

TP24 – Nettoyage de données

Techniques Avancées

Ressource : Plane_crashes.xlsx

1. Crowdsourcing

- a. Crowdcrafting
- b. Amzone Mechanic turc

2. Bases de données :

- a. Ouvrir le fichier Plane_crashes.xlsx dans google spreadsheet.
- b. Dans la feuille full_dataset, les noms des passagers et l'âge sont uniques, les autres valeurs sont toutes liées à un autre aspect. Cette feuille peut être divisée en plusieurs autres feuilles :
 - i. Une nouvelle feuille (Accidents) basée sur les accidents montre des informations liées à l'accident.
 - ii. Une feuille (Aircrafts) sur les avions montre des informations relatives à l'avion impliqué dans les accidents.

Note: Avec ce découpage, vous aurez moins d'éléments possibles à nettoyer.