X THE PLACE OF USEFUL LEARNING

University of Strathclyde Glasgow

Database Fundamentals – CS990

Database and Web Systems Development - CS952

Image source: https://www.strath.ac.uk/branding/

Course Content

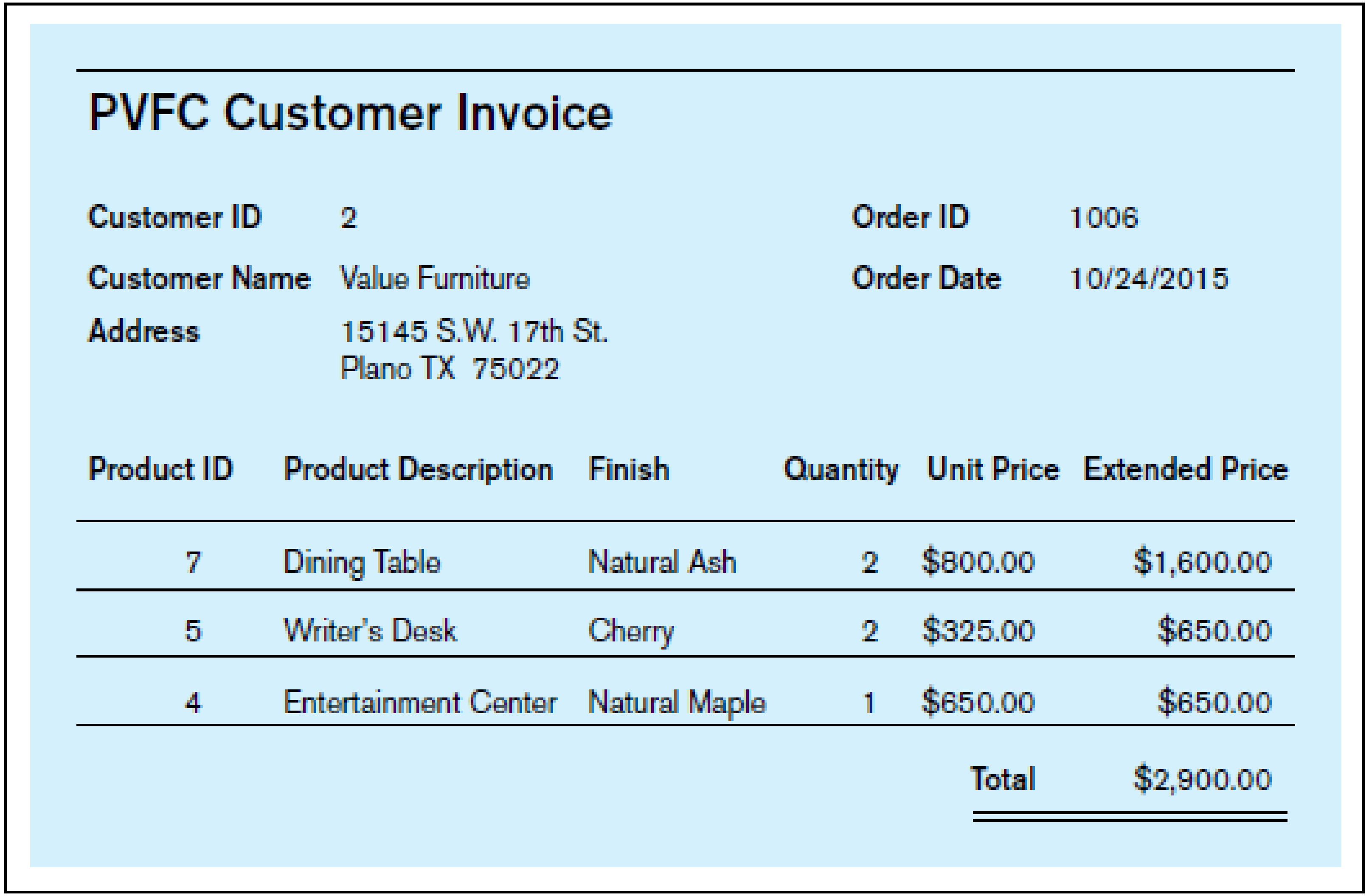
1. Introduction to Relational Databases *(Introduction + Relational Model)*
2. Data Modelling - *(Entity Relationship Modelling + The Enhanced Entity Relationship Model)*
3. Database Design and SQL - *(Logical modelling + Introduction to SQL)*
4. Further SQL - *(Advanced SQL queries + Creating tables with SQL)*
5. Normalisation - *(Normalisation to second normal form + Third normal form)*

Reference material

- Database, A Practical Approach to Design, Implementation, and Management (Thomas Connolly•Carolyn Begg)
- 387 – 412

