

CALYXIS MAVIE Observatory : Analyzing Accidents of Everyday Life

A dynamic tool for analyzing and visualizing accidents of everyday life
(AcEL) in France

BABIN Paul / DIOUF Mariama / GASSE Armand-Valentin / BENOIT Nathaël

Calyxis Marion Dupuy, Mme Belim, M. Bureau et M. Garnier



Contents

- **Presentation of the context and requirements** 2
- **Organisation and planning of the work** 5
- **Proposed approach and demonstration of tasks already completed** 6



Objective of this project

Prevention strategies by identifying:

- The most at-risk profiles
- The circumstances under which accidents occur
- The factors that increase or reduce the likelihood and severity of accidents



Project Overview



Public Health Issue

AcEL include falls, burns, poisonings at home or during leisure



MAVIE Dataset

Volunteer profiles, accident details, household information



Target Users

Health professionals, researchers, prevention organisations

Project Organisation



Team Structure

Four developers with specialised roles



Agile Methodology

Scrum with daily meetings



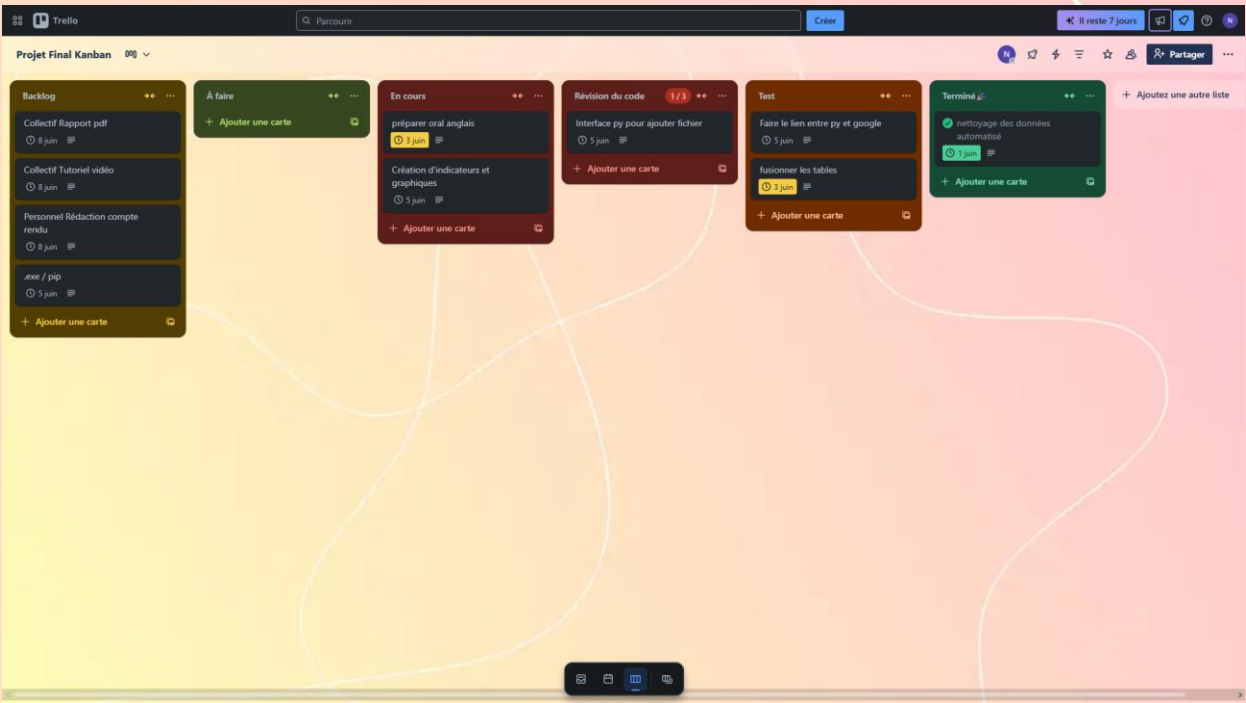
Project Management

Trello for task tracking, deadlines and userstory



Internal Deadlines

Specific targets for efficient progress





Technical Approach



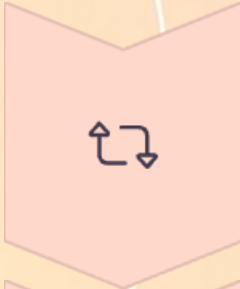
Data Cleaning

Python for simplicity and file management



Integration Challenge

Problem with power BI automatisation



Online Storage with Cloud

File updated on BigQuery (Google Cloud)



