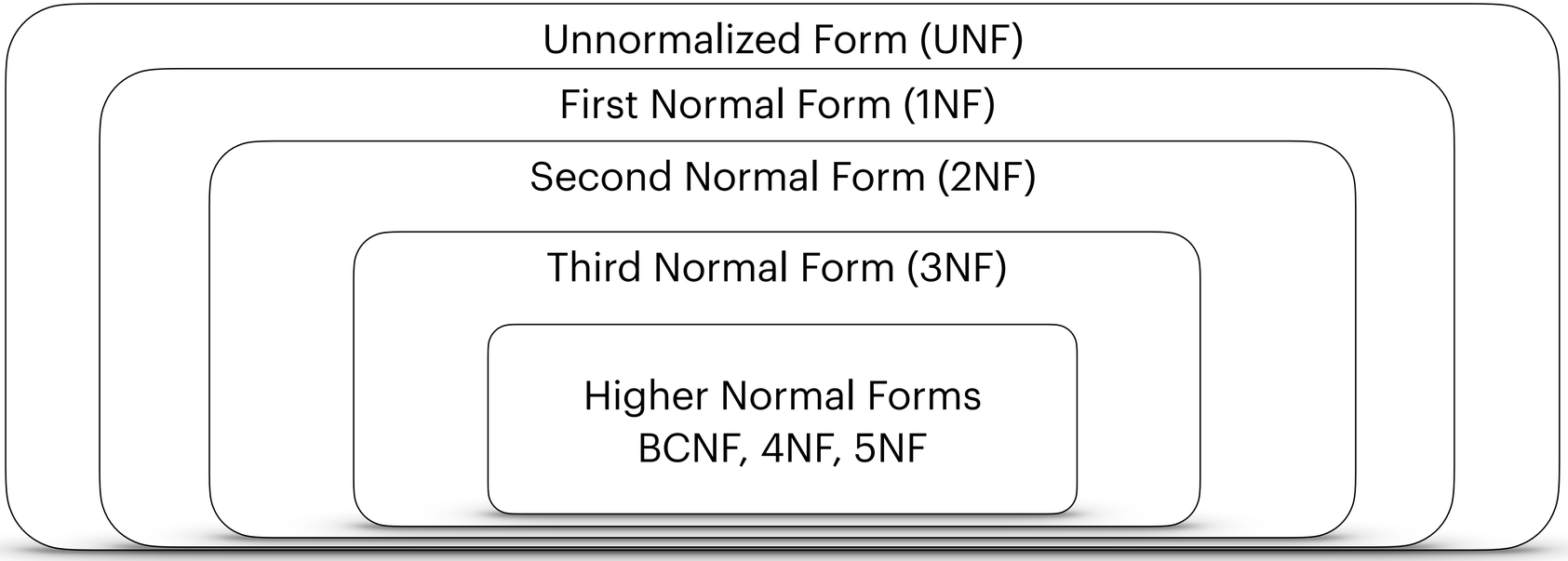
# Normalization Process

- Normalization Process is converting a relation from less restricted form to more restricted form.
- The level of restriction, is called Normal Form.



# Unnormalized Form (UNF)

- A table that contains one or more repeating groups.

| FirstName | LastName | DoB | Position | Department | StoreID |
|---|---|---|---|---|---|
| John | Ford | 1998/2/12 | Manager<br>Vice President | HR | #1506<br>#1545 |
| Anne | Brand | 2001/3/12 | Intern<br>Assistant | Marketing | #1546<br>#1506 |
| David | Biden | 2000/2/20 | Assistant | Sales | #1524 |
| William | Potter | 2001/9/12 | Senior Manager | HR | #1506 |

# First Normal Form (1NF)

- A relation in which the intersection of each row and column contains one and only one value.
- This is the requirement of a relation.
  - Each cell of relation contains exactly one value.
- Relation R (A, B, C, D, E, F, G) has following FDs:
  - FD1: A, B, C → D, E, F, G
  - FD2: A → D
  - FD3: B, C → E
  - FD4: F → G
- Relation R is in 1NF.

