### Analyse et conception

**Module 06 – Analyse des données** 

### **Objectifs**

- Apprendre à concevoir une base de données
- La Modélisation Conceptuelle des Données (MCD)
  - Les entités
  - Les liens
  - Les attributs
  - Les identifiants
  - Les liens N:N
  - Les structures complexes
- La Modélisation Relationnelle des Données (MRD)
- La Modélisation Physique des Données

#### L'analyse des données Les 3 étapes de modélisation

## Informations du domaine Modélisation Conceptuelle Vue de l'entreprise Le schéma Entités / Association

Vue du système

Base de données opérationnelle

Modélisation

Relationnelle

Modélisation

Physique

#### • Occurrence de donnée :

• Un objet, un évènement, un lieu, une personne..., une chose identifiable sans ambiguïté

#### Définition d'une entité :

- Une classe ou catégorie d'objets
- Quelque chose de <u>significatif pour l'entreprise et son applicatif</u>
- Quelque chose de <u>significatif sur lequel l'entreprise a besoin de stocker de l'information</u>

#### Définition d'un attribut

- Noms utilisés pour décrire l'entité
- Constitue le moyen de stocker de l'information au sujet de l'entité

#### L'analyse des données Règles de représentation des entités

- Boîte : rectangle à coins arrondis
- Nom de l'entité au singulier, en majuscule
- Synonymes
- Noms des attributs en minuscule

SOCIÉTÉ DÉPARTEMENT

ABONNEMENT

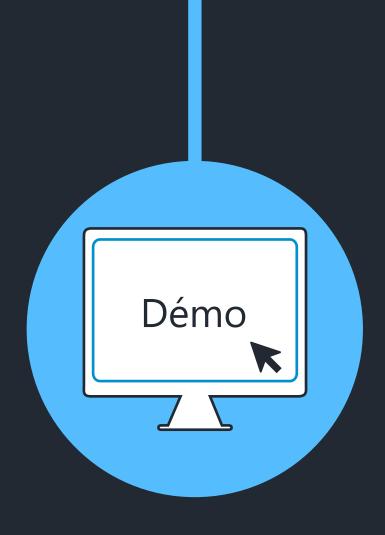
**EMPLOYÉ** 

nom prénom date naissance

#### Exemple : gestion d'une société de formation

« Je suis le responsable d'une société de formation et je souhaite acquérir un logiciel me permettant de gérer mon activité. Notre centre de formation propose des cours de techniques de développement, animés par des formateurs. Nous donnons des cours, chacun d'entre eux ayant un code, un nom et un prix. Introduction à Unix et Langage C sont deux de nos cours les plus appréciés.

Les cours varient par leur durée, de 1 à 4 jours. Jean Dumoulin et Claire Desroches sont deux de nos meilleurs formateurs. Nous avons besoin du nom et du numéro de téléphone de chaque formateur. Les clients peuvent choisir plusieurs cours et beaucoup le font. Le stagiaire Didier Dulong a pris tous les cours que nous proposons. Nous souhaitons également avoir le nom et le numéro de téléphone de chaque client. »



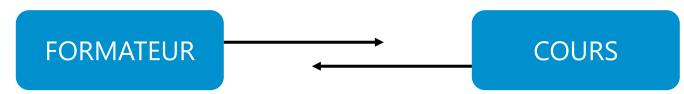
### Installer Datamodeler



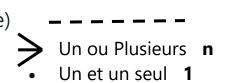
## Modéliser les entités

#### L'analyse des données Modéliser les liens

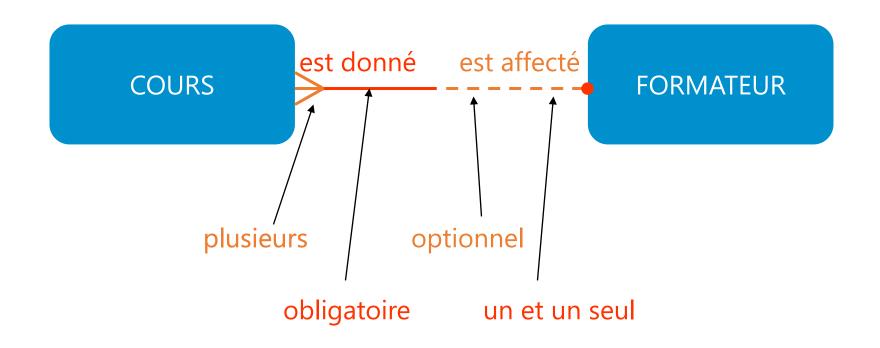
- Un lien:
  - « Manière dont <u>une entité</u> est reliée avec <u>une autre entité</u> »
  - Les règles de gestion qui relient les besoins informationnels de l'entreprise
  - Ce qu'une chose a à voir avec une autre
  - Une association nommée dans les deux sens entre deux entités
- Association toujours bidirectionnelle:



- Règles de représentation :
  - Un trait entre deux entités
  - Nom du lien en minuscule
  - Optionalité (cardinalité minimale)
  - Degré (cardinalité maximale)

