

# Normalization

1

## Why Normalization?

- To develop a logical data model for relational database systems in order to create an accurate representation of the data, its relationships, and constraints.
- *Normalization* is a method for organizing data elements in a database into tables.

2

## Normalization Avoids...

- **Duplication of Data** – The same data is listed in multiple lines of the database
- **Insert Anomaly** – A record about an entity cannot be inserted into the table without first inserting information about another entity – Cannot enter a customer without a sales order
- **Delete Anomaly** – A record cannot be deleted without deleting a record about a related entity. Cannot delete a sales order without deleting all of the customer's information.
- **Update Anomaly** – Cannot update information without changing information in many places. To update customer information, it must be updated for each sales order the customer has placed

3

## Normalization Stages...

Normalization is a three stage process:

- After the first stage, the data is said to be in first normal form (1NF)
- After the second, it is in second normal form (2NF)
- After the third, it is in third normal form (3NF)

4



## First Normal Form (1NF) QA

**Questions: What if we did not normalize the Database to First Normal Form?**

**Answer: Repetition of Data – So, Header data (customer no., customer name, customer address, etc) is repeated for every line in sales order.**

9

## Normalization: Second Normal Form (2NF)

- Remove Partial Dependencies.
- **Functional Dependency:** The value of one attribute in a table is determined entirely by the value of another.
- **Partial Dependency:** A type of functional dependency where an attribute is functionally dependent on only part of the primary key (primary key must be a composite key).
- Create separate table with the functionally dependent data and the part of the key on which it depends. Tables created at this step will usually contain descriptions of resources.

10

## Second Normal Form Example

- The new table will contain the following fields:  
ItemNo, Description
- All of these fields except the primary key will be removed from the original table. The primary key will be left in the original table to allow linking of data:  
SalesOrderNo, ItemNo, Qty, UnitPrice
- Never treat price as dependent on item. Price may be different for different sales orders (discounts, special customers, etc.)
- Along with the unchanged table below, these tables make up a database in second normal form:  
SalesOrderNo, Date, CustomerNo, CustomerName, CustomerAdd, ClerkNo, ClerkName

11

## Second Normal Form (2NF) QA

**Question: What if we did not normalize the Database to Second Normal Form?**

**Answer:**

1. Repetition of Data – Description would appear every time we had an order for the item
2. Delete Anomalies – All information about inventory items is stored in the SalesOrderDetail table. Delete a sales order, delete the item.
3. Insert Anomalies – To insert an inventory item, must insert sales order.
4. Update Anomalies – To change the description, must change it on every SO.

12

## Normalization: Third Normal Form (3NF)

- Remove transitive dependencies.
- **Transitive Dependency:** A type of functional dependency where an attribute is functionally dependent on an attribute other than the primary key. Thus its value is only indirectly determined by the primary key.
- Create a separate table containing the attribute and the fields that are functionally dependent on it. Tables created at this step will usually contain descriptions of either resources or agents. Keep a copy of the key attribute in the original file.

13

## Third Normal Form Example

- The new tables would be:  
CustomerNo, CustomerName, CustomerAdd  
ClerkNo, ClerkName
- All of these fields except the primary key will be removed from the original table. The primary key will be left in the original table to allow linking of data as follows:  
SalesOrderNo, Date, CustomerNo, ClerkNo
- Together with the unchanged tables below, these tables make up the database in third normal form.  
ItemNo, Description  
SalesOrderNo, ItemNo, Qty, UnitPrice

14

## Third Normal Form (3NF) QA

**Question:** What if we did not normalize the Database to Third Normal Form?

**Answer:**

- Repetition of Data – Detail for Cust/Clerk would appear on every SO
- Delete Anomalies – Delete a sales order, delete the customer/clerk
- Insert Anomalies – To insert a customer/clerk, must insert sales order.
- Update Anomalies – To change the name/address, etc, must change it on every SO.

15

## Completed Tables in Third Normal Form

Customers: CustomerNo, CustomerName, CustomerAdd

Clerks: ClerkNo, ClerkName

Inventory Items: *ItemNo*, Description

Sales Orders: SalesOrderNo, Date, CustomerNo, ClerkNo

SalesOrderDetail: SalesOrderNo, ItemNo, Qty, UnitPrice

16

## Exercise: Normalize ...

### Student

St_ID	St_Last	St_Address	St_City	St_POB	Course_No	Course_Name	Credits	Grade	F_ID	F_Last
0091111	Koko	XYZ Str.	Amman	1111	1902321	Database Management	3	C	1	Al-Sayyed
					1903415	Database Tools	3	D		
0082222	Soso	ABC Str.	Amman	1212	1902321	Database Management	3	A	4	Al-Zghool
0076666	Lolo	Main Str.	Amman	3333	1903410	Information Systems	3	B	3	Zamzeer
					1901666	Computer Application	3	D		
0071234	Shosho	University Str.	Amman	9944	1903415	Database Tools	3	B	2	Rababah
0088888	Dodo	WWW Str.	Amman	7878	1901666	Computer Application	3	C	4	Al-Zghool
2059922	Fofo	High Str.	Amman	3573	1902321	Database Management	3	A	1	Al-Sayyed