

What's covered here?

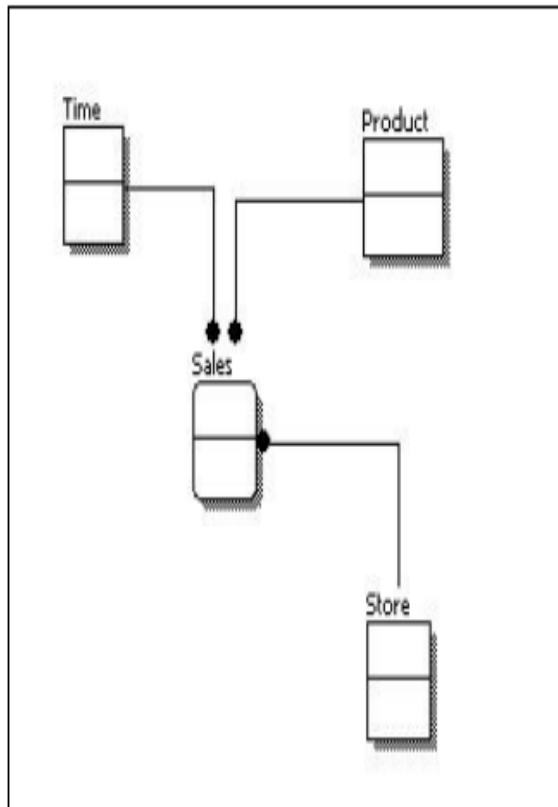
- Normalization

Purpose of Normalization

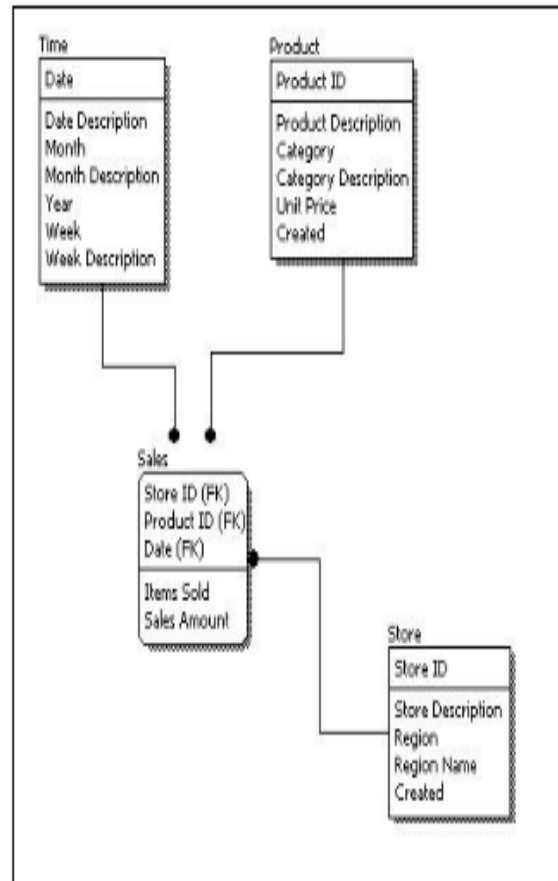
- Produce a set of suitable entities that support the data requirements of an enterprise
- Minimize data redundancy
- Provide reporting flexibility
- Maintain data integrity