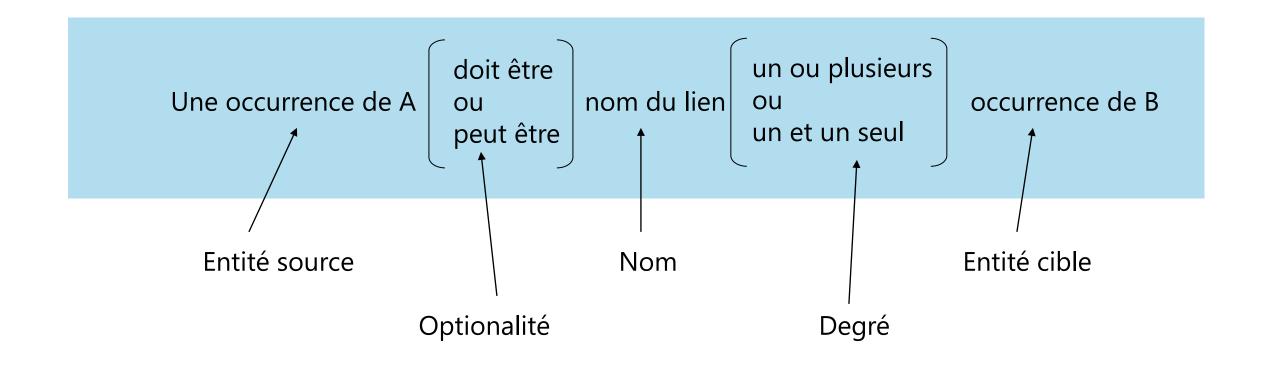


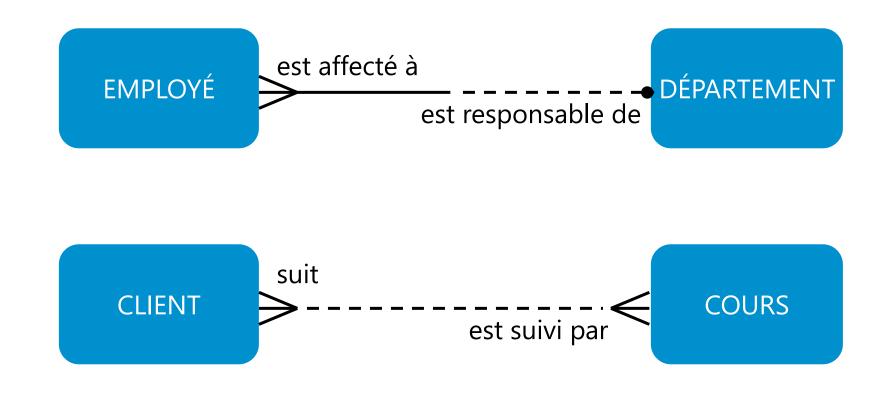
Obligatoire : doit être **1** Optionnel : peut être **0** 



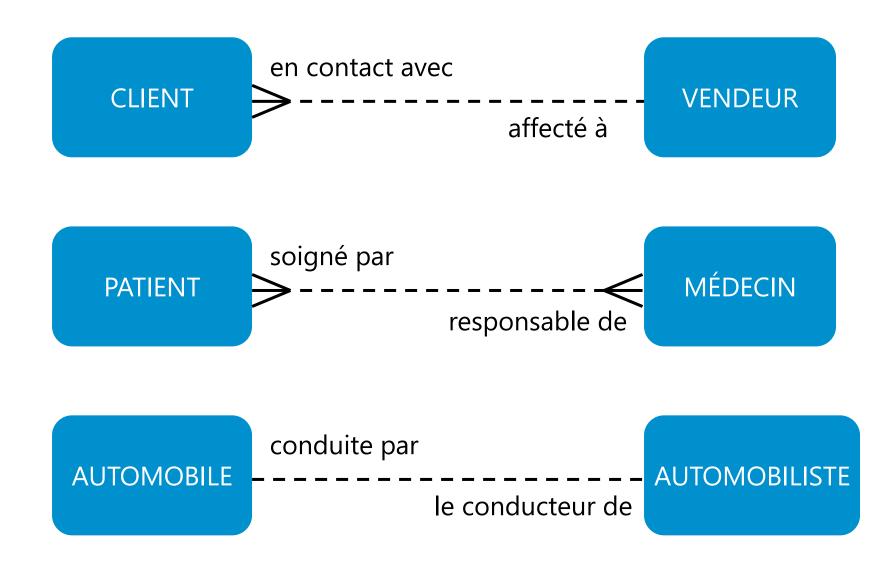
- Chaque cours doit être donné par un et un seul formateur
- Chaque formateur peut être affecté à un ou plusieurs cours

Lire le lien dans les deux sens en respectant la syntaxe suivante pour s'assurer de sa définition





#### L'analyse des données Différents types de liens





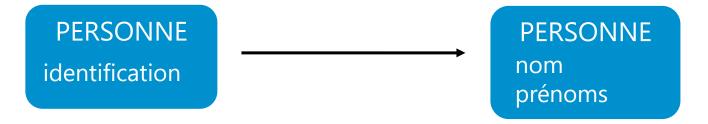
### Modéliser les liens

#### L'analyse des données Modéliser les attributs

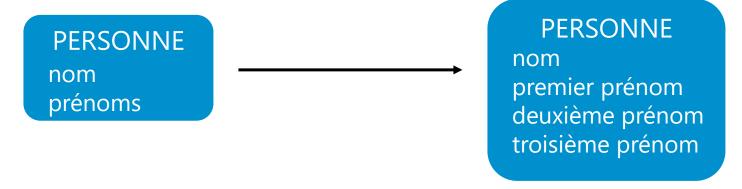
• Les attributs sont des informations qui caractérisent une entité et qui portent les valeurs décrivant chaque occurrence

- Ils décrivent l'entité :
  - En la qualifiant : nom qualifie un employé
  - En l'identifiant : numéro badge identifie un employé
  - En la classifiant : fonction classifie un employé
  - En la quantifiant
  - En exprimant son état
- Le nom d'un attribut doit être clair et concis

• Composant élémentaire significatif



Les attributs agrégés doivent être décomposés



Un attribut ne doit contenir qu'une information de base

Attribut possédant des attributs

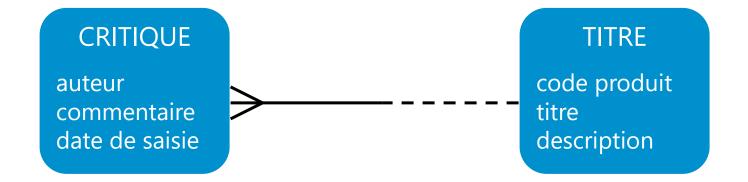
TITRE

code produit

titre

description

détails de critique



### L'analyse des données **Modéliser les attributs**

- Identifier les données dupliquées ou calculées
  - Données dupliquées :
    - Données qui apparaissent dans plusieurs entités différentes du MCD
  - Données calculées :
    - Totaux
    - Cumuls, maximum, minimum, moyennes
    - Calculs