- Modern Database Management (12 Edition) (Jeff Hoffer et al.)
- 214 -223

Pine Valley Furniture Company

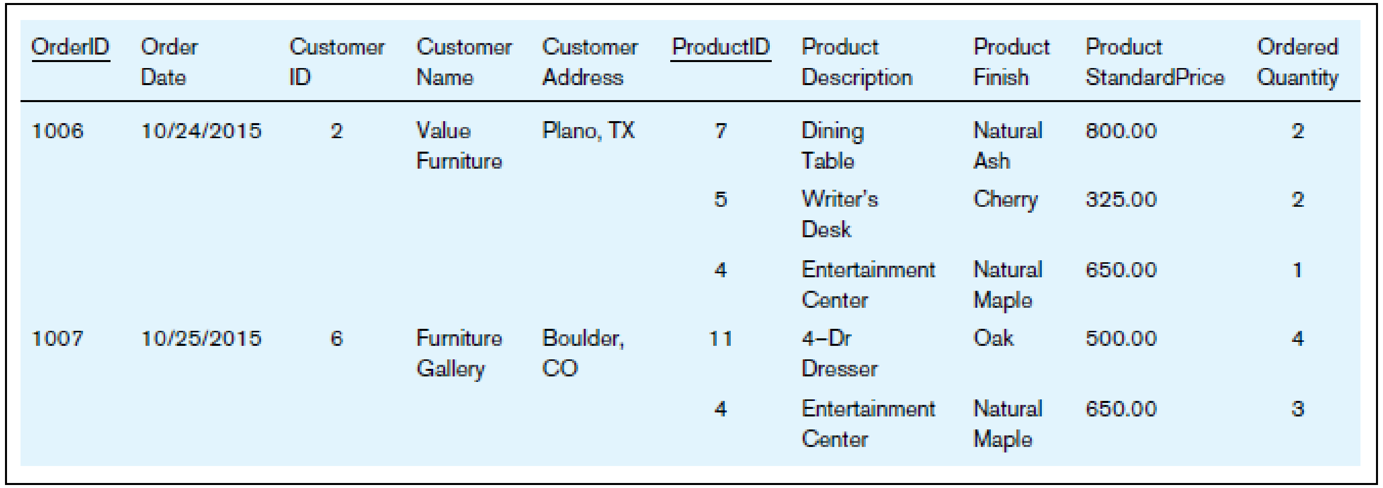


PVFC Customer Invoice

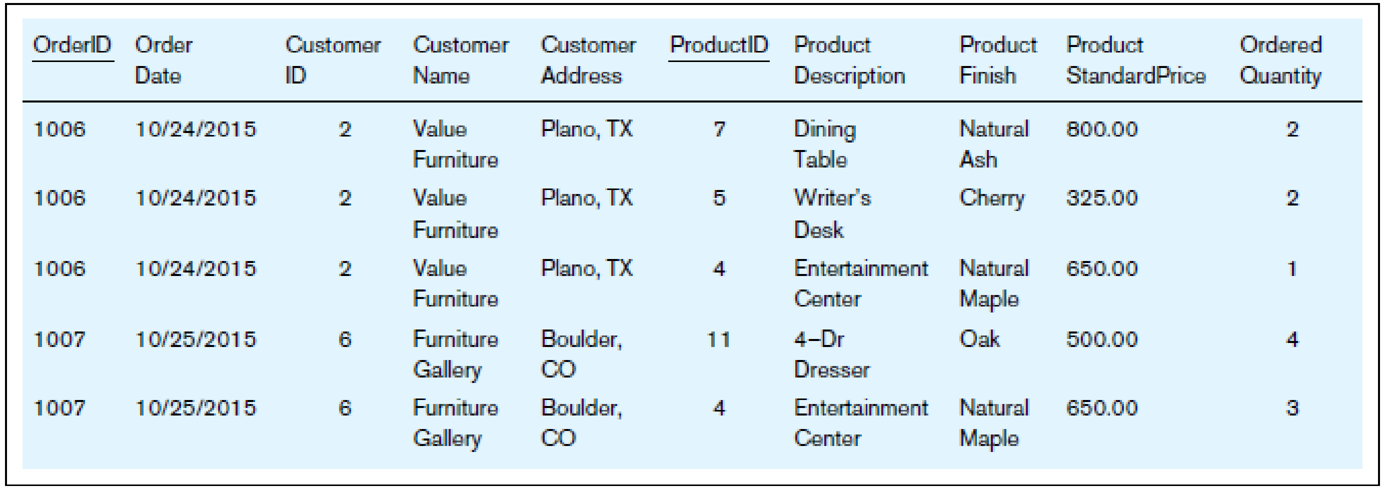
Customer ID	2	**Order ID**	1006
Customer Name	Value Furniture	**Order Date**	10/24/2015
Address	15145 S.W. 17th St. Plano TX 75022		

Product ID	Product Description	Finish	Quantity	Unit Price	Extended Price
7	Dining Table	Natural Ash	2	$800.00	$1,600.00
5	Writer's Desk	Cherry	2	$325.00	$650.00
4	Entertainment Center	Natural Maple	1	$650.00	$650.00
				Total	$2,900.00

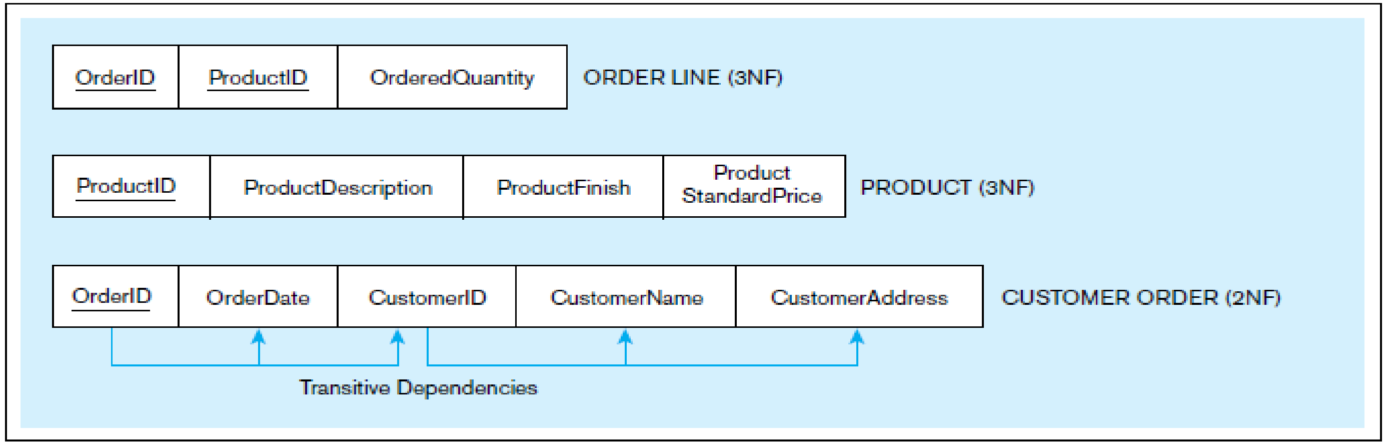
Modern Database Management *Twelfth Edition* Jeffrey A. Hoffer

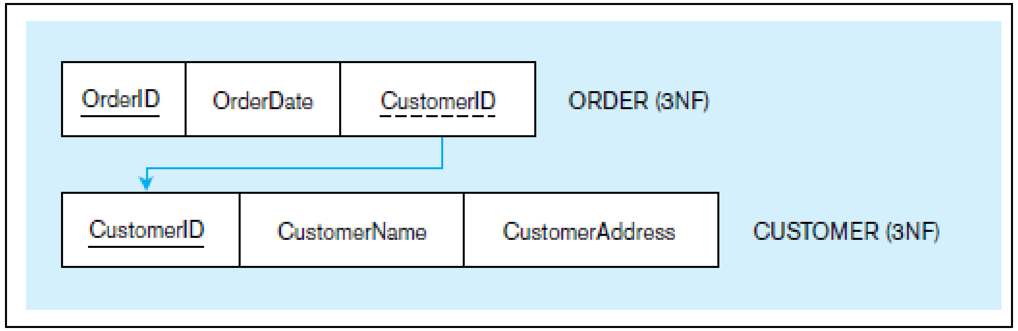


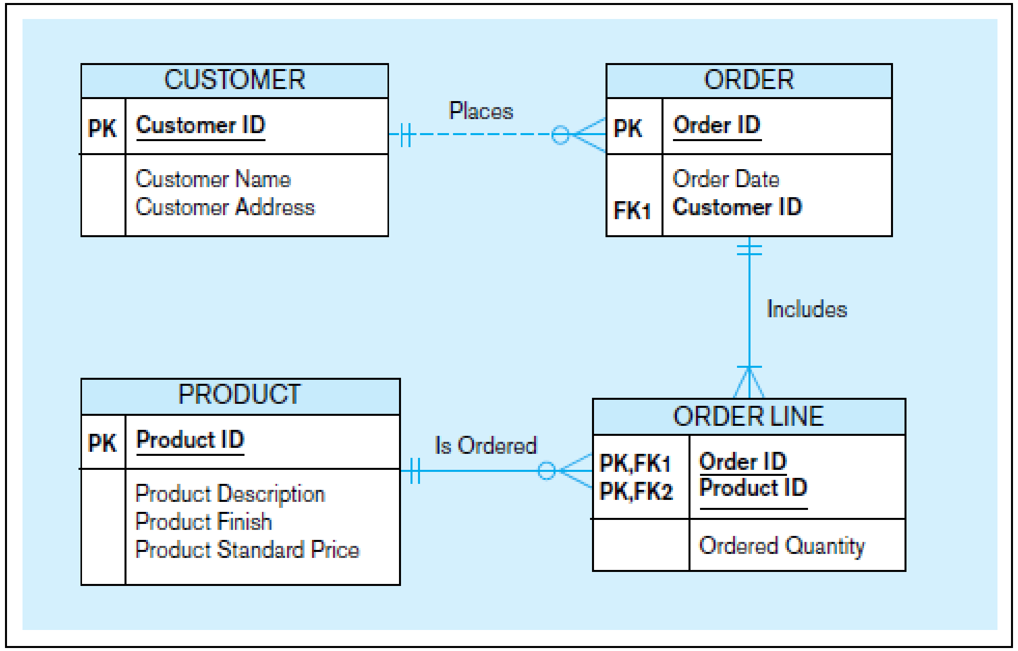
OrderID	Order Date	Customer ID	Customer Name	Customer Address	ProductID	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2015	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
					5	Writer's Desk	Cherry	325.00	2
					4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	11	4–Dr Dresser	Oak	500.00	4
					4	Entertainment Center	Natural Maple	650.00	3



