



# ***Numérique et sciences de l'information***

***Tous niveaux  
Module No 08***

## ***Structures et Bases de Données***





# Programme

- **CM#8 Structures et bases de données**
- Modélisation des données
- Données structurées et non structurées
- Structures de données
- Identifier les liens entre objets
- Entités, associations et propriétés
- Cardinalités
- Modèle Conceptuel de Données (MCD)
- Base de données. Le modèle relationnel.
- Du modèle des données à l'implantation physique.
- SQL.





# Programme

- TD Séance #4 : Manipuler les structures de données et interroger une base de données
- Exercice #1 : Comptage des éléments d'un tableau
- Exercice #2 : Implanter une base de données : un parc automobile
- Exercice #3 : Quels sont les étudiants (nom et prénom) qui possèdent ou ont possédé une voiture de collection ?
- Exercice #4 : Quelles sont les marques de voitures achetées par les étudiants après le 10/1/2010 inclus ?
- Exercice #5 : Dans quelles villes habitent les étudiants qui ont des voitures ?
- Exercice #6 : Quelles sont les dates d'achat des voitures dont la puissance est supérieure ou égale à 7CV ?
- Exercice #7 : Nom des étudiants qui habitent la même ville que DUPONT ?
- Exercice #8 : Dans quelles villes habitent les étudiants qui possèdent une voiture dont le prix est supérieur au prix moyen ?





# Quelques questions

Quelles différences entre données structurées et non structurées ?

Quels modèles de données permettent la construction des bases de données exploitées par les applications ?

Qu'est ce qui caractérise un tri par fusion ?

Quelle démarche pour construire et exploiter une base de données ?





# Modélisation des données

- Les processus s'appuient sur des données (propriétés) qui décrivent les objets de l'organisation
- **Processus** : « order to cash » (traitement de la commande client)
  - **Objets** : client, commande, produit, facture, magasin
  - **Propriétés** : code client, adresse de facturation client, catégorie produit, code postal magasin (saisis), montant HT facture (calculée)
- **Processus** : « Résolution équation du second degré »
  - **Objet** : Equation
  - **Propriétés** : coefficients  $a$ ,  $b$ ,  $c$  (saisies), déterminant, racines (calculées)

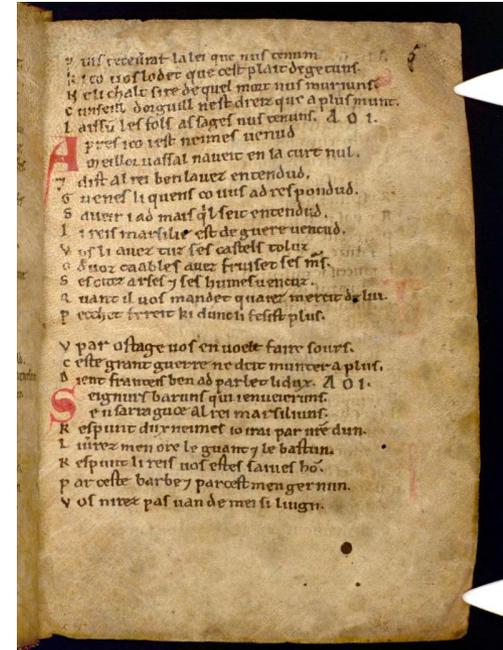




# Données structurées et non structurées

	A	B	C	D	E	F	G
1	<b>NAS</b>	<b>NOM</b>	<b>PRENOM</b>	<b>SEXE</b>	<b>TITRE</b>	<b>SALAIRE</b>	<b>CATEGORIE</b>
2	555 555 555	Thibault	Yvon	M	Administrateur	27 000 \$	3
3	222 222 222	Dupuis	Josée	F	Vendeur	22 500 \$	2
4	666 666 666	Smith	Alex	M	Vendeur	18 000 \$	1
5	777 777 777	Crosby	Julian	M	Administrateur	27 000 \$	3
6	888 888 888	Allard	Jocelyne	F	Secrétaire	27 000 \$	3
7	111 111 111	Savoie	Jean	M	Vendeur	31 500 \$	4
8	444 444 444	Bibeau	Martin	M	Secrétaire	22 500 \$	2
9	999 999 999	Allard	Benoit	M	Ouvrier	22 500 \$	2
10	333 333 333	Gingras	Marc	M	Administrateur	40 500 \$	4
11	000 000 000	Lalonde	Karl	M	Ouvrier	31 500 \$	4
12	123 456 789	St-Pierre	Aline	F	Secrétaire	22 500 \$	2
13	249 456 456	Bibeau	Rita	F	Administrateur	27 000 \$	3
14	343 456 987	Cardinal	Paul	M	Ouvrier	20 000 \$	2
15	345 456 324	Thibault	Gratien	M	Administrateur	32 000 \$	4
16	456 434 234	Dupuis	Carole	F	Vendeur	22 900 \$	2

**Données structurées** : on peut accéder directement à la donnée grâce à son adresse (dans un tableau, grâce au No de ligne et au No de colonne)



**Données non structurées** : on ne peut accéder à une donnée qu'en ayant lu toutes les données précédentes.

On peut donner un minimum de structure à des données non structurées au moyen d'index et de métadonnées.





# Structures de données

- **Chaîne** : une chaîne est une succession d'éléments de même type. L'étymologie latine catena a donné le terme concaténation. (Chaîne de caractères)
- **Liste linéaire séquentielle** (Liste des départements d'un pays)
- **Piles** (containers empilés sur une plate-forme)
- **File d'attente** (personnes à un guichet)
- **Tableau** (ou **matrice**) : emploi du temps (colonnes) des professeurs (lignes)
- **Réseau** (ou **arbre**) : nomenclature d'une pièce



# → Modéliser les données d'une organisation

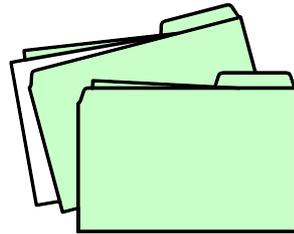
- Les informaticiens ont besoin des diagrammes décrivant les processus pour écrire les programmes (Que faire ? Quand le faire ? Qui le fait ?)
- Ils ont aussi à concevoir des bases de données (Avec quoi le faire ?)
- Au sein d'UML, ils disposent du diagramme de classes.
- Ce diagramme, adapté au concept d'objet, est difficile à concevoir et délicat à manipuler.
- Nous allons aborder le problème de la modélisation des données de manière un peu plus simple,
- En nous appuyant sur le MCD de la méthode Merise.
- Diagramme de classe d'UML et MCD de Merise reposent tous deux sur le concept de diagramme entité-association de Chen.



# → Identifier les liens entre objets

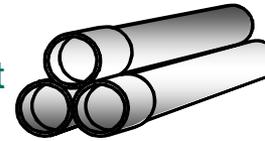


Client



Commande

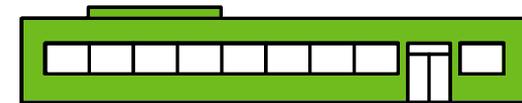
Produit



Expédition



Réglement



Magasin



Compte Client



Facture

Bon de Commande  
de réapprovisionnement

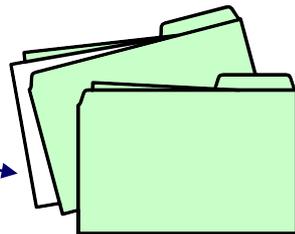


# → Identifier les liens entre objets



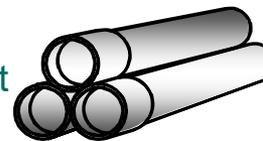
Client

Le Client passe une commande



Commande

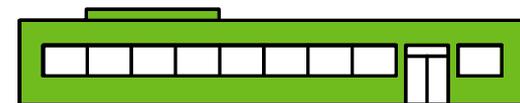
Produit



Expédition



Réglement



Magasin



Compte Client

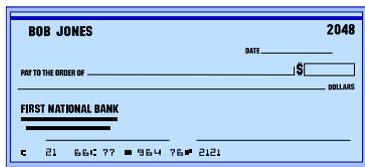
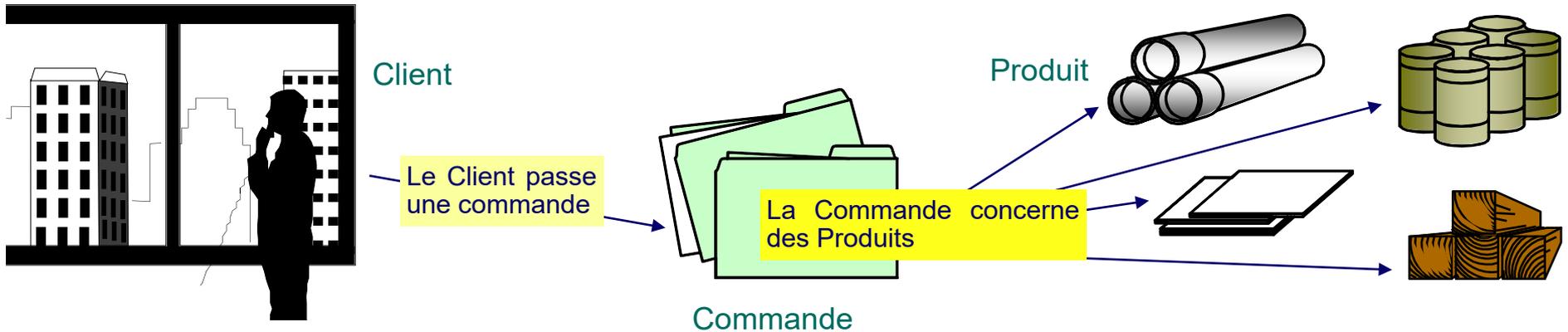


