

Entité, attribut, occurrence

Un fichier de réservations tient très bien dans un tableur... jusqu'au jour où il devient ingérable. En le démontant, tu découvres les trois mots qui servent à décrire n'importe quelles données : entité, attribut et occurrence.

4TTR

5TTR

6TTR



Découverte

Avant de créer une base de données, il faut savoir **ce qu'on va y ranger**. Le meilleur moyen de le comprendre, c'est de partir d'un fichier qui fonctionne mal et de chercher pourquoi.



Télécharge `tickets.csv` dans la sidebar et ouvre-le avec SmoothCsv (un tableur tel que Excel ou LibreOffice Calc fonctionne aussi)



Objectifs

À la fin de cet article, tu seras capable de :

1. Repérer dans un tableau les informations répétées, mal écrites ou manquantes.
2. Découper un tableau en groupes d'informations qui parlent chacun d'une seule chose.
3. Distinguer une entité, un attribut et une occurrence.
4. Repérer les entités dans un texte qui décrit un besoin.
5. Dessiner une entité dans un schéma entité-association.

Un fichier qui semble suffire

Une association organise des soirées quiz à Namur, à Liège et à Wavre. Pour suivre ses réservations, elle tient un fichier `tickets.csv`. Chaque ligne correspond à un ticket vendu.



NOUVELLE NOTION : LE FICHIER CSV

Un fichier **CSV** est un simple fichier texte qui contient un tableau. Chaque ligne du fichier est une ligne du tableau, et les colonnes sont séparées par un caractère convenu, ici le point-virgule `;`. Tous les tableurs savent l'ouvrir.

Le fichier compte **60 lignes** et **12 colonnes**. Voici les colonnes, avec la valeur de la première ligne :

Colonne	Valeur sur la première ligne
ticket_id	T001
nom_client	Dupont
prenom_client	Marie
email_client	marie.dupont@email.be
telephone_client	0471 12 34 56
evenement	Soirée Quiz Pop Culture
date_evenement	2026-10-03
lieu	Salle des Fêtes
ville	Namur
prix_ticket	12
statut_ticket	payé
code_ticket	QZ-1001

À première vue, tout est là. On sait qui a acheté quoi, pour quelle soirée, où, et à quel prix. Alors où est le problème ?

La chasse aux anomalies

Travaille à deux, pendant dix minutes. Réponds aux questions suivantes en fouillant le fichier.

1. Combien de tickets Marie Dupont a-t-elle achetés ? Combien de fois son numéro de téléphone est-il écrit ?
2. Combien d'orthographes différentes trouves-tu pour la soirée « Blind Test Années 90 » ?
3. Combien de façons différentes d'écrire une date y a-t-il ?
4. Un prix n'est pas écrit en chiffres. Sur quelle ligne ?
5. Un client n'a pas de numéro de téléphone. Sur quelle ligne ?
6. Quelles autres différences d'écriture repères-tu dans les colonnes **lieu** et **ville** ?

💡 MÉTHODE : LE FILTRE AUTOMATIQUE

Sélectionne la ligne des titres, puis active le filtre (*Données > Filtrer* dans Excel, *Données > Plus de filtres > AutoFiltre* dans LibreOffice). La petite flèche de chaque colonne affiche **toutes les valeurs différentes** de cette colonne. Une faute de frappe y saute aux yeux.

► Voir ce qu'il fallait trouver

Toutes ces anomalies ne se ressemblent pas. On peut les ranger en trois familles.

Première famille : la même information est recopiée encore et encore. L'adresse e-mail de Marie Dupont apparaît 6 fois. Le nom « Salle des Fêtes » et la ville « Namur » apparaissent sur toutes les lignes de la soirée quiz.

NOUVELLE NOTION : LA REDONDANCE

Il y a **redondance** quand une même information est écrite à plusieurs endroits. Chaque copie est une occasion de se tromper, et chaque modification doit être faite partout.

En anglais : GB *redundancy*.

Deuxième famille : la même chose est écrite de plusieurs façons. Liège et Liège désignent la même ville. 10/10/2026 et 2026-10-10 désignent le même jour. Un humain le devine ; un programme, non. Si tu demandes à un programme de compter les tickets vendus à Liège, il en oubliera un.

NOUVELLE NOTION : L'INCOHÉRENCE

Il y a **incohérence** quand des données qui devraient être identiques ne le sont pas. Les données se contredisent, et plus rien ne dit quelle version est la bonne.

En anglais : GB *inconsistency*.

Troisième famille : des informations manquent ou ont un format impossible à exploiter. Un téléphone vide, un prix écrit quinze : impossible de calculer la recette de la soirée.

