does not vary lot by lot within the same Marz)
- ✓ FD4 says that the price of a lot is determined by its area regardless of which Marz it is in. (Assume that this is the price of the lot for tax purposes.)



## Example (cont.)

- The LOTS relation schema violates the general definition of 2NF because Tax\_rate is partially dependent on the candidate key {Marz, Lot#}, due to FD3.



# General Definition of 3NF

## Definition based on primary key:

A relation schema  $R$  is in **3NF** if it satisfies 2NF *and* no nonprime attribute of  $R$  is transitively dependent on the primary key.

## General Definition:

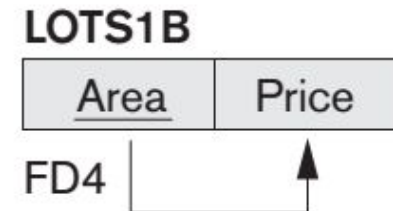
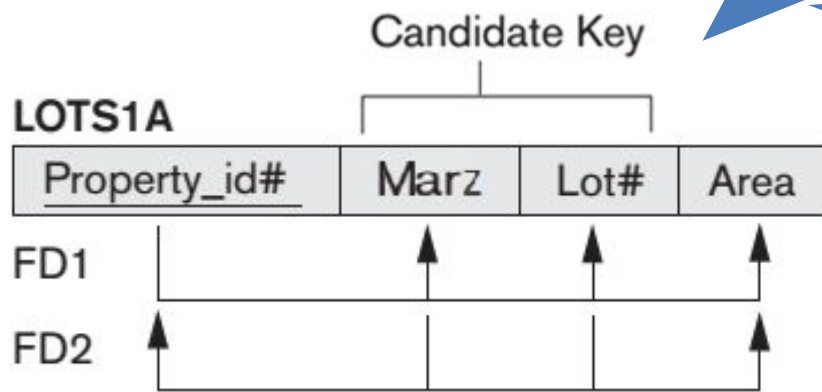
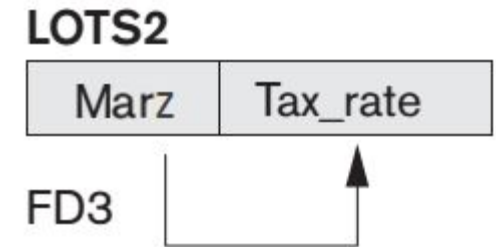
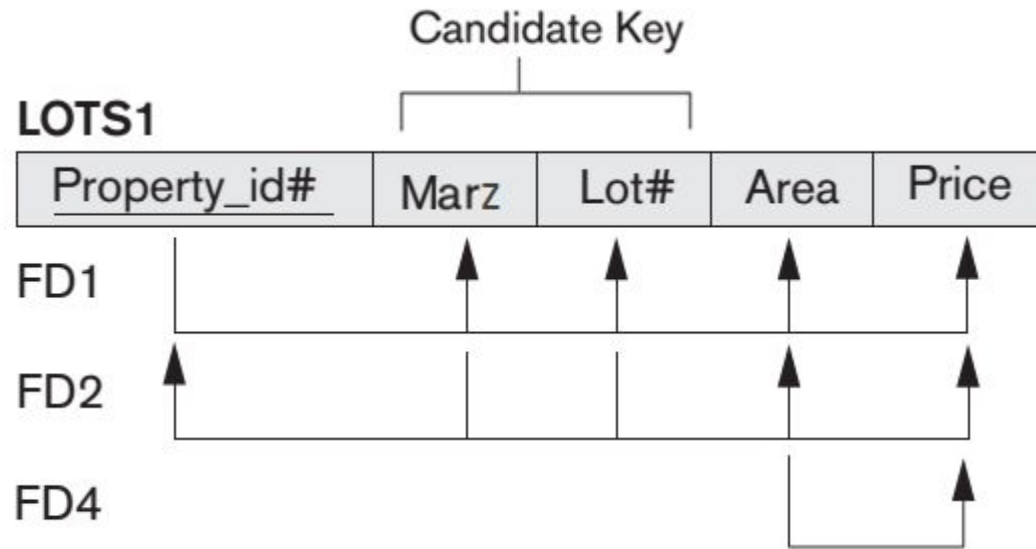
A relation schema  $R$  is in **3NF** if every nonprime attribute of  $R$  meets both of the following conditions:

- It is fully functionally dependent on every key of  $R$ .
- It is nontransitively dependent on every key of  $R$ .