Facture

Bon de Commande de réapprovisionnement



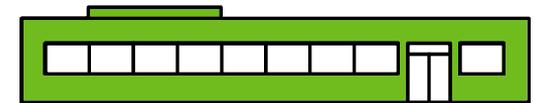
# → Identifier les liens entre objets



Réglement



Expédition



Magasin



Compte Client



Facture

Bon de Commande de réapprovisionnement



# → Identifier les liens entre objets



Client

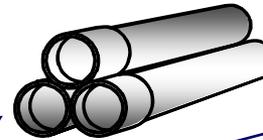
Le Client passe une commande



Commande

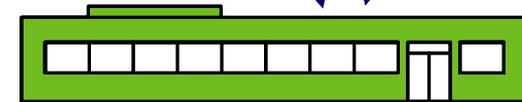
La Commande concerne des Produits

Produit

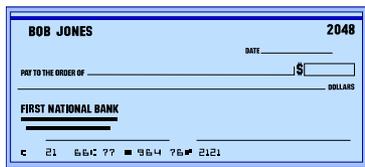


Expédition

Les Produits sont stockés dans un Magasin



Magasin



Réglement



Compte Client



Facture

Bon de Commande de réapprovisionnement



# → Identifier les liens entre objets



Client

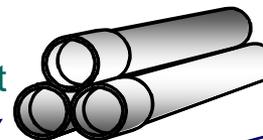
Le Client passe une commande



Commande

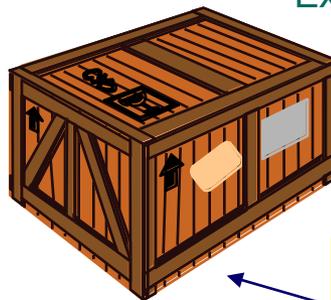
La Commande concerne des Produits

Produit

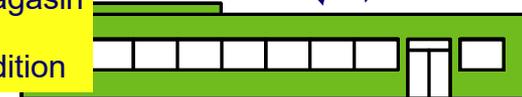


Expédition

Les Produits sont stockés dans un Magasin



Le Magasin monte l'Expédition



Magasin



Réglement



Compte Client

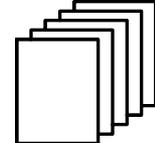
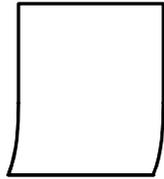
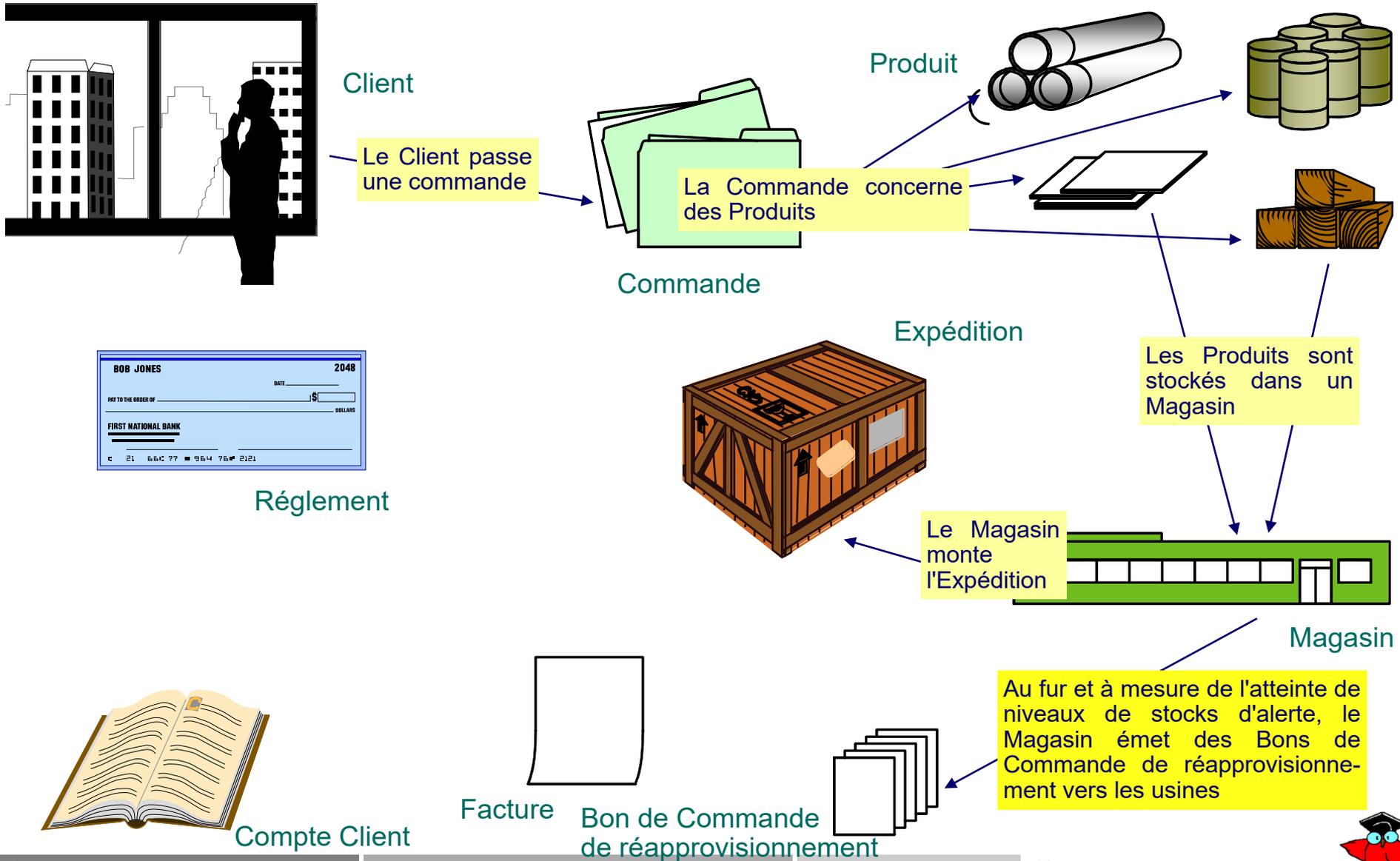