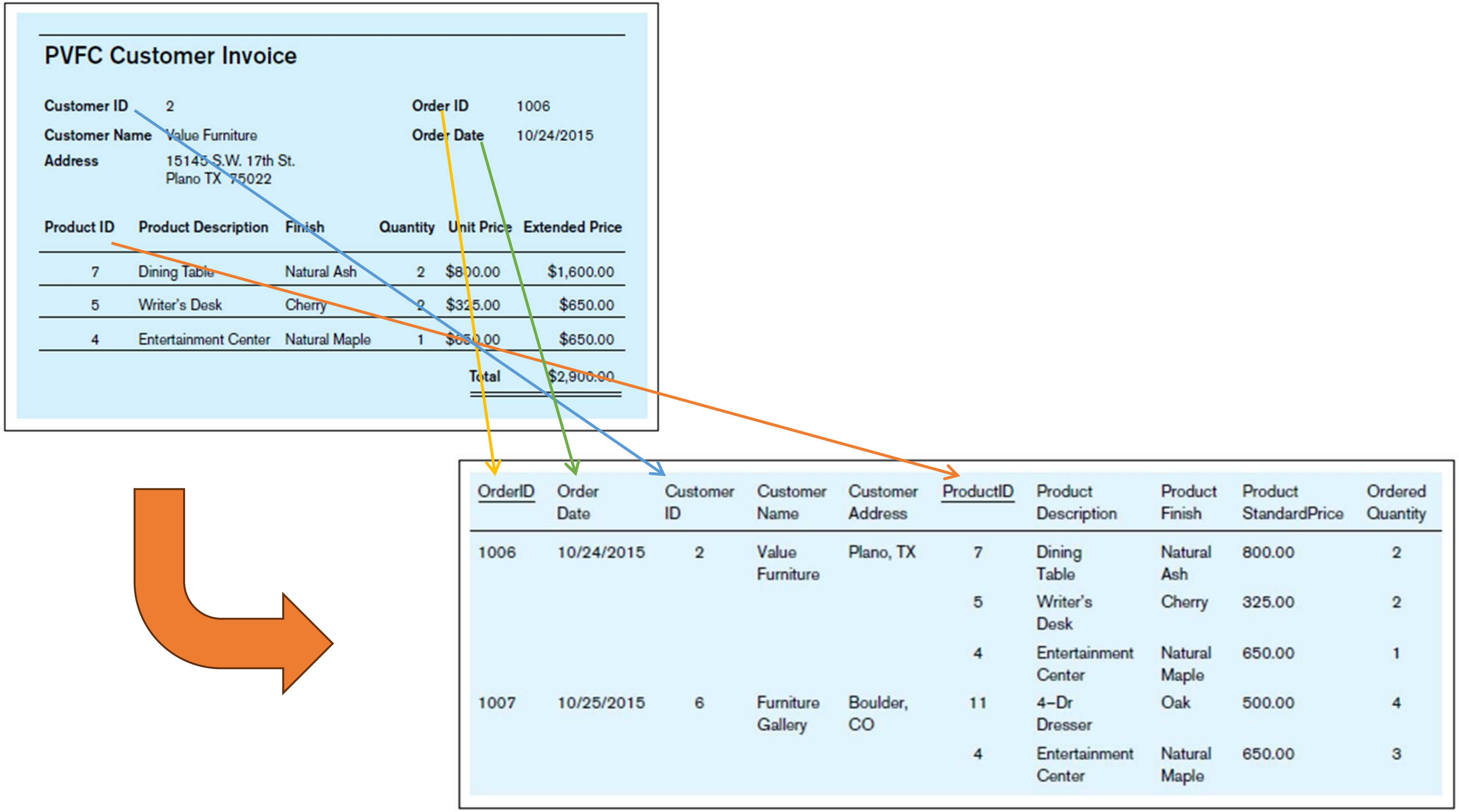
OrderID	Order Date	Customer ID	Customer Name	Customer Address	ProductID	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2015	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	5	Writer's Desk	Cherry	325.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	11	4–Dr Dresser	Oak	500.00	4
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	4	Entertainment Center	Natural Maple	650.00	3







Step 0: Represent the View in Tabular Form



PVFC Customer Invoice

Customer ID	2	**Order ID**	1006
Customer Name	Value Furniture	**Order Date**	10/24/2015
Address	15145 S.W. 17th St. Plano TX 75022		

Product ID	Product Description	Finish	Quantity	Unit Price	Extended Price
7	Dining Table	Natural Ash	2	$800.00	$1,600.00
5	Writer's Desk	Cherry	2	$325.00	$650.00
4	Entertainment Center	Natural Maple	1	$650.00	$650.00
				Total	$2,900.00

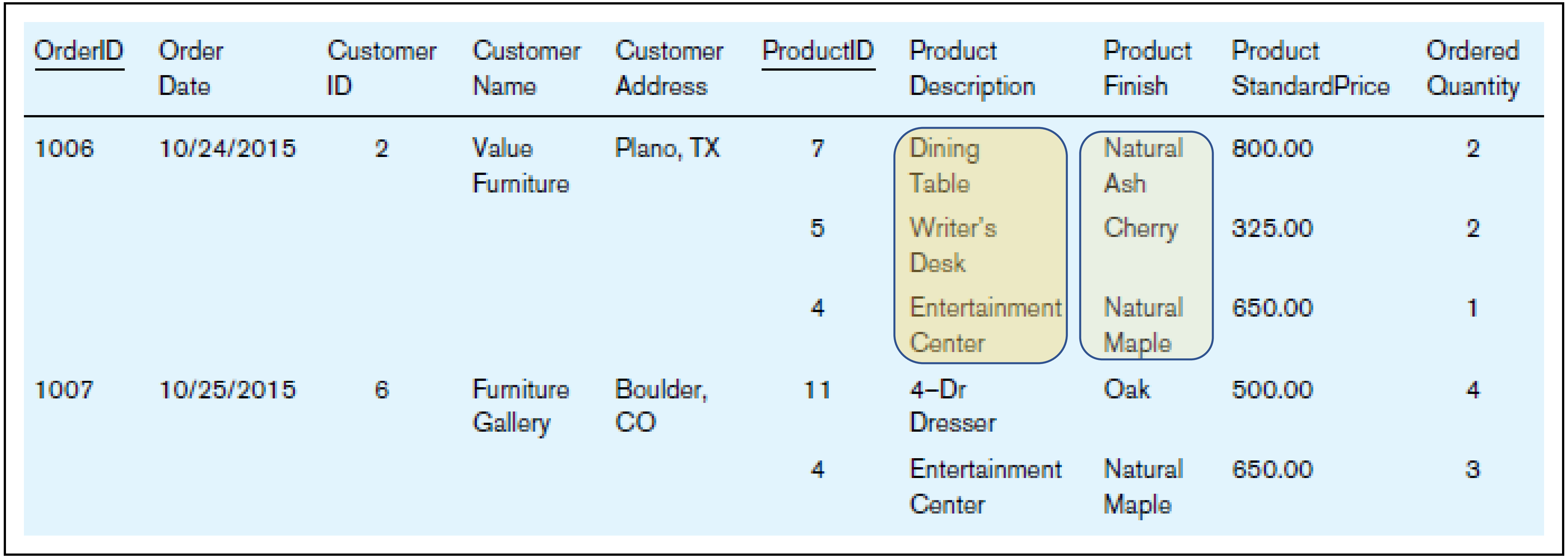
OrderID	Order Date	Customer ID	Customer Name	Customer Address	ProductID	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2015	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
					5	Writer's Desk	Cherry	325.00	2
					4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	11	4–Dr Dresser	Oak	500.00	4
					4	Entertainment Center	Natural Maple	650.00	3