Google Looker Studio

Dashboard linked with BigQuery for visualisation

Conclusion

Difficulties

- Graphics platform
- List / dictionary manipulation in data cleaning

Completed

- File deposit interface
- Automatic data cleaning
- Individual factor graphs

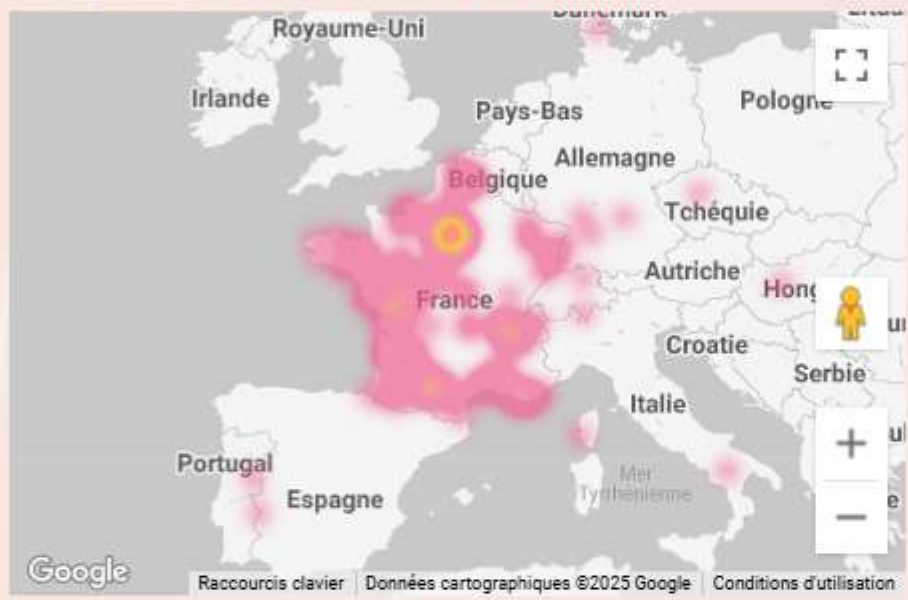
In Progress

- Tutorial creation
- Project debriefings (individual and in group)
- Factor comparison

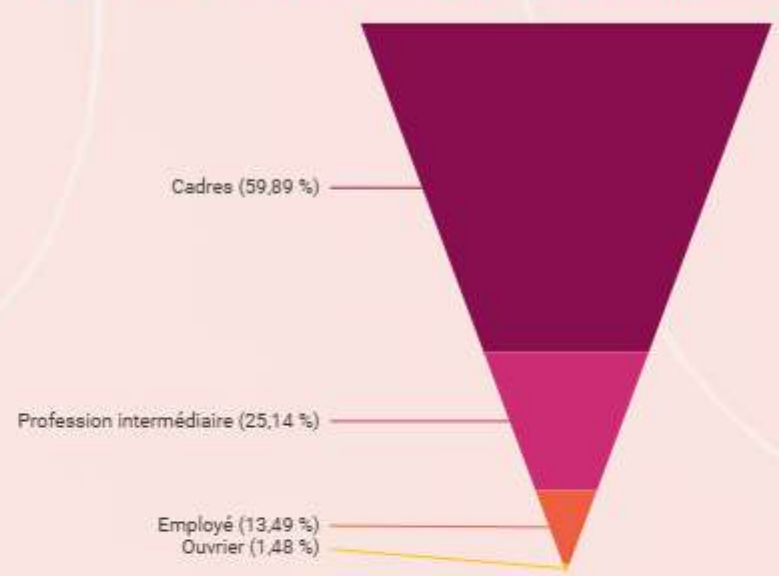


Portrait Sociodémographique des Répondants : Profils, Emplois et Répartition Mondiale

