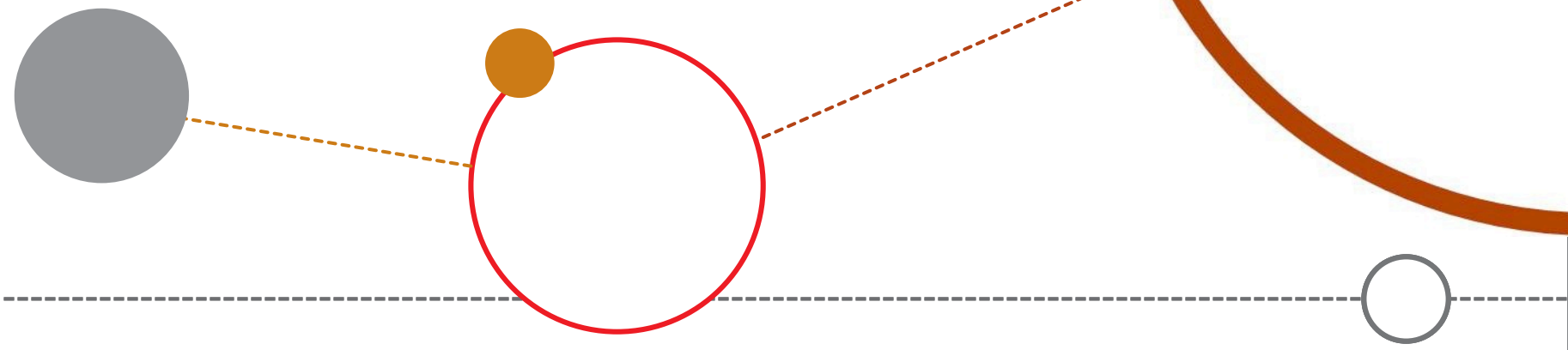
UNIVERSITY OF WEST LONDON
The Career University

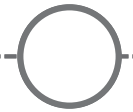
Entity Relationship Diagram 3/3

School of Computing and Engineering



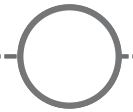
Learning Objectives

- Understand Normalisation technique following three process:
  - 1NF
  - 2NF
  - 3NF

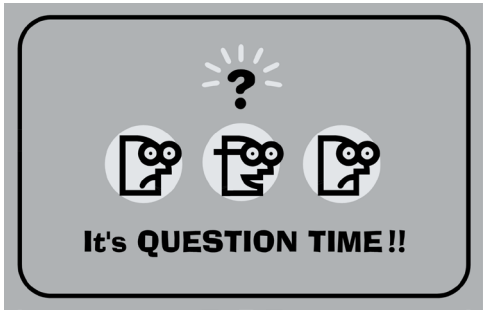


Lecture Plan

- Recap from the last lecture 5min
- Normalisation technique 35min
- Class Exercise 15min



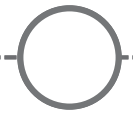
Recap from the previous lecture



- What is an identifier?
- Identifier can uniquely identify one instance of an entity.
- All entities have one attribute as their identifier.

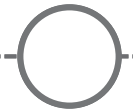
Normalisation

- ***Normalisation*** is a technique that can help analysts validate the data models.
- It is a process whereby a series of rules are applied to a logical data model to determine how well formed it is.
- The result of normalisation process is that the data attributes are arranged to form stable yet flexible relations for the data model.
- There are three steps of normalisation.
  - 1NF: Remove repeating groups/attributes
  - 2NF: Remove Partial Dependencies
  - 3NF: Remove Transitive Dependencies