Database Design Approach

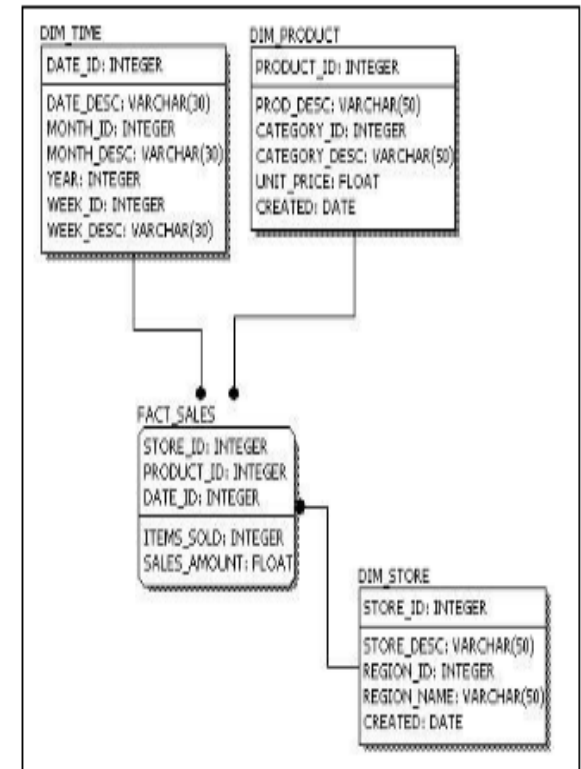
Conceptual Model Design



Logical Model Design



Physical Model Design



Functional Dependency

Employee ID	Employee Name
101	John
102	Sara
103	Daniel
104	John
105	Chris

Functionally dependent on Employee ID

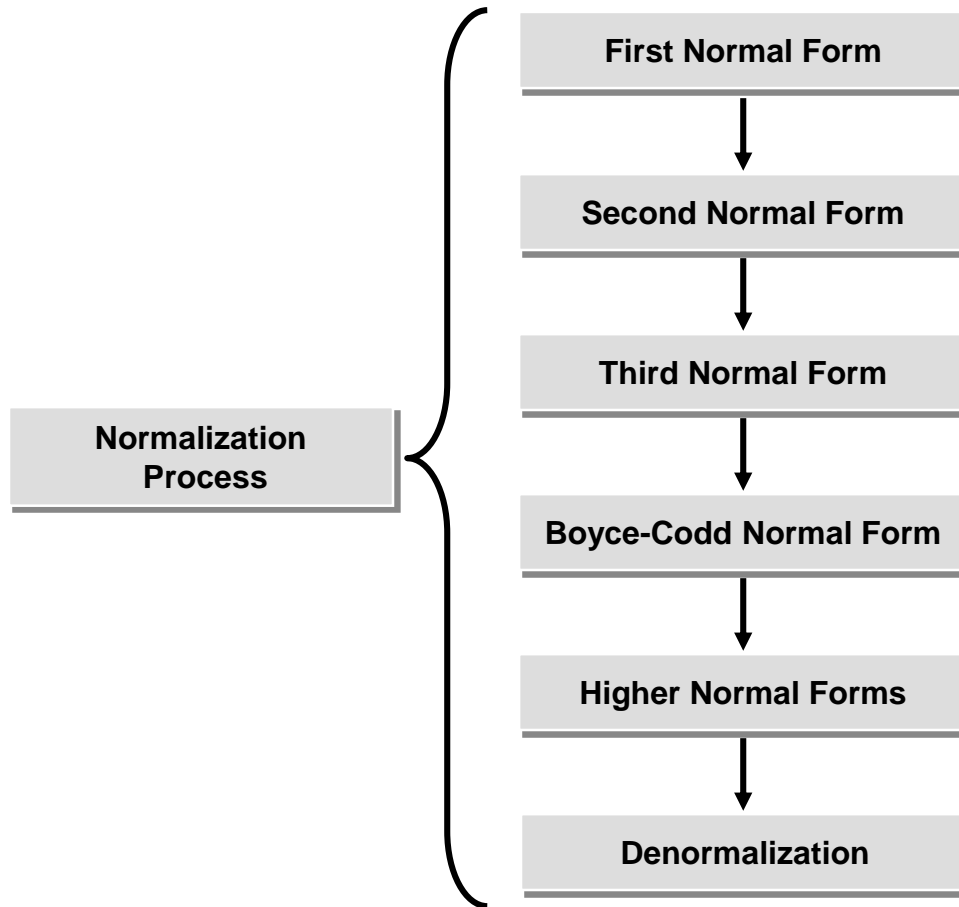
Employee ID → Employee Name

Employee Name ~~→~~ Employee ID

Lossless-join and Dependency Preservation Properties

- Two important properties of decomposition.
 - *Lossless-join property* enables us to find any instance of the original relation from corresponding instances in the smaller relations.
 - *Dependency preservation property* enables us to enforce a constraint on the original relation by enforcing some constraint on each of the smaller relations.