La dimension géographique



La situation professionnelle : des cadres en majorité



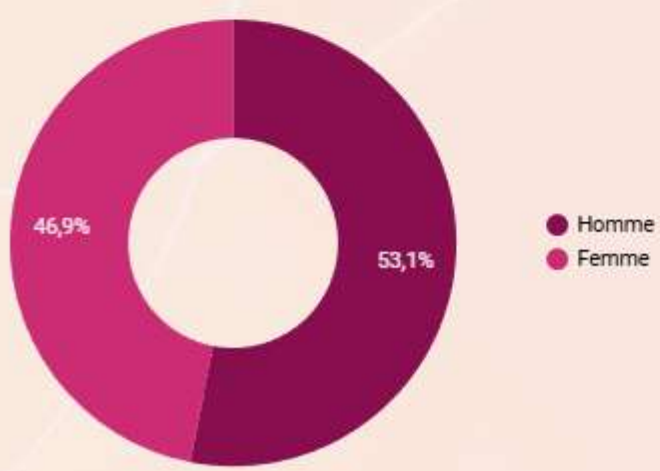
622
répondants

57
ans en moyenne

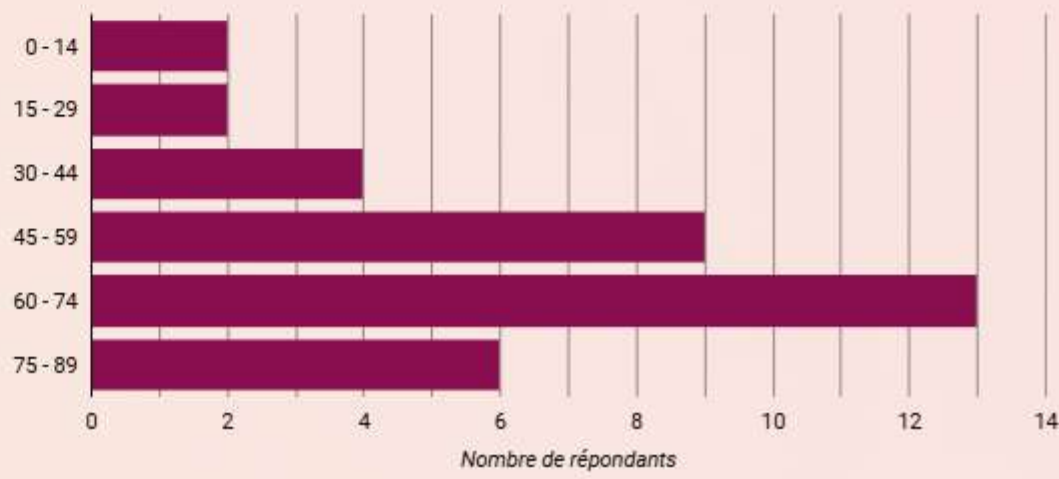
83,76 %
de sportifs

31,54 %
de fumeurs

Quelle répartition du sexe ?



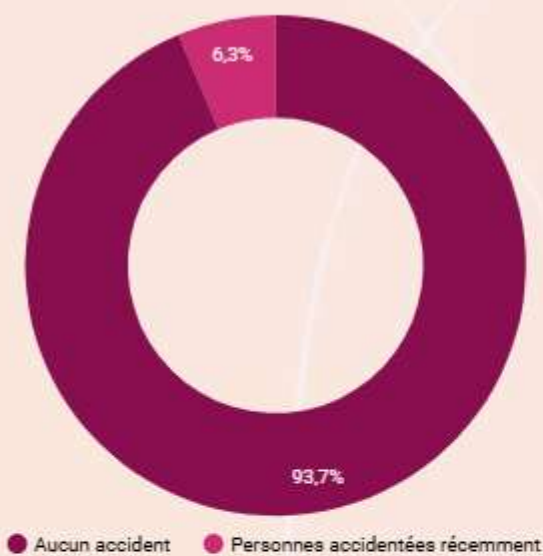
Répartition en classe d'âge des répondants



Accidents de la Vie Courante : un profil mis en évidence



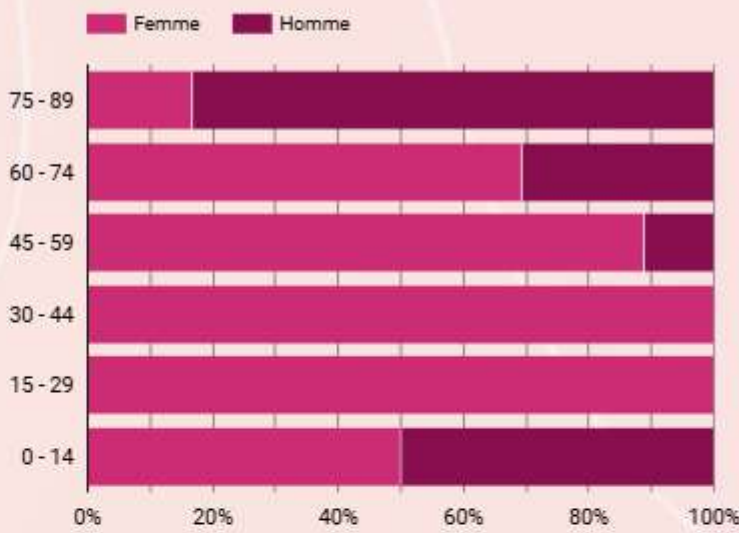
Part des accidentés parmi les répondants



