



## Accidents routiers en France

**Cursus concerné : Data Scientist**

**Niveau de difficulté : 07/10**

### Description du projet :

*L'objectif de ce projet est d'essayer de prédire la gravité des accidents routiers en France. Les prédictions seront basées sur les données historiques.*

*C'est un problème parfait pour traiter l'ensemble des étapes d'un projet de Data Science. Une première étape est d'étudier et appliquer des méthodes pour nettoyer le jeu de données. Une fois le jeu de données propre, une deuxième étape est d'extraire les caractéristiques qui semblent être pertinentes pour estimer la gravité d'accidents. On pourra créer un modèle prédictif, une fois l'entraînement du modèle effectué, nous allons comparer notre modèle avec les données historiques. Ensuite, à partir de ces résultats, on peut développer un scoring des zones à risque en fonction des informations météorologiques, l'emplacement géographique (coordonnées GPS, images satellite, ...) ....*

### Ressources à consulter :

- **Données :**
  - [Bases de données annuelles des accidents corporels de la circulation routière - Années de 2005 à 2021 - data.gouv.fr](https://www.data.gouv.fr/datasets/bases-de donnees-annuelles-des-accidents-corporels-de-la-circulation-routiere-annnees-de-2005-a-2021/)
  - <https://www.kaggle.com/ahmedlahlou/accidents-in-france-from-2005-to-2016>
- **Bibliographie :**
  - [Bases de données annuelles des accidents corporels de la circulation routière - Années de 2005 à 2021 - data.gouv.fr](https://www.data.gouv.fr/datasets/bases-de donnees-annuelles-des-accidents-corporels-de-la-circulation-routiere-annnees-de-2005-a-2021/)

### Conditions de validation du projet :

- un **rapport** d'exploration, de data visualisation et de pre-processing des données ;
- un **rapport** de modélisation ;
- un **rapport** final et le **GitHub** associé.