1) Introduction : Thierry

2) Jeu de données : Donovan

Expliquer le pre-processing et l’utilisation du test statistiques Mann-Whitney

3) Phase Pré-traitement : JM

4) Analyse et data visualisation : Donovan

5) Modélisation des modèles ML et analyses des résultats : Sophie

6) Présentation des prédictions et estimations des pénalités : JM

7) Conclusions et limites : Sophie

**C) Visualisation de l'évolution des émissions de CO2 par année.**

**F) Corrélation hors segment électrique et hybride**

**1) Analyse des données 2010 – 2023**

**2) ML de 2017 à 2023**

**1) Garder les données communes et utiles à notre analyse etc nox ()**

**2) Certains données n’avaient pas de sens (immatriculation)**

**3) Toutes les données non rempli #N/A etc**

**1) On n’a pas gardé 2001 à 2009 car peu de donnée et on conservé 2010 à 2015 car concordance sur la gamme et le type de véhicule (carroserrie)**

Test statistiques :

Pourquoi nous avons utilisé Man W ?

Loi normale affichage graphique Qplot : HO et H1

Uniquement graphiques moyennes avec résultats

* Expliquer pourquoi nous ne gardons que les 9 premiers constructeurs pour la partie ML.
* Pourquoi nous sommes passées des marques vers les constructeurs