Répartition des accidentés par sexe



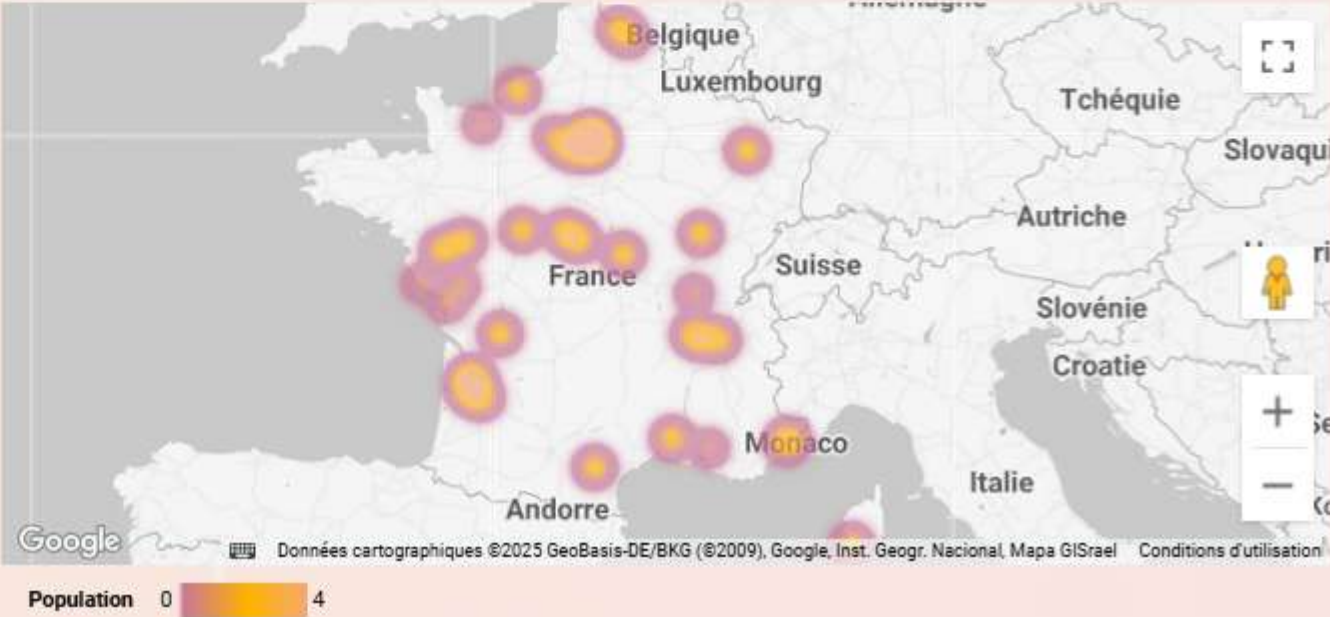
Répartition des accidentés par classe d'âge



Répartition des accidentés selon leur emploi

Situation de l'emploi	Répartition
Retraité(e) ou pré-retraité(e)	100%
Occupe un emploi	10%
En invalidité/longue maladie	5%
Au foyer	1%
Chômeur(se) non indemnisé(e)	1%
Retraite	1%
maladie	1%
Retraité en cumul emploi retraite	1%
Chômeur(se) indemnisé(e)	1%
Allocataire du RSA	1%

Communes les plus accidentées





Lieu de l'accident en fonction du genre

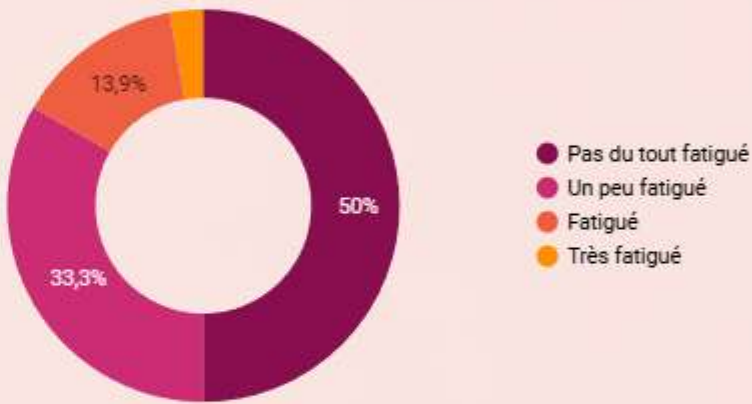
Lieu de l'accident	Homme	Femme	Répartition
Habitat, dans mon propre foyer	34,78 %	65,22 %	63,89 %
Habitat, chez une autre personne	33,33 %	66,67 %	8,33 %
Pleine nature	33,33 %	66,67 %	8,33 %
sur un trottoir	0 %	100 %	2,78 %

1 - 10 / 10 < >

A quel moment se déroulent les accidents ?



Etat de fatigue 48h avant l'accident



Répartition du type de logement des personnes accidentées



Source d'énergie : un impact sur les accidents ?

