



PROJET GESTION DE BASES DE DONNEES

MASTER 2 AUTOMATIQUE, ROBOTIQUE PARCOURS
INGÉNIERIE DES SYSTÈMES INTELLIGENTS
2023-2024

Lynda FEDDAK
Alhussein JAMIL

Professeur :
Mr. Lamine BOUGUEROUA

7 janvier 2024

Table des matières

1	Introduction	4
2	Description de la base de données	5
2.1	Les Données Clés des Tables Principales dans la Base de Données	5
2.2	Les données ajoutées à la base de donnée	6
3	Modélisation	8
4	Résultats	9
4.1	Visualisation des graphiques avec Matplotlib	9
4.2	Application de visualisation avec Tkinter	9
5	Conclusion générale	10

Table des figures

1	Le Modèle Logique des Données (MLD)	8
2	Gravité des accidents en fonction des conditions atmosphériques, des conditions d'éclairage et du type de route	9
3	Types de collision les plus fréquents	10

1 – Introduction

Ce projet représente la mise en œuvre d'une solution pour exploiter des données volumineuses, en particulier celles recueillies à partir du site data.gouv.fr. La démarche inclut la collecte et le nettoyage rigoureux des données avant de les stocker dans une base de données (BDD) adéquate. La modélisation du stockage, la préparation de l'environnement, et l'enregistrement des données traitées dans la BDD sont des étapes cruciales de cette initiative.

On utilise dans ce projet une base de données exhaustive sur les accidents routiers, accessible via le lien [Bases de données annuelles des accidents corporels de la circulation routière](#). Cette base de données, couvrant les années 2005 à 2022, fournit des informations détaillées sur les accidents corporels de la circulation routière en France. En analysant ces données, le projet vise à extraire des insights précieux pour améliorer la sécurité routière, comprendre les tendances des accidents, et développer des approches préventives pour réduire les risques sur les routes. Une approche basée sur les données de ce portail gouvernemental offre une opportunité unique d'apporter des solutions significatives en matière de sécurité routière.

En outre, l'application de visualisation et d'exploitation des données, développée notamment avec Python, offre une interface pour analyser les tendances, identifier des motifs et formuler des recommandations en matière de sécurité routière. Cette approche holistique promet des insights significatifs pour améliorer la sécurité sur nos routes.

Le projet Python, la base de données et une présentation détaillée sont accessibles sur le repository GitHub disponible via le lien suivant : [Projet Gestion de Bases De Données](#).¹

1. Lien vers le repo : <https://github.com/alhussein-jamil/ProjetGestionDeBasesDeDonnees>

2 – Description de la base de données

Le projet s'appuie sur une base de données exhaustive sur les accidents corporels de la circulation routière en France, on s'est concentré sur l'année 2022. Ces données, provenant du fichier national des accidents corporels de la circulation (Fichier BAAC) géré par l'Observatoire national interministériel de la sécurité routière (ONISR), sont recueillies par les forces de l'ordre sur le lieu des accidents. La base extrait des fiches d'analyse des accidents corporels, regroupant des informations détaillées sur la localisation, les caractéristiques, les véhicules impliqués et les victimes. Les bases de données annuelles, constituées à partir du fichier BAAC, sont désormais annuelles, composées de quatre fichiers au format CSV : Caractéristiques, Lieux, Véhicules et Usagers.

La base de données Etalab sur les accidents corporels de la circulation d'une année spécifique est organisée en quatre rubriques, chacune représentée par un fichier au format CSV :

- 1 - Caractéristiques** : Cette rubrique décrit les circonstances générales entourant l'accident.
- 2 - Lieux** : Elle détaille le lieu principal de l'accident, même en cas d'intersection complexe.
- 3 - Véhicules** : Cette rubrique fournit des informations sur les véhicules concernés dans l'accident.
- 4 - Usagers** : Elle donne des détails sur les personnes impliquées dans l'accident.

