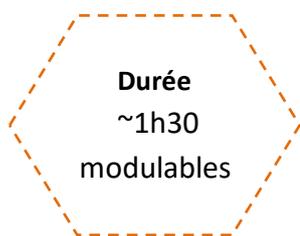


# 01. Introduction aux données structurées



Thématique  
**DONNEES  
STRUCTUREES**



## Description de l'activité

Dans cette activité, on découvre le thème Données structurées à travers une vidéo tirée du Mooc SNT.

## Objectifs pédagogiques ou compétences

Objectifs généraux	Objectifs intermédiaires	Compétences
<b>Notions de cours</b>	<ul style="list-style-type: none"><li>- Savoir ce qu'est une donnée</li><li>- Savoir qui peut récupérer des données et comment</li><li>- Comprendre le fonctionnement global du traitement automatique des données</li></ul>	<ul style="list-style-type: none"><li>- Analyser une demande et identifier les données pertinentes pour y répondre</li><li>- Comprendre une vidéo et en extraire les informations pertinentes</li></ul>

## Matériel et outils

- Fiche élève à imprimer
- 1 poste par élève pour visionner les vidéos

## Tags

#données structurées ;

# Déroulé de l'activité

## Introduction : (~30 minutes)

- **Présenter les objectifs de la séance (contenu théorique et productions attendues) (2-3 minutes)**
- **Introduire la thématique des données : (~30 minutes)**

Pour lancer la thématique des données, on propose aux élèves une première activité de classement, à faire en petits groupes de 4 ou 5 élèves. L'objectif est de mieux leur faire comprendre le besoin de classer les informations pour automatiser le travail.

- Dans un premier temps, les élèves définissent 5 critères, et remplissent un tableau sans noter le nom de la personne.
- Ensuite, chaque groupe échange son tableau avec un autre, chacun devant retrouver qui correspond aux descriptions.
- On vérifie si les propositions sont correctes (on peut à ce moment réunir les 2 groupes et les laisser discuter plutôt que de tout vérifier ensemble), puis on discute de l'efficacité des critères choisis.
- L'enseignant.e peut demander aux élèves si les critères proposés permettent de cibler efficacement une personne dans un plus grand groupe (à l'échelle du lycée, de la ville, de la planète, ...)

Ensuite, les élèves retournent avec les mêmes groupes et affinent les critères pour permettre d'identifier la personne correspondant aux mieux à chaque cas spécial. Cette partie a pour objectif de rappeler que les types de données à cibler et traiter sont nombreuses, en fonction de multiples besoins. Les élèves peuvent présenter lors de la correction la personne du groupe la plus à même d'être concernée.

Enfin, on conclut sur un petit point sur le traitement automatisé des données. Les élèves répondent avec leur groupe ou en binôme, sans effectuer de recherches, le but étant de faire le point sur leurs connaissances.

## Visionner la vidéo et répondre aux questions (~20-40 minutes)

Les élèves visionnent la vidéo et rédigent un texte contenant une liste de notions. Avant correction, on peut laisser un temps de vérification des réponses en binôme, ou alors demander aux élèves d'échanger leurs copies et de corriger ou compléter les réponses des camarades.

L'enseignant.e proposera pour finir une correction globale.

## Conclusion (15-20 minutes)

- **Bilan de la séance :** (5 minutes)

Pour clôturer la séance, on peut revenir sur les principales difficultés rencontrées pendant l'activité. Éventuellement, il est possible de finir sur un court échange autour :

- **L'omniprésence des données** (5 minutes)

On peut, dans un premier temps, demander aux élèves ce que l'omniprésence et l'augmentation de l'usage des données impliquent, notamment en insistant sur les thématiques qui les concernent.