Step 1: Convert to First Normal Form

A relation is in 1NF if two constraints both apply

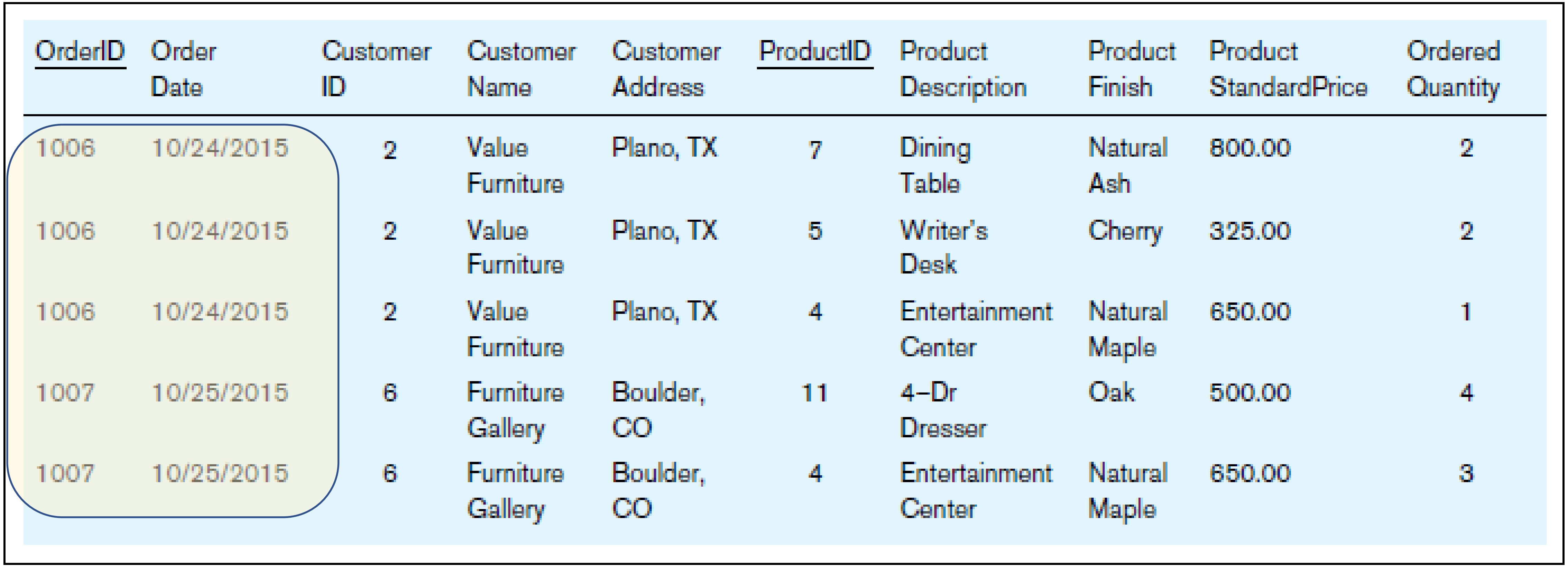
- *No Repeating Groups*
- *A Primary Key has been defined*

A relation that has a primary key and in which there are no repeating groups.



OrderID	Order Date	Customer ID	Customer Name	Customer Address	ProductID	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2015	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
					5	Writer's Desk	Cherry	325.00	2
					4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	11	4–Dr Dresser	Oak	500.00	4
					4	Entertainment Center	Natural Maple	650.00	3

Repeated Groups Removed



OrderID	Order Date	Customer ID	Customer Name	Customer Address	ProductID	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2015	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	5	Writer's Desk	Cherry	325.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	11	4–Dr Dresser	Oak	500.00	4
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	4	Entertainment Center	Natural Maple	650.00	3

Functional Dependencies

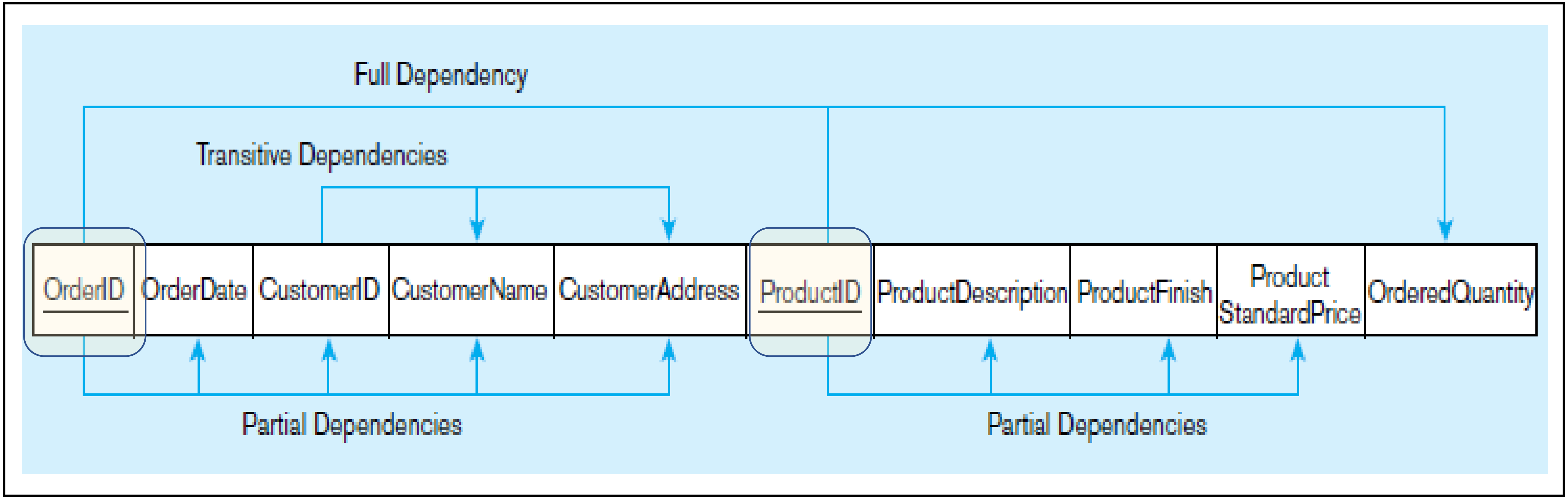
- For any relation R, ***attribute B*** is functionally dependent on **attribute A** if, for every valid instance of A, that value of A uniquely determines the value of B *(Dutka and Hanson, 1989).*
- **Typical examples of functional dependencies are the following:**

  1. SSN → *Name, Address, Birthdate:* A person's name, address, and birth date are functionally dependent on that person's Social Security number (in other words, there can be only one Name, one Address, and one Birthdate for each ***SSN***).