Chaque variable à l'intérieur d'une table est interconnectée avec les variables des autres tables. L'identifiant unique de l'accident **Num_Acc** présent dans ces quatre tables établit un lien entre toutes les variables décrivant un accident. En cas d'accident impliquant plusieurs véhicules, il est également possible de relier chaque véhicule à ses usagers grâce à la variable **id_vehicule**. Cette structure permet une analyse holistique des données en reliant de manière cohérente toutes les composantes essentielles d'un accident routier.

2.1 – Les Données Clés des Tables Principales dans la Base de Données

CARACTERISTIQUES	LIEUX
Num_Acc (Numéro d'identifiant de l'accident)	Num_Acc
jour mois	catr (Catégorie de route)
an (Année)	voie (Numéro de la route)
hrmn (Heure et minutes)	V1 (Numéro de route)
lum (Conditions d'éclairage)	circ (Régime de circulation)
dep (Département)	nbv (Nombre de voies de circulation)
com (Commune)	vosp (Voie réservée)
agg (Localisation)	prof (Profil en long)
int (Intersection)	plan (Tracé en plan)
atm (Condi atmosphériques)	lartpc (Largeur du terre-plein central)
col (Type de collision)	larout (Largeur de la chaussée)
adr (Adresse postale)	surf (État de la surface)
lat (Latitude)	infra (Aménagement - Infrastructure)
Long (Longitude)	situ (Situation de l'accident)

TABLE 1 – Colonnes des tables CARACTERISTIQUES et LIEUX

VÉHICULES	USAGERS
Num_Acc	Num_Acc
id_vehicule	id_usager
Num_Veh	id_vehicule
senc (Sens de circulation)	num_Veh
catv (Catégorie du véhicule)	place (Place occupée dans le véhicule)
obs (Obstacle fixe heurté)	catu (Catégorie d'utilisateur)
obsm (Obstacle mobile heurté)	grav (Gravité de blessure)
choc (Point de choc initial)	sexe (Sexe de l'utilisateur)
manv (Manœuvre principale avant l'accident)	An_nais (Année de naissance de l'utilisateur)
motor (Type de motorisation du véhicule)	trajet (Motif du déplacement)
occutc (Nombre d'occupants dans le transport)	secu1
secu1	secu2
secu2	secu3
secu3	locp (Localisation du piéton)
	actp (Action du piéton)
	etatp (Etat du piéton)

TABLE 2 – Colonnes des tables VÉHICULES et USAGERS

2.2 – Les données ajoutées à la base de donnée

Nous avons réalisé un travail complémentaire sur cette base de données, impliquant l'intégration de tables spécifiques pour clarifier le sens de certaines colonnes des tables principales (Caractéristiques, Lieux, Véhicules et Usagers). Ces colonnes contiennent des identifiants, et chaque identifiant possède une signification particulière. Par exemple, dans la table Usagers, la colonne "sexe" contient les valeurs 1 et 2, où 1 représente "Masculin" et 2 représente "Féminin".

Voici l'ensemble des tables ajoutées, chacune comportant deux colonnes : la première pour l'identifiant (clé primaire de la table associée, agissant également comme clé étrangère dans la table principale, telle que la table Usagers dans notre exemple), et la deuxième pour la signification de cet identifiant :

La table **USAGERS** est connectée aux tables suivantes :

- **Catégorie d'utilisateur (catu)** : Catégorie d'utilisateur (conducteur, passager, piéton).
- **Gravité de blessure (grav)** : Gravité de la blessure subie par l'utilisateur.
- **Sexe (sexe)** : Sexe de l'utilisateur.
- **Motif du déplacement (trajet)** : Motif du déplacement au moment de l'accident.
- **Équipement de sécurité (secu1, secu2, secu3)** : Présence et utilisation de l'équipement de sécurité.
- **Localisation du piéton (locp)** : Localisation du piéton dans l'accident.
- **Action du piéton (actp)** : Action du piéton au moment de l'accident.
- **État du piéton (etatp)** : Précision sur la présence ou non d'autres personnes avec le piéton accidenté.