### Pistes de discussion :

- **Tendance à la personnalisation :** L'omniprésence des données permet une personnalisation accrue des expériences utilisateur, que ce soit dans la publicité ciblée, les recommandations de produits, ou les services sur mesure.
- **Innovation et transformation numérique :** Les entreprises utilisent les données pour créer de nouveaux produits, services et modèles commerciaux.
- **Protection de la vie privée :** L'omniprésence des données soulève des préoccupations en matière de protection de la vie privée. Les individus sont de plus en plus conscients de la manière dont leurs données personnelles sont collectées, utilisées et partagées, ce qui a entraîné des débats sur la réglementation et la confidentialité des données.
- **Responsabilité et éthique :** L'omniprésence des données soulève des questions éthiques sur la collecte et l'utilisation des données, ainsi que sur la transparence et la responsabilité des organisations et des gouvernements dans leur gestion.
- **Explosion des volumes de données :** Le volume de données générées et collectées à l'échelle mondiale augmente exponentiellement. Cette explosion est en grande partie due à la numérisation croissante de nos vies et de nos activités.
- **Diversité des sources de données :** Les données proviennent de nombreuses sources, notamment des appareils connectés, des médias sociaux, des capteurs IoT, des transactions en ligne, des bases de données, etc.
- **Analyse et traitement des données :** L'omniprésence des données a conduit à une demande croissante d'outils et de compétences en analyse de données et en intelligence artificielle. Les entreprises utilisent des techniques telles que l'apprentissage automatique pour extraire des informations utiles à partir de ces vastes ensembles de données.
- **Rôle dans la recherche scientifique :** Les données sont essentielles pour la recherche scientifique dans de nombreux domaines, de la médecine à l'astronomie en passant par la biologie. Elles permettent d'analyser des phénomènes complexes et d'obtenir des informations précieuses. Ces données les aident à mieux comprendre leurs clients et ainsi, adapter leurs produits et services, prendre des décisions stratégiques, et anticiper les tendances du marché.
- **Défis de gestion des données :** La gestion efficace des données est devenue un défi majeur pour de nombreuses organisations. Cela inclut la collecte, le stockage, la sécurisation et la gestion de la qualité des données.

- **Les métiers en lien (10 minutes)**

On peut également évoquer les principaux métiers en lien avec la collecte et le traitement des données pour rendre ce domaine plus concret pour les élèves. Voici quelques exemples :

**Data Science et Analyse de Données :**

- **Data Scientist** : Analyse des données, création de modèles prédictifs et recommandations basées sur les données.
- **Analyste de données** : Extraction, nettoyage et analyse des données pour aider à la prise de décision.
- **Data Analyst** : Concentré sur l'analyse de données pour résoudre des problèmes spécifiques.

**Gestion des Données :**

- **Data Manager** : Gestion de la collecte, du stockage et de la qualité des données.
- **Data Steward** : Responsable de l'intégrité et de la gestion des données au sein d'une organisation.

**Business Intelligence :**

- **Analyste BI (Business Intelligence)** : Création de tableaux de bord, rapports et visualisations pour aider à la prise de décision.
- **Architecte BI** : Conception et gestion de l'infrastructure de données pour la BI.

**Ingénierie des Données :**

- **Ingénieur en données** : Conception, construction et maintenance des flux de données.
- **Ingénieur Big Data** : Gestion des données à grande échelle, y compris le stockage et le traitement de données massives.

**Statistiques et Modélisation :**

- **Statisticien** : Application de techniques statistiques avancées pour l'analyse des données.
- **Modélisateur de données** : Création de modèles mathématiques pour prédire des événements futurs.

**Protection des Données :**

- **Responsable de la protection des données (Data Protection Officer - DPO)** : Assure la conformité aux réglementations sur la protection des données et la sécurité des données.
- **Analyste en sécurité des données** : Sécurisation des données contre les menaces et les violations.

**Intégration et Migration des Données :**

- **Consultant en intégration de données** : Aide à l'intégration de données provenant de différentes sources.
- **Ingénieur en migration de données** : Responsable du transfert de données d'un système à un autre.

**Intelligence Artificielle et Machine Learning :**

- **Ingénieur en apprentissage automatique** : Création de modèles d'apprentissage automatique pour l'analyse prédictive.
- **Chercheur en IA** : Travaux de recherche pour développer de nouvelles techniques d'IA.

**Gestion de Projet en Données :**

- **Chef de projet en données** : Gestion de projets liés à la collecte et à l'analyse des données.

**Analyse de Marché et Marketing :**

- **Analyste de marché** : Collecte de données pour évaluer les tendances du marché.
- **Analyste marketing numérique** : Utilisation des données pour optimiser les campagnes marketing en ligne.

# Introduction aux données structurées

## Fiche activité - *Correction*

### Introduction – Créer des données et les exploiter

- **Données générales**

Avec votre groupe, vous allez choisir **5 critères permettant**, à n'importe quelle autre personne de la classe (en dehors de votre groupe) de deviner qui est la personne correspondant aux données.

**Exemples de critères :**

- Date de naissance
- Ville de résidence
- Taille
- Couleurs de cheveux
- Sexe
- Vêtements et/ou bijoux
- Centres d'intérêts

- **Données spécifiques**

Maintenant, allons un peu plus loin ... Pour chaque cas, donnez une liste de plusieurs critères permettant de cibler au mieux leurs candidats idéaux, et ce à l'échelle de votre ville. Dans votre groupe, qui serait concerné ?

- Le club de football local cherche des personnes sportives pour relancer le club amateur des seniors. Ce club ne se retrouve que les jeudis de 18h à 20h, et organise ses matchs les samedis matin.