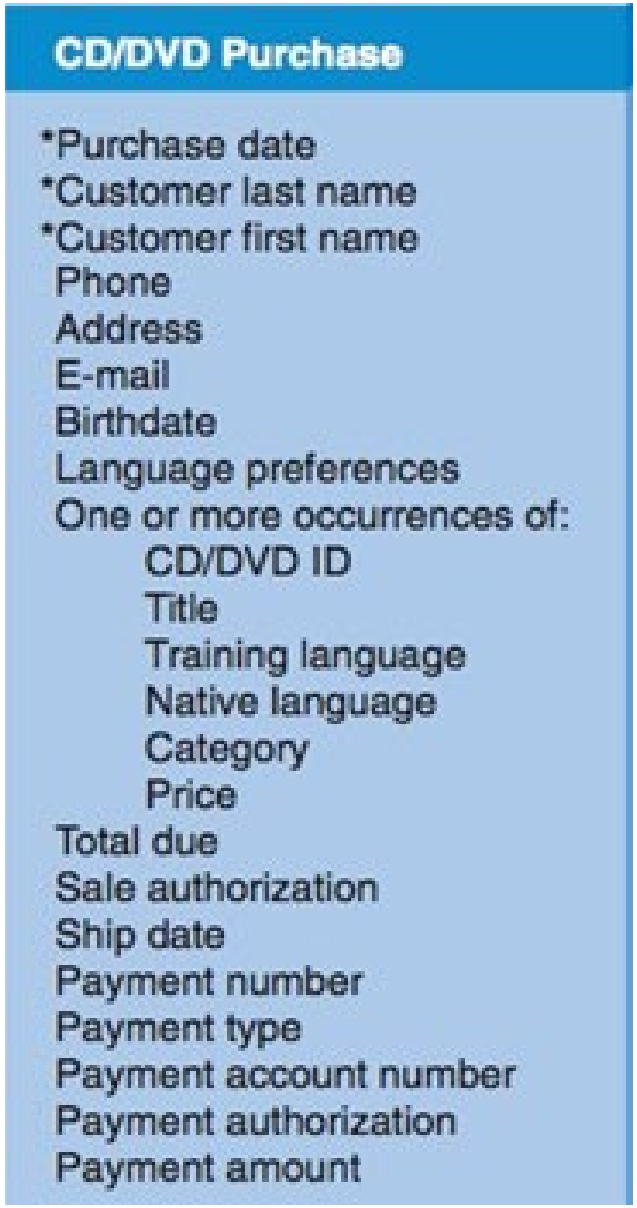
First Normal Form – 1NF

- A logical data model is in first normal form if it does not contain attributes that have repeating values for a single instance of an entity.
- Repeating attributes should be removed.
- Create another entity to place the repeating attributes.
- Every attribute in an entity should have only one value per instance for the model to pass 1NF.