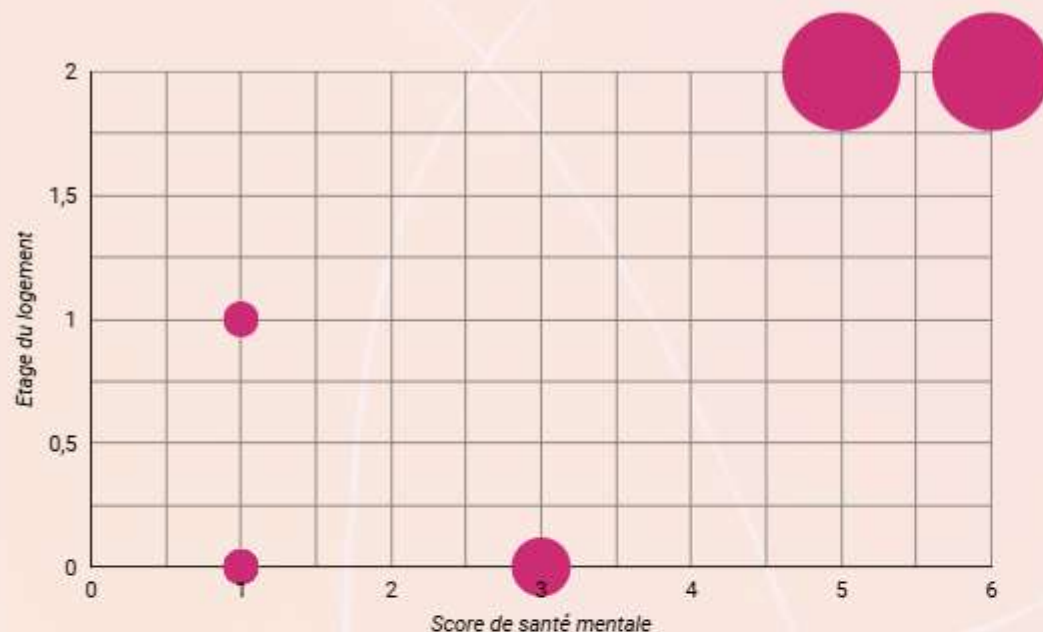
Sources d'énergie	Répartition
Gaz de ville ou de réseau	100%
Gaz de ville ou de réseau, Electricité	83%
Electricité	33%
Electricité, Bois	33%
Pompe à chaleur, géothermie	33%
Fioul (mazout) domestique, Electricité	33%
Gaz de ville ou de réseau, Climatisation réver...	17%
Gaz de ville ou de réseau, Electricité, Bois	17%

1 - 15 / 15 < >

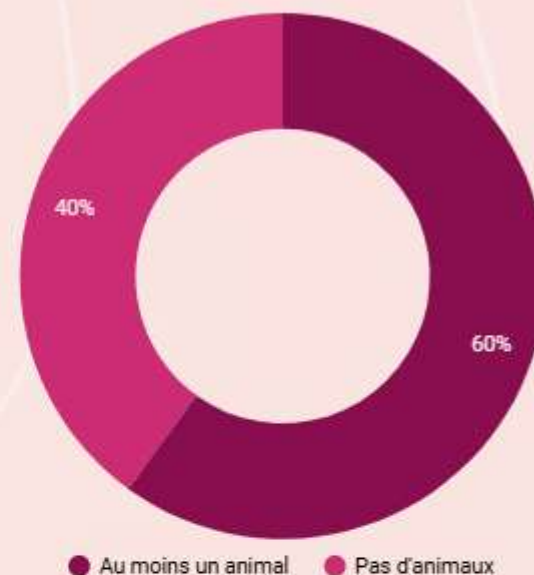
Accidents domestiques : influence de l'espace, le mode de vie et même les animaux



Accidents selon l'étage et le bien-être mental



Les animaux ont-ils un impact sur le nombre d'accidents ?



Répartition des répondants selon leur situation familiale

Nombre personnes d...	Mode de vie	Répartition
4	Seul(e) avec enfant	2,86 %
3	En couple avec enfant(s)	10 %
2	Seul(e) avec enfant	5,71 %
2	En couple avec enfant(s)	2,86 %
1	En couple sans enfant	45,71 %
1	Seul(e) sans enfant	5,71 %
1	Seul(e) avec enfant	4,29 %
null	Seul(e) sans enfant	22,86 %

1 - 8 / 8 < >

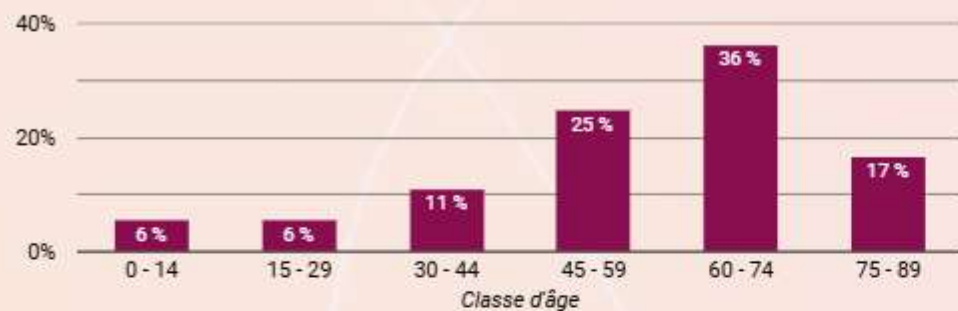
Les petites surfaces moins touchées ?