  2. **VIN** → *Make, Model, Color:* The make, model, and the original color of a vehicle are functionally dependent on the vehicle identification number (as above, there can be only one value of Make, Model, and Color associated with each ***VIN***).

  3. **ISBN** → *Title, FirstAuthorName, Publisher:* The title of a book, the name of the first author, and the publisher are functionally dependent on the book's international standard book number (ISBN).

Functional Dependencies



Depends on full key is **ONLY** OrderedQuantity

All other are either partial dependencies or transitive dependencies

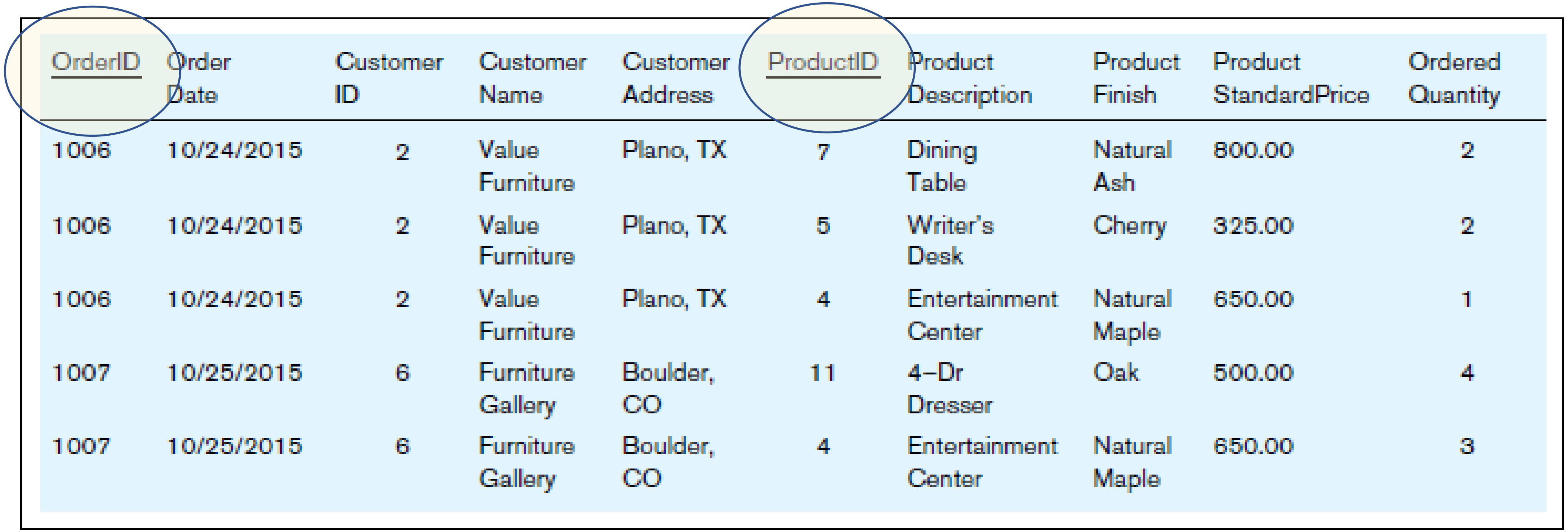
Determinants (Four)

- OrderID → OrderDate, CustomerID, CustomerName, CustomerAddress
- CustomerID → CustomerName, CustomerAddress
- ProductID → ProductDescription, ProductFinish, ProductStandardPrice

  OrderID, ProductID → OrderedQuantity
- Why do we know these are the functional dependencies?
  - These business rules come from the organization. We know these from studying the nature of the Pine Valley Furniture Company business

Candidate Key

- Only candidate key for INVOICE
- Attributes OrderID and ProductID (because there is only one row in the table for any combination of values for these attributes



OrderID	Order Date	Customer ID	Customer Name	Customer Address	ProductID	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2015	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	5	Writer's Desk	Cherry	325.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	11	4–Dr Dresser	Oak	500.00	4
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	4	Entertainment Center	Natural Maple	650.00	3

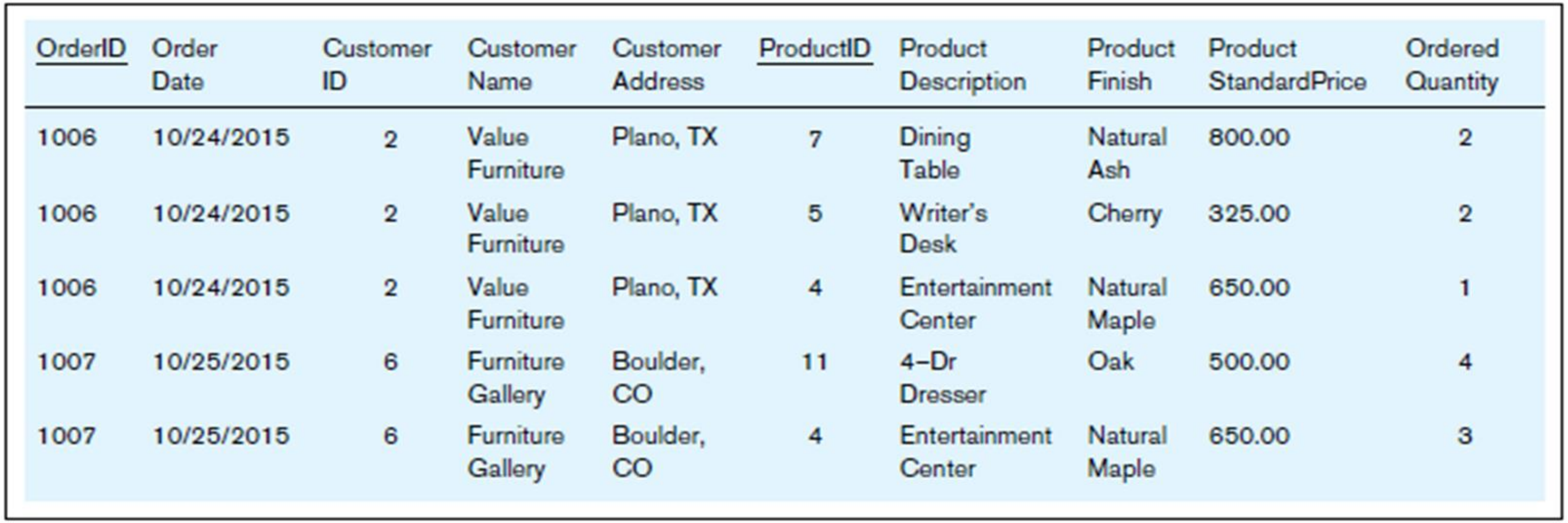
Anomalies in 1NF

- still contain considerable redundancy.
  - For example, *CustomerID, CustomerName, and CustomerAddress* for Value Furniture are recorded in three rows (at least) in the table.