## Alternative Definition:

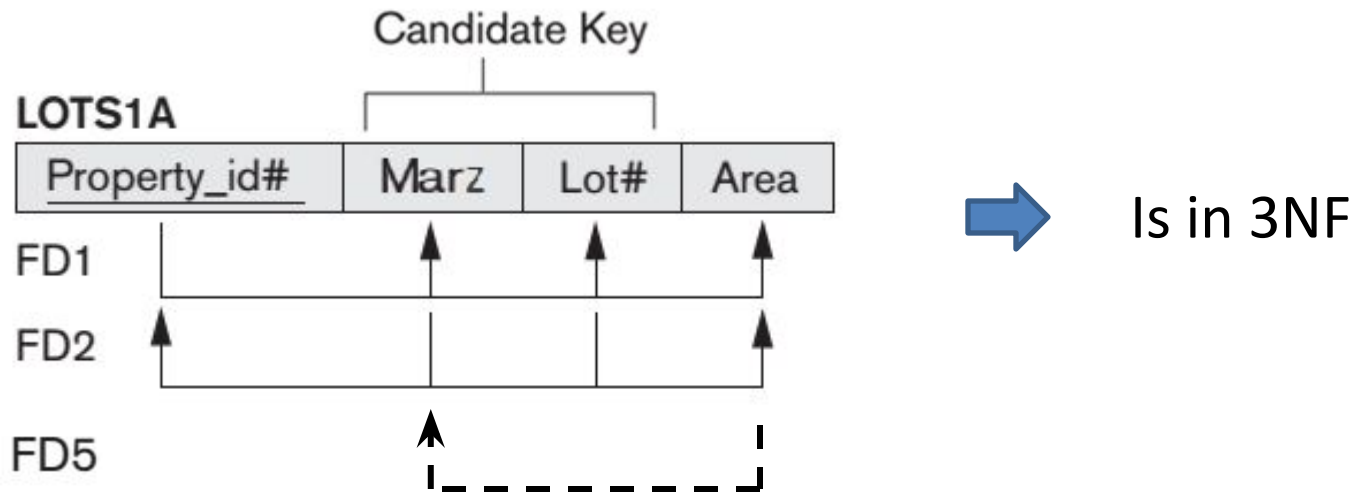
A relation schema  $R$  is in **3NF** if, whenever a *nontrivial* functional dependency  $X \rightarrow A$  holds in  $R$ , either (a)  $X$  is a superkey of  $R$ , or (b)  $A$  is a prime attribute of  $R$ .

# Example





# Boyce-Codd Normal Form (BCNF)



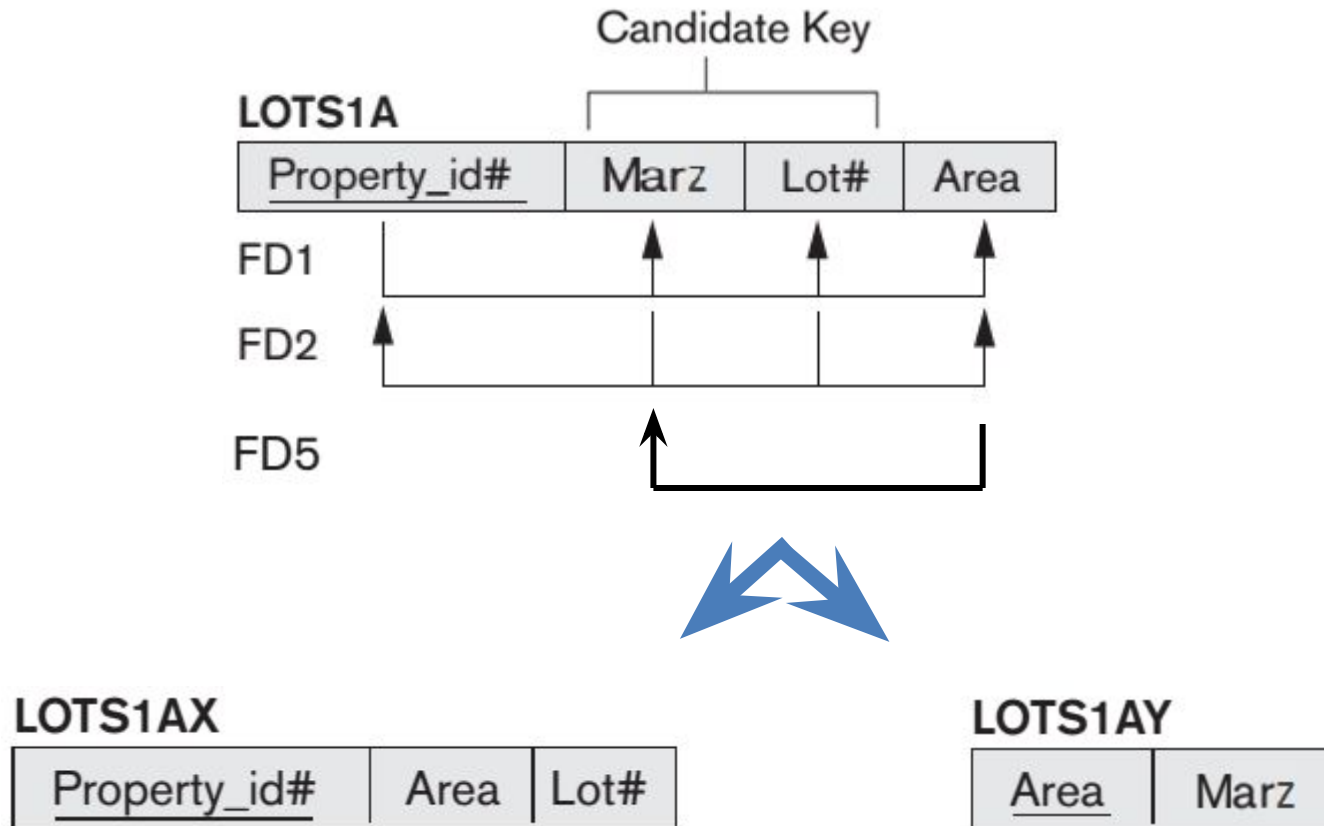
Let's imagine that:

- We have only two Marzes: Kotayk and Shirak
- Lot sizes in Kotayk marz are only 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0 hectares
- Lot sizes in Shirak marz are restricted to 1.1, 1.2, ..., 1.9, and 2.0 hectares

✓ In such a situation we would have new FD5: Area → County\_name

□ FD5 is a source of redundancy

# Boyce-Codd Normal Form (BCNF)



**Definition.** A relation schema  $R$  is in **BCNF** if whenever a *nontrivial* functional dependency  $X \rightarrow A$  holds in  $R$ , then  $X$  is a superkey of  $R$ .