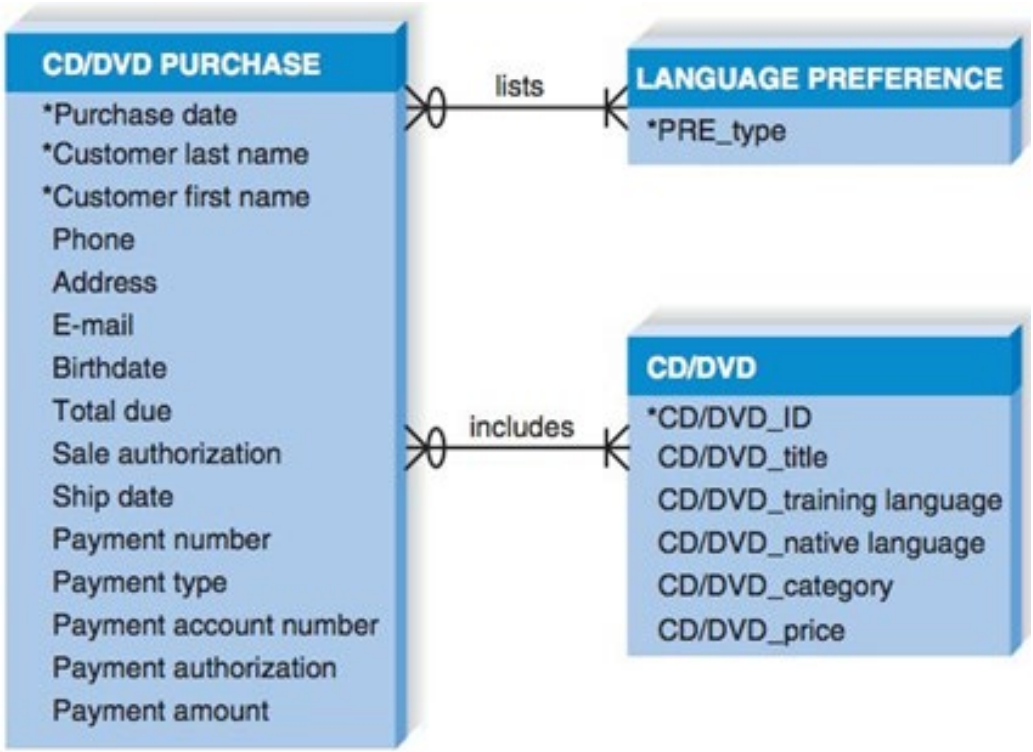


The first repeating group of attributes are about *CD/DVD* which should be removed and placed in a new entity called CD/DVD.

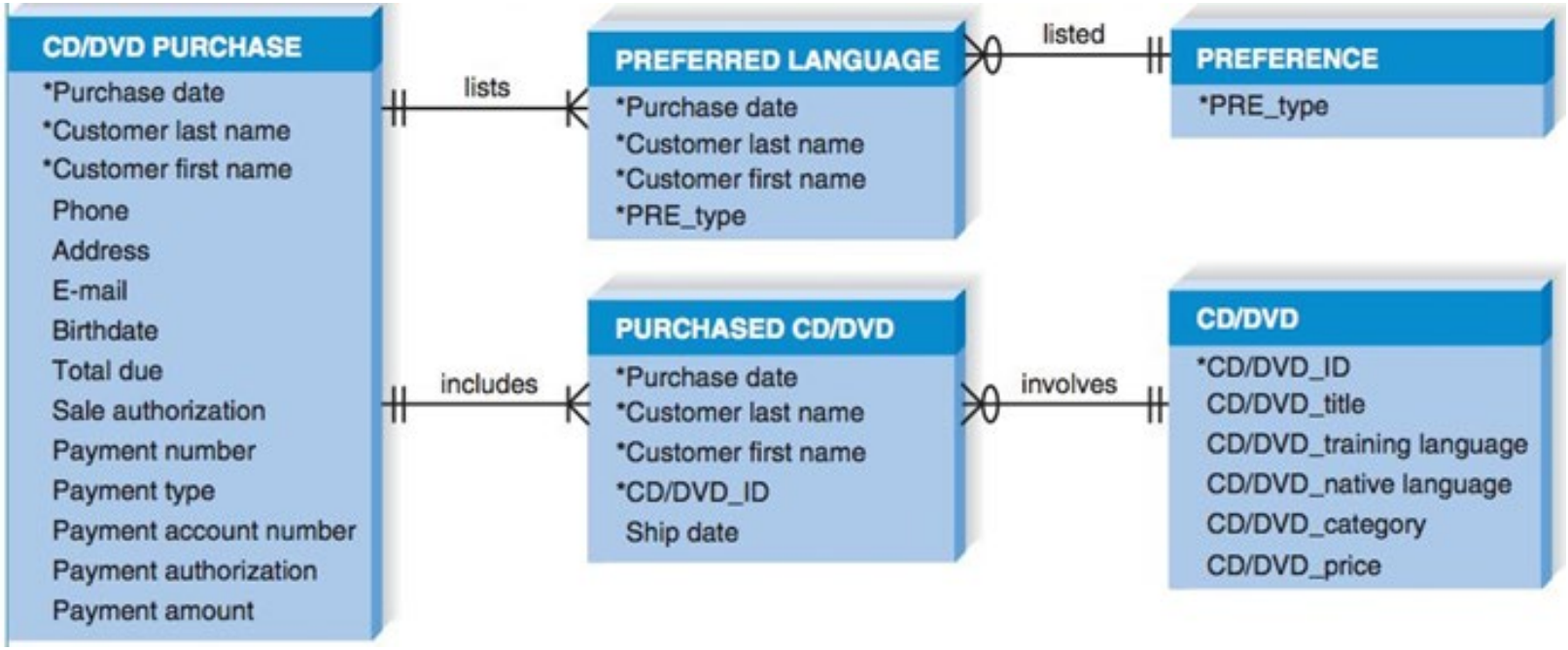
Language preferences attribute includes the languages of interest to the customer (e.g. French, Italian, Germany and etc). This is also a repeating attribute as many different preferences may be captured for each instance of a sale.