#### L'analyse des données Règles de représentation des attributs

- Nom au singulier, en minuscule et à l'intérieur d'une entité
- Attribut obligatoire
  - Une valeur doit être stockée pour chaque occurrence de l'entité
  - Représenté par \*
- Attribut optionnel
  - Une valeur peut être stockée pour chaque occurrence de l'entité
  - Représenté par O

#### **EMPLOYÉ**

- \* numéro badge
- \* nom
- \* prénom
- O date de naissance



## Modéliser les attributs

#### L'analyse des données Identifier des domaines de valeurs

- Ensemble de valeurs qui s'applique à un ou plusieurs attributs
- Les domaines sont utilisés pour des informations qui changent rarement
- Exemples :
  - Les mois de l'année
  - Les unités de mesure
  - Les états acceptables pour une occurrence



## Créer un domaine de valeur statique

#### L'analyse des données Modéliser les identifiants

- Chaque occurrence de l'entité doit être identifiée de manière unique
- Un identifiant est une combinaison d'attributs et de liens qui sert à identifier chaque occurrence de l'entité de manière unique
- Les identifiants peuvent être déterminés à n'importe quel moment de la phase d'analyse
- Chaque entité du schéma Entités / Associations doit avoir un identifiant principal avant la transposition en schéma relationnel
- Il existe différents types d'identifiants :
  - Identifiant principal simple
  - Identifiant principal composé d'attributs
  - Identifiant principal composé d'attributs et de liens
  - Identifiant principal composé de liens
  - Identifiant secondaire

### L'analyse des données **Identifiant principal simple**

#### **CLIENT**

#\* numéro

Attribut simple

Préfixé par un #



## L'identifiant principal simple

### L'analyse des données Identifiant principal composé d'attributs

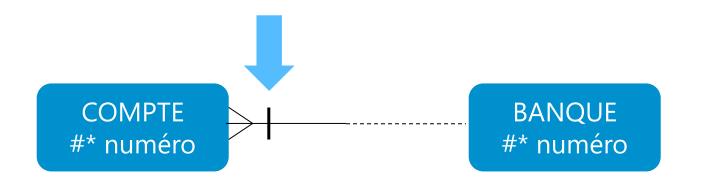
```
#* nom
#* prénom
#* date naissance
#* adresse
```

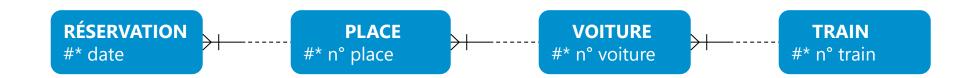
Attributs multiples



# L'identifiant principal composé d'attributs

#### Identifiant principal composé d'attributs et de liens



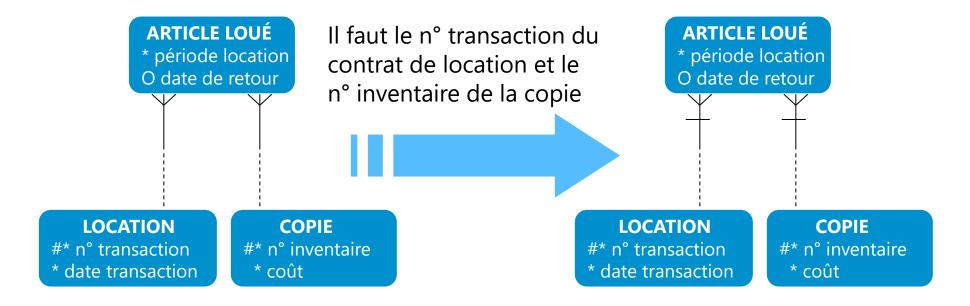




### L'identifiant principal composé d'attributs et de liens

### L'analyse des données **Identifiant principal composé de liens**

Que faut-il connaître pour identifier une occurrence d'un article loué ?





## Identifiant principal composé de liens

• Une entité peut être identifiée de manières différentes

numéro nom, prénom

numéro téléphone

#### **EMPLOYÉ**

```
# * numéro
```

 $(U)_1$  \* nom

(U)<sub>1</sub> \* prénom

\* date naissance

\* adresse

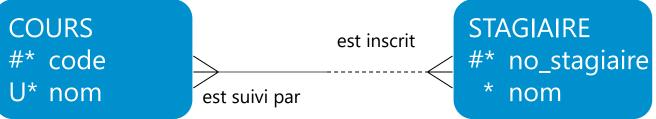
(U)<sub>2</sub> \* numéro tél



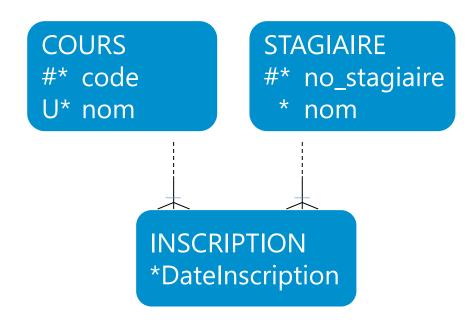
## Les identifiants secondaires

#### L'analyse des données Résolution des liens N:N

• Peut-on connaître la date d'inscription de chaque stagiaire au cours PHP ?