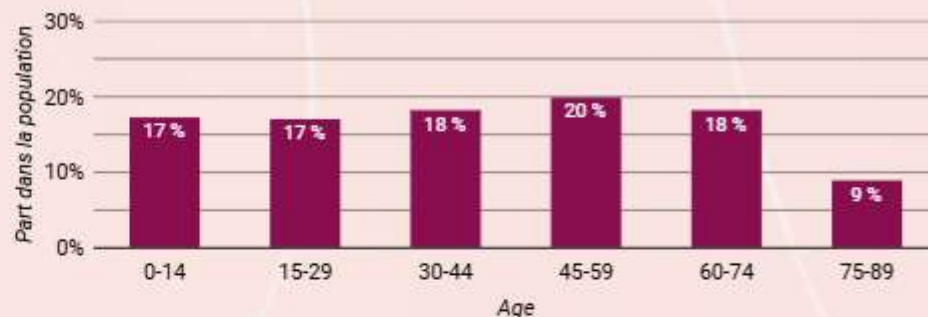
Enquêtes sur les accidents : l'échantillon vraiment représentatifs des Français ?



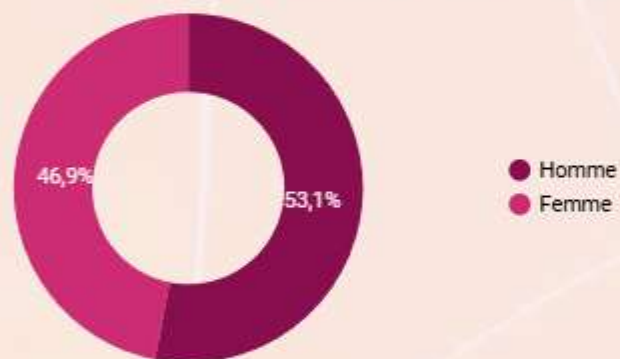
La répartition des répondants selon leur âge



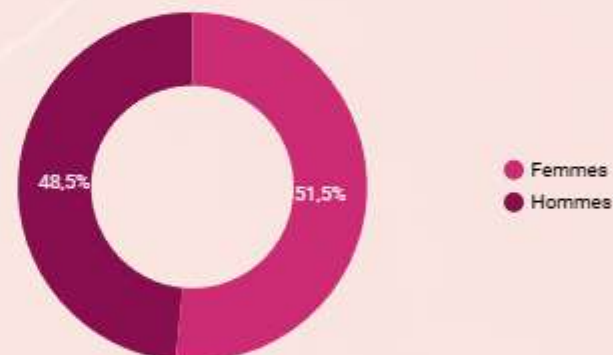
Une homogénéité marquée pour la population française



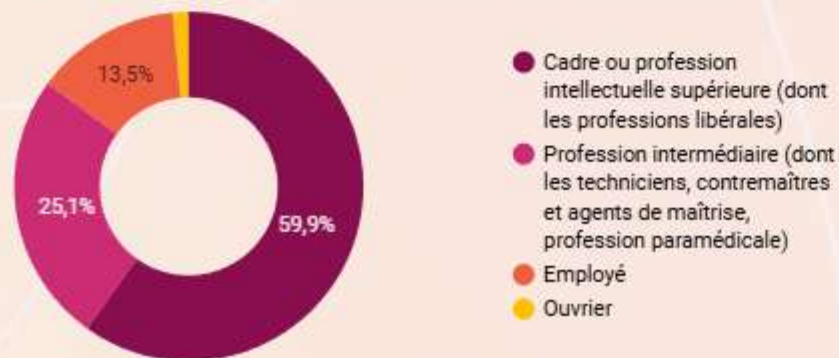
La répartition du sexe dans l'échantillon



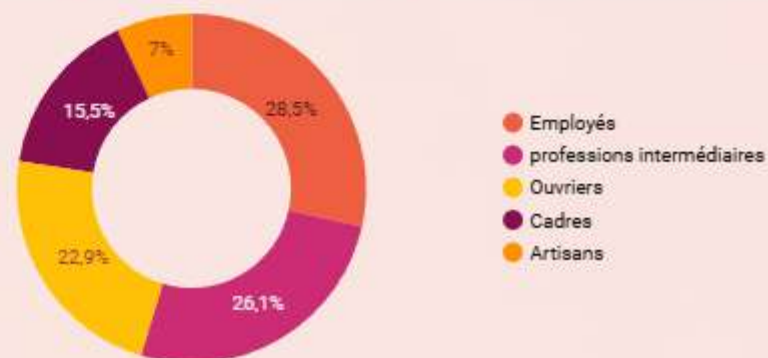
Une majorité de femmes dans la population