LA REDONDANCE REND LES INCOHÉRENCES POSSIBLES.

Si l'adresse de Marie Dupont n'était écrite qu'une seule fois, elle ne pourrait pas être écrite de deux façons différentes. Recopier une information, c'est prendre le risque de la recopier mal.

Une ligne raconte plusieurs choses

Regarde la première ligne du fichier, colonne par colonne. Les informations ne parlent pas toutes de la même chose :

1	ticket_id	T001	}	
2	statut_ticket	payé		le ticket
3	code_ticket	QZ-1001	}	
4	nom_client	Dupont	}	
5	prenom_client	Marie		la cliente
6	email_client	marie.dupont@email.be		
7	telephone_client	0471 12 34 56	}	
8	evenement	Soirée Quiz Pop Culture	}	

9	date_evenement	2026-10-03		la soirée
10	prix_ticket	12		
11	lieu	Salle des Fêtes		le lieu
12	ville	Namur		

Une seule ligne mélange **quatre sujets** : un ticket, une cliente, une soirée et un lieu. C'est exactement pour ça que les informations se répètent. Chaque fois que Marie Dupont achète un ticket, on recopie tout ce qu'on sait d'elle. Chaque fois qu'on vend un ticket pour la soirée quiz, on recopie tout ce qu'on sait de la soirée.

💡 OÙ RANGER LE PRIX ?

Toutes les places d'une même soirée coûtent le même prix : 12 euros pour la soirée quiz, 15 euros pour le blind test. Le prix décrit donc **la soirée**, pas le ticket. Si l'association proposait des tarifs réduits pour les étudiants, il décrirait le ticket. **C'est le besoin qui décide.**

Imagine maintenant quatre listes séparées : une liste des clients, une liste des soirées, une liste des lieux, une liste des tickets. Le téléphone de Marie Dupont ne serait écrit **qu'une seule fois**, dans la liste des clients. S'il change, on le corrige à un seul endroit.

Entité, attribut, occurrence

Tu viens de faire ce que fait tout informaticien avant de créer une base de données : découper les informations en groupes qui parlent chacun d'une seule chose. Ces groupes ont un nom.

📖 NOUVELLE NOTION : L'ENTITÉ

Une **entité** est une catégorie de choses dont on veut garder une liste : des clients, des soirées, des lieux, des tickets. On l'écrit au singulier et en majuscules : **CLIENT** , **EVENEMENT** , **LIEU** , **TICKET** .

En anglais : GB **entity**.

Pour savoir si tu tiens une entité, pose-toi une seule question :

🧠 MON APPLICATION DOIT-ELLE GARDER UNE LISTE DE... ?

Une liste de clients ? Oui : **CLIENT** est une entité. Une liste d'adresses e-mail ? Non : l'adresse e-mail n'existe pas seule, elle décrit un client.

Chaque entité est décrite par des informations.

📖 NOUVELLE NOTION : L'ATTRIBUT

Un **attribut** est une information que l'on garde à propos de chaque élément d'une entité. Le nom, le prénom, l'adresse e-mail et le téléphone sont des attributs de **CLIENT** .

En anglais : GB **attribute**.

Enfin, une entité est une catégorie générale. Marie Dupont, elle, est une cliente bien réelle.

NOUVELLE NOTION : L'OCCURRENCE

Une **occurrence** est un élément concret d'une entité. Marie Dupont est une occurrence de **CLIENT**. La « Soirée Quiz Pop Culture » du 3 octobre est une occurrence de **EVENEMENT**.

En anglais : GB **occurrence**, ou parfois *instance*.

Mot	Ce que c'est	Exemple
Entité	une catégorie de choses dont on garde une liste	CLIENT
Attribut	une information sur chaque élément de cette catégorie	email
Occurrence	un élément concret de cette catégorie	Marie Dupont

Entité ou attribut ?

La frontière n'est pas toujours évidente. Prends « le lieu » d'une soirée. Est-ce un attribut de **EVENEMENT**, comme la date ? Ou une entité à part entière ?

Regarde le fichier : un lieu possède lui-même une information, sa ville. Et l'association pourrait vouloir garder d'autres informations sur chaque lieu, comme son adresse ou le nombre de places. Le lieu mérite donc sa propre liste : c'est une **entité**.

SI UNE INFORMATION POSSÈDE ELLE-MÊME DES INFORMATIONS, C'EST PROBABLEMENT UNE ENTITÉ.

L'adresse e-mail d'un client ne se décrit pas davantage : c'est un attribut. Un lieu a une ville, une adresse, une capacité : c'est une entité.

