Clean data, IMPORTANT

Première ligne = en-tête

Ligne = observation

Variable le long des colonnes

* **Une info par cellule**

!!! une majuscule et une minuscule sont différent

Quand on nettoyé des données, JAMAIS MODIFIER -> TOUJOURS FAIRE UN NOUVEAU FICHIER/SHEET

1. nouveau fichier-copie de sauvegarde
2. faire un readme ET DOCUMENTER TOUT LE NETTOYAGE (metadonné)
3. Nettoyer

Nom de variable devrait commencer par des lettes, des chiffres et des \_ PAS DE CHARACTER SPECIAUX ( ,

Manque d’info ? dans excel 🡪 laisser la cellule vide, elle changera auto vers la chose dans le language (n/A dans R, etc), JAMAIS METTRE 0

Oui .txt, .xcl ,etc. cest pas mal mais ya mieux, on prefere des fichier ouvrable dans un bloque texte 🡪 exporter en .CSV (comma seperated value), .TSV (tab …)