The Normalization Process



First Normal Form

- Remove multi-valued attributes
- Remove composite attributes
- Remove repeating groups
- Remove many-to-many relationship

Associative Entity

Each term students enroll in many courses

Students

Student Name	Contact	Courses
Amity Rogers	661-555-9848	Math103
Lara Steele	661-555-8192	Math103
Phyllis Frank	661-555-0342	Phy111
Zena Lane	661-555-0934	Phy111
Brenda Michael	661-555-4881	Chem101
Amity Rogers	661-555-9848	Chem101
Sandra Dale	661-555-6270	Bio121
Phyllis Frank	661-555-0342	Bio121
Zena Lane	661-555-0934	Math103
Ora Stevenson	661-555-3503	Bio121
Rina Lowery	661-555-3682	Phy111
Brenda Michael	661-555-4881	Bio121

Courses

Courses	Student Name
Math103	Amity Rogers
	Lara Steele
	Kalia Wilcox
Phy111	Phyllis Frank
	Zena Lane
	Ora Stevenson
Chem101	Rina Lowery
	Brenda Michael
	Amity Rogers
Bio121	Sandra Dale
	Phyllis Frank

Enrollments

Course ID	Student ID
C01	101
C01	109
C01	112
C01	110
C01	105
C02	123
C02	105
C02	107
C02	129
C03	129
C03	121
C03	101
C03	109
C03	112
C04	110
C04	123
C04	117
C04	121

Students

Student ID	Student Name	Contact
101	Amity Rogers	661-555-9848
109	Lara Steele	661-555-8192
112	Kalia Wilcox	661-555-8294
110	Sandra Dale	661-555-6270
123	Phyllis Frank	661-555-0342
105	Zena Lane	661-555-0934
117	Ora Stevenson	661-555-3503
129	Rina Lowery	661-555-3682
121	Brenda Michael	661-555-4881

Courses

Course ID	Course Description
C01	Math103
C02	Phy111
C03	Chem101
C04	Bio121

With bridge table

The First Normal Form

Logically repeating attributes

Employees

Before 1NF

Employee ID	Employee Name	Project 1 ID	Project 2 ID	Project 3 ID
123	Dave Thompson	112-253-01	112-478-03	
124	Burt Rios	112-568-08	113-235-19	165-203-35
125	Alice Parker	106-564-12		

The First Normal Form

Employees

EmployeeID	Employee First Name	Employee Last Name
123	Dave	Thompson
124	Burt	Rios
125	Alice	Parker

Projects

Project ID
112-253-01
112-478-03
112-568-08
113-235-19
165-203-35
106-564-12

Deployment

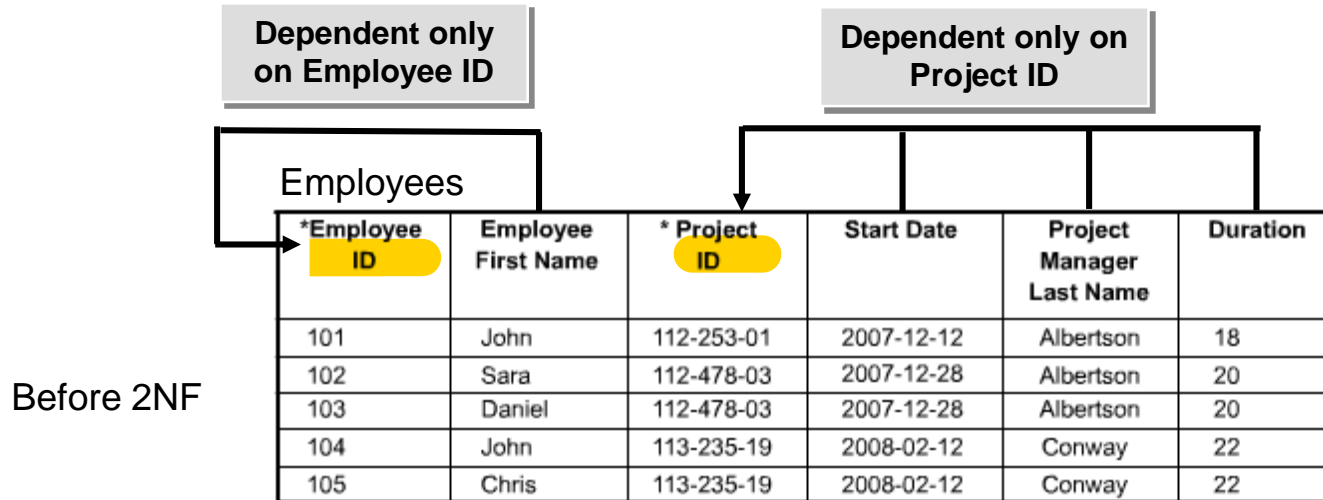
Employee ID	Project ID
123	112-253-01
123	112-478-03
124	112-568-08
124	113-235-19
124	165-203-35
125	106-564-12