La répartition des professions dans l'échantillon



Une majorité d'employés et de professions intermédiaires



Pour commencer, pour réaliser nos analyses, nous nous sommes focalisés sur les facteurs individuels des répondants. Ainsi, nous pouvons remarquer que la moyenne d'âge des répondants se situe autour des 57 ans, et qu'il y a davantage d'hommes interrogés que de femmes, avec 53% d'hommes. Aussi, nous pouvons remarquer qu'une grande partie des répondants habitent dans l'Île-de-France et plus de la moitié des interrogés sont des cadres ou exercent une profession intellectuelle supérieure. De plus, nous pouvons observer que la majorité des répondants sont des retraités, puisqu'ils représentent deux tiers de la population interrogée. Effectivement, ce sont surtout des individus âgés entre 60 et 75 ans qui ont répondu au questionnaire. Comme nous pouvons aussi constater à travers ces graphiques, ce sont les femmes qui ont subi le plus d'accident de la vie courante (acVC), soit 69%.

En effet, ces dernières ont subi plus d'acVC que les hommes, aussi bien dans leur propre habitat, qu'en pleine nature, et que dans l'habitat d'une autre personne, où elles représentent deux tiers des répondants dans ces trois types de lieu. Plus globalement, seulement 36 des 622 répondants ont subi un accident de la vie courante récemment, soit uniquement 6%. Pour autant, on peut remarquer une variation importante du nombre d'accidents de la vie courante en fonction du moment de la journée. En effet, 39 % des personnes interrogées ont eu un accident de la vie courante au moment de l'après-midi.

Étonnamment, nous avons pu constater que la moitié des répondants ayant eu des accidents de la vie courante, n'étaient pas fatigués dans les 48 heures précédant l'accident. Et il y a une hausse importante du nombre d'acVC, lorsqu'une personne habite un logement avec deux étages, bien qu'elles aient une santé mentale satisfaisante. Par la suite, nous avons pu observer que les animaux pouvaient diminuer le nombre d'accidents. Effectivement, la proportion de personnes ayant eu un accident et ayant un ou des animaux est bien moins importante que celle qui n'ont pas d'animaux domestiques (40%).

Ainsi, les facteurs augmentant le plus le nombre d'accidents dans la vie courante sont le moment de la journée, le nombre d'étages de l'habitat, ainsi que le manque d'animaux. Ensuite, bien que les femmes subissent davantage d'accidents de la vie courante que les hommes, le faible nombre de répondant ne nous permet pas d'affirmer que le sexe a un lien avec le nombre d'accidents. En effet, grâce au test du chi-deux d'indépendance nous avons pu constater qu'il n'y avait presque aucun lien entre le sexe et le nombre d'accidents de la vie courante. Nous pouvons aussi supposer que la fatigue n'a pas ou peu d'impact sur ce nombre d'accident, car le test du chi-deux est lui aussi peu significatif. En revanche, à l'aide de ce test, nous avons trouvé un lien significatif entre le nombre d'accidents et si l'individu est fumeur ou non, car le chi-deux calculé est supérieur au chi-deux théorique.

Cependant, nous pouvons observer que les caractéristiques des répondants sont assez différentes comparées à l'ensemble de la population française. En effet, nous pouvons constater qu'il y a une forte hétérogénéité des tranches d'âge des répondants, tandis que les classes des Français sont quant à elles homogènes, ces classes étant chacune représentée par environ 17 % des Français. De plus, on peut remarquer qu'il y a davantage d'hommes que de femmes au sein des personnes interrogées (seulement 46,9% de femmes), contrairement à l'ensemble de la France, où la population est davantage féminisée (51,5% de femmes). Enfin, en ce qui concerne la catégorie socioprofessionnelle des répondants, vous constatez que plus de la moitié de ces

derniers sont des cadres ou exerce une profession intellectuelle supérieure (57%). En revanche, la population française est davantage constituée d'employés, précisément 28,5%. Ainsi, nous pouvons admettre que les personnes ciblées pour ce sondage ne sont pas représentatives de la population française.

Pour conclure, nous conseillons aux populations d'avoir un ou plusieurs animaux domestiques, nous leur proposons d'avoir un habitat sans étage. Par ailleurs, nous leur conseillons aussi de ne pas fumer ou bien d'arrêter de fumer, dans le but de réduire ces accidents de la vie courante. Néanmoins, ces conseils sont à considérer avec précautions, car la population interrogée est loin d'être représentative de la population française.

Compte rendu personnel BENOIT Nathaël– SAÉ Analyse de données, reporting et datavisualisation

Apprentissage critique validé : Comprendre les structures algorithmiques de base et leur contexte d'usage.