Repérer les entités dans un texte

Dans la vraie vie, personne ne te donne un fichier à découper. On te décrit un besoin, avec des phrases. Voici celui de l'association :

« Un client achète des tickets. Chaque ticket donne accès à une soirée. Une soirée se déroule dans un lieu, à une date précise, et chaque soirée a son prix. »

Pour trouver les entités, cherche les phrases construites sur le modèle **sujet – verbe – complément**. Le sujet et le complément sont souvent des entités.

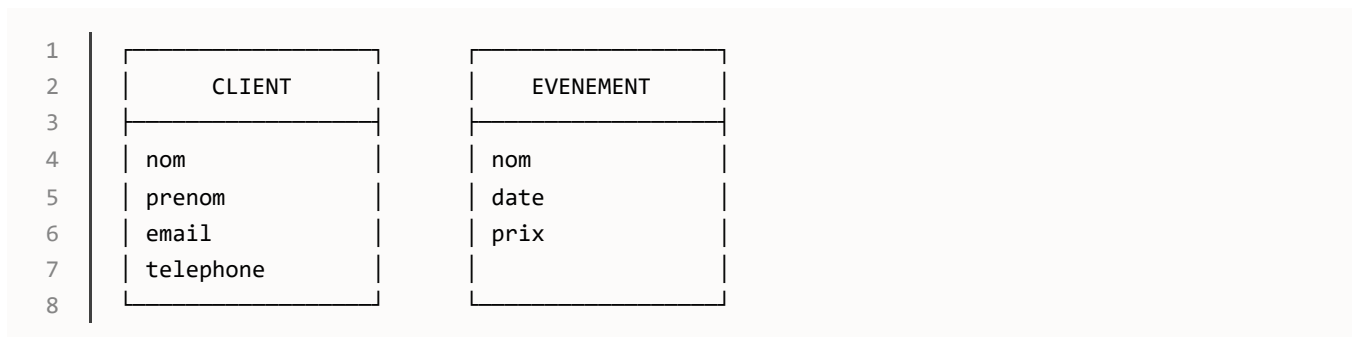
Sujet	Verbe	Complément
un client	achète	des tickets
chaque ticket	donne accès à	une soirée
une soirée	se déroule dans	un lieu

On retrouve nos quatre entités : **CLIENT**, **TICKET**, **EVENEMENT** et **LIEU**. Ce qui reste – « une date précise », « son prix » – décrit une soirée : ce sont des attributs.

Et les verbes ? Ils indiquent comment les entités sont liées entre elles : un client *achète* un ticket. Relier des entités demande un outil de plus, qu'on garde pour plus tard. Pour l'instant, retiens seulement que **les noms donnent les entités** et que **les verbes donnent les liens**.

Dessiner une entité

Pour réfléchir à une entité, on la dessine dans un rectangle. Son nom se place en haut, en majuscules et au singulier. Ses attributs se placent en dessous, en minuscules, sans accent et sans espace.



Ces rectangles sont les premières pièces du dessin que tu vas apprendre à réaliser tout au long du parcours : le **schéma entité-association**. C'est lui qu'on construit en premier, sur papier, pour analyser n'importe quel projet de base de données.

NOUVELLE NOTION : LE SCHÉMA ENTITÉ-ASSOCIATION

Le **schéma entité-association** est le dessin qui représente les données d'un domaine. Chaque **entité** est un rectangle qui contient ses attributs. Les **associations**, c'est-à-dire les liens entre les entités, se dessinent dans un losange qui porte un verbe : un client *achète* un ticket. Pour l'instant, tes schémas ne contiennent que des entités. Les associations viendront quand tu relieras plusieurs entités.

En anglais : GB **entity-relationship diagram**, souvent abrégé ERD.

Un schéma entité-association complet sert à décrire tout ce que la base devra retenir. Les informaticiens appellent ce qu'il décrit le **MCD**.

NOUVELLE NOTION : LE MCD

Le **MCD**, ou **modèle conceptuel de données**, est la description des entités d'un domaine, de leurs attributs et des associations entre elles. On le représente par un schéma entité-association, dessiné avant de toucher à un logiciel : il décrit *ce qu'on veut retenir*, pas *comment l'ordinateur va le ranger*.

En anglais : GB *conceptual data model*.

NOUVELLE NOTION : MERISE

Merise est une méthode française de conception des bases de données, créée dans les années 1970. Elle fixe les règles de dessin du schéma entité-association utilisées dans ce parcours : entités dans des rectangles, noms au singulier en majuscules, associations dans des losanges.

Merise est un nom propre : il ne se traduit pas.

Une entité à la fois

Découper le fichier des tickets a fait apparaître quatre entités liées entre elles. C'est beaucoup pour commencer.