Facture

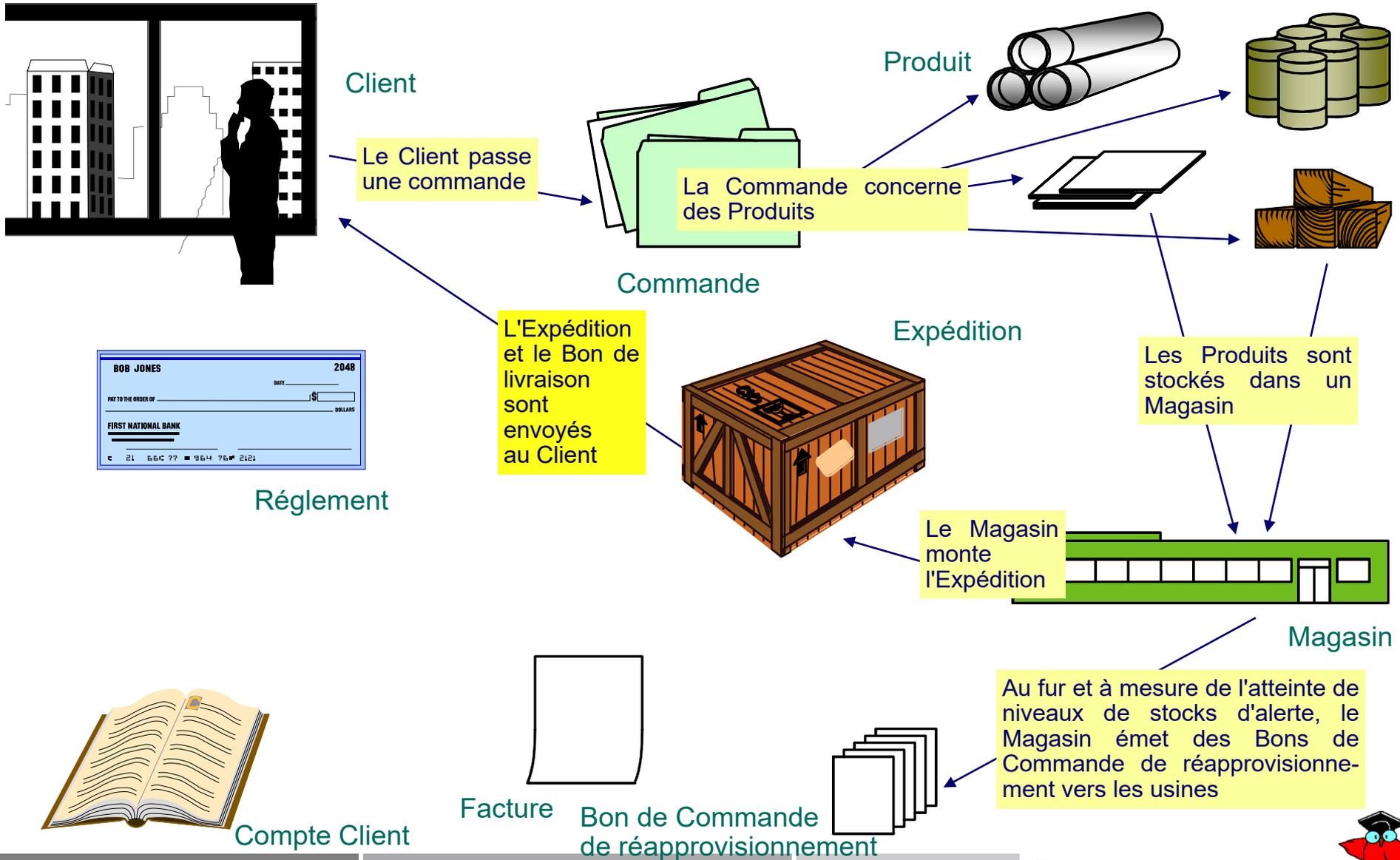
Bon de Commande de réapprovisionnement



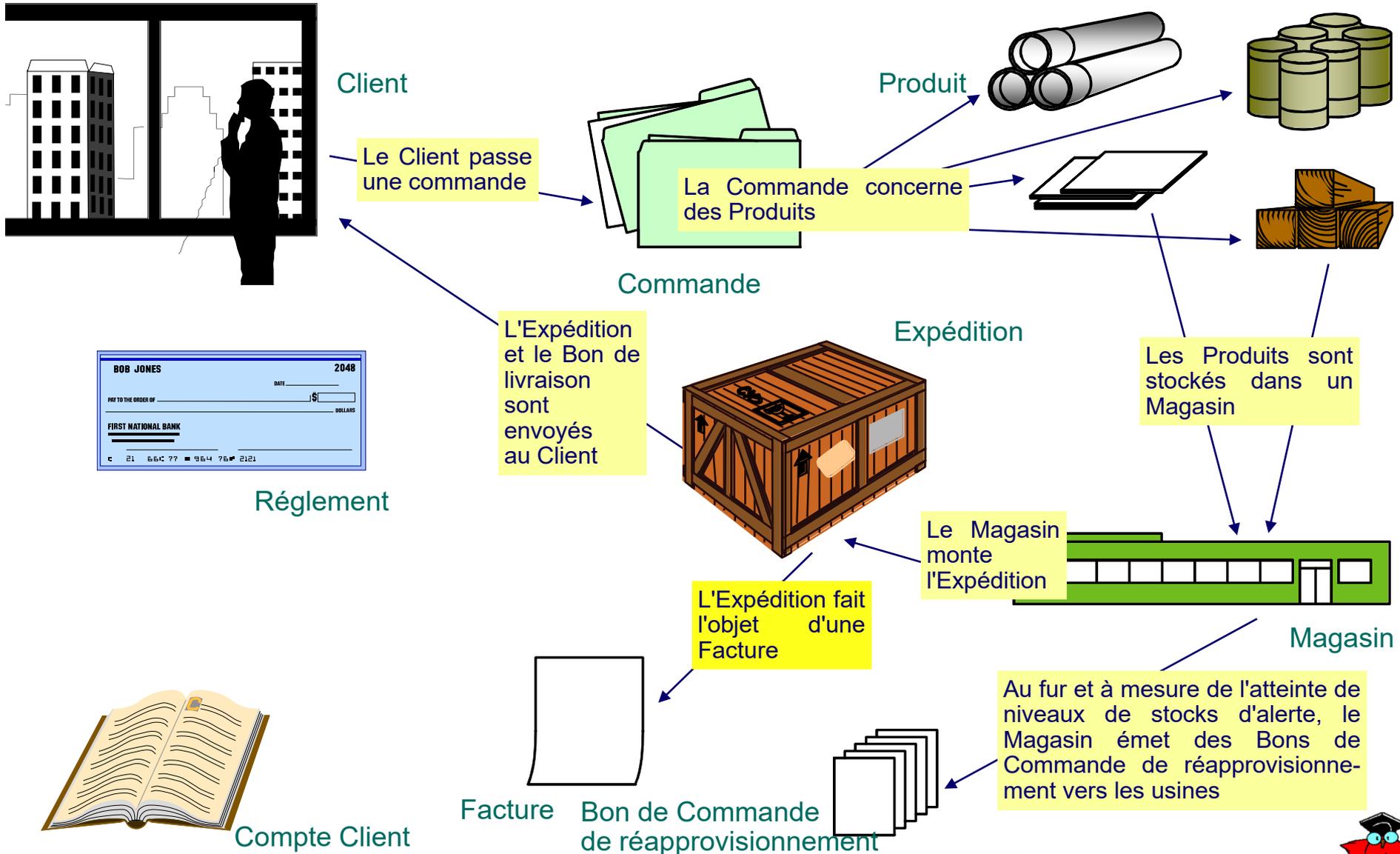
# → Identifier les liens entre objets



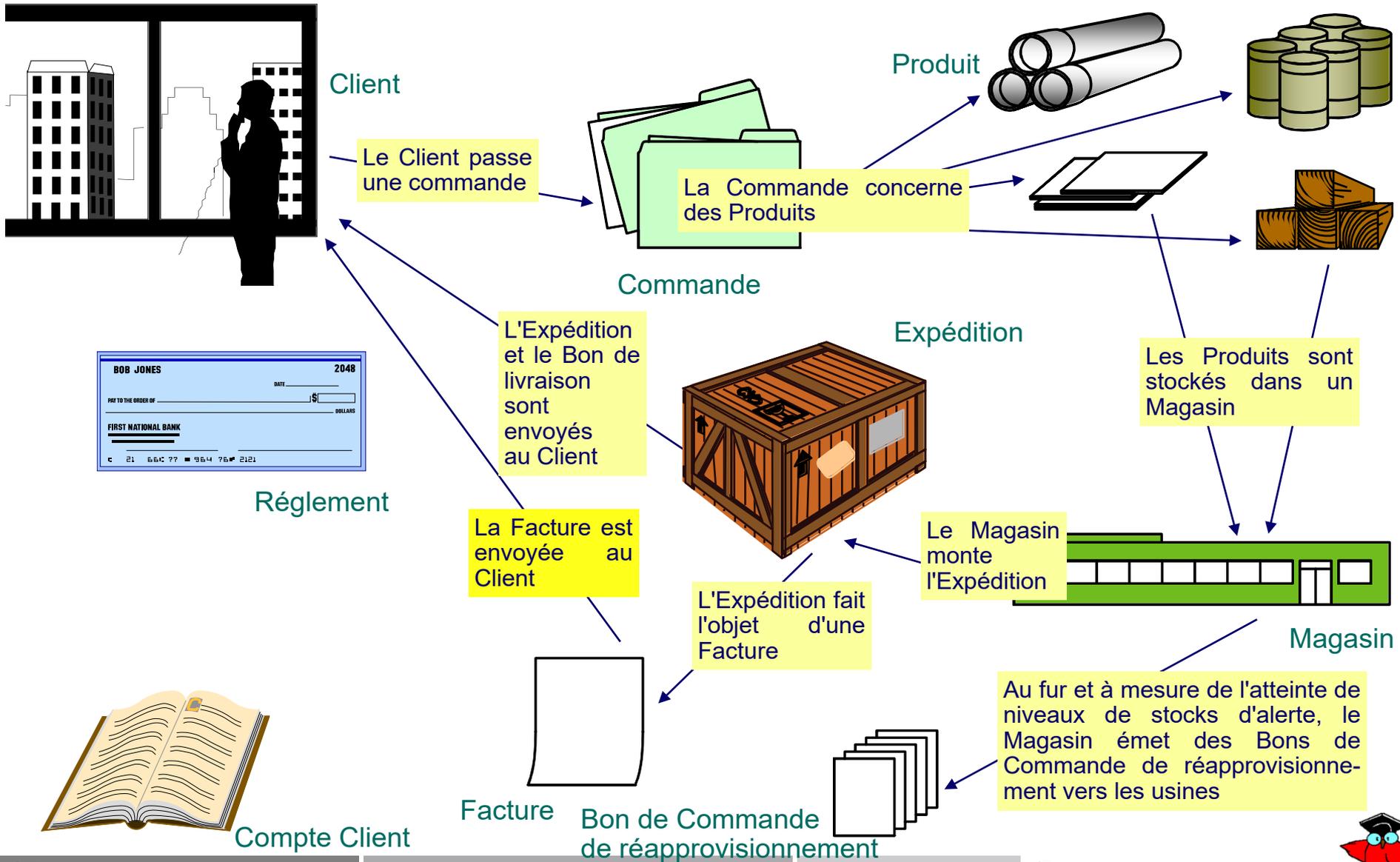
# → Identifier les liens entre objets



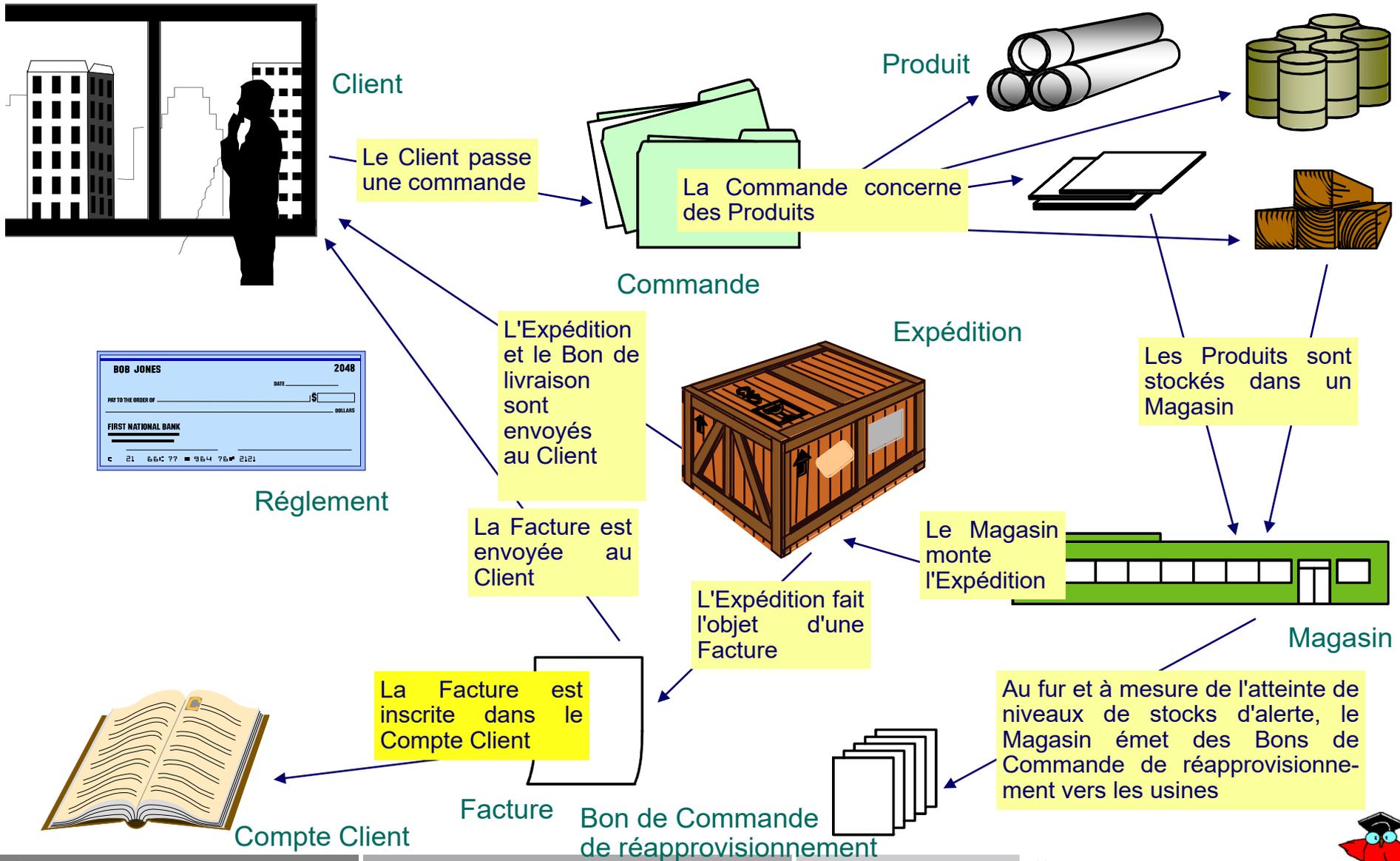
# → Identifier les liens entre objets



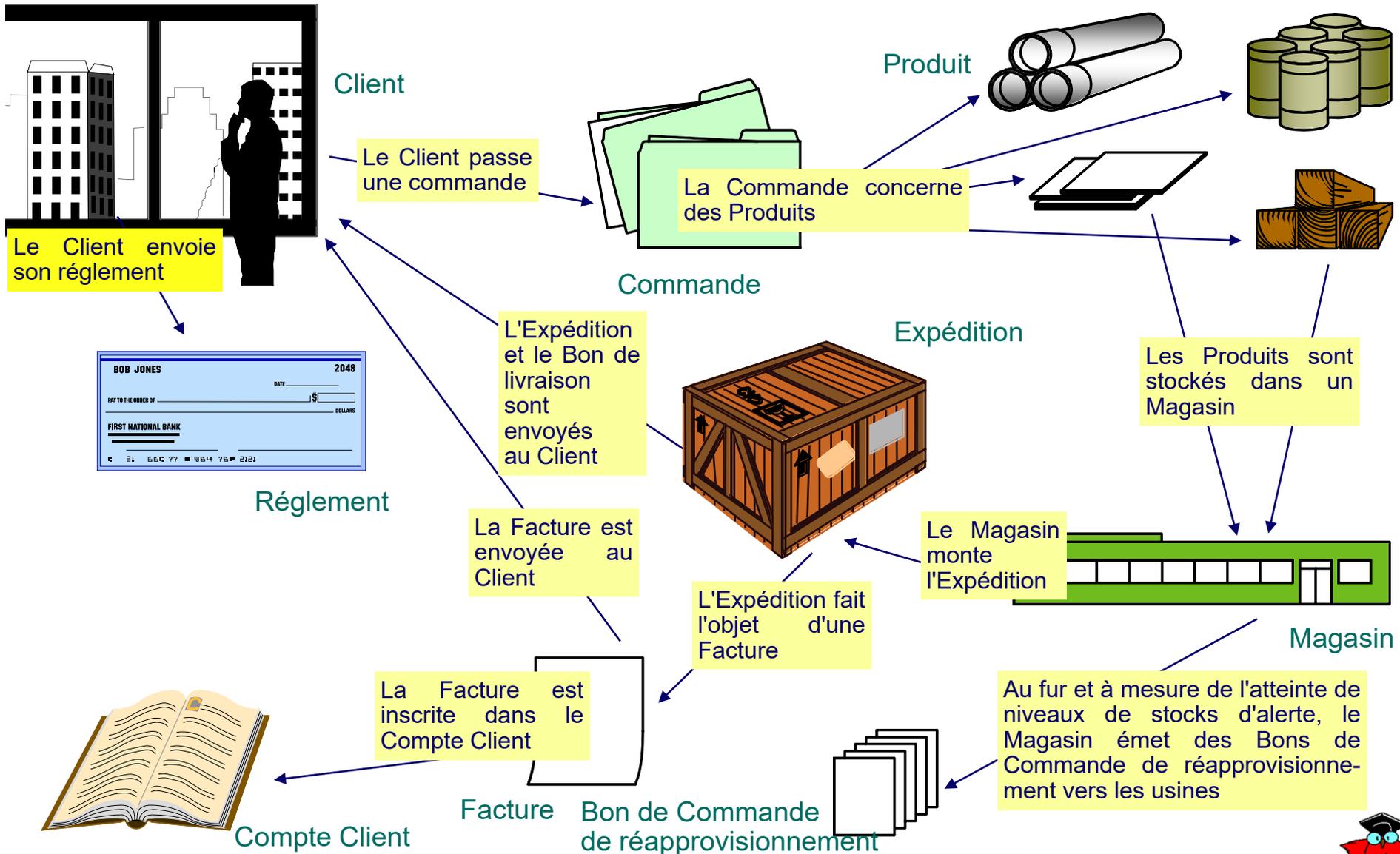
# → Identifier les liens entre objets



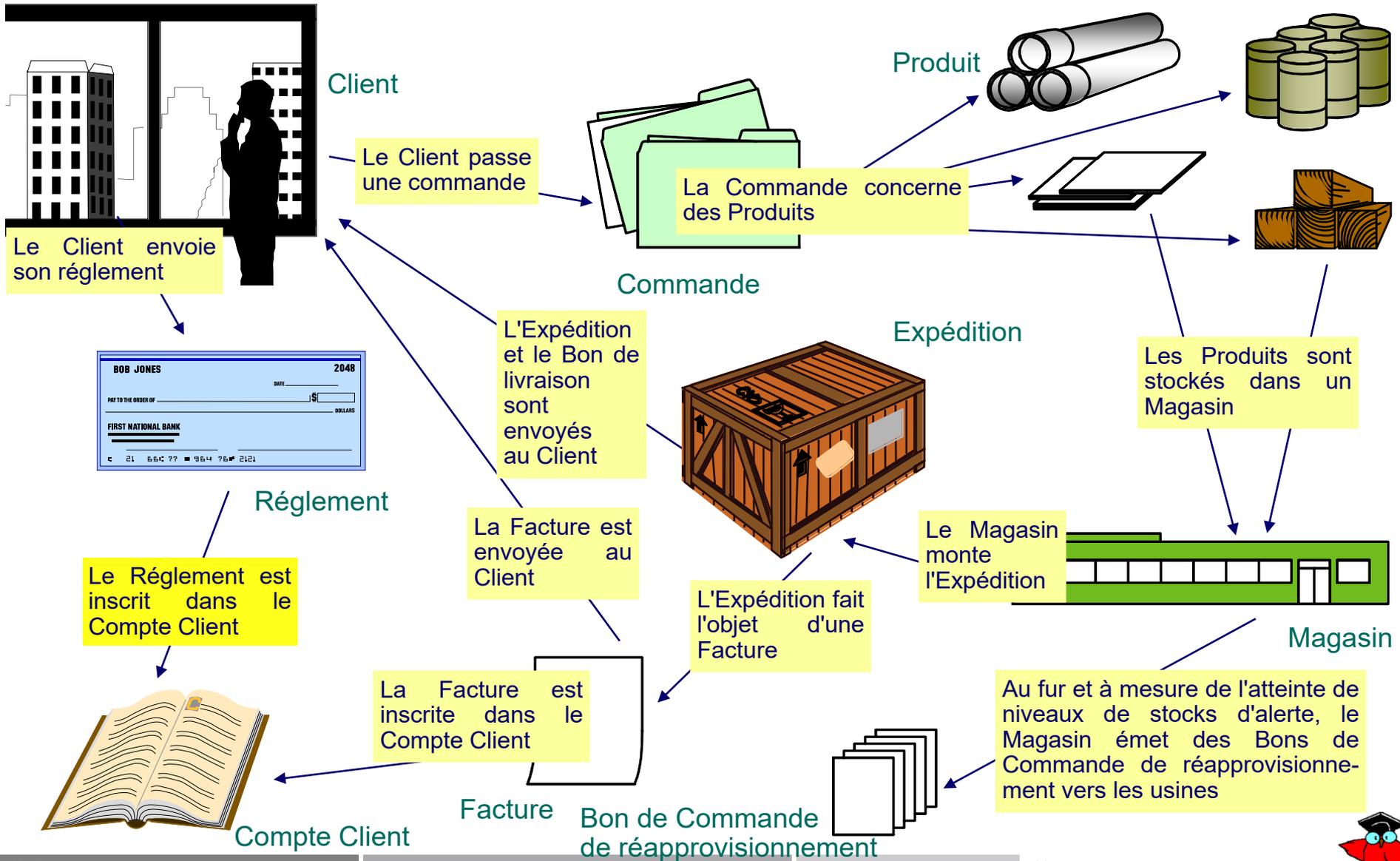
# → Identifier les liens entre objets



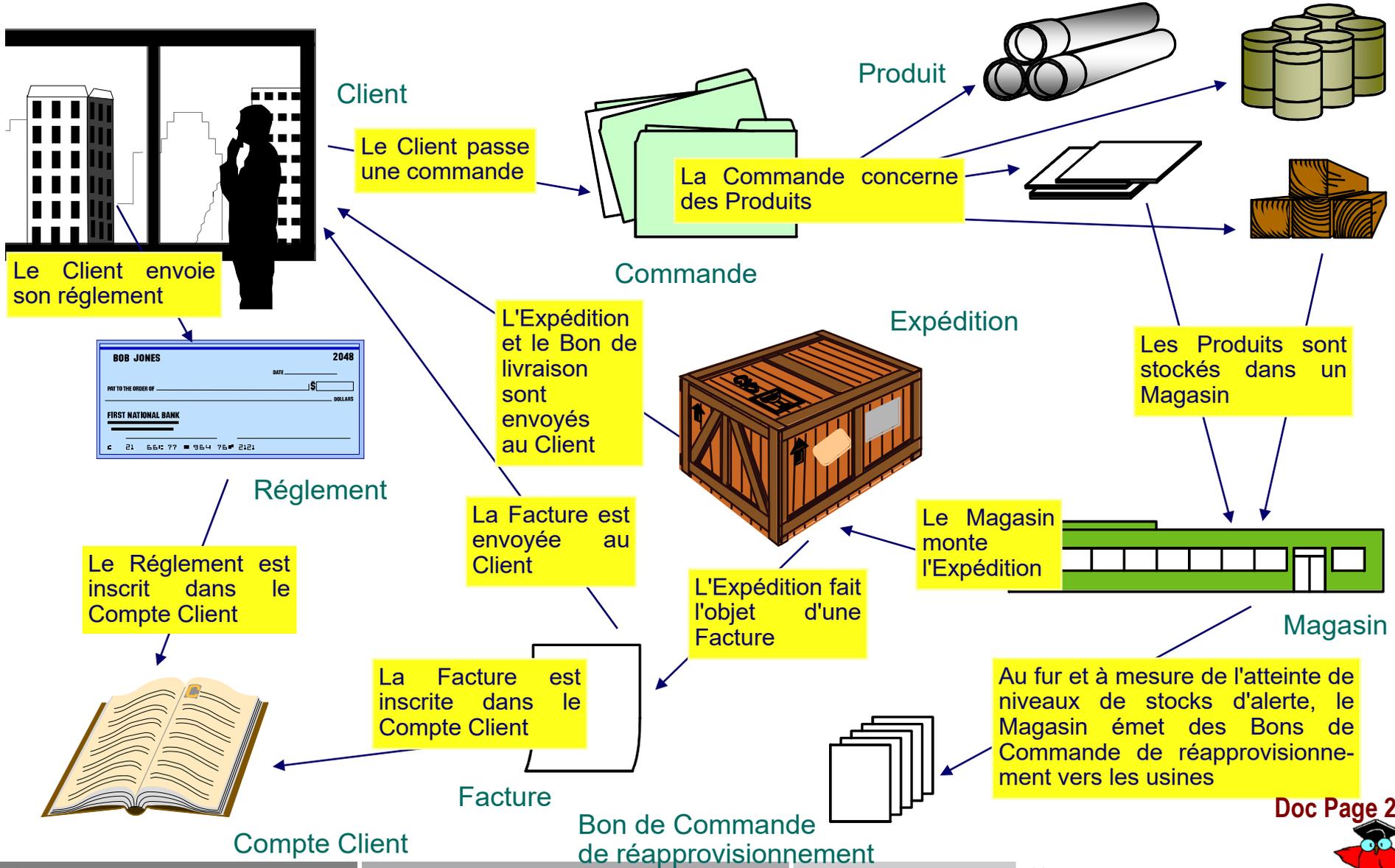