Au sein du groupe, je me suis particulièrement investi dans la conception et le développement de l'application principale en Python, dont le rôle était de permettre le traitement automatisé, le nettoyage et l'analyse de fichiers Excel contenant des données pour l'observatoire MaVie. Ce travail a été structuré autour de plusieurs objectifs : fiabiliser l'import des données, automatiser les corrections nécessaires, et faciliter la visualisation via un tableau de bord dynamique.

L'apprentissage critique que j'ai choisi de valider est pleinement justifié par la nature du travail que j'ai réalisé. Le projet s'appuie sur une organisation modulaire du code, avec des fonctions spécifiques dédiées à des tâches précises : extraction de données, nettoyage de texte, conversion de types, création de colonnes booléennes, gestion des erreurs, etc. J'ai notamment mis en œuvre des structures de contrôle (conditions, boucles) pour traiter automatiquement les valeurs erronées, gérer les fichiers vides ou invalides, ou encore pour adapter le comportement de l'interface graphique selon les actions de l'utilisateur.

Un exemple concret est la fonction `traitement_donnees()` qui centralise la logique du programme. Elle applique plusieurs traitements algorithmiques sur des tableaux de données : suppression conditionnelle de colonnes, fusion de DataFrames, transformation de chaînes de caractères en formats exploitables (dates, booléens), ou encore génération de nouvelles colonnes selon des critères complexes.

De plus, la construction de l'interface graphique via Tkinter repose sur la compréhension des événements, des états logiques et de la gestion du multithreading pour ne pas bloquer l'interface lors des traitements lourds.

En parallèle de ma contribution technique, j'ai également endossé le rôle de Scrum Master au sein du groupe. J'ai animé des réunions quotidiennes (daily meetings) pour suivre l'avancement de chacun, anticiper les difficultés, et ajuster la répartition des tâches. J'ai mis en place un tableau Kanban partagé, permettant de visualiser en temps réel l'état d'avancement du projet (« Backlog », « À faire », « En cours », « Révision du code », « Test », « Terminé »). Cela a favorisé la coordination de l'équipe, la responsabilisation individuelle et un suivi agile du projet.

En résumé, ce projet m'a permis d'approfondir la compréhension des structures algorithmiques de base dans un contexte concret et de mettre en pratique des outils d'organisation inspirés de la méthodologie agile. J'ai ainsi développé à la fois mes compétences techniques et mes capacités à piloter efficacement un travail collaboratif, ce qui valide pleinement cet apprentissage critique.

Compte rendu personnel BABIN Paul – SAÉ Analyse de données, reporting et datavisualisation

Apprentissage critique validé : connaître la syntaxe des langages et savoir l'utiliser.

Avec Nathaël, nous nous sommes davantage concentrés sur le développement de l'application et du stockage des données donc j'ai décidé de choisir un apprentissage en rapport avec la programmation

L'apprentissage critique que j'ai donc choisi de valider est : "connaître la syntaxe des langages et savoir l'utiliser". Ce projet m'a confronté à la nécessité de maîtriser les bases du langage Python, mais aussi à savoir mobiliser efficacement cette syntaxe pour répondre à des besoins concrets : manipulation de fichiers (utilisation de dataframes), gestion d'erreurs, organisation du code en fonctions, ou encore interaction avec des modules comme tkinter, pandas, google-authenticator ou subprocess. Ces derniers sont utiles mais compliqués à comprendre et utiliser, surtout s'ils nous sont inconnus. Je me suis rendu compte que nous pouvions tout faire avec Python à l'aide de bibliothèques externes. Il suffit donc de connaître la meilleure bibliothèque, que nous devons donc apprendre à utiliser, pour réaliser nos objectifs.

Un exemple marquant est la création du script principal, dans lequel j'ai dû structurer notre code de manière claire et fonctionnelle : définir des fonctions dédiées à chaque tâche (import de fichier, nettoyage, ouverture d'un lien web), gérer l'état des boutons dans l'interface, mais aussi assurer le bon enchaînement des actions déclenchées par l'utilisateur. La syntaxe que le développement de l'interface requiert est primordiale et complexe. En effet, tout doit être ordonné et réfléchi pour que l'interface donne ce qu'on veut. J'ai appris à adopter une logique rigoureuse dans l'écriture de mon code pour éviter les erreurs de syntaxe ou de logique et pour simplifier la relecture en cas d'éventuels futurs bugs.

Nous avons essayé de générer un .exe sans succès et nous nous sommes rabattus sur un .bat. Cependant, c'était quelque chose d'inconnu pour nous et le fait que la moindre petite erreur de ma part faisait que le programme ne fonctionnait pas m'a fait réaliser l'importance de respecter la bonne syntaxe.

En résumé, ce projet m'a permis non seulement d'acquérir une meilleure maîtrise de la syntaxe Python, mais surtout d'apprendre à utiliser cette syntaxe avec pertinence dans un cadre applicatif. J'ai gagné en précision et en rigueur dans ma manière d'écrