

# TP 1 — Prise en main de Tableau Public

## Contexte

On va travailler avec **EM-DAT** (Emergency Events Database), une base qui recense les catastrophes naturelles mondiales depuis 1900. C'est maintenu par le Centre de Recherche sur l'Épidémiologie des Désastres (CRED) de l'UCLouvain.

Le dataset : 17 360 catastrophes, 46 variables (localisation, type, dates, victimes, dégâts économiques...), couverture mondiale sur plus de 120 ans.

## Objectifs

À la fin du TP, vous saurez utiliser Tableau Public pour connecter des données Excel, les nettoyer un minimum, créer des visualisations (barres, lignes, cartes) et publier tout ça en ligne.

**Durée estimée :** 1h30

---

## Partie 1 — Installation

### Tableau Public

Allez sur <https://public.tableau.com/>, téléchargez l'application (il faut donner son email), installez-la et créez un compte si ce n'est pas déjà fait.

### Le dataset

Téléchargez `catnat.xlsx` [ici](#).

---

## Partie 2 — Connexion des données

### Importer le fichier

Ouvrez Tableau Public, cliquez sur "Microsoft Excel" et sélectionnez `catnat.xlsx`. Glissez la feuille "EM-DAT Data" dans la zone centrale. Vous voyez un aperçu des données en bas.

### Le volet de gauche

Cliquez sur "Feuille 1". À gauche, vous avez vos variables divisées en deux catégories :

**Dimensions** (en bleu) : `Country`, `Region`, `Subregion`, `Disaster Type`, `Disaster Subtype`, etc.

**Mesures** (en vert) : `Total Deaths`, `No. Injured`, `Total Affected`, `Total Damage ('000 US$)`, `Magnitude` ...

C'est la distinction classique entre variables qualitatives et quantitatives.

---

## Partie 3 — Préparation des données

### 3.1 Vérifier les types de données

1. Revenez sur l'écran de connexion (onglet "Source de données" en bas)
2. Observez les icônes au-dessus de chaque colonne :
3. **Abc** = Texte
4. **#** = Nombre
5.  = Date
6.  = Géographique
7. Vérifiez que `Start Year` est bien en **Nombre** (pas en Date)
8. Cliquez sur l'icône d'une colonne pour changer son type si nécessaire

## 3.2 Les valeurs nulles

Retournez sur "Feuille 1", glissez `Disaster Type` vers **Lignes** et `Total Deaths` vers **Colonnes**. Vous verrez que certaines catastrophes n'ont pas de valeur pour les décès.

Attention : `null`  $\neq$  zéro. Ça veut juste dire qu'on n'a pas l'info. Une catastrophe avec `Total Deaths` = null, c'est pas forcément une catastrophe sans victime, c'est juste qu'on ne sait pas.

## 3.3 Gérer les valeurs nulles

**Option 1 — Filtrer les nuls :**

1. Faites glisser `Total Deaths` vers **Filtres**
2. Cliquez sur "**Valeurs spéciales**" en bas de la fenêtre
3. Sélectionnez "**Non-null**" pour exclure les valeurs manquantes
4. Cliquez OK

**Option 2 — Remplacer les nuls par zéro (champ calculé) :**

1. Menu **Analyse** → **Créer un champ calculé**
2. Nommez-le `Total Deaths (sans null)`
3. Entrez la formule : `IFNULL([Total Deaths], 0)`
4. Cliquez OK
5. Utilisez ce nouveau champ à la place de `Total Deaths`

## 3.4 Repérer les anomalies

1. Nouvelle feuille
2. Faites glisser `Start Year` vers **Colonnes**
3. Faites glisser `Total Deaths` vers **Lignes**
4. Faites glisser `DisNo.` vers **Détail** (carte des Repères)
5. Observez les points extrêmes
6. Survolez-les pour identifier les événements

**Question :** Ces valeurs extrêmes sont-elles des erreurs ou des catastrophes réellement exceptionnelles ?

---

## Partie 4 — Premières visualisations

### Exercice 1 — Catastrophes par type

Créez une nouvelle feuille. Glissez `Disaster Type` vers **Lignes** et `DisNo.` vers **Colonnes**. Faites un clic droit sur `DisNo.` dans Colonnes → Mesure → **Comptage (distinct)**. Triez par ordre décroissant.

Quel type de catastrophe revient le plus souvent ? Combien de séismes sont enregistrés ?

---

### Exercice 2 — Évolution dans le temps

Nouvelle feuille. `Start Year` vers **Colonnes**, `DisNo.` vers **Lignes** (en comptage distinct). Vous avez une courbe.

Si vous voulez colorer par type de catastrophe, glissez `Disaster Subgroup` sur **Couleur**.

Question intéressante : la courbe monte clairement au fil du temps. Est-ce qu'il y a vraiment plus de catastrophes, ou est-ce qu'on les enregistre juste mieux ?

---

### Exercice 3 — Impact par région

Nouvelle feuille, `Region` vers **Lignes**, `Total Deaths` vers **Colonnes**. Triez.

Vous pouvez aussi ajouter `Total Affected` dans **Colonnes** pour comparer décès et personnes affectées. Est-ce que c'est les mêmes régions qui ressortent ?

---

### Exercice 4 — Carte mondiale

Nouvelle feuille, double-cliquez sur `Country`. Tableau fait une carte automatiquement. Glissez `DisNo.` sur **Couleur** (comptage distinct) et changez la palette si vous voulez (Couleur → Modifier les couleurs).

Essayez de filtrer sur un seul type (ex: "Earthquake" dans **Filtres**). Ça correspond aux zones sismiques qu'on connaît ?

---

## Exercice 5 — Top 10 des pires catastrophes

Nouvelle feuille, glissez `Event Name` vers **Lignes**. Tableau va vous avertir qu'il y a trop de valeurs → "Filtrer puis ajouter".

Dans la fenêtre de filtre, onglet "**Premiers**", sélectionnez "**Par champ**" : Premiers → 10 → Total Deaths → Somme. OK.

Glissez `Total Deaths` vers **Colonnes** et `Country` vers **Couleur** pour voir d'où viennent ces catastrophes. Triez.

La numéro 1, c'est quoi ?

---

## Partie 5 — À vous de jouer

Créez une ou deux visualisations selon ce qui vous intéresse. Quelques idées :

### Dégâts économiques

1. Nouvelle feuille
2. Faites glisser `Disaster Type` vers **Lignes**
3. Faites glisser `Total Damage ('000 US$)` vers **Colonnes**
4. Triez par ordre décroissant
5. (Optionnel) Faites glisser `Total Damage ('000 US$)` sur **Couleur** pour un dégradé

### Évolution des décès par décennie

1. Nouvelle feuille
2. Menu **Analyse** → **Créer un champ calculé**
3. Nommez-le `Décennie` et entrez la formule : `INT([Start Year]/10)*10`
4. Cliquez OK
5. Dans le volet Données, clic droit sur `Décennie` → **Convertir en dimension**
6. Faites glisser `Décennie` vers **Colonnes**
7. Faites glisser `Total Deaths` vers **Lignes**

### Comparaison des sous-régions d'un continent

1. Nouvelle feuille

2. Faites glisser **Region** vers **Filtres**
3. Cochez un seul continent (ex : "Asia") → OK
4. Faites glisser **Subregion** vers **Lignes**
5. Faites glisser **DisNo.** vers **Colonnes** (comptage distinct)
6. Triez par ordre décroissant

## Relation magnitude / victimes (scatter plot)

1. Nouvelle feuille
  2. Faites glisser **Magnitude** vers **Colonnes**
  3. Faites glisser **Total Deaths** vers **Lignes**
  4. Faites glisser **DisNo.** vers **Détail** (carte des Repères)
  5. Faites glisser **Disaster Type** vers **Couleur**
  6. (Optionnel) Filtrez sur "Earthquake" pour isoler les séismes
- 

## Partie 6 — Publication

Renommez vos feuilles avec des noms clairs (double-clic sur l'onglet), puis Fichier → Enregistrer sur Tableau Public. Connectez-vous et nommez votre classeur "TP1 - Catastrophes Naturelles - [Votre Nom]".

Notez l'URL pour la rendre si besoin.

---

## Pour les curieux 😊

- Ajoutez des **titres** explicites à chaque visualisation
  - Personnalisez les **info-bulles** (ce qui s'affiche au survol)
  - Explorez les **filtres interactifs**
  - Consultez la galerie Tableau Public pour vous inspirer
- 

## Ressources

- Source des données : [EM-DAT](#)
- Glossaire EM-DAT : [Documentation](#)