La table **CARACTERISTIQUES** est connectée aux tables suivantes :

- **Lumière (lum)** : Conditions d'éclairage lors de l'accident.
- **Localisation (agg)** : Localisation de l'accident.
- **Intersection (int)** : Type d'intersection où l'accident s'est produit.
- **Conditions atmosphériques (atm)** : Conditions météorologiques au moment de l'accident.
- **Type de collision (col)** : Nature de la collision.

La table **LIEUX** est connectée aux tables suivantes :

- **Catégorie de route (catr)** : Classification de la route.

- **Régime de circulation (circ)** : Régime de circulation sur la route.
- **Voie réservée (vosp)** : Existence d'une voie réservée, indépendamment de l'emplacement de l'accident.
- **Profil en long (prof)** : Déclivité de la route à l'endroit de l'accident.
- **Tracé en plan (plan)** : Tracé de la route.
- **État de la surface (surf)** : État de la surface de la route.
- **Aménagement - Infrastructure (infra)** : Aménagement ou infrastructure à proximité de l'accident.
- **Situation de l'accident (situ)** : Situation générale de l'accident.

La table **VÉHICULES** est connectée aux tables suivantes :

- **Sens de circulation (senc)** : Direction du véhicule.
- **Catégorie du véhicule (catv)** : Catégorie du véhicule impliqué.
- **Obstacle fixe heurté (obs)** : Présence d'un obstacle fixe heurté.
- **Obstacle mobile heurté (obsm)** : Présence d'un obstacle mobile heurté.
- **Point de choc initial (choc)** : Point de contact initial lors de la collision.
- **Manœuvre principale (manv)** : Manœuvre principale effectuée avant l'accident.
- **Type de motorisation (motor)** : Type de motorisation du véhicule.

3 – Modélisation

Voici le Modèle Logique des données décrivant les relations entre les tables et les contraintes d'intégrité :

