



DataScientest • com

## Accidents routiers en France

**Cursus concernés : Data Scientist**

**Difficulté : 07/10**

### Description détaillée

*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 à partir de l'historique des matchs les caractéristiques qui semblent être pertinentes pour estimer la gravité d'accidents. Ensuite, à partir de ses résultats, l'objectif est de travailler un scoring des zones à risque en fonction des informations météorologique, l'emplacement géographique (coordonnées GPS, images satellite, ...) ...*

*Une fois l'entraînement du modèle effectué, nous allons comparer notre modèle avec les données historiques.*

### Data

<https://www.data.gouv.fr/en/datasets/bases-de-donnees-annuelles-des-accidents-corporels-de-la-circulation-routiere-annees-de-2005-a-2019/>

<https://www.kaggle.com/ahmedlhlou/accidents-in-france-from-2005-to-2016>

### Bibliographie

Liste des papiers / blogpost / repo github relatif au projet.

- [Bases de données annuelles des accidents corporels de la circulation routière - Années de 2005 à 2020 - data.gouv.fr](#)

### Conditions de validation du projet

- Exploration des données (vision d'ensemble sur les données)
- Visualisations commentées et analysées
- Rapport technique + programmes + Github
- Une démo (Streamlit) pour donner le score de prédiction