1. The new entities have M:N relationships with CD/DVD Purchase entity.
2. We care resolve this by using intersection entity.

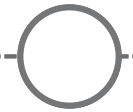


3. A new intersection entity was inserted between CD/DVD Purchase and CD/DVD. Also between CD/DVD Purchase and Preference.
4. Attribute Ship date was moved to Purchased CD/DVD because the various CD/DVDs in a purchase may ship at different dates.

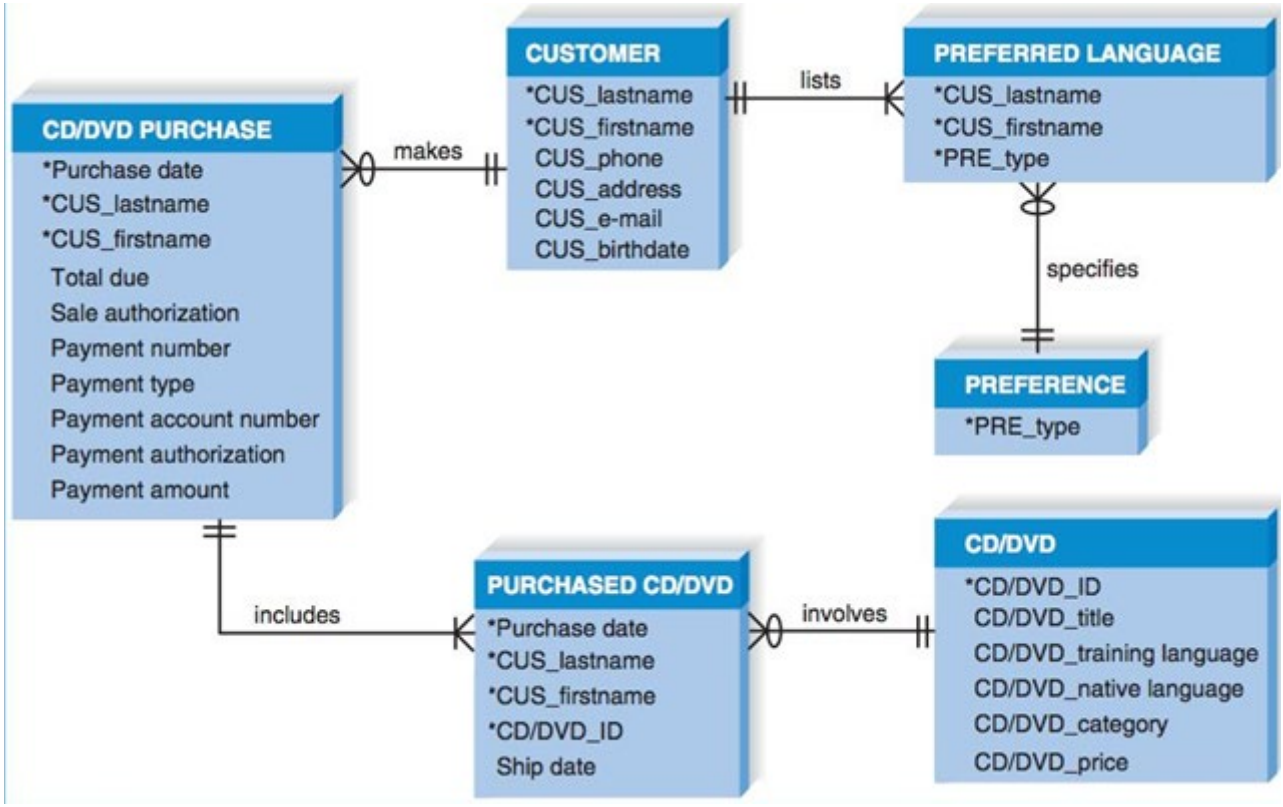


Second Normal Form – 2NF

- Once the data model is in 1NF we check if the model is in 2NF.
- All entities should contain attributes that are dependent on the whole identifier this means that the value of all attributes that serve as identifier can determine the value for all of the other attributes for an instance in an entity.
- Non-identifier attributes are dependent on only part of the identifier (partial dependency) and these attributes belong in another entity.

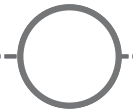


- The three attributes from CD/DVD Purchase used identifiers were moved to a new entity (Customer). The problem was that some of the attributes were dependent on the customer last name and first name but had no dependency on purchase date.

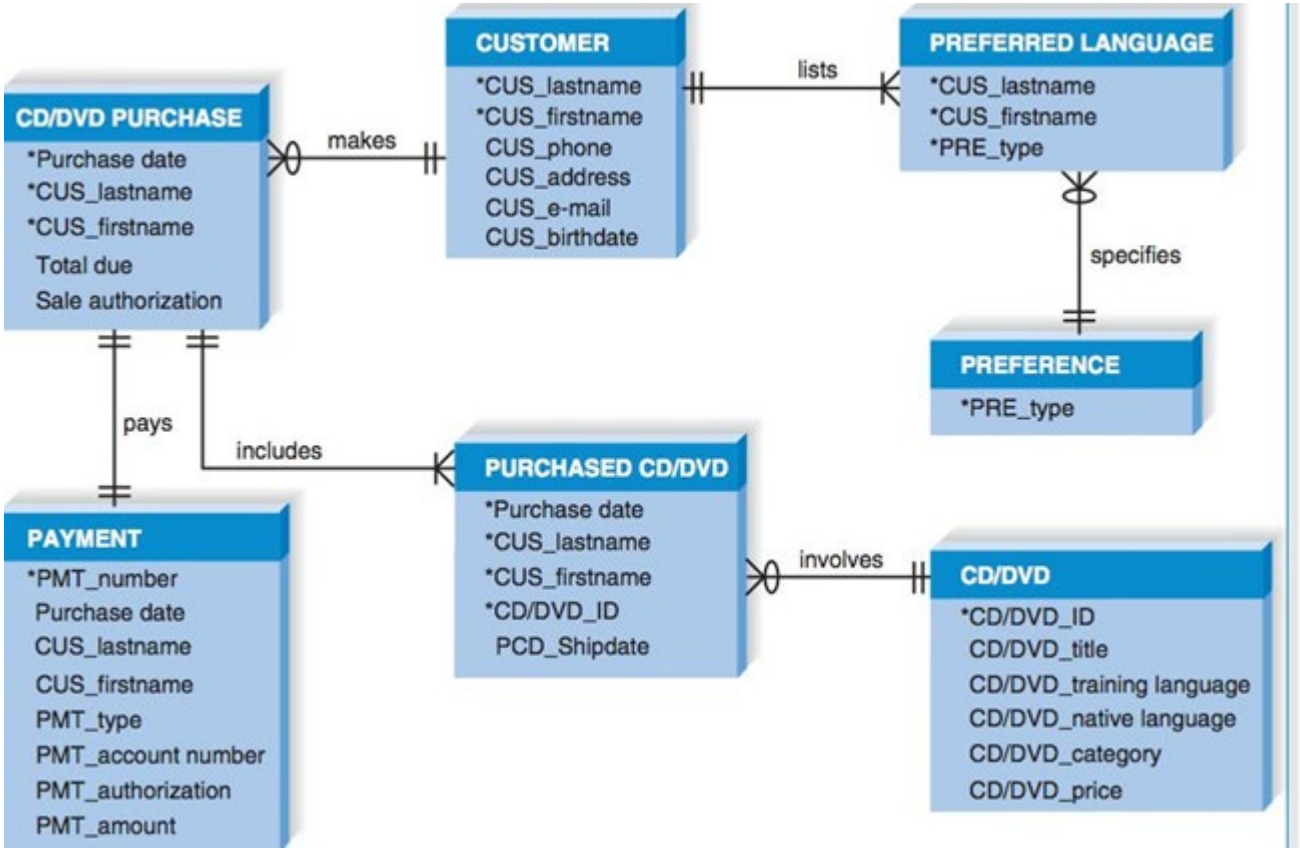


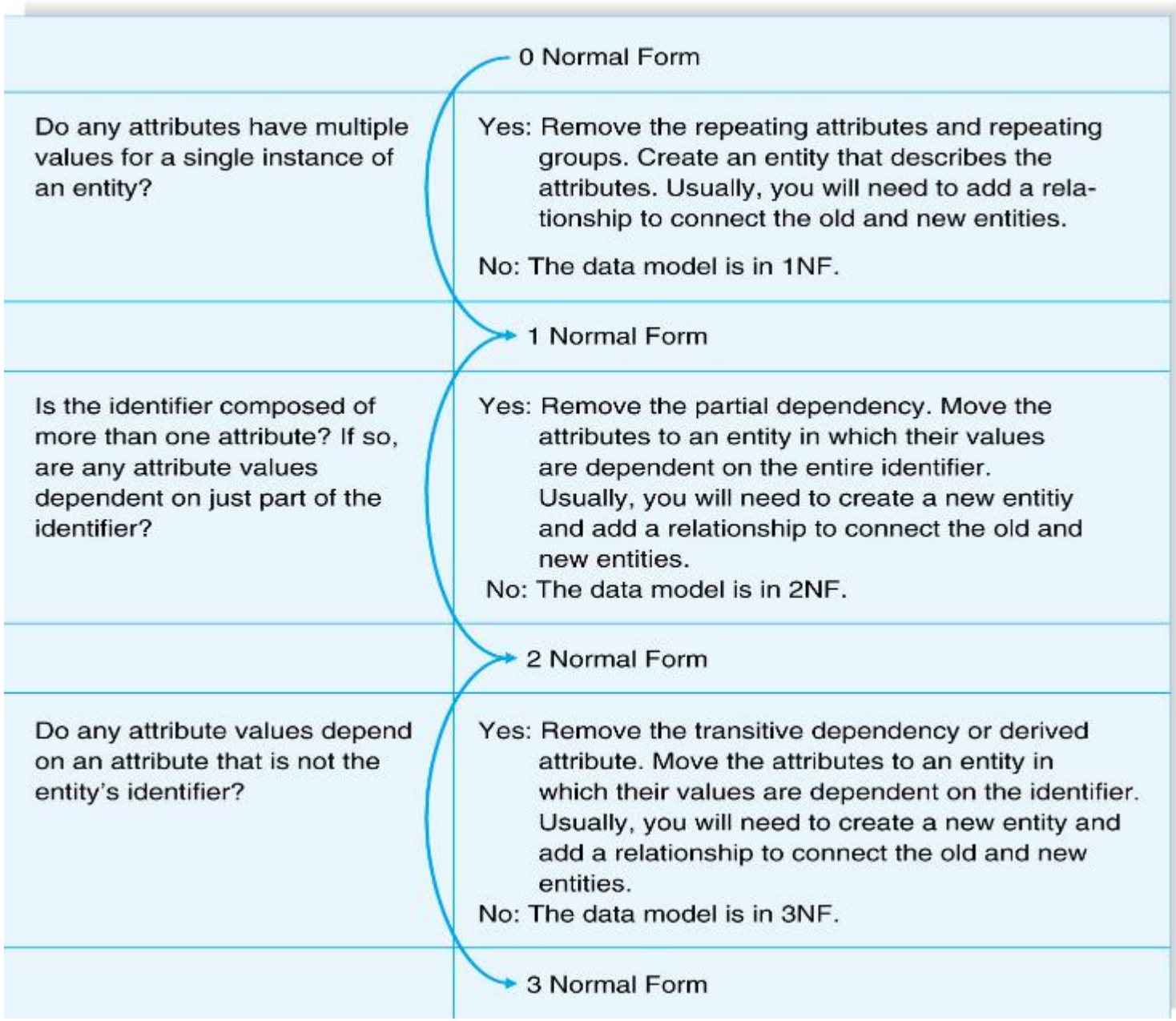
Third Normal Form – 3NF

- The model should be in 2NF. No attributes is dependent on a non-identifier attribute (transitive dependency).
- If there are attributes dependent on a non-identifier attributes, then they should be removed and placed in a new entity.
- There are attributes in CD/DVD Purchase that depend on the payment number not the CD/DVD purchase date and customer first, last names.
- The payment type, account number, authorisation and amount depend on the payment number (a non-identifying attribute)

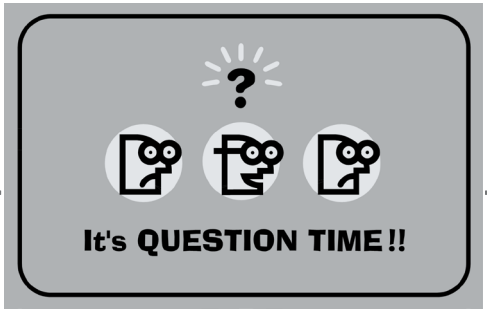


- We create a new entity (Payment) and move payment attributes to it.




0 Normal Form
Do any attributes have multiple values for a single instance of an entity?
Yes: Remove the repeating attributes and repeating groups. Create an entity that describes the attributes. Usually, you will need to add a relationship to connect the old and new entities.
No: The data model is in 1NF.
1 Normal Form
Is the identifier composed of more than one attribute? If so, are any attribute values dependent on just part of the identifier?
Yes: Remove the partial dependency. Move the attributes to an entity in which their values are dependent on the entire identifier. Usually, you will need to create a new entitiy and add a relationship to connect the old and new entities.
No: The data model is in 2NF.
2 Normal Form
Do any attribute values depend on an attribute that is not the entity's identifier?
Yes: Remove the transitive dependency or derived attribute. Move the attributes to an entity in which their values are dependent on the identifier. Usually, you will need to create a new entity and add a relationship to connect the old and new entities.
No: The data model is in 3NF.
3 Normal Form

Activity 1



Pretend that you have been asked to build a system that tracks student involvement in activities around campus. You have been given a file with information that needs to be imported into the system, and the file contains the following fields:

- Student Social Security number (identifier)
- Activity 1 code (identifier)
- Activity 1 description
- Activity 1 start date
- Activity 1 years with activity
- Activity 2 code
- Activity 2 description
- Activity 2 start date
- Activity 3 code
- Activity 3 description
- Activity 3 start date
- Activity 3 years with activity
- Student last name
- Student first name
- Student birthdate
- Student age
- Student advisor name
- Student advisor phone

Normalize the file. Show how the logical data model would change as you move from 1NF to 2NF to 3NF.

Activity 2

- Apply 1NF, 2NF and 3NF to the following model of Sales Report:

References

- Dennis, A., and Wixom, B. H, “Systems Analysis and Design”, 5th Edition, John Wiley & Sons (2013), Chapter Six.