Création d'une entité d'intersection

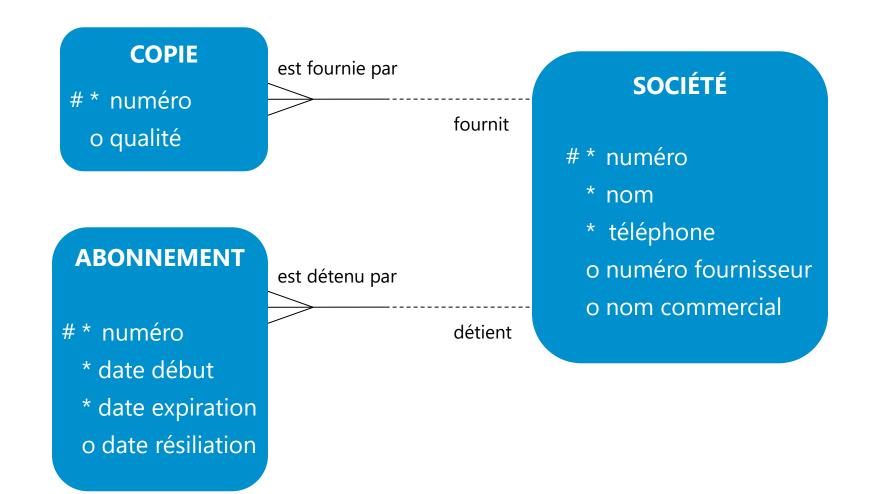


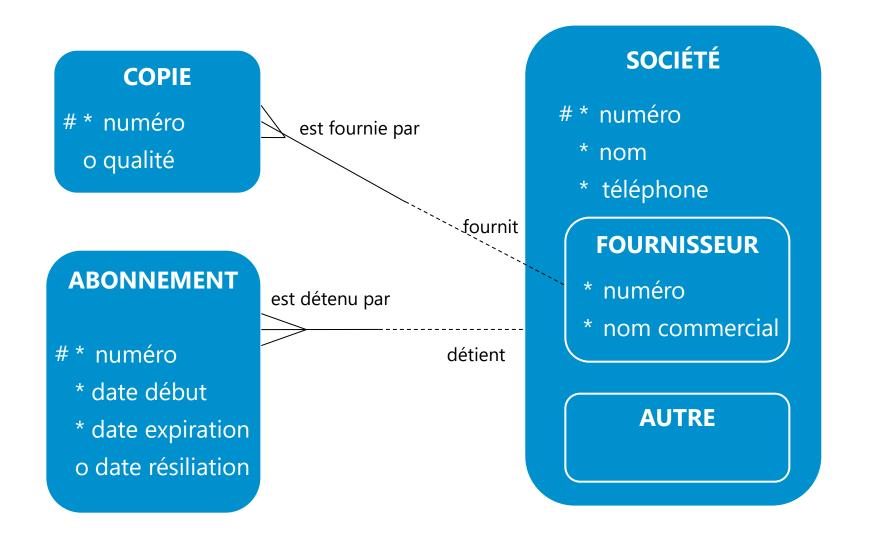


## La résolution des liens N:N



## Location vidéos parties 1 et 2

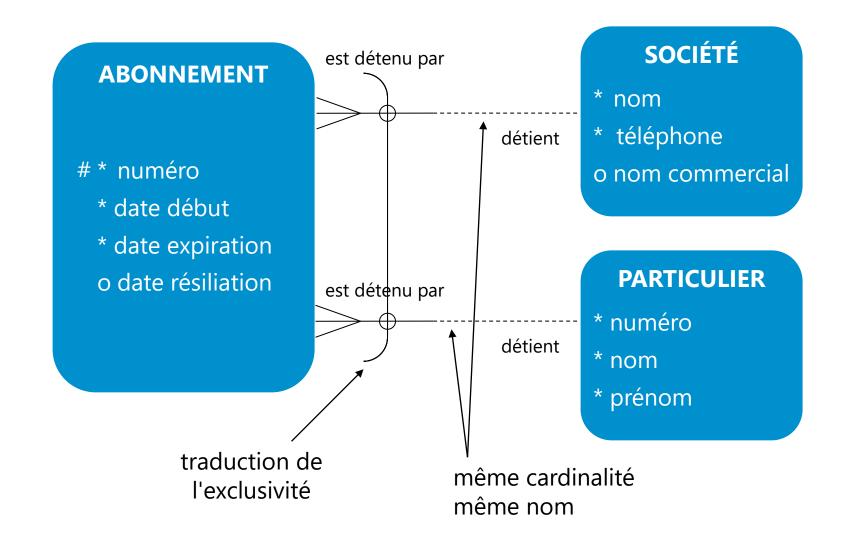




- Une entité peut être décomposée en plusieurs sous-types mutuellement exclusifs.
- Les attributs et les liens qui sont communs aux entités spécifiques sont définis au niveau de l'entité générique. Une entité spécifique hérite automatiquement de tous les attributs et liens de l'entité générique.
- Les entités spécifiques peuvent avoir leurs propres attributs et leurs propres liens. Il faudra déterminer le caractère optionnel de ces attributs et de ces liens.