- can lead to anomalies such as the following:
  - *a) Insertion anomaly*
  - *b) Deletion anomaly*
  - *c) Update anomaly*

Insert anomaly

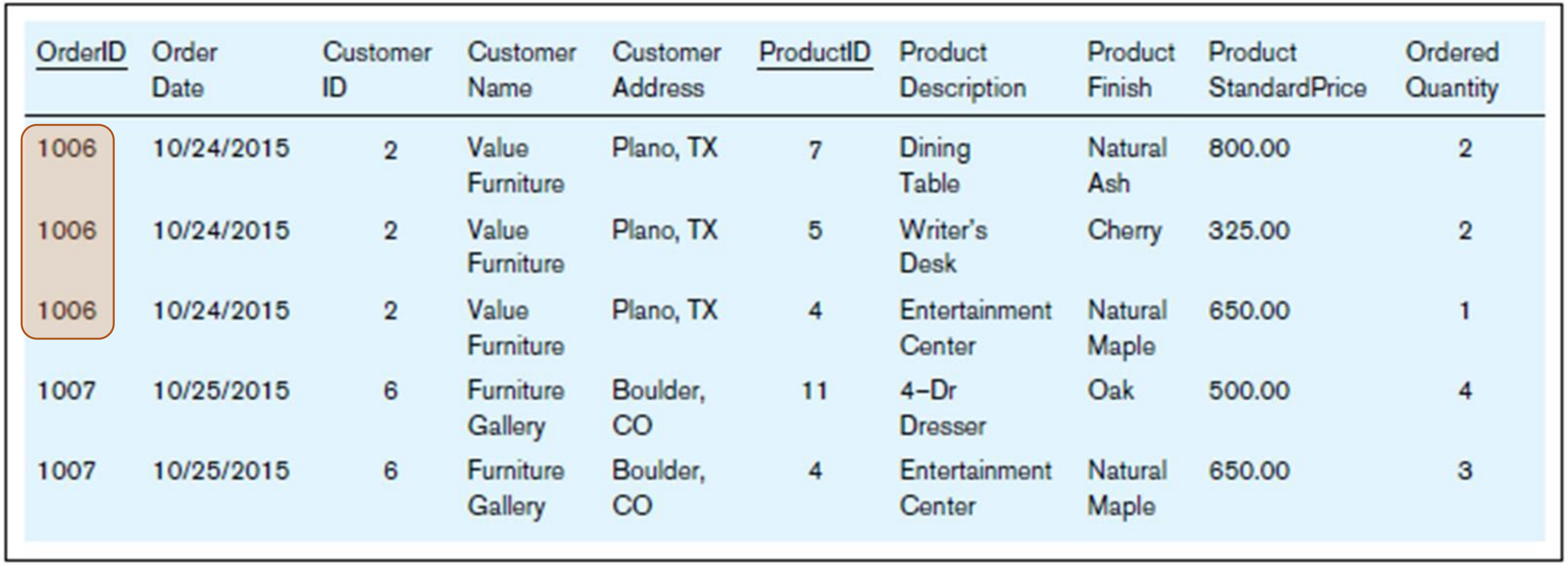
- With this table structure, the company is not able to introduce a new product (say, ***Breakfast Table*** with ***ProductID 8***) and add it to the database before it is ordered the first time:
- No entries can be added to the table without both *ProductID* and *OrderID*.



OrderID	Order Date	Customer ID	Customer Name	Customer Address	ProductID	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2015	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	5	Writer's Desk	Cherry	325.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	11	4–Dr Dresser	Oak	500.00	4
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	4	Entertainment Center	Natural Maple	650.00	3

Deletion Anomaly

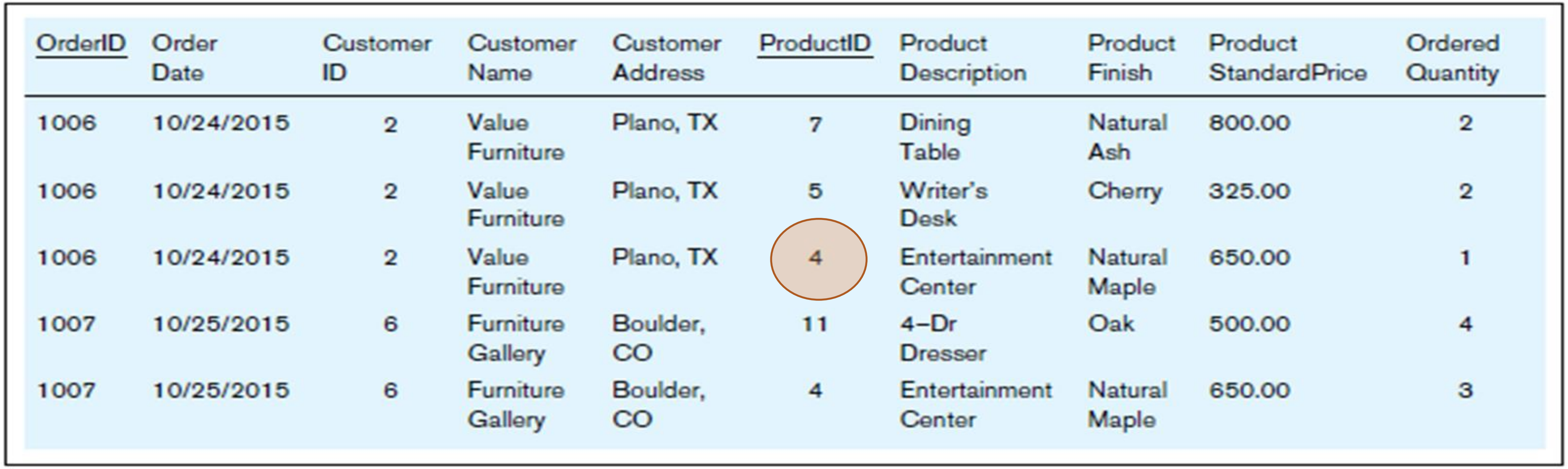
- If a customer calls and requests that the **Dining Table** be deleted from her *OrderID 1006*, this row must be deleted from the relation, and we lose the information concerning this item's finish (*Natural Ash*) and *price ($800.00).*



OrderID	Order Date	Customer ID	Customer Name	Customer Address	ProductID	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2015	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	5	Writer's Desk	Cherry	325.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	11	4–Dr Dresser	Oak	500.00	4
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	4	Entertainment Center	Natural Maple	650.00	3

Update Anomaly

- If Pine Valley Furniture (as part of a price adjustment) increases the price of the Entertainment Center (*ProductID 4*) to $750.00, this change must be recorded in all rows containing that item.
- There are two such rows in Figure:



OrderID	Order Date	Customer ID	Customer Name	Customer Address	ProductID	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2015	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	5	Writer's Desk	Cherry	325.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	11	4–Dr Dresser	Oak	500.00	4
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	4	Entertainment Center	Natural Maple	650.00	3