**Pistes de critères :** âge, proximité de la ville, sports pratiqués et fréquence, disponibilité sur les horaires mentionnés.

- Le dernier samedi soir du mois prochain, un quiz musical sur le thème du rap/hip-hop des années 2000-2010 a lieu. Les questions sont basées sur la reconnaissance des titres et artistes, mais aussi sur la reconnaissance des paroles.

**Pistes de critères :** styles de musiques écoutés et leur année de création, connaissance des titres et artistes, connaissance des paroles, gestion du stress, disponibilité le jour du quiz.

- Une entreprise d'e-commerce proposant des chaussures et accessoires en cuir végétal à destination des enfants (0-10 ans) souhaite mieux cibler ses clients pour augmenter ses ventes.

**Pistes de critères :** convictions (via leurs recherches web ou leurs profils sur les RS par exemple), âge, avoir des enfants ou petits-enfants, revenus, lieu d'habitation.

- Suite à un départ, une équipe d'e-sport française (genre : MOBA) souhaite recruter leur prochain membre.

**Pistes de critères :** rang dans le jeu, années de pratique, prix ou tournois gagnés, âge, fréquence de jeu, connaissance des tactiques, gestion du stress, disponibilité, capacité à communiquer, capacité à gérer la frustration.

- **Traitement automatique des données :**

- Qu'est-ce qu'une donnée ?

Une donnée est une représentation symbolique d'informations sous une forme brute ou non traitée. Elle peut prendre diverses formes, telles que des nombres, des textes, des images, des sons, etc. Les données peuvent être collectées, stockées, et utilisées pour diverses finalités, notamment pour l'analyse, la prise de décision, la recherche, et bien d'autres domaines.

- Qui peut récupérer des données ? Comment ?

Les données peuvent être récupérées par divers acteurs, notamment :

- **Les individus :** Ils peuvent collecter des données personnelles à travers des enquêtes, des formulaires en ligne, ou en utilisant des dispositifs de suivi.
- **Les organisations :** Les entreprises, les gouvernements, et les organismes à but non lucratif collectent souvent des données pour des études de marché, des rapports statistiques, ou pour des services publics.
- **Les machines :** Les dispositifs automatisés, tels que les capteurs, les caméras, et les ordinateurs, peuvent recueillir des données en continu.

La collecte de données peut se faire par le biais de sondages, de questionnaires, de capteurs, de dispositifs de suivi, d'applications mobiles, de formulaires en ligne, et de nombreuses autres méthodes.

- Quels sont les avantages du traitement automatique des données ?

- **Rapidité :** Les machines peuvent traiter d'énormes volumes de données beaucoup plus rapidement que les humains.
- **Précision :** Les erreurs humaines sont réduites, ce qui garantit des résultats plus précis.
- **Analyse avancée :** À partir de très grandes quantités de données, les algorithmes peuvent détecter des tendances et des modèles complexes dans les données.
- **Automatisation :** Les machines / programmes peuvent effectuer la collecte et l'analyse des données (plus rapidement et en fonction des cas, plus efficacement), et ainsi nous permettre d'économiser du temps.
- **Prise de décision éclairée :** Le traitement automatique des données peut fournir des informations essentielles pour la prise de décisions stratégiques dans les domaines tels que les affaires, la médecine, la recherche scientifique, etc.

## 1. Données, comment les manipuler ?

Visionnez la vidéo et prenez des notes. En vous appuyant sur la vidéo précédente et sur vos connaissances personnelles, rédigez un texte permettant de répondre aux questions suivantes :

- Qu'est-ce que l'Informatique ?
- Qu'est-ce qu'une donnée ?
- Pourquoi avoir plusieurs données pour définir un objet ?
- Pourquoi structurer des données ?
- Comment structurer des données ?
- Que sont les métadonnées ? A quoi servent-elles ?
- Que sont les bases de données ? A quoi servent-elles ?

<https://www.youtube.com/watch?v=IJgcZ2DEs0>



### Proposition de correction :

#### Qu'est-ce que l'Informatique ?

L'informatique est le domaine qui traite de la manipulation et du traitement de l'information à l'aide d'ordinateurs et de logiciels. Cela comprend le stockage, la récupération, la transformation et la gestion des données pour effectuer diverses tâches.

#### Qu'est-ce qu'une donnée ?

Une donnée est une valeur décrivant un objet ou une information. Les données peuvent être de différents types, tels que du texte, des valeurs numériques, des dates, etc. Pour identifier un objet de manière unique et effectuer des opérations, plusieurs données de catégories différentes sont souvent nécessaires.