## Entité générique – Entité spécifique

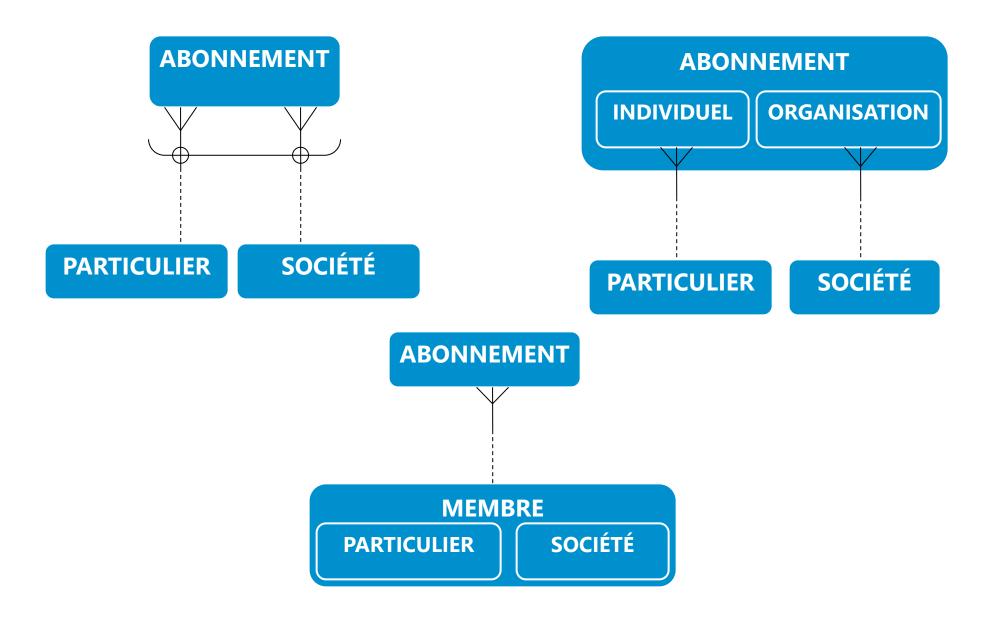


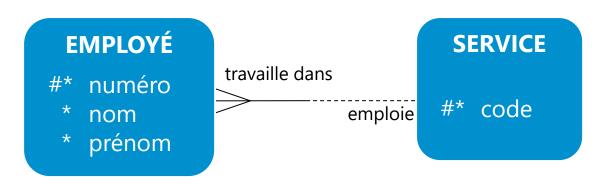
- Tous les liens côté arc doivent avoir obligatoirement le même caractère optionnel : soit ils sont tous optionnels, soit ils sont tous obligatoires
- Un lien ne peut être que dans un seul arc
- Tous les liens côté arc doivent être rattachés à la même entité
- Il n'y a pas de nombre limite de liens à inclure dans un arc
- Les liens inclus dans le même arc ont le même nom



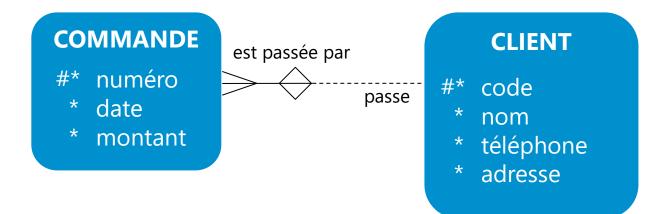
## Créer un arc

#### L'analyse des données Entité générique et liens exclusifs





Un employé peut changer de service. Le lien est donc modifiable.



Lien obligatoire unique qui ne peut pas changer. Le lien est non modifiable dans le temps.



## Liens transférables/non transférables

#### L'analyse des données Résumé des symboles

#### <u>Côté attribut</u>

- \* obligatoire
- o optionnel
- # identifiant
- U identifiant secondaire

#### <u>Côté lien</u>

obligatoire source optionnel source

> multiple cible

unique cible

identifiant fils / père

exclusif



## Location vidéos parties 3 à 5

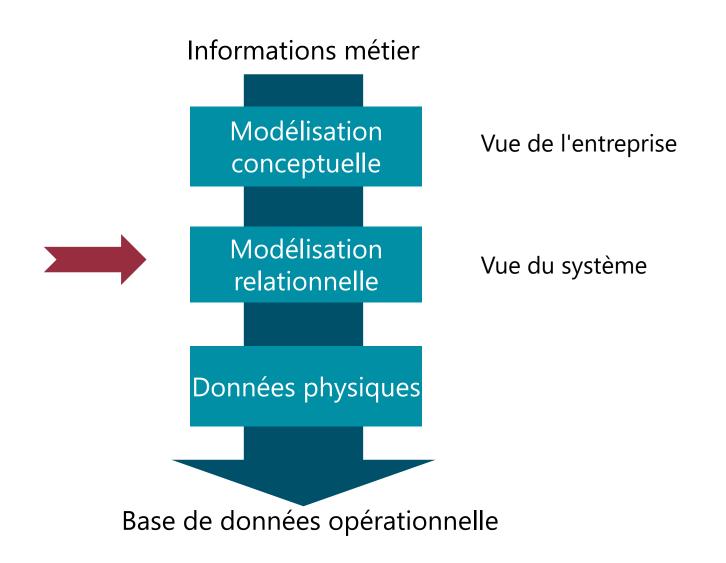


## Drills MCD



### Cas Intendance

#### L'analyse des données La modélisation relationnelle

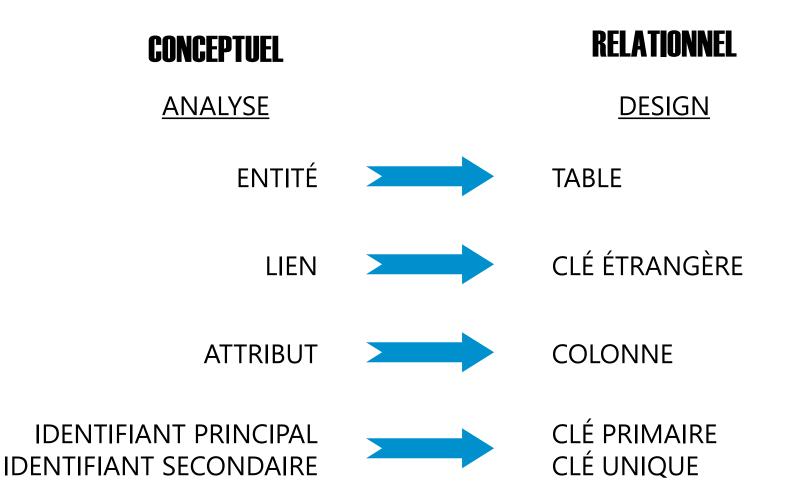


#### L'analyse des données Les phases de la transposition

- Transposer les entités simples en tables
- Transposer les attributs en colonnes
- Transposer les identifiants primaires et secondaires en clés primaires et uniques
- Transposer les liens en clés étrangères
- Choisir une option de transposition pour les arcs
- Choisir une option de transposition pour les entités génériques / spécifiques

- Avant de transposer le MCD en modèle relationnel (MRD), les contrôles qualité suivants doivent être réalisés :
  - Toutes les entités doivent avoir une description
  - Toutes les entités doivent avoir un nom
  - Tous les liens N:N doivent être résolus
  - Toutes les entités doivent avoir un identifiant
  - L'information volumétrique doit être définie
  - Les règles d'intégrité référentielle doivent être définies
  - Le format et le caractère optionnel des attributs définis

#### L'analyse des données Revue de la terminologie



	Nom colonne			
EMPLOYÉ	Type clé			
# * numéro (U) <sub>1</sub> * nom	Null ou Not Null			
(U) <sub>1</sub> * prénom (U) <sub>2</sub> * tél o fonction	Exemple de			
o forfetion	données			

#### Nom de la table : <u>EMPLOYÉS</u>

	Nom colonne	NUMÉRO	NOM	PRÉNOM	TÉL	FONC
EMPLOYÉ	Type clé					
# * numéro (U) <sub>1</sub> * nom	Null ou Not Null	<u>NN</u>	<u>NN</u>	<u>NN</u>	<u>NN</u>	
(U) <sub>1</sub> * prénom (U) <sub>2</sub> * tél o fonction	Exemple					
	de données					

 $(U)_1$ 

(U)<sub>1</sub> \*

 $(U)_2$  \*

#### L'analyse des données Transposer les identifiants

**FONC** 

#### Nom de la table : <u>EMPLOYÉS</u>

NOM

PRÉNOM

TÉL

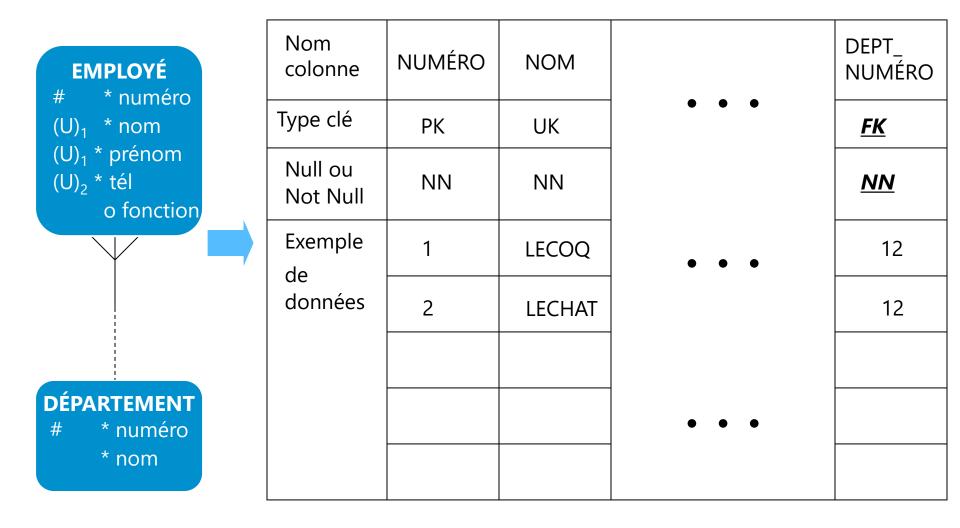
EMPLOYÉ	Type clé	<u>PK</u>	<u>UK1</u>	<u>UK1</u>	<u>UK2</u>
# * numéro (U) <sub>1</sub> * nom (U) <sub>1</sub> * prénom (U) <sub>2</sub> * tél o fonction	Null ou Not Null	<u>NN</u>	<u>NN</u>	<u>NN</u>	<u>NN</u>
	Exemple				
	de données				

NUMÉRO

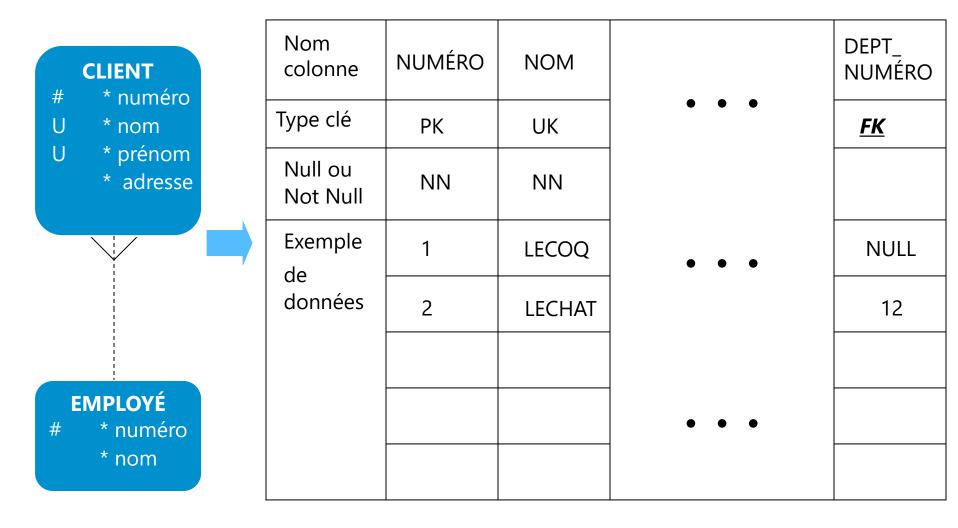
Nom

colonne

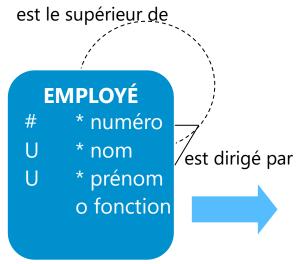
#### **Transposer les liens 1:N obligatoires**



#### **Transposer les liens 1:N optionnels**



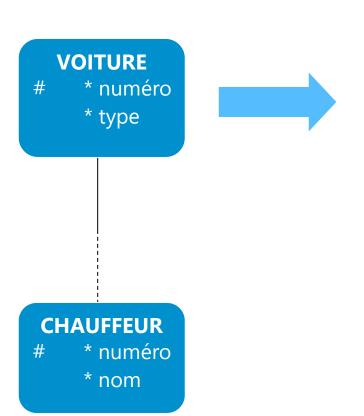
#### **Transposer les liens 1:N récursifs**