Step 2: Convert to Second Normal Form

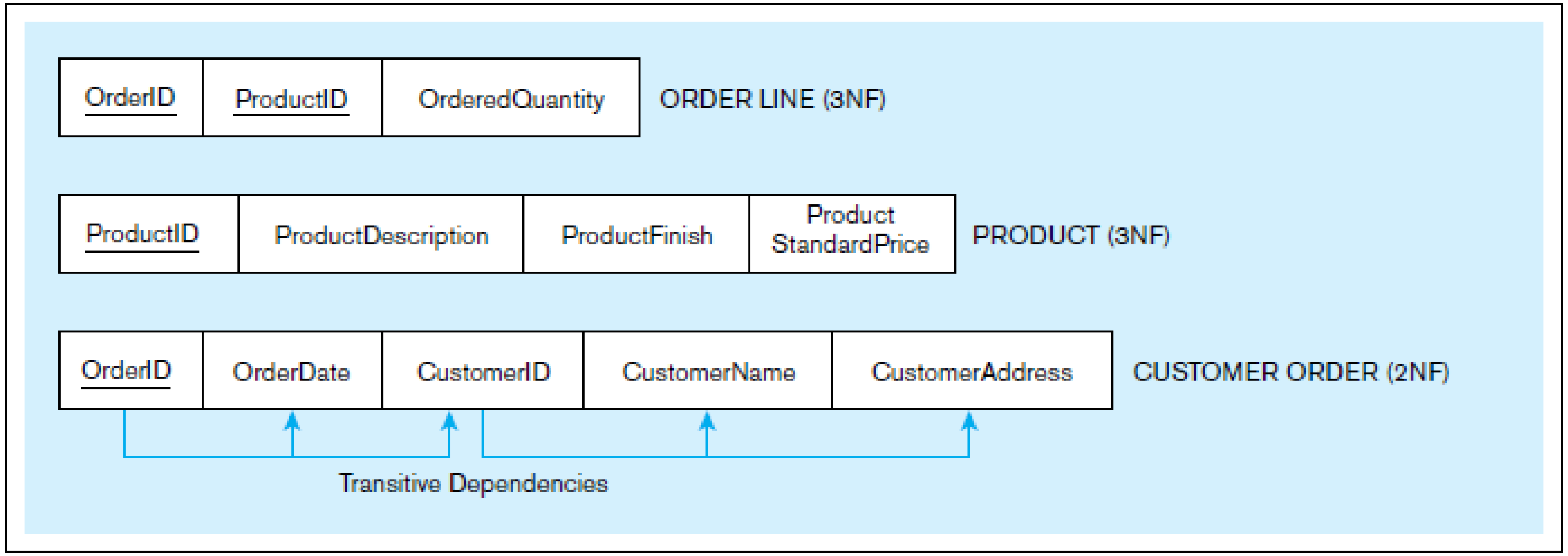
- A relation is in second normal form (2NF) if it is in ***first normal form*** and contains ***no partial functional dependencies***.
- A partial functional dependency exists when ***a non-key attribute*** *is* ***functionally dependent*** on part (but not all) of the primary key.

- Partial dependencies
  - **OrderID** → *OrderDate, CustomerID, CustomerName, CustomerAddress*
  - **ProductID** → *ProductDescription, ProductFinish, ProductStandardPrice*

  - *The first of these partial dependencies (for example) states that the date on an order is uniquely determined by the order number and has nothing to do with the ProductID.*

Conversion with partial dependencies to second normal form.

- Following steps are required:
  - *Create a **new relation** for each primary key attribute (or combination of attributes) that is a **determinant** in a partial dependency. That attribute is the primary key in the new relation.*
  - *Move the non-key attributes that are only dependent on this primary key attribute (or attributes) from the old relation to the new relation.*



OrderID	ProductID	OrderedQuantity	ORDER LINE (3NF)

ProductID	ProductDescription	ProductFinish	Product StandardPrice	PRODUCT (3NF)

OrderID	OrderDate	CustomerID	CustomerName	CustomerAddress	CUSTOMER ORDER (2NF)

In 2NF, if "One of the following conditions"

- Second normal form if any one of the following conditions applies:
  1. The primary key consists of only one attribute (e.g., the attribute ProductID in the PRODUCT relation). By definition, there cannot be a partial dependency in such a relation.
  2. No non-key attributes exist in the relation (thus all of the attributes in the relation are components of the primary key). There are no functional dependencies in such a relation.
  3. Every non-key attribute is functionally dependent on the full set of primary key attributes (e.g., the attribute *OrderedQuantity* in the *ORDER LINE relation.*

- *ORDER LINE and PRODUCT are in third normal form.*

however

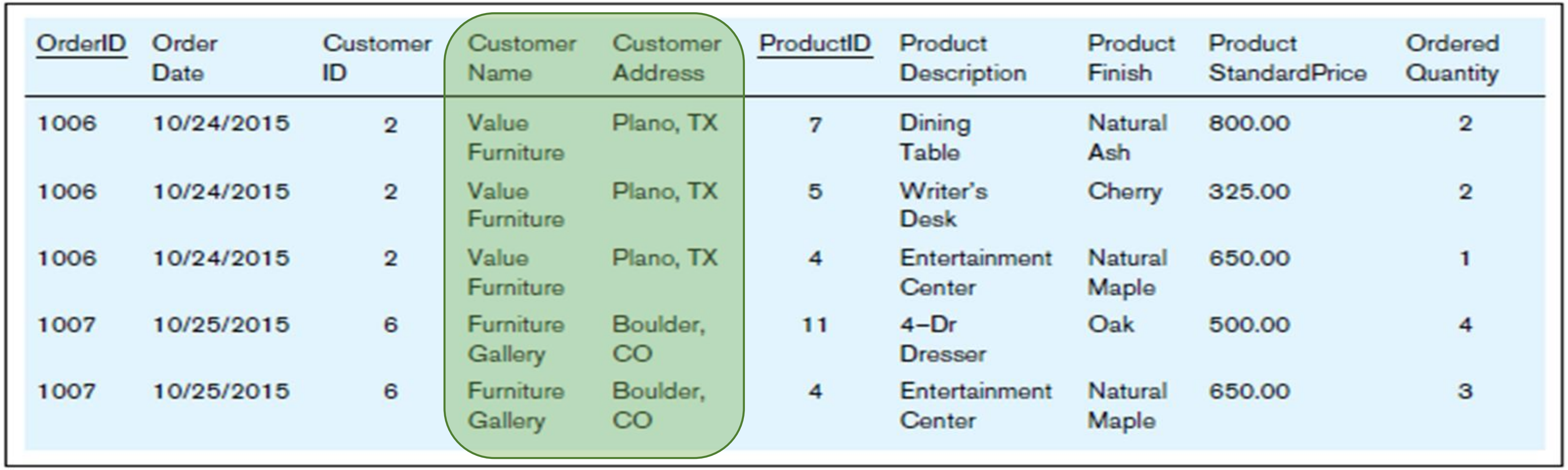
- *CUSTOMER ORDER contains transitive dependencies*

Step 3: Convert to Third Normal Form

- …if it is in second normal form and no ***transitive dependencies*** exist.
- A ***transitive dependency*** in a relation is a functional dependency between the primary key and one or more non-key attributes that are dependent on the primary key via another non-key attribute.
- For example, there are two transitive dependencies in the *CUSTOMER ORDER* relation.
- OrderID → CustomerID → CustomerName
- OrderID → CustomerID → CustomerAddress