# UNF to 1NF

- We need to eliminate the repeating values.

| FirstName | LastName | DoB | Position | Department | StoreID |
|---|---|---|---|---|---|
| John | Ford | 1998/2/12 | Manager | HR | #1506 |
| John | Ford | 1998/2/12 | Vice President | HR | #1545 |
| Anne | Brand | 2001/3/12 | Intern | Marketing | #1546 |
| Anne | Brand | 2001/3/12 | Assistant | Marketing | #1506 |
| David | Biden | 2000/2/20 | Assistant | Sales | #1524 |
| William | Potter | 2001/9/12 | Senior Manager | HR | #1506 |

# Second Normal Form (2NF)

- A relation is in 2NF if:
  - It is in 1NF
  - Every non-primary-key attribute is fully functionally dependent on the primary key.
  - In other words, if there is any non-primary-key attribute is partially functionally dependent on the primary key, a relation is not in 2NF.

# Example

- Relation R (A, B, C, D, E, F, G) has following FDs:
  - FD1: A, B, C → D, E, F, G
  - FD2: A → D
  - FD3: B, C → E
  - FD4: F → G
- Is R in 2NF?
- No.
  - Because FD2: A is a proper subset of primary key (A, B, C), and A → D. Thus, D is partially dependent on (A, B, C).
  - Also FD3: (B, C) is a proper subset of primary key (A, B, C), and B, C → E. Thus, E is partially dependent on (A, B, C).

# 1NF to 2NF

- We need to eliminate the partial FDs.
- Since in relation R, FD2, and FD3 lead to partial FDs, we need to move them out to new relations.
- Determinants will be primary keys in new relations; and will be foreign keys in original relations.

# 1NF to 2NF

- Relation R (A, B, C, D, E, F, G):
  - FD1: A, B, C → D, E, F, G
  - FD2: A → D
  - FD3: B, C → E
  - FD4: F → G
- Relation R (A(fk), B(fk), C(fk), ~~D~~, ~~E~~, F, G):
  - FD1: A, B, C → ~~D, E,~~ F, G
  - ~~FD2: A → D~~
  - ~~FD3: B, C → E~~
  - FD2: F → G
- Relation R1 (A, D):
  - FD1: A → D
- Relation R2 (B, C, E):
  - FD1: B, C → E

# Third Normal Form (3NF)

- A relation is in 3NF if:
  - It is in 2NF
  - No non-primary-key attribute is transitively dependent on the primary key.
  - In other words, if there is any non-primary-key attribute is transitively dependent on the primary key, a relation is not in 3NF.

# Example

- Relation R (A, B, C, F, G) has following FDs:
  - FD1: A, B, C → F, G
  - FD2: F → G
- Is R in 3NF?
- No.
  - Because FD2: G is a non-primary-key attribute, A, B, C → F, and F → G. Thus, G is is transitively dependent on (A, B, C).

# 2NF to 3NF

- We need to eliminate the transitive FDs.
- Since in relation R, FD2 lead to transitive FDs, we need to move them out to new relations.
- Determinants will be primary keys in new relations; and will be foreign keys in original relations.

# 2NF to 3NF

- Relation R (A, B, C, F, G):
  - FD1: A, B, C → F, G
  - FD2: F → G

- Relation R (A(fk), B(fk), C(fk), F(fk), ~~G~~):
  - FD1: A, B, C → F, ~~G~~
  - ~~FD2: F → G~~

- Relation R3 (F, G):
  - FD1: F → G

# Practice

- Let's do more practice in Lab3.

# Assignment

- Let's do the assignment to assess your understanding of Functional Dependencies, Normal Forms, and Normalization Process.

# Congratulations

- Now you finished Module 3!
- You should be comfortable to lay out the Entity Relationship Model, convert it to a Relational Model, and normalize it to 3NF.
- Now you will be able to implement it as a database!
- Next module, we are going to do a case study.
- See you soon!