#### Pourquoi avoir plusieurs données pour définir un objet ?

Il faut avoir plusieurs données pour définir un objet car une seule donnée ne suffit généralement pas à identifier de manière unique cet objet. Par exemple, la durée d'une chanson en minutes et secondes est une donnée numérique,

mais elle ne suffit pas à identifier la chanson. Il est donc nécessaire de combiner plusieurs données de catégories différentes pour obtenir une identification précise.

### **Pourquoi structurer des données ?**

Structurer des données est nécessaire pour organiser les informations de manière à les rendre utilisables et accessibles. Cela implique de regrouper les données dans des tables, où chaque colonne correspond à une catégorie de données, afin de faciliter leur traitement, leur recherche et leur analyse.

### **Comment structurer des données ?**

Pour structurer des données, on crée généralement une table dans laquelle on regroupe toutes les informations relatives à un objet. Chaque colonne de la table correspond à une catégorie de données, et chaque ligne à un objet. On peut remplir manuellement la table ou utiliser des programmes automatisés, appelés "moulinettes", pour indexer les données.

### **Que sont les métadonnées ? À quoi servent-elles ?**

Les métadonnées sont des informations qui décrivent le contenu d'un fichier numérique ou d'une base de données. Elles servent à fournir des détails sur les données, tels que leur origine, leur format, leur structure, etc. Les métadonnées facilitent la recherche, l'organisation et la compréhension des données.

### **Que sont les bases de données ? À quoi servent-elles ?**

Les bases de données sont des ensembles structurés de données organisées dans des tables. Elles servent à stocker, gérer et manipuler efficacement des informations. Les bases de données permettent de rechercher, trier, analyser et récupérer des données de manière efficace, ce qui les rend essentielles pour de nombreuses applications, notamment la gestion de l'information, la recherche, et bien d'autres domaines.

## **2. Êtes-vous d'accord ?**

Vous allez donner votre fiche à un.e autre élève, qui vous fera ses retours ci-dessous. Vous allez vous-même faire vos retours à un.e autre élève.

# Introduction aux données structurées

## Fiche activité

### Introduction – Créer des données et les exploiter

- **Données générales**

Avec votre groupe, vous allez choisir **5 critères permettant**, à n'importe quelle autre personne de la classe (en dehors de votre groupe) de deviner qui est la personne correspondant aux données.

	Critère 1	Critère 2	Critère 3	Critère 4	Critère 5
Elève 1					
Elève 2					
Elève 3					
Elève 4					
Elève 5					

- **Données spécifiques**

Maintenant, allons un peu plus loin ... Pour chaque cas, donnez une liste de plusieurs critères permettant de cibler au mieux leurs candidats idéaux, et ce à l'échelle de votre ville. Dans votre groupe, qui serait concerné ?

- Le club de football local cherche des personnes sportives pour relancer le club amateur des seniors. Ce club ne se retrouve que les jeudis de 18h à 20h, et organise ses matchs les samedis matin.

.....

.....

.....

.....

.....

.....

.....

- Le dernier samedi soir du mois prochain, un quiz musical sur le thème du rap/hip-hop des années 2000-2010 a lieu. Les questions sont basées sur la reconnaissance des titres et artistes, mais aussi sur la reconnaissance des paroles.

.....

.....

.....

.....

.....

.....

.....

- Une entreprise d'e-commerce proposant des chaussures et accessoires en cuir végétal à destination des enfants (0-10 ans) souhaite mieux cibler ses clients pour augmenter ses ventes.

.....

.....

.....

.....

.....

.....

.....

- Suite à un départ, une équipe d'e-sport française (genre : MOBA) souhaite recruter leur prochain membre.

.....

.....

.....

.....

.....

.....

.....

- **Traitement automatique des données :**

- Qu'est-ce qu'une donnée ?

.....

.....

.....

.....

.....

.....

.....

.....

- Qui peut récupérer des données ? Comment ?

.....

.....

.....

.....

.....

.....

.....

.....

- Quels sont les avantages du traitement automatique des données ?

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

## 1. Données, comment les manipuler ?

Visionnez la vidéo et prenez des notes. En vous appuyant sur la vidéo précédente et sur vos connaissances personnelles, rédigez un texte permettant de répondre aux questions suivantes :

- Qu'est-ce que l'Informatique ?
- Qu'est-ce qu'une donnée ?
- Pourquoi avoir plusieurs données pour définir un objet ?
- Pourquoi structurer des données ?
- Comment structurer des données ?
- Que sont les métadonnées ? A quoi servent-elles ?
- Que sont les bases de données ? A quoi servent-elles ?

<https://www.youtube.com/watch?v=IJgcZ2DEs0>



.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....



Handwriting practice area consisting of 20 horizontal dotted lines.

