Cas d’usage : Planification du raccordement électrique de bâtiments

Une ville étant touchée par des intempéries, endommageant plusieurs infrastructures essentielles pour le raccordement des foyers au réseau électrique. En tant que data scientist, le maire de ville fait appel à ma boite afin de proposer une planification pour les travaux de rétablissement de la fibre en priorisant le rétablissement de la connexion pour le plus grand nombre de citoyens de cette ville avec budget le plus faible possible.

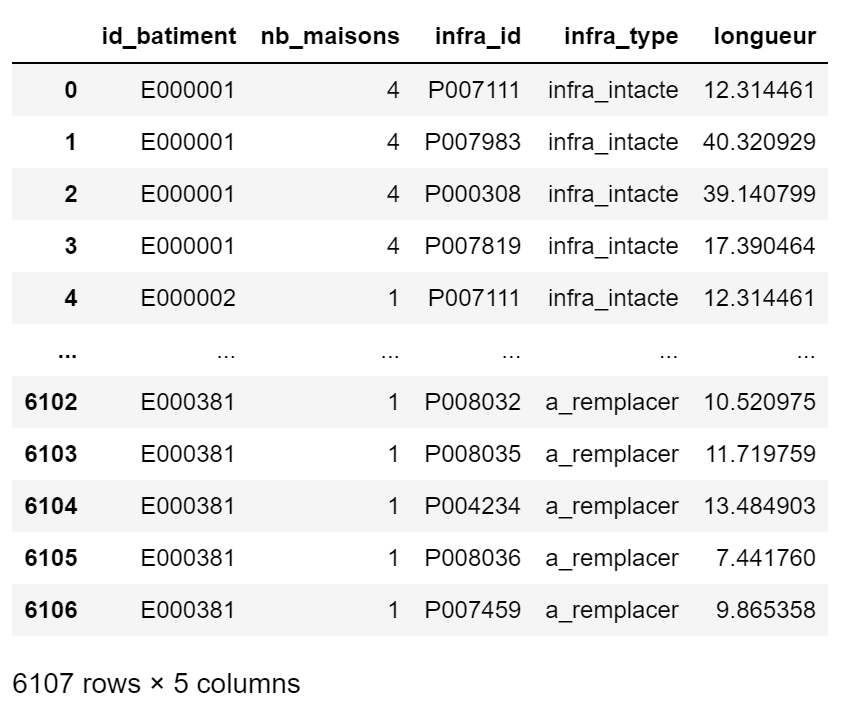
De ce fait, ma mission sera donc de créer un plan de raccordement qui favorise les bâtiments qui sont les plus simples à raccorder, ce qui permet à la fois de minimiser les coûts de l’intervention et maximiser le nombre de prises raccordées.

Dans un premier temps, on va étudier les différentes métriques possibles (difficulté infra, difficulté bâtiment, radd, add, lt, …), puis on abordera le plan de raccordement des bâtiments avec l’ordre de priorité.

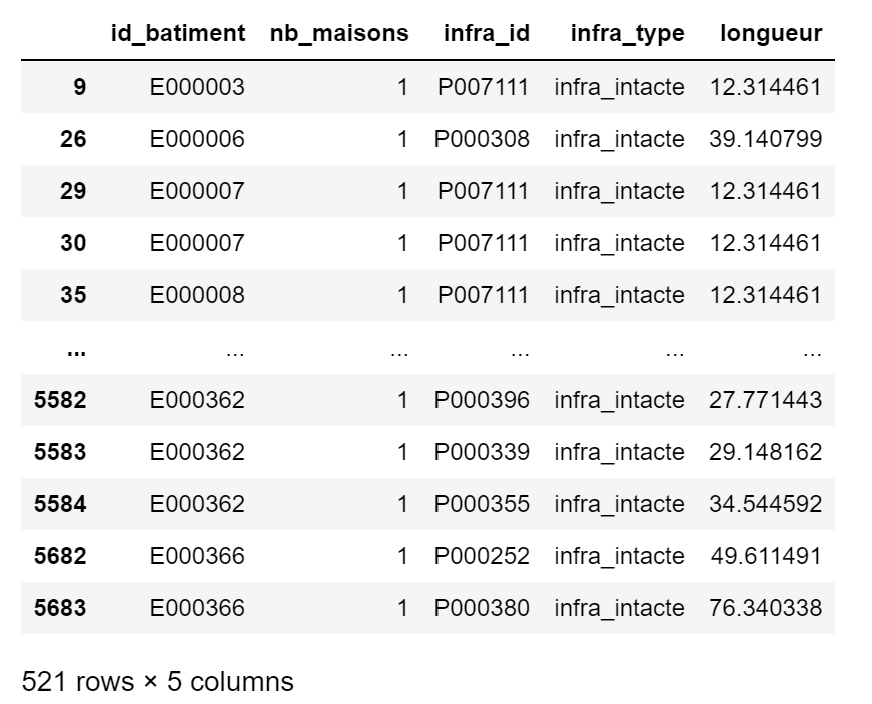
**Analyse de données**

Dans cette partie, nous allons examiner le fichier csv afin de comprendre les données et effectuer éventuellement des traitements.

Ce graphique donne la structure des données. Notre jeu de donnée contient 6107 individus (bâtiments). En effet, pour chaque bâtiment, on a son identifiant, nombre de maisons, identifiant de l’infrastructure, le type d’infrastructure (intacte ou à remplacer) ainsi que sa longueur.



Ce graphique ci-dessous présente les données dupliquées de notre jeu de donné. Par consequent, une première analyse consiste à supprimer les doublons. En effet, on voit qu’il y a 521 doublons.



Dans la suite du projet, nous avons supprimé les doublons afin de garder que les infrastructures qui ont été endommagées par les intempéries.

**Choix de la métrique**

**Difficulté(infra) = (longueur / nombre de maison qu’elle permet de raccorder)** : On cherchera à calculer la difficulté de l’infrastructure. Puisqu’on cherche à raccorder les foyers qui ont été endommagées, on prendra que les infrastructures intactes. Par consequent, on aura comme cheminement :

* Si le type d'infrastructure est "infra\_intact", la difficulté est définie à 0.
* Sinon, la difficulté est calculée en fonction de la longueur de l'infrastructure divisée par le nombre de maisons associées

**Difficulté(bâtiment) = somme (difficultés des infras qui permettent de le raccorder) :** Le calcule la difficulté d’un u bâtiment sera la somme les difficultés de toutes les infrastructures contenues dans la liste list\_infras.

**Plan de raccordement**

Le plan de raccordement établit selon la métrique permet de raccorder les bâtiments les plus simple selon le niveau de difficulté de l’infrastructure. En effet, une infrastructure qui permet le rétablissement plusieurs bâtiments est priorisée au détriment d’une infrastructure permettant de raccorder une ou deux bâtiments.

Les deux graphiques ci-dessous présentent les bâtiments à prioriser en fonction de leur difficulté (la somme des difficultés de toutes les infrastructures associées à chaque bâtiment. On peut donc affirmer que les bâtiments E000194, E000195, E000196, E000225, E000258, E000259 sont ceux les plus faciles à raccorder car leur niveau de difficulté est plus faible soit 1.17339629… Par ailleurs, E000194 et E000003 sont les plus difficiles car leur niveau de difficulté est de 54.54138948…

