



Calyxis

PÔLE D'EXPERTISE DU RISQUE

Compte rendu personnel

Théo GARDERE

Introduction

Dans le cadre de la **SAÉ Analyse de données, Reporting et Datavisualisation**. Ce projet a été réalisé en partenariat avec Calyxis, un pôle d'expertise national spécialisé dans l'analyse, la prévention et la gestion des risques. Le travail s'appuie sur les données issues de l'étude MAVIE, un programme de recherche mené par l'Inserm (Institut national de la santé et de la recherche médicale), qui vise à mieux comprendre les accidents de la vie courante (AcVC) en France.

Objectifs du projet

L'objectif de cette SAÉ était de concevoir un outil interactif, automatisé et ergonomique capable de répondre aux besoins du commanditaire. Plus précisément, cet outil devait permettre :

- de comparer les caractéristiques des volontaires de l'étude MAVIE à celles de la population française
- d'analyser les accidents survenus (lieu, gravité, circonstances...)
- d'identifier les facteurs de risque ou de protection
- et de proposer des recommandations de prévention pertinentes, s'appuyant sur des visualisations claires et des indicateurs statistiques compréhensibles.

Travaux attendus

Pour atteindre ces objectifs, plusieurs livrables étaient demandés :

- un tableau de bord interactif intégrant des analyses statistiques et des visualisations dynamiques
- un système permettant l'automatisation du traitement des données, même en cas d'import de nouveaux fichiers
- un rapport comparatif entre les volontaires MAVIE et la population française
- une organisation de groupe selon la méthode SCRUM, avec des user stories, des sprints et un tableau Kanban assurant le suivi du projet.

Développement

Dans le cadre de cette SAÉ, j'ai principalement été en charge de deux aspects techniques du projet :

1. Le nettoyage des données via un script Python

2. La création d'un tableau de bord interactif en Tkinter, permettant la visualisation graphique des résultats.

Nettoyage des données

J'ai développé un script Python dont l'objectif était d'améliorer la qualité des données utilisées. Pour cela, j'ai mis en place des traitements automatisés visant à :

- Supprimer les doublons ;
- Détecter et corriger les erreurs (valeurs aberrantes ou manquantes)
- Uniformiser les formats pour faciliter l'analyse.

Ce travail m'a permis de mieux comprendre que la qualité des données conditionne la validité des analyses statistiques, et que des données non préparées peuvent conduire à des interprétations erronées.

Conception du tableau de bord (Tkinter)

J'ai également travaillé sur la conception d'une interface graphique avec Tkinter, incluant un tableau de bord interactif. Mon objectif était de proposer une visualisation claire et cohérente des résultats (graphiques à barres, camemberts, heatmaps, etc.).

Cette étape m'a permis de réfléchir à la façon dont on peut rendre des données complexes compréhensibles par un utilisateur non expert.

Conclusion

Ce projet m'a permis de mobiliser plusieurs compétences techniques et organisationnelles. D'un point de vue pratique, j'ai réalisé un nettoyage rigoureux des données à l'aide de Python, en identifiant et en supprimant les doublons et erreurs. J'ai également conçu une interface graphique (tableau de bord) avec Tkinter, permettant une visualisation claire et pertinente des résultats statistiques. Ces réalisations m'ont permis de mieux comprendre les enjeux de la qualité des données et l'importance de leur présentation.

Au-delà des aspects techniques, cette SAÉ m'a surtout permis de travailler efficacement en groupe. J'ai appris à m'organiser dans le cadre d'un projet long et structurant, à échanger avec mes coéquipiers, à répartir les tâches selon les compétences de chacun, et à maintenir une cohérence dans notre travail collectif.