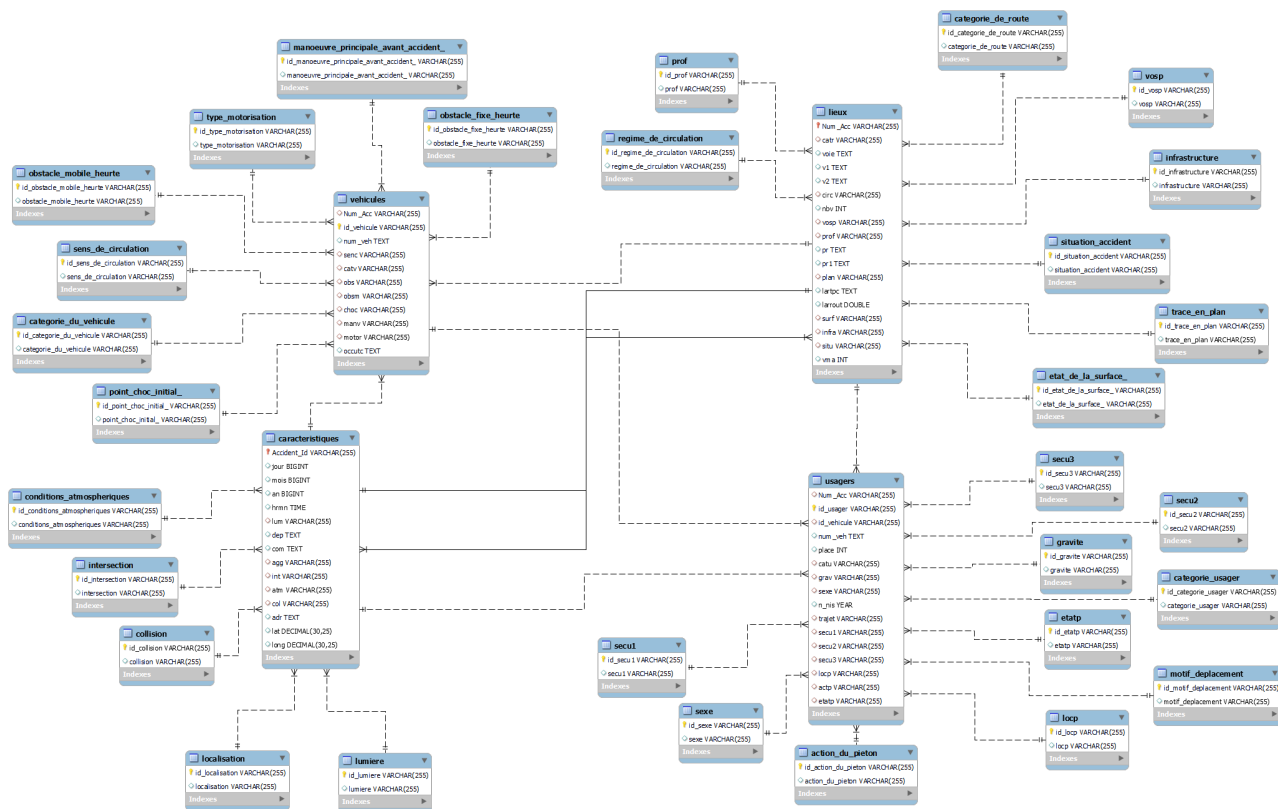


FIGURE 1 – Le Modèle Logique des Données (MLD)

4 – Résultats

4.1 – Visualisation des graphiques avec Matplotlib

Severity of Accidents across lighting conditions, weather conditions and road types

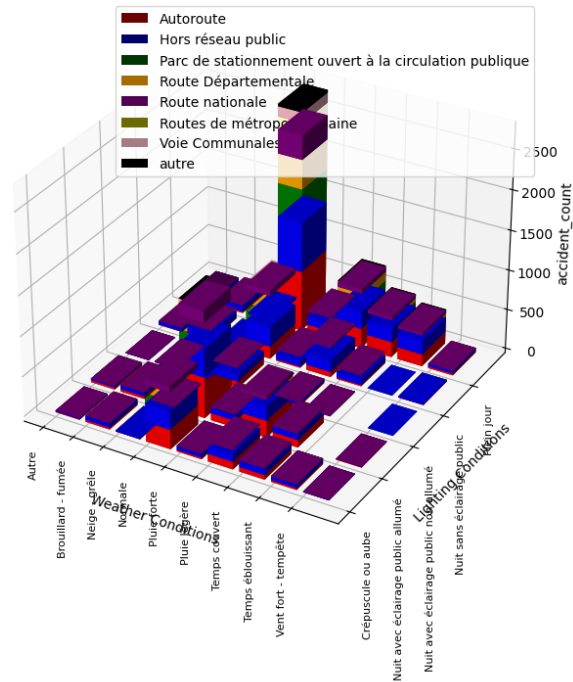


FIGURE 2 – Gravité des accidents en fonction des conditions atmosphériques, des conditions d'éclairage et du type de route

4.2 – Application de visualisation avec Tkinter

Voici un exemple de visualisation avec l'interface qui permet de sélectionner les paramètres (zone en rouge) et d'afficher simultanément la requête associée sur le graphique.

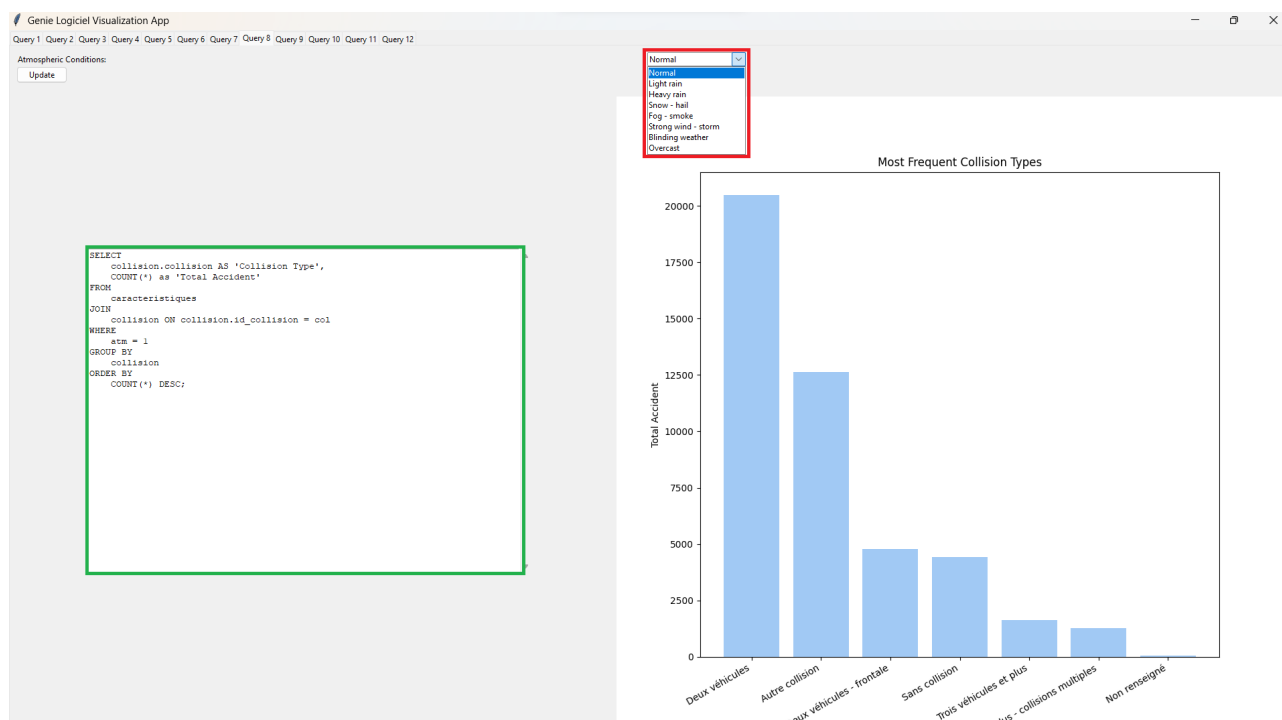


FIGURE 3 – Types de collision les plus fréquents

5 – Conclusion générale

La conduite de ce projet a impliqué une démarche méthodique, débutant par une exploration minutieuse de la base de données et des liens entre les différentes données. Nous avons appliqué des pratiques de nettoyage rigoureuses pour assurer la qualité des informations traitées.

Notre compréhension approfondie des questions potentielles que notre application pourrait résoudre a guidé le déroulement du projet. La décomposition en grandes parties, la création de tâches parallèles et la répartition des responsabilités en fonction des expertises individuelles ont contribué à une gestion efficace du projet.

L'expérimentation avec divers outils tels que tkinter, matplotlib, etc., a été cruciale pour orienter notre modèle vers la faisabilité et l'efficacité. L'intégration des enseignements issus du prototypage a enrichi le processus, conduisant à la création d'outils personnalisables et universels et à une simplification de la procédure d'assemblage.

La construction progressive de l'application à partir de sous-modules a été suivie d'une phase minutieuse de débogage pour résoudre tout conflit potentiel. Nous avons accordé une attention particulière à l'aspect visuel de l'application, travaillant sur la présentation pour offrir une expérience utilisateur à la fois fluide et intuitive. Ce projet a non seulement renforcé nos compétences techniques, mais il a également jeté les bases d'une solution pratique répondant aux besoins identifiés.