# → Identifier les liens entre objets



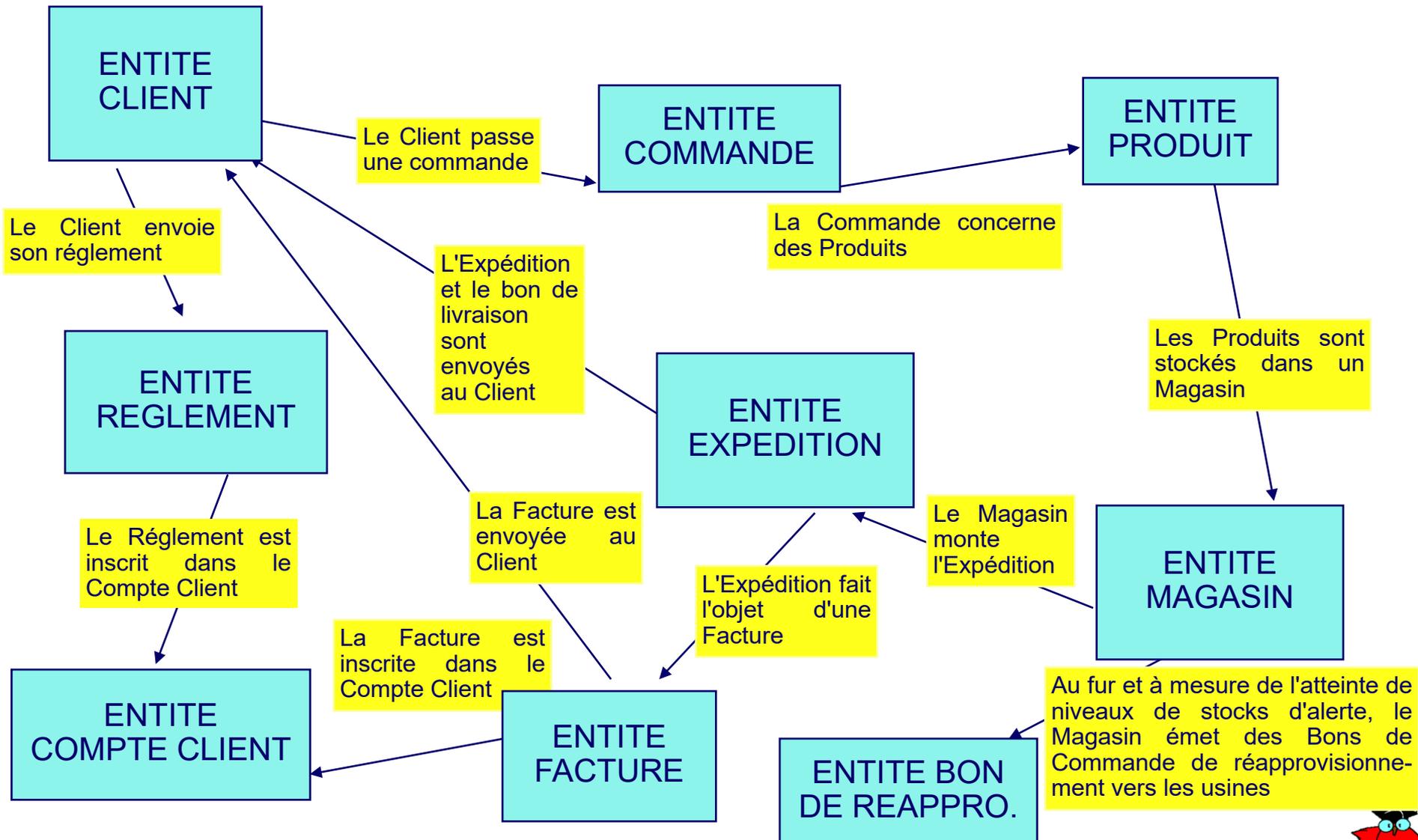
# → Identifier les liens entre objets



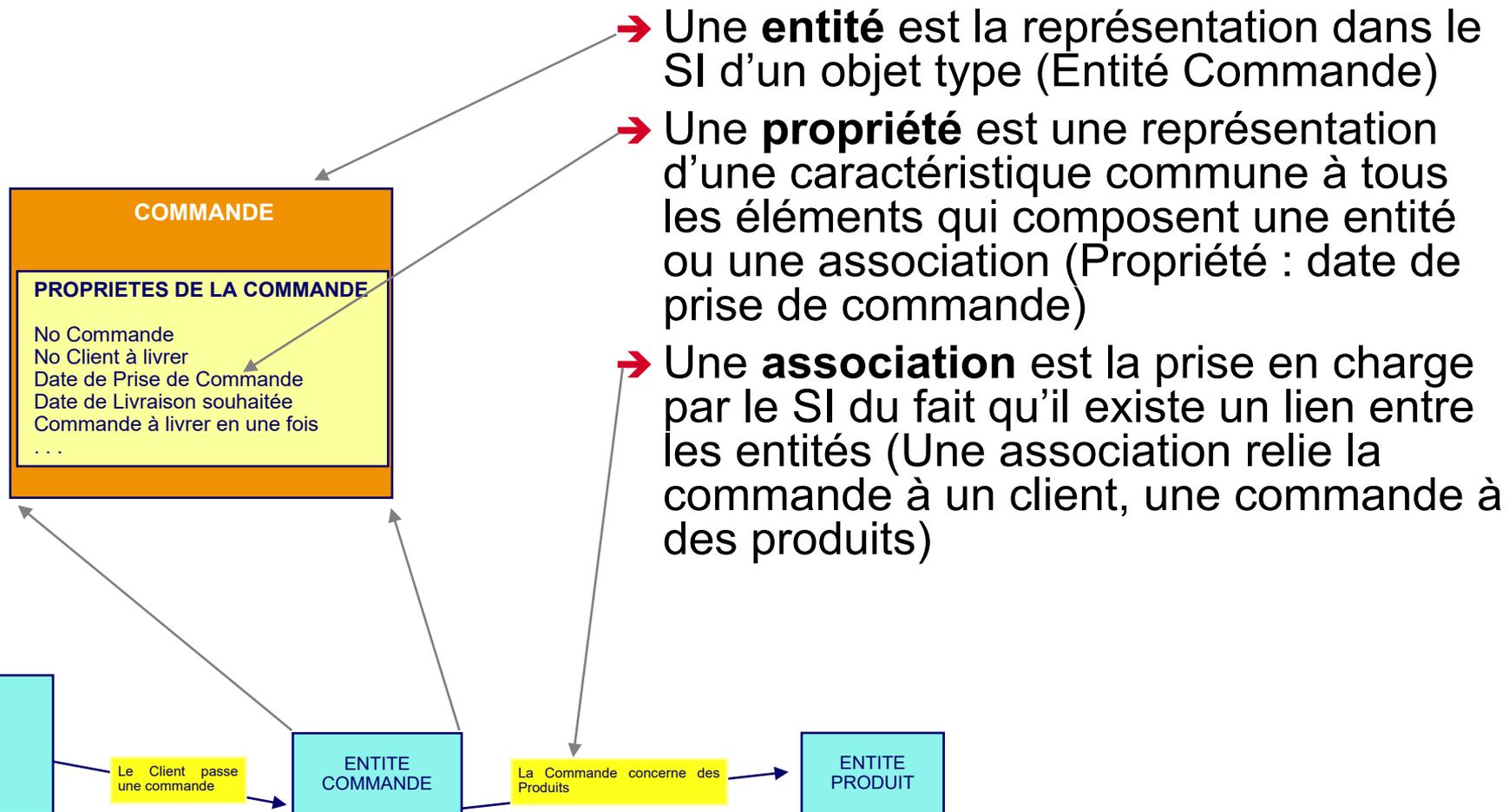
# → Identifier les liens entre objets



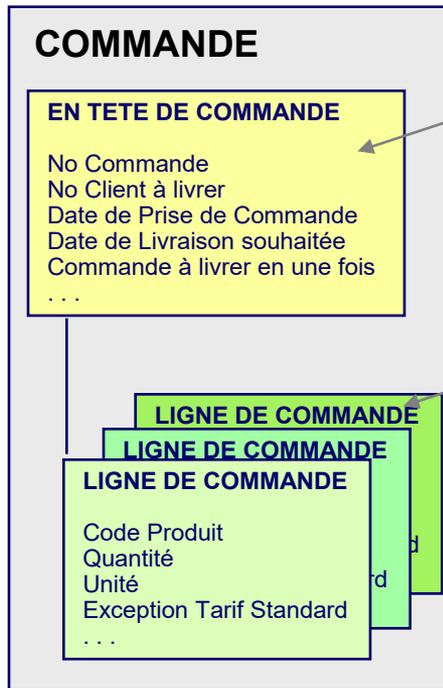
# → Modéliser les entités



# → Entités, associations et propriétés



# → Entités, associations et propriétés



- Dans ce cas, **l'entité** Commande représentera en fait l'en-tête de commande
- La ligne de commande sera représentée par **l'association** qui relie l'entité Commande à l'entité Produit
- Une ligne de commande associe un produit donné à une commande donnée



# → Entités, associations et propriétés

## COMMANDE

### CommandeNo

CommandeReque le  
Date souhaitée livraison  
CmmandeALivrerEnUneFois

## Commander

Quantité  
Unité  
Exception Tarif Standard

## PRODUIT

### ProduitNo

ProduitCategorie  
ProduitIntitule

1,n

0,n

→ Le bon schéma

→ Deux entités

– L'entité Commande

– L'entité Produit

→ Une association

– L'association Commander

– Que représentent les chiffres 0,n et 1,n ?