Dans ce niveau du parcours, tu vas travailler avec **une seule entité** : les élèves d'une école. Tu vas apprendre à la transformer en table, à identifier chacune de ses lignes et à l'interroger. Relier plusieurs entités viendra ensuite, quand tu sauras parfaitement ranger une seule d'entre elles.



Exercices

 Ces exercices existent aussi **EN LIGNE**, corrigés automatiquement, dans l'[atelier SQL](#) — avec les tableaux affichés à côté des questions. Le parcours [Comprendre l'article](#) te fait d'abord réviser les notions et la chasse aux anomalies.

Exercice 1 — Entité ou attribut ? ★☆☆☆☆

Pour chaque élément, indique s'il s'agit plutôt d'une **entité** ou d'un **attribut**, puis justifie en une phrase.

Élément	Entité ou attribut ?	Pourquoi ?
livre		

Élément	Entité ou attribut ?	Pourquoi ?
couleur de la couverture d'un livre		
commande		
prix d'un article		
adresse e-mail d'un client		
client		
fournisseur		
auteur d'un livre		

💡 Pour « auteur d'un livre », il n'y a pas une seule bonne réponse. Demande-toi ce que l'application doit savoir sur les auteurs.

Exercice 2 – Découper un autre tableau ★★☆☆☆

Un garage note ses réparations dans ce tableau :

1	date	plaque	marque	nom_client	telephone_client	mecanicien	travail
2	-----	-----	-----	-----	-----	-----	-----
3	2026-09-02	1-ABC-123	Peugeot	Légrand	0470 11 22 33	Karim	vidange
4	2026-09-02	2-XYZ-987	Renault	Simon	0485 44 55 66	Julie	pneus avant
5	2026-09-05	1-ABC-123	Peugot	Légrand	0470 11 22 33	Karim	freins
6	2026-09-09	1-DEF-456	Toyota	Légrand	0470 112233	Julie	contrôle techr

1. Repère une redondance et une incohérence. Nomme-les avec le bon mot.
2. Découpe les colonnes en groupes qui parlent chacun d'une seule chose.
3. Donne un nom d'entité à chaque groupe.
4. Pour l'entité qui décrit les voitures, donne deux occurrences.

Exercice 3 – Du texte aux entités ★★☆☆☆

Lis ce besoin :

« Un centre de formation propose des formations. Chaque formation a un titre, une durée en heures et un prix. Un formateur anime des sessions. Chaque session a lieu à une date précise et concerne une formation. Des stagiaires s'inscrivent aux sessions. »

1. Applique la méthode sujet – verbe – complément et liste les entités.
2. Pour chaque entité, donne au moins trois attributs. Invente ceux qui ne sont pas dans le texte, s'il le faut.
3. Dessine deux de ces entités sous forme de rectangles.
4. Donne deux occurrences de l'entité qui décrit les formations.

Exercice 4 – Le serveur Discord ★★★★★

« Sur notre serveur Discord, chaque membre a un pseudo et une date d'arrivée. Les membres écrivent des messages dans des salons. Chaque salon a un nom et un sujet. Certains membres ont un rôle : modérateur, animateur ou invité. »

1. Liste les entités certaines.
2. « Rôle » est-il une entité ou un attribut du membre ? Donne un argument pour chaque réponse, puis dis ce qui te ferait trancher.
3. Dessine l'entité qui décrit les messages avec au moins trois attributs.

Exercice 5 – Ton propre domaine ★★★★★

Choisis un domaine que tu connais bien : ta collection de jeux vidéo, un club de sport, une bibliothèque de mangas...

1. Écris un besoin de trois ou quatre phrases, comme celui de l'association.
2. Repère au moins trois entités avec la méthode sujet – verbe – complément.
3. Dessine chaque entité avec ses attributs.
4. Échange ta feuille avec un voisin : trouve-t-il les mêmes entités que toi ?



À retenir

- La **redondance**, c'est la même information écrite à plusieurs endroits. Elle rend les **incohérences** possibles.
- Une ligne qui parle de plusieurs choses à la fois est le signe que plusieurs entités sont mélangées.
- Une **entité** est une catégorie de choses dont on garde une liste ; on l'écrit au singulier, en majuscules.
- Un **attribut** est une information sur chaque élément d'une entité.
- Une **occurrence** est un élément concret d'une entité.
- Dans un texte, les sujets et les compléments donnent les entités ; les verbes donnent les liens.
- On analyse un domaine en dessinant un **schéma entité-association** : il représente le **MCD**, selon les règles de la méthode **Merise**.

Suite

Tu sais maintenant repérer une entité. Dans l'article [Une entité devient une table](#), tu transformes l'entité **ELEVE** en une vraie table, dans une vraie base de données.