Issues with Transitive dependencies

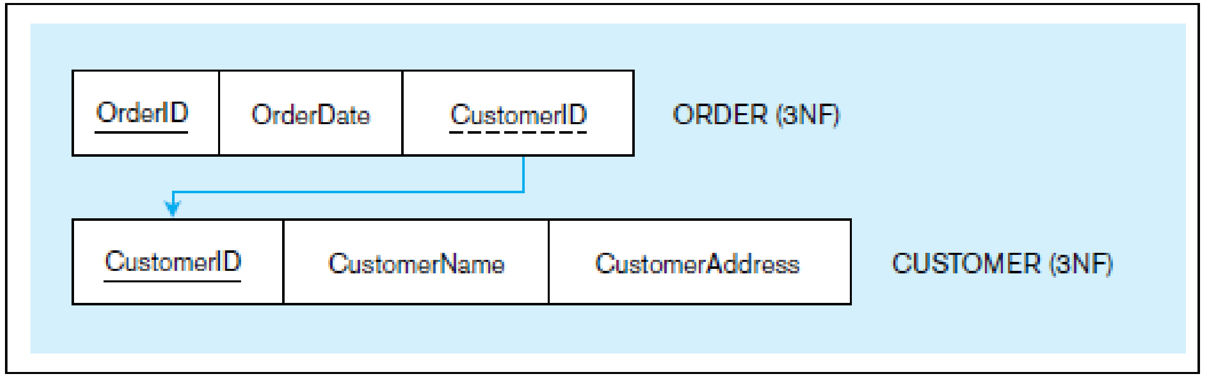
- Transitive dependencies create **unnecessary redundancy** that may lead to the type of anomalies discussed earlier.
- For example, the transitive dependency in *CUSTOMER ORDER* (**Figure below**) requires that a customer's name and address be re-entered every time a customer submits a new order, regardless of how many times they have been entered previously.



OrderID	Order Date	Customer ID	Customer Name	Customer Address	ProductID	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2015	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	5	Writer's Desk	Cherry	325.00	2
1006	10/24/2015	2	Value Furniture	Plano, TX	4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	11	4–Dr Dresser	Oak	500.00	4
1007	10/25/2015	6	Furniture Gallery	Boulder, CO	4	Entertainment Center	Natural Maple	650.00	3

Removing Transitive Dependencies

- You can easily remove transitive dependencies from a relation by means of a three-step procedure:
  1. For each *non-key attribute* (or set of attributes) that is a determinant in a relation, create a new relation. That attribute (or set of attributes) becomes the *primary key* of the new relation.
  2. Move all of the *attributes* that are functionally dependent only on the primary key of the new relation from the old to the new relation.
  3. Leave the attribute that serves as a primary key in the new relation in the old relation to serve as a foreign key that allows you to associate the two relations.

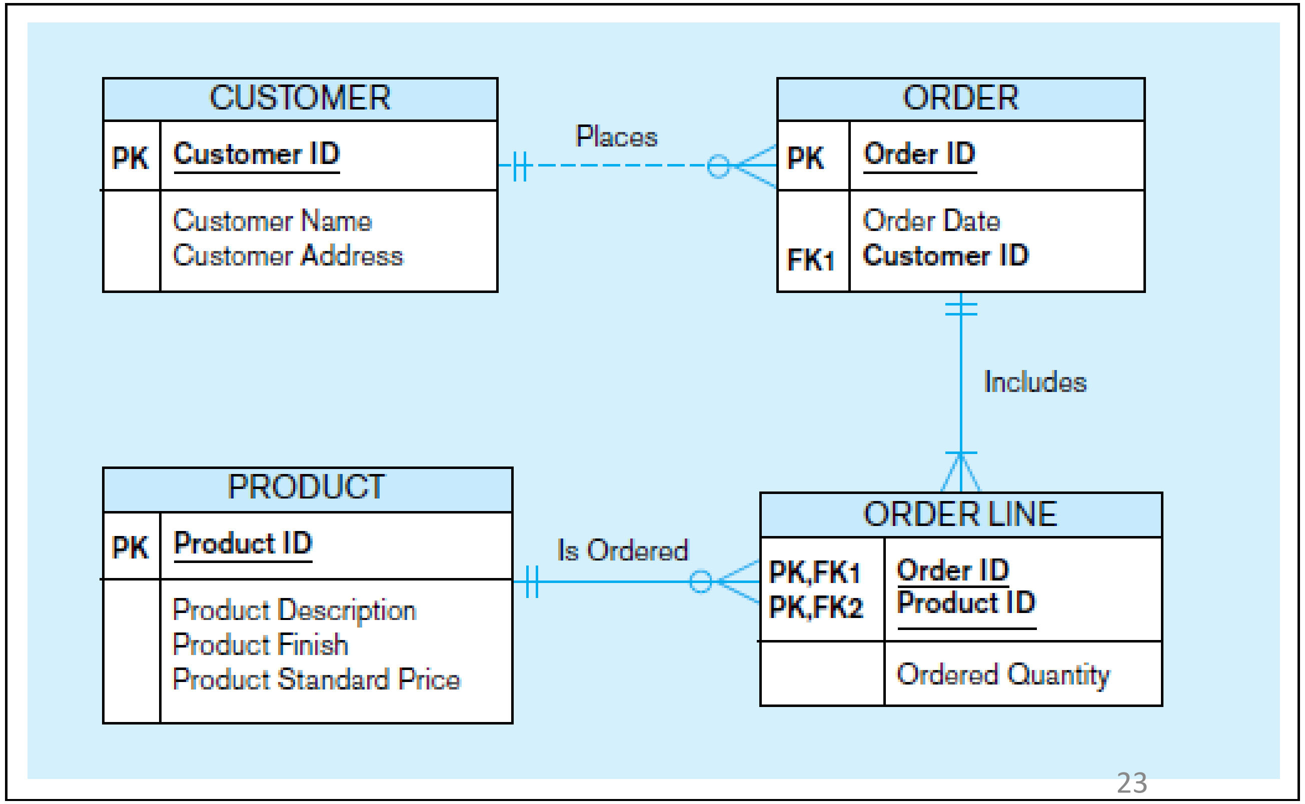


Final schema

- Normalizing the data in the INVOICE view has resulted in the creation of four relations in third normal form:
  - **CUSTOMER**
  - **PRODUCT**
  - **ORDER**
  - **ORDER LINE**

- A **relational schema** showing these four relations and their associations

https://en.wikipedia.org/wiki/Mickey_Mouse



CUSTOMER
PK Customer ID
Customer Name
Customer Address

ORDER
PK Order ID
Order Date
FK1 Customer ID

PRODUCT
PK Product ID
Product Description
Product Finish
Product Standard Price

ORDER LINE
PK,FK1 Order ID
PK,FK2 Product ID
Ordered Quantity

Course Content

1. Introduction to Relational Databases *(Introduction + Relational Model)*
2. Data Modelling - *(Entity Relationship Modelling + The Enhanced Entity Relationship Model)*
3. Database Design and SQL - *(Logical modelling + Introduction to SQL)*
4. Further SQL - *(Advanced SQL queries + Creating tables with SQL)*
5. Normalisation - *(Normalisation to second normal form + Third normal form)*

Thank you.

X THE PLACE OF USEFUL LEARNING

University of Strathclyde Glasgow

Database Fundamentals – CS990

Database and Web Systems Development - CS952

Image source: https://www.strath.ac.uk/branding/