Implicitement :

#CommandeNo

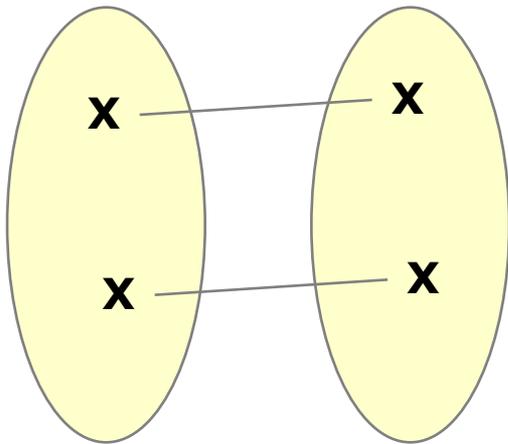
#ProduitNo

Un identifiant de la ligne de commande formé par la concaténation des deux clefs primaires identifiant une commande et un produit

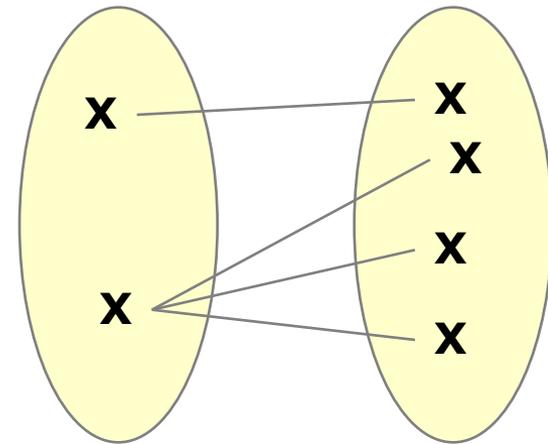




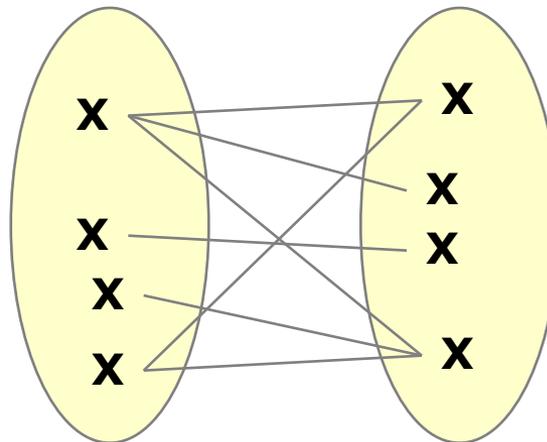
# Les cardinalités



Association 1-1



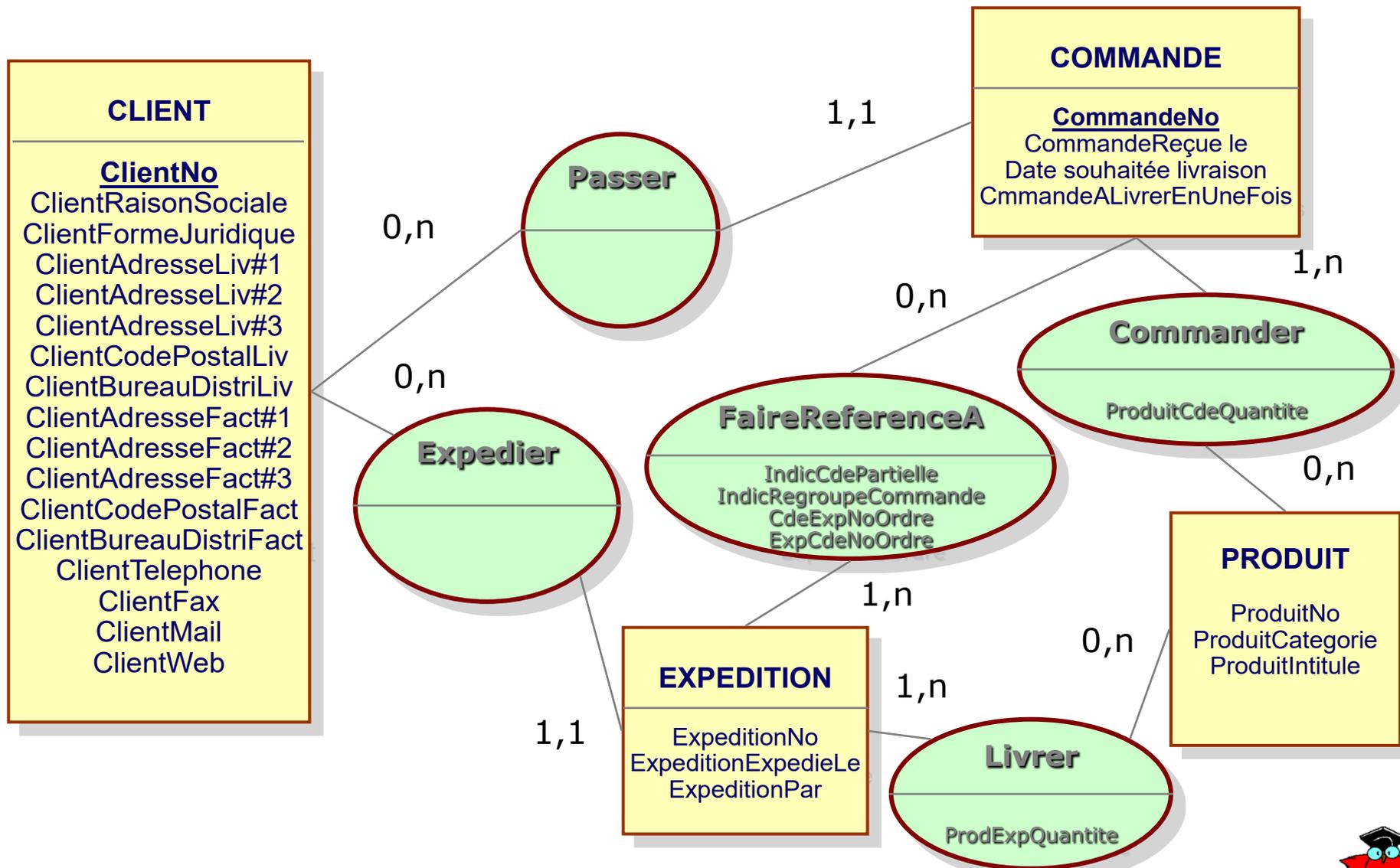
Association 1-n (hiérarchique)



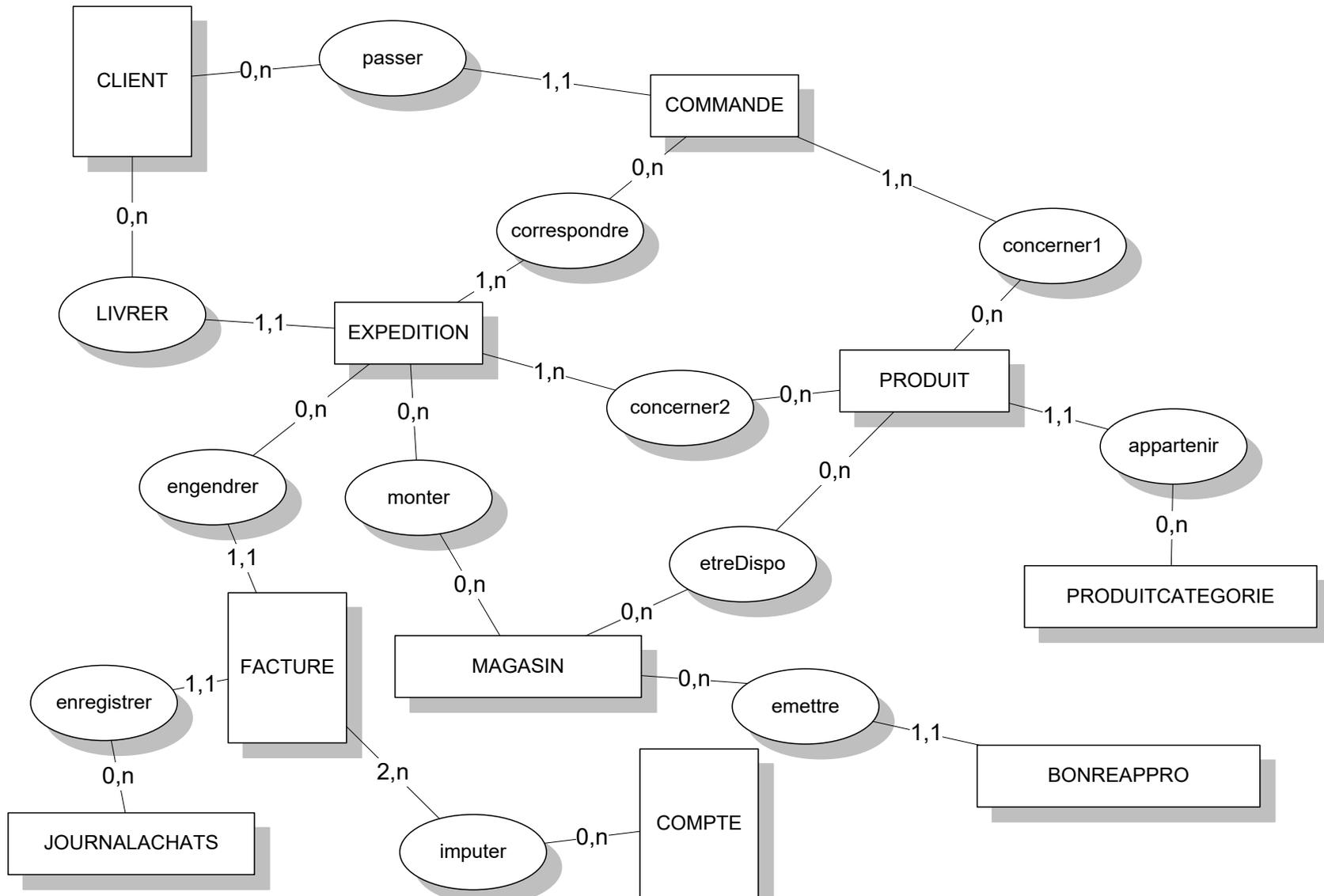
Association n-p (non hiérarchique)



# Modéliser entités et associations : le MCD



# → Un MCD avec un EDI (Microsoft Visio)



# → Les tables du modèle relationnel

Table des CLIENTS

ClientNo	Raison sociale	Forme juridique...
01	MyBestCustomer	SA
02	Dupont	SARL

Table des COMMANDES

CommandeNo	ClientNo	Reçue le	Livraison souhaitée le...
1001	01	05/11/08	22/12/08
1002	02	05/11/08	12/12/08
1005	01	10/11/08	05/01/09
1012	01	12/11/08	09/01/09

Table des LignesDeCommande

CommandeNo	ProduitNo	Qté
1001	01	20
1001	03	30
1002	02	100

Table des PRODUITS

ProduitNo	ProduitCatégorie	ProduitIntitule
01	07	Rondelle
02	07	Vis
03	09	Interrupteur



# → Le MRD (modèle relationnel des données)

CLIENT (ClientNo, ClientRaisonSociale, ClientFormeJuridique, ClientAdresseLiv#1, ClientAdresseLiv#2, ClientAdresseLiv#3, ClientCodePostalLiv, ClientBureauDistriLiv, ClientAdresseFact#1, ClientAdresseFact#2, ClientAdresseFact#3, ClientCodePostalFact, ClientBureauDistriFact, ClientTelephone, ClientFax, ClientMail, ClientWeb)

COMMANDE (CommandeNo, CommandeReçue le, Date souhaitée livraison, CommandeALivrerEnUneFois, #ClientNo)

PRODUIT (ProduitNo, ProduitCategorie, ProduitIntitule)

EXPEDITION (ExpeditionNo, ExpeditionExpedieLe, ExpeditionPar, #ClientNo)

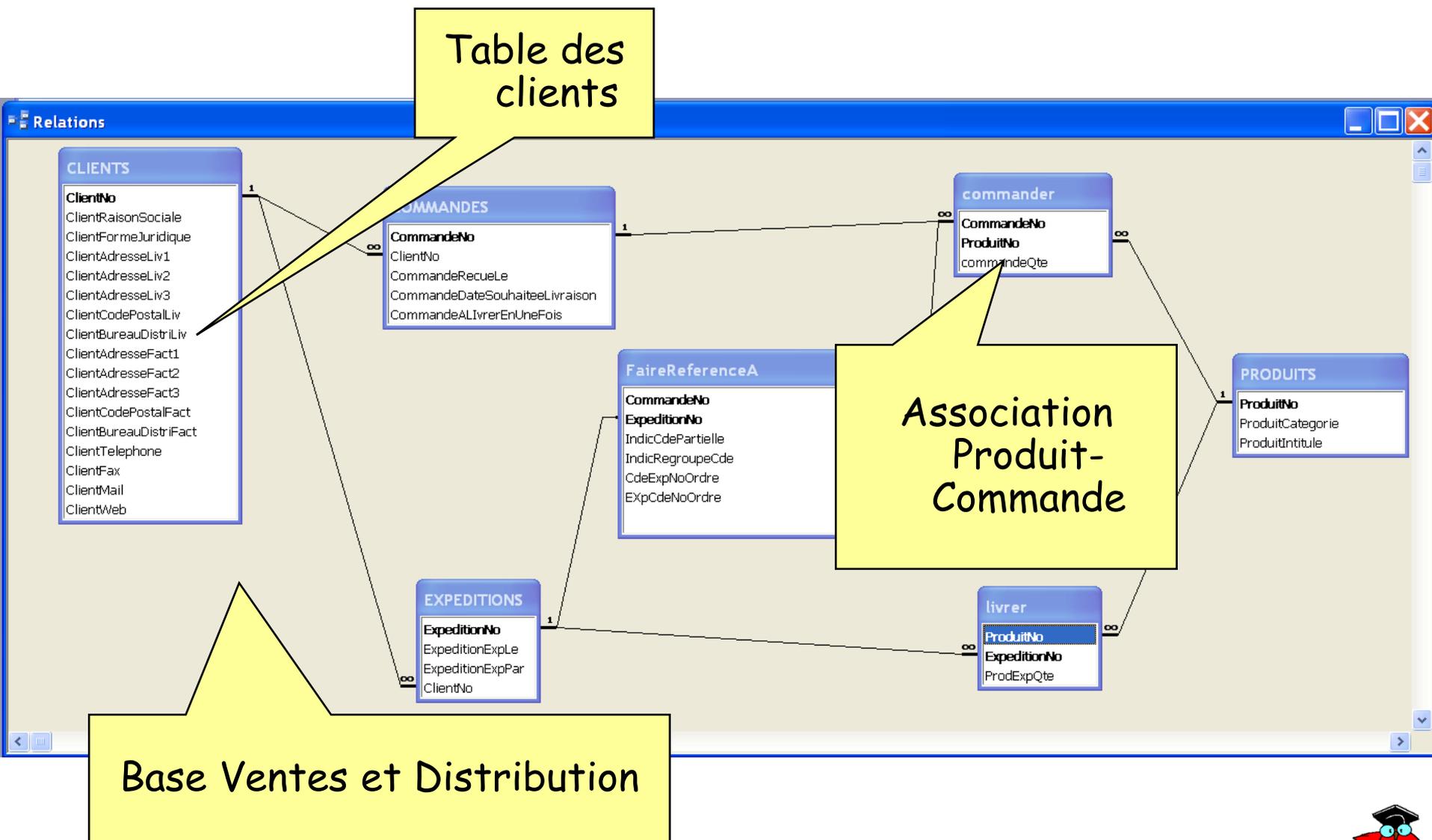
FaireReferenceA (#CommandeNo, #ExpeditionNo, IndicCdePartielle, IndicRegroupeCommande, CdeExpNoOrdre, ExpCdeNoOrdre)

Commander (#CommandeNo, #ProduitNo, ProduitCdeQuantite)

Livrer (#ExpeditionNo, #ProduitNo, ProdExpQuantite)



# → Implantation d'une base relationnelle (Access)



# → Implantation d'une base relationnelle (Access)

baseReference : Base de données (format de...)

Ouvrir Modifier Nouveau

### CLIENTS : Table

	ClientNo	ClientRaisonSociale	ClientFormeJur	ClientAdresseLiv1	ClientAdresseLiv2	ClientAdresse
+	1	Oranaise de Bâtiments	SA		8, rue de la Méditerranée	
+	2	Alger TP	SA		18, Bd de la corniche	
+	3	Blida Constructions	SARL		5, allée du désert	
+	4	Les nouvelles pyramides SA	SA		80, avenue de collines	
+	5					
*	(NuméroAuto)					

Enr : 1 sur 5

### COMMANDES : Table

	CommandeNo	ClientNo	CommandeRecueLe	CommandeDateSouh	CommandeALivrerE
+	1	4	26/11/2008	10/12/2008	<input type="checkbox"/>
+	2	3	26/11/2008	12/12/2008	<input checked="" type="checkbox"/>
+	3	2	27/11/2008	20/12/2008	<input type="checkbox"/>
+	4	4	27/11/2008	05/01/2009	<input checked="" type="checkbox"/>
*	(NuméroAuto)	0			<input type="checkbox"/>

Enr : 1 sur 4

### PRODUITS : Table

	ProduitNo	ProduitCategorie	ProduitIntitule
+	1	1	Rondelle AM6 Diam 12 (Produit No 1, de catégorie1)
+	2	1	Vis Parker V23 Diam 7 (Produit No 2, de catégorie 1)
+	3	2	Contacteur X34 24V (Produit No 3, de catégorie 2)
*	(NuméroAuto)	0	

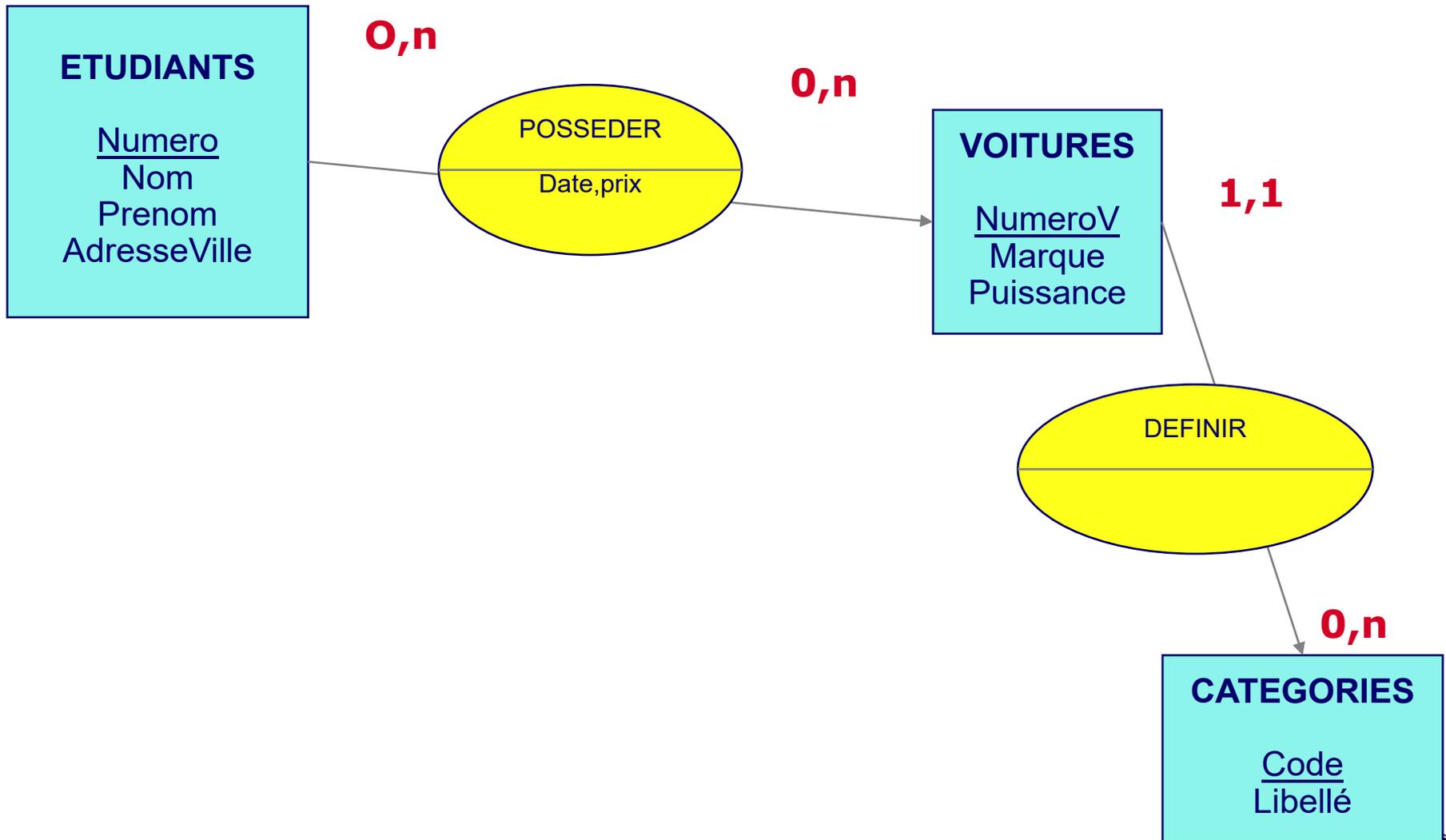


# → Big Data

- Sous les effets d'une interconnexion toujours plus grande et d'un coût de stockage de plus en plus bas, nous sommes entrés dans une ère de déluge informationnel.
- Chaque acteur économique (entreprise, particulier ou administration) génère et sauvegarde perpétuellement des données plus ou moins structurées : géolocalisation par smartphone, fichiers de santé, transactions marchandes en ligne, commentaires sur les réseaux sociaux, photos sur Flickr, microblogs, capteurs RFID/NFC (sans contact), communications machine to machine, etc.
- McKinsey estime qu'en 2010, les entreprises auraient stocké 7 Eo (exaoctets= $10^{18}$  octets) supplémentaires de données, et les particuliers 6 Eo.
- Un exaoctet représentant 4 000 fois le contenu de la Bibliothèque du Congrès des Etats-Unis.
- L'idée du **Big Data** est de transformer toutes ces données brutes en mine d'or. Elle est en partie liée à celle de **l'informatique décisionnelle**.
- Le modèle relationnel ne s'applique plus. Il faut trouver des méthodes pour s'y ramener.

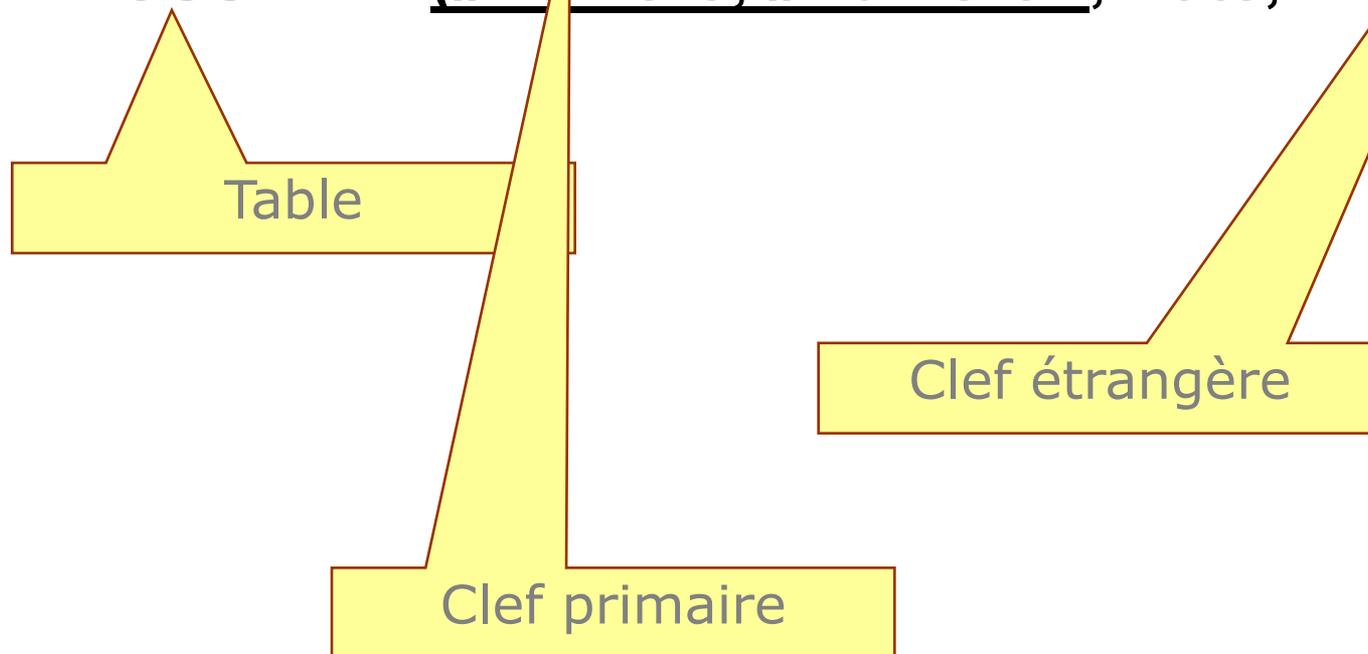


# → Un exemple pour s'entraîner (MCD)



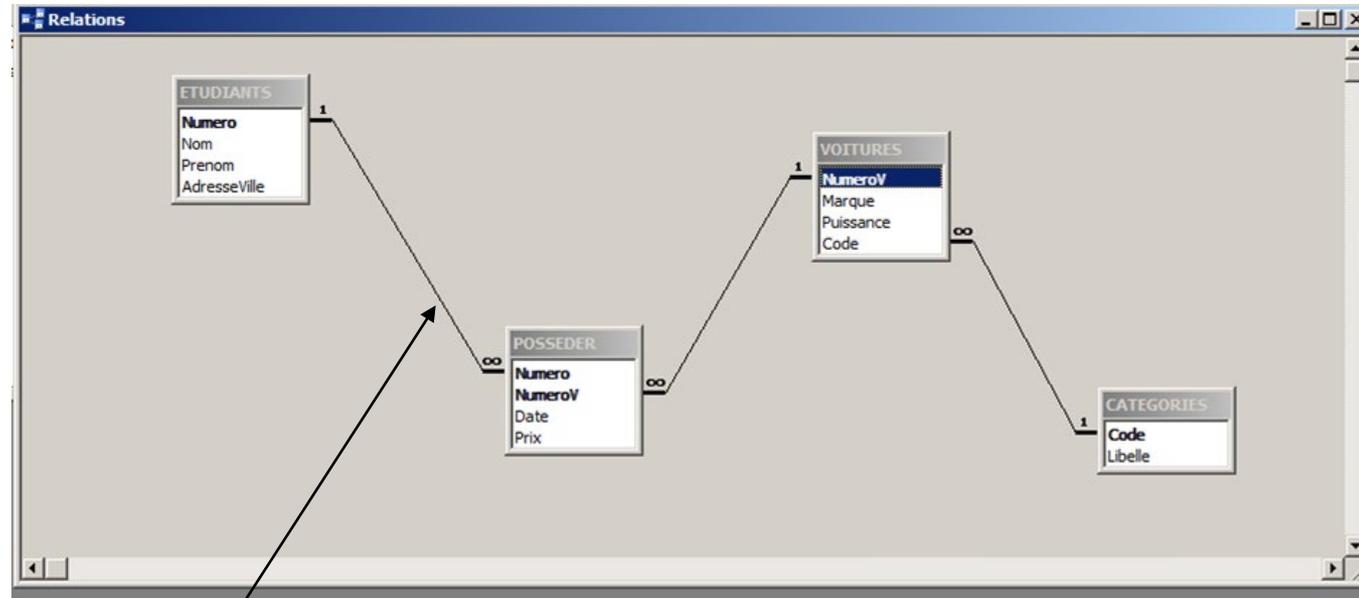
# → Un exemple pour s'entraîner (MRD)

- ❑ **ETUDIANTS** (Numero, Nom, Prenom, AdresseVille)
- ❑ **VOITURES**(NumeroV, Marque, Puissance, #Code)
- ❑ **CATEGORIES**(Code, Libelle)
- ❑ **POSSEDER**(#Numero, #NumeroV, Date, Prix)



# → Un exemple pour s'entraîner (Construction tables)

- Etudiants
- Voitures
- Catégories
- Posséder



- *Définition des relations (au sens Access) entre tables (entités et associations non hiérarchiques)*



# → Un exemple pour s'entraîner (SQL)

- Le **langage SQL** est aujourd'hui le standard d'accès aux bases de données relationnelles. Il comporte quatre opérations essentielles.
- La recherche (**SELECT**) permet de retrouver les lignes (tuples) vérifiant les critères qualifiés en arguments.
- L'insertion (**INSERT**) permet d'ajouter des tuples dans une relation.
- La suppression (**DELETE**)
- La mise à jour (**UPDATE**)
- SQL est normalisé par l'ISO.





# Un exemple pour s'entraîner (Clause SELECT)

- SELECT      liste d'attributs
- FROM        table(s)
- WHERE       condition
- GROUP BY    liste des attributs de regroupement
- HAVING       clause
- ORDER BY    liste des attributs de tri





# Un exemple pour s'entraîner (Clause SELECT)

- ❑ Toutes les informations de l'étudiant dont le No est 5
- ❑ **SELECT**
- ❑ **????**
- ❑ **FROM**
- ❑ **??????**
- ❑ **WHERE**
- ❑ **????????????;**





# Un exemple pour s'entraîner (Clause SELECT)

- ❑ **Toutes les informations de l'étudiant dont le No est :**
- ❑ **SELECT**
- ❑ **Numero, Nom, Prenom, AdresseVille**
- ❑ **FROM**
- ❑ **Etudiants**
- ❑ **WHERE**
- ❑ **Numero = 5;**





# Un exemple pour s'entraîner (Clause SELECT)

- A quelle date "Larfi" a t il acquis une "VW Golf "?





# Un exemple pour s'entraîner (Clause SELECT)

- ❑ A quelle date "Larfi" a t il acquis une "VW Golf"??
- ❑ SELECT
- ❑     Date
- ❑ FROM
- ❑     Voitures, Etudiants, Posseder
- ❑ WHERE
- ❑     Prenom="Larfi"
- ❑     AND Marque="VW Golf"
- ❑     AND Etudiants.Numero = Posseder.Numero
- ❑     AND Voitures.NumeroV = Posseder.NumeroV;

JOINTURE



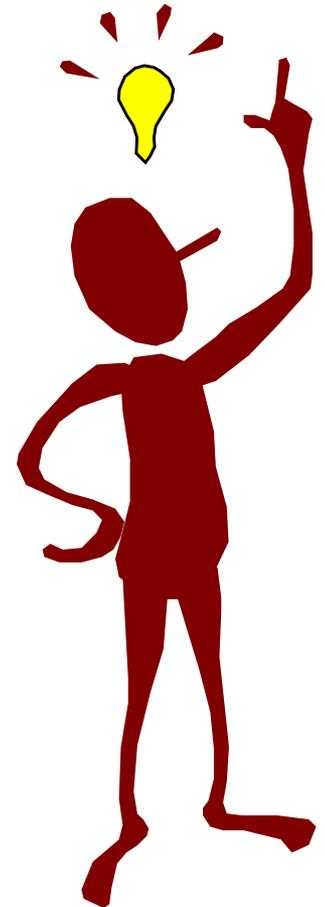
# → Un exemple pour s'entraîner (Exercices)

- ❑ Quels sont les étudiants (nom et prénom) qui possèdent ou ont possédé une voiture de collection ?
- ❑ Quelles sont les marques de voitures achetées par les étudiants après le 10/1/2004 inclus ?
- ❑ Dans quelles villes habitent les étudiants qui ont des voitures ?
- ❑ Quelles sont les dates d'achat des voitures dont la puissance est supérieure ou égale à 7CV ?
- ❑ Nom des étudiants qui habitent la même ville que DUPONT ?
- ❑ Dans quelles villes habitent les étudiants qui possèdent une voiture dont le prix est supérieur au prix moyen ?



# → Quelles réponses à nos questions ?

- Quelles différences entre données structurées et non structurées ?
- On peut accéder directement à une donnée structurée au moyen de son adresse (dans un tableau, grâce au No de ligne et au No de colonne).
- On ne peut accéder à une donnée non structurée qu'en ayant lu toutes les données précédentes.



# → Quelles réponses à nos questions ?

- Qu'est ce qui caractérise un tri par fusion ?
- L'objectif de l'algorithme de tri par fusion est de construire la liste qui représente la fusion des deux listes triées dans l'ordre croissant de sorte que cette liste résultante soit triée dans l'ordre croissant.



# → Quelles réponses à nos questions ?

- Quels modèles de données permettent la construction des bases de données exploitées par les applications ?
- Ce sont les MCD (Modèle conceptuel de données) et le MRD (Modèle Relationnel de Données).



# → Quelles réponses à nos questions ?

- Quelle démarche pour construire et exploiter une base de données ?
- Pour construire : ce sont les phases d'analyse (diagramme de flux, dictionnaire des données, diagramme des Dépendances fonctionnelles et normalisation), de conception (MCD) et de déploiement au sein d'un SGBD (MRD).
- Pour exploiter : le langage SQL

