

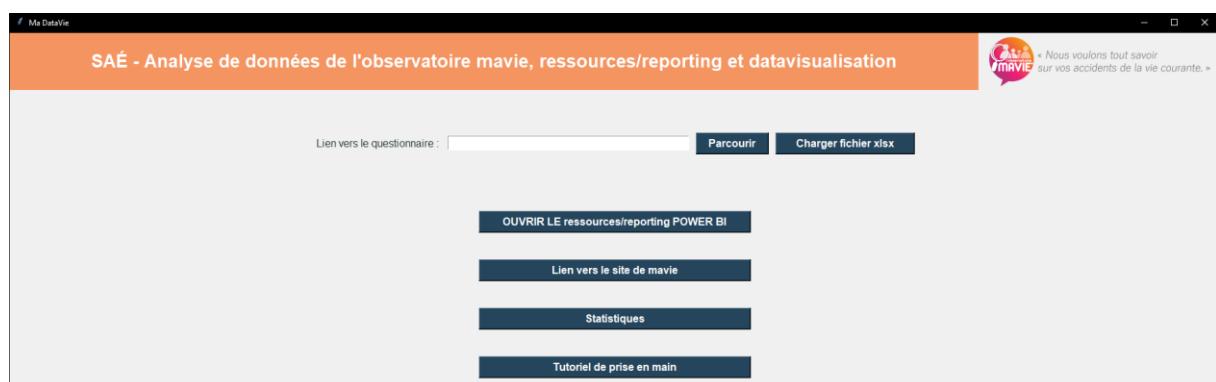


## Compte rendu de projet

BUT science des données : MaVie / Analyse de données  
sur les ACVC

Dans ce du projet, l'objectif de ma partie était de concevoir une interface graphique visant à fusionner des fichiers afin de mettre à jour une base de données, rendant ainsi les données plus accessibles pour l'analyse. Le jeu de données initial était constitué d'un fichier Excel comprenant deux feuilles distinctes : l'une contenant les informations des utilisateurs et l'autre les réponses à un questionnaire sur les accidents de la vie courante et leurs caractéristiques.

Pour faciliter l'expérience utilisateur lors de la visualisation des données, j'ai développé une interface graphique en utilisant R, avec une esthétique harmonisée aux couleurs du site internet de notre commanditaire, MaVie. Cette interface offre plusieurs fonctionnalités, notamment des boutons permettant d'accéder à notre analyse de données PowerBI, au site internet de MaVie, à un tutoriel d'utilisation de l'interface, ainsi qu'à des statistiques de base sur le jeu de données.



La zone de texte en haut de l'interface permet à l'utilisateur de charger le contenu d'un fichier en entrant le lien vers le fichier des réponses au questionnaire, qui est ensuite chargé dans l'environnement R.

J'ai mis en place des mécanismes de gestion des erreurs pour garantir que seuls les fichiers au format XLS ou XLSX puissent être chargés, avec une notification d'erreur personnalisée si le format du fichier Excel ne correspond pas à celui des réponses de MaVie. De plus, j'ai développé une fonction qui est appelée uniquement en l'absence d'erreurs. Cette fonction prend en paramètre le lien vers la base de données au format CSV, les nouvelles données extraites du fichier Excel chargé, ainsi que les colonnes constituant les clés primaires. Elle filtre les doublons dans la base de données, en veillant à ce qu'un seul utilisateur puisse avoir un seul ID\_volontaire et qu'un même utilisateur ne puisse avoir qu'un accident à la même date et heure. Une fois les doublons supprimés, la base de données est mise à jour.

Pour l'onglet "Statistiques", j'ai mis en place une interface permettant d'effectuer des tests statistiques de base pour évaluer la relation entre la variable "accident" et d'autres variables des utilisateurs, telles que le sexe, l'âge et la situation familiale.

SAE - Analyse de données de l'observatoire mavie, reporting et datavisualisation

Veuillez choisir la variable à croiser avec la variable 'Accident'

Variable: Sexe  Exécuter l'analyse

Table de contingence:

	Femme	Homme
Non	230	335
Oui	22	11

Résultats:

	Test	Valeur
1	Chi2	7.5851e7237
2	P-value	Chi2 0.005885050
3	V de Cramer	0.112624270
4	P-value Fisher	0.005701651
5	Rapport des Chances	0.343903581

Informations supplémentaires sur la population

Age moyen : 65.8  
Age médian : 69  
Age moyen des accidentés : 59.53

Le rapport des chances compare les chances d'un événement entre deux groupes. Un rapport supérieur à 1 indique des chances plus élevées dans le premier groupe, tandis qu'un rapport inférieur à 1 indique des chances plus faibles.

Informations supplémentaires

Nombre d'inscrits dans la BDD : 707  
Nombre de réponses aux questionnaires : 602  
Nombre total d'accidents : 33

Une fonction récupère les données de la base de données à chaque ouverture de cet onglet pour garantir leur actualité, puis effectue les tests statistiques, dont les résultats sont ensuite affichés. L'onglet fournit également des informations sur le jeu de données, telles que l'âge moyen et médian des utilisateurs, ainsi que le nombre total d'inscrits et d'accidents.

Ce projet m'a permis de renforcer mes compétences en R, notamment en matière de développement d'interfaces utilisateur avec Tkinter, ainsi qu'en traitement de données.