# Hélène/Orphée – Codebook

\*Data Analytics 2021\*

**Executive Analysis**

This document describes the dataset used for this project, including general info ad variables (data type, units of measurement, etc.) This document is in French.

1) Information sur la nature, l’origine du dataset, la manière d’être collectée, etc

En France, chaque accident de la circulation fait l’objet d’une saisie d’informations descriptives par les forces de l’ordre (police, gendarmerie), s’il a impliqué au moins un véhicule et a fait au moins une victime ayant nécessité des soins.

Ces informations alimentent les BAAC : bases de données annuelles des accidents corporels de la circulation routière.

Ces bases de données sont mises à disposition du public dans une version simplifiée, sur le site [www.data.gouv.fr](http://www.data.gouv.fr) (plateforme internet des données publiques françaises libres d’utilisation.)

Nous étudierons la base de données répertoriant les accidents corporels de la circulation intervenus durant l’année 2019 en France métropolitaine, dans les départements et territoires d’outre-mer.

Elle comprend des informations concernant les caractéristiques de l’accident, son lieu, les véhicules impliqués et leurs victimes, et n’occulte que les données spécifiques qui porteraient atteinte à la protection de la vie privée.

Elle est composée de 4 fichiers (Caractéristiques – Lieux – Véhicules – Usagers) au format csv, que nous avons fusionnés en un seul fichier ‘victime.csv’.

2) Notre projet

Classification de la gravité des accidents pour les victimes : victime indemne, victime légèrement blessée, victime hospitalisée, victime décédée.

2) Informations sur les variables :

NB colonnes contentant l'information retenue pour prévoir le degré de gravité des accidents de la route en 2019, en fonction de NB indicateurs.

| Nom de la Variable | Description | Exemple | Unité | Influence supposée) sur la variable cible (avant étude) | Type de donnée initial  (Catégoriel / Numérique) | Type de distribution | % de valeurs manquates | Outliers | Transformations mathématiques potentiellement intéressantes |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Num\_Acc. | Numéro d'identifiant de l’accident | 201900000001 | nombre | Néant | N |  | 0 | 0 | 0 |
| id\_vehicule | Identifiant unique du véhicule repris pour chacun des usagers occupant ce véhicule (y compris les piétons qui sont rattachés aux véhicules qui les ont heurtés) – Code numérique. | 138 306 524 |  | Néant | C |  |  |  |  |
| Num\_Veh | Identifiant du véhicule repris pour chacun des usagers occupant ce véhicule (y compris les piétons qui sont rattachés aux véhicules qui les ont heurtés) – Code alphanumérique | B01 |  | Néant | C |  |  |  |  |
| Place | Place occupée dans le véhicule par l'usager au moment de l'accident (voir schéma dans le référentiel officiel) | 2 |  | Forte | N |  |  |  |  |
| catu | Catégorie de victime (conducteur, passager, piéton) | 2 |  | Forte | C |  |  |  |  |
| grav | Gravité de blessure de l'usager, les usagers accidentés sont classés en trois catégories de victimes plus les indemnes : 1 – Indemne 2 – Tué 3 – Blessé hospitalisé 4 – Blessé léger | 3 |  | CIBLE | C |  |  |  |  |
| sexe | Sexe de la victime | 1 |  | Medium | C |  |  |  |  |
| an\_nais | Année de naissance de la victime | 1979 |  | Faible | N |  |  |  |  |
| trajet | Motif du déplacement au moment de l’accident (travail, achats, loisirs…) | 2 |  | Faible | N |  |  |  |  |
| secu1 | Utilisation d’équipement de sécurité (airbag, casque, ceinture…) | 1 |  | Forte | N |  |  |  |  |
| secu2 | Utilisation d’équipement de sécurité (airbag, casque, ceinture…) | 1 |  | Forte | N |  |  |  |  |
| secu3 | Utilisation d’équipement de sécurité (airbag, casque, ceinture…) | 1 |  | Forte | N |  |  |  |  |
| locp | Localisation du piéton (passage piéton, chaussée…) | 1 |  | Forte | N |  |  |  |  |
| actp | Action du piéton (traversant, courant…) | 1 |  | Medium | C |  |  |  |  |
| etatp | Précise si le piéton accidenté était seul, accompagné ou en groupe | 3 |  | Faible | C |  |  |  |  |
| senc | Sens de circulation (croissant ou décroissant) | 2 |  | Forte | C |  |  |  |  |
| catv | Catégorie du véhicule (voiture, vélo, tracteur…) | 2 |  | Forte | N |  |  |  |  |
| obs | Obstacle fixe heurté (poteau, arbre…) | 3 |  | Forte | N |  |  |  |  |
| obsm | Obstacle mobile heurté (piéton, animal…) | 4 |  | Medium | N |  |  |  |  |
| choc | Point de choc initial du véhicule (avant, côté…) | 3 |  | Forte | N |  |  |  |  |
| manv | Manoeuvre principale avant l’accident (dépassement, marche arrière…) | 5 |  | Medium | N |  |  |  |  |
| motor | Type de motorisation du véhicule (hydrocarbures, hybride…) | 1 |  | Faible | N |  |  |  |  |
| occutc | Nombre d’occupants dans le transport en commun | 21 |  | Néant | N |  |  |  |  |
| jour | Jour de l’accident | 05 |  | Néant | N |  |  |  |  |
| mois | Mois de l’accident | 12 |  | Medium | N |  |  |  |  |
| an | Année de l’accident | 2019 |  | Néant | N |  |  |  |  |
| hrmn | Heure et minutes de l'accident | 17:21 |  | Medium | C |  |  |  |  |
| lum | Conditions d’éclairage dans lesquelles l'accident s'est produit (plein jour, aube, nuit…) | 1 |  | Forte | C |  |  |  |  |
| dep | Code INSEE du département | 59 |  | Faible | C |  |  |  |  |
| com | Code INSEE de la commune | 59220 |  | Faible | C |  |  |  |  |
| agg | Localisation en agglomération ou hors agglomération | 1 |  | Medium | C |  |  |  |  |
| int | Intersection | 2 |  | Medium | N |  |  |  |  |
| atm | Conditions atmosphériques (vent, brouillard, pluie…) | 3 |  | Forte | N |  |  |  |  |
| col | Type de collision (frontale, latérale…) | 5 |  | Forte | N |  |  |  |  |
| adr | Adresse postale si accident survenu en agglomération | Autoroute A86, |  | Faible | C |  |  |  |  |
| lat | Latitude | 48,8962100 |  | Faible | C |  |  |  |  |
| Long | Longitude | 2,4701200 |  | Faible | C |  |  |  |  |
| catr | Catégorie de route (autoroute, route nationale…) | 2 |  | Medium | N |  |  |  |  |
| voie | Numéro de la route | 86 |  | Néant | C |  |  |  |  |
| V1 | Indice numérique du numéro de route (exemple : 2 bis, 3 ter etc.) | 0.0 |  | Néant | C |  |  |  |  |
| V2 | Lettre indice alphanumérique de la route | A |  | Néant | C |  |  |  |  |
| circ | Régime de circulation (à sens unique, bidirectionnelle, à chaussées séparées…) | 3 |  | Fort | C |  |  |  |  |
| nbv | Nombre total de voies de circulation | 10 |  | Medium | N |  |  |  |  |
| vosp | Existence d’une voie réservée aux cycles, indépendamment du fait que l’accident ait lieu ou non sur cette voie | 0 |  | Faible | C |  |  |  |  |
| prof | Déclivité de la route à l'endroit de l'accident | 2 |  | Medium | C |  |  |  |  |
| pr | Numéro du point de repère de rattachement (numéro de la borne amont) | 303 |  | Néant | C |  |  |  |  |
| pr1 | Distance en mètres à la borne amont). | 282 |  | Néant | C |  |  |  |  |
| lartpc | Largeur du terre-plein central s’il existe (en m) | 50 |  | Néant | N |  |  |  |  |
| larrout | Largeur de la chaussée | 125 |  | Medium | N |  |  |  |  |
| surf | Etat de la surface (flaque, neige…) | 1 |  | Forte | N |  |  |  |  |
| infra | Infrastructure éventuelle (pont, voie ferrée…) | 1 |  | Forte | N |  |  |  |  |
| situ | Situation de l’accident (sur chaussée, sur trottoir…) | 2 |  | Forte | N |  |  |  |  |
| vma | Vitesse maximale autorisée sur le lieu et au moment de l’accident | 120 |  | Forte | N |  |  |  |  |