After 1NF

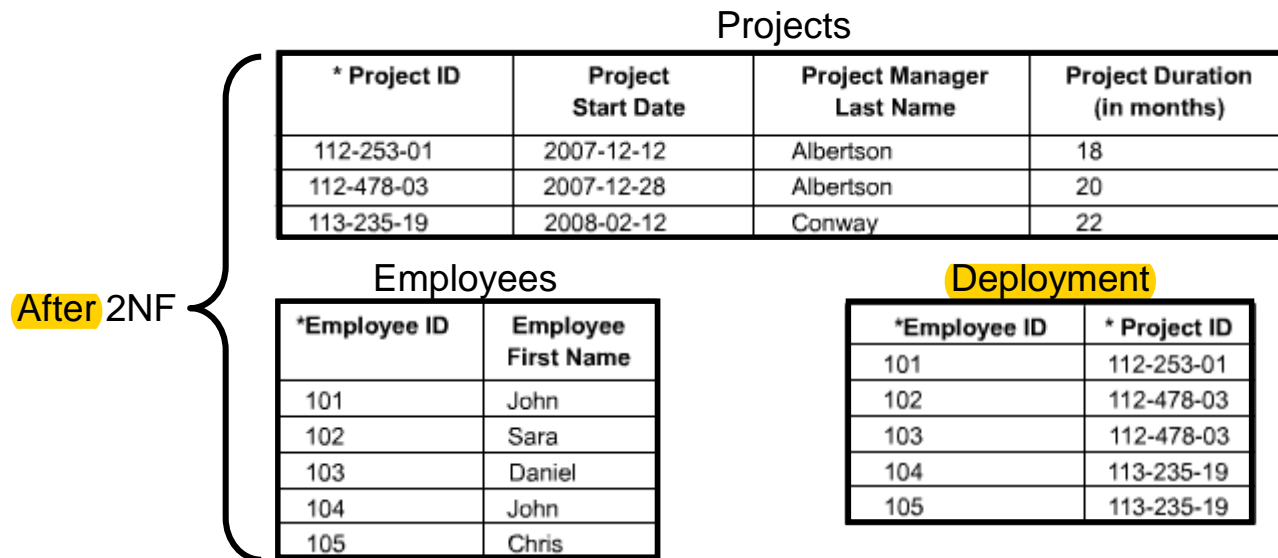
Second Normal Form

- Remove partial dependencies

The Second Normal Form



The Second Normal Form



Third Normal Form

- Remove transitive dependencies

Transitive Dependency

Manager first and last name are dependent on Manager ID, a non-key field and not on the Employee ID, the key attribute

*Employee ID	Employee First Name	Employee Last Name	Manager ID	Manager First Name	Manager Last Name
101	John	Allen	201	Paul	Wright
102	Sara	Duran	203	Marsha	Albertson
103	Daniel	Friedman	234	Paul	Wright
104	John	Baker	556	Cindy	Conway
105	Chris	Spencer	556	Cindy	Conway

Primary key

Non-key

Non-key

Non-key

Non-key

Non-key

The Third Normal Form

*Employee ID	Employee First Name	Employee Last Name	Manager ID	Manager First Name	Manager Last Name	Manager Mid Initial
101	John	Allen	201	Paul	Wright	R
102	Sara	Duran	203	Marsha	Albertson	A
103	Daniel	Friedman	234	Paul	Wright	K
104	John	Baker	556	Cindy	Conway	B
105	Chris	Spencer	556	Cindy	Conway	B

Before 3NF

Deleting the details for Daniel Friedman
will remove details of Paul Wright as well

The Third Normal Form

After 3NF

Removing transitive dependency
prevents loss of data

*Employee ID	Employee First Name	Employee Last Name	Manager ID
101	John	Allen	201
102	Sara	Duran	203
103	Daniel	Friedman	234
104	John	Baker	556
105	Chris	Spencer	556

Manager ID	Manager First Name	Manager Last Name	Manager Mid Initial
201	Paul	Wright	R
203	Marsha	Albertson	A
234	Paul	Wright	K
556	Cindy	Conway	B