Nom colonne	NUMÉRO	NOM		EMP_ NUMÉRO
Type clé	PK	UK		<u>FK</u>
Null ou Not Null	NN	NN		
Exemple de	1	LECOQ	• • •	12
données	12	LECHAT		24
	24	LAPIE		NULL
			• • •	

#### **Transposer les liens 1:1 obligatoires**





Nom colonne	NUMÉRO	TYPE	CHAUF_ NUMÉRO
Type clé	PK		FK,UK
Null ou Not Null	NN	NN	<u>NN</u>
Exemple données	1	Zoë	24
	2	Mégane	5

#### **Transposer les liens 1:1 optionnels**

Nom de la table : <u>VOITURES</u>

<b>VOITURE</b> # * numéro * type		
	<u>OU</u>	
CHAUFFEUR  # * numéro  * nom		

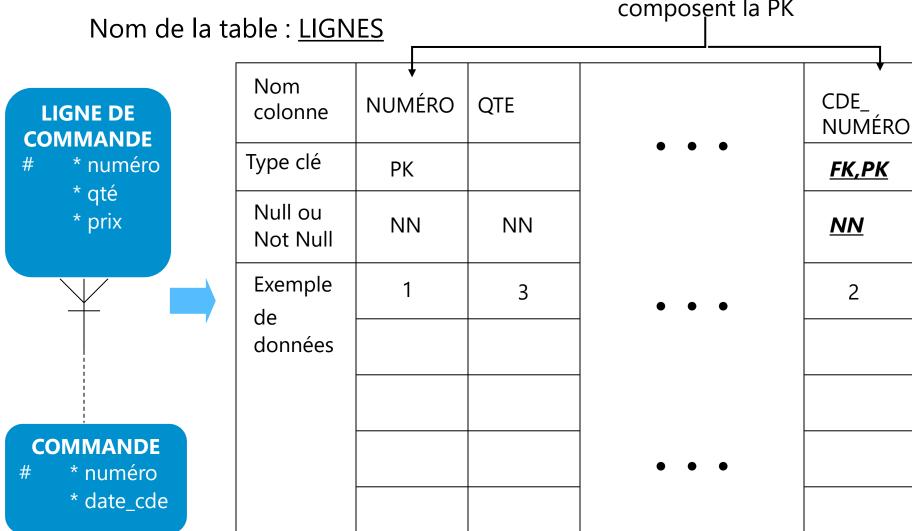
Nom colonne	NUMÉRO	TYPE	CHAUF_ NUMÉRO
Type clé	PK		FK,UK
Null ou Not Null	NN	NN	
Exemple données			

Nom de la table : <u>CHAUFFEURS</u>

Nom colonne	NUMÉRO	NOM	VOIT_ NUMÉRO
Type clé	PK		FK,UK
Null ou Not Null	NN	NN	
Exemple données			

#### L'analyse des données Transposer un lien identifiant

ces deux colonnes composent la PK



#### Transposition des entités génériques et spécifiques

#### TITRE

- # \* code
  - \* titre
  - \* description

#### **FILM**

- \* catégorie
- \* durée
- \* noir et blanc

#### **JEU VIDÉO**

- \* catégorie
- \* support
- \* mémoire mini

#### Choix possibles:

- une seule table
- deux tables séparées
- trois tables séparées

#### L'analyse des données Transposition en une seule table

#### **TITRE**

# \* code
 \* titre

#### **FILM**

- \* catégorie
- \* durée

#### **JEU VIDÉO**

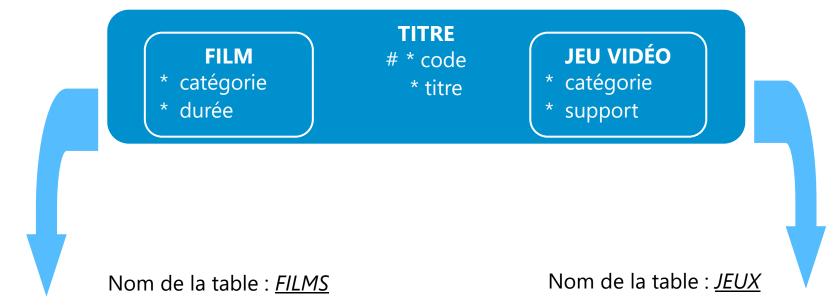
- \* catégorie
- \* support

Nom de la table : <u>TITRES</u>

Nom colonne	CODE	TITRE	TYPE	FILM_ CAT	FILM_ DUR	JEU_ CAT	JEU_ SUP
Type clé	PK						
Null ou Not Null	NN	NN	NN				
Exemple de	255	ET	F	XX	1h30		
données	142	SU.	J			А	CD
	314	BA.	F	уу	1h30		

Colonne discriminante

#### L'analyse des données Transposition en tables séparées



Nom colonne	CODE	TITRE	FILM _CAT	FILM _DUR
Type clé	PK			
Null ou Not Null	NN	NN	NN	NN
Exemple de	255	ET	xx	1h30
données	314	BA.	уу	1h30

Nom colonne	CODE	TITRE	JEU _CAT	JEU _SUP
Type clé	PK			
Null ou Not Null	NN	NN	NN	NN
Exemple de	132	GP	xx	CAS
données	481	MC.	уу	CD

#### L'analyse des données Trois tables séparées

Nom de la table : FILMS

Nom colonne	CODE	FILM _CAT	FILM _DUR
Type clé	PK		
Null ou Not Null	NN	NN	NN
Exemple de	255	XX	1h30
données	314	уу	1h30

Nom de l	table	: <i>JEUX</i>
----------	-------	---------------

Nom colonne	CODE	JEU _CAT	JEU _SUP
Type clé	PK		
Null ou Not Null	NN	NN	NN
Exemple de	132	XX	CD
données			



3 tables liées : TITRES, FILMS et JEUX Nom de la table : <u>TITRES</u>

Nom colonne	CODE		CODE_ FILM	CODE_ JEU
Type clé	PK	• • •	<u>FK1,UK1</u>	<u>FK2,UK2</u>
Null ou Not Null	NN			
Exemple de	10	• • •	255	
données	20			132
	30		314	

#### L'analyse des données Ou trois tables séparées

Nom de la table : FILMS

Nom colonne	CODE	FILM _CAT	FILM _DUR
Type clé	PK,FK		
Null ou Not Null	NN	NN	NN
Exemple de	255	XX	1h30
données	314	уу	1h30

Nom	ae	ıa	table	: <u>JEUX</u>

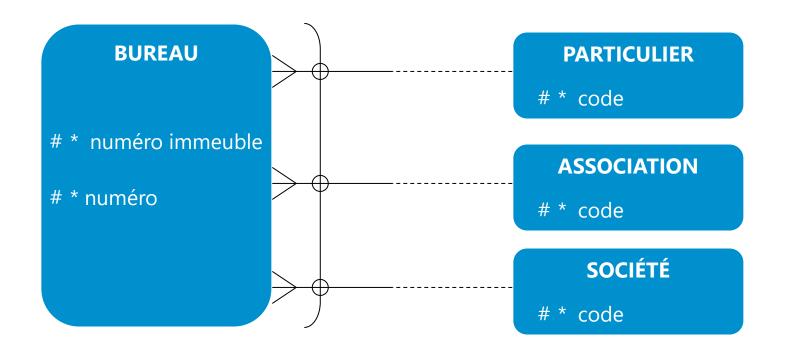
Nom colonne	CODE	JEU _CAT	JEU _SUP
Type clé	PK,FK		
Null ou Not Null	NN	NN	NN
Exemple de	132	XX	CD
données			

Nom de la table : *TITRES* 

3 tables liées : TITRES, FILMS et JEUX

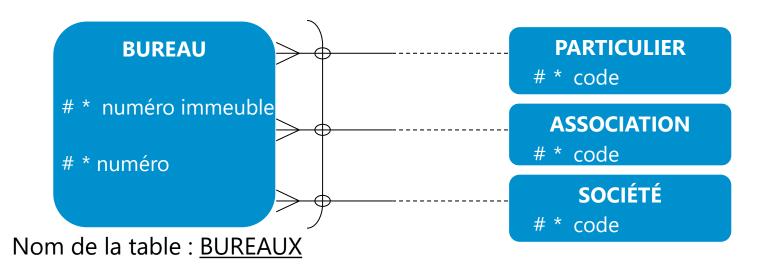
Nom colonne	CODE		
Type clé	PK	• • •	
Null ou Not Null	NN		
Exemple de	132	• • •	
données	255		
	314		

#### L'analyse des données Transposition des arcs



Une entité participe à un arc lorsqu'elle a des liens avec plusieurs autres entités de manière exclusive

#### L'analyse des données Transposition explicite d'un arc

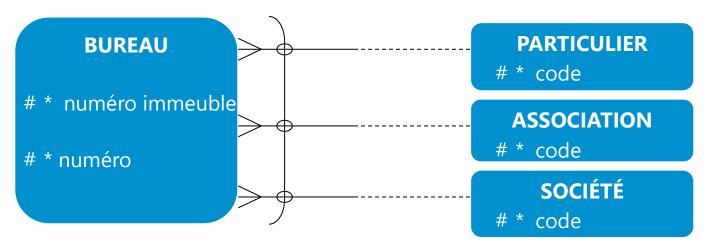


Nom colonne	IMM_NUM	BUR_NUM	PART_CODE	ASS_CODE	STE_CODE
Type clé	PK	PK	FK1	FK2	FK3
Null ou Not Null	NN	NN			
Exemple de	25	9		B312	
données	12	11	0154		
	14	9			1609

colonnes naturelles

+ contrainte check à écrire

#### L'analyse des données Transposition générique d'un arc



#### Nom de la table : <u>BUREAUX</u>

Nom colonne	IMM_NUM	BUR_NUM	LOC_NUM	LOC_TYP
Type clé	PK	PK	"FK"	
Null ou Not Null	NN	NN	NN	NN
Exemple de	25	9	B312	А
données	12	11	0154	Р
	14	9	1609	S



colonnes naturelles

colonne générique colonne type

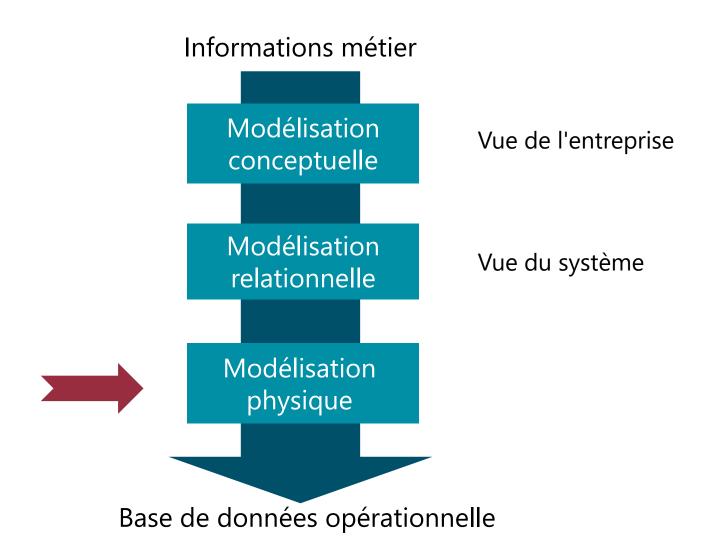
#### L'analyse des données Affiner le niveau relationnel

- Ajout de tables, d'attributs logiques afin d'améliorer les performances
- Ajout de règles logiques (compteurs, contraintes, notion de liens non transférables)
- Ajout de déclencheurs (triggers) de base de données afin d'automatiser des comportements



# Cas vidéo : la transposition

#### L'analyse des données La modélisation physique des données



### L'analyse des données Modèle physique des données

- Le MPD correspond à la traduction du Modèle Relationnel des Données dans le langage SQL de la base de données cible
- Nous utilisons pour cette phase une partie du SQL nommé DDL (Data Definition Language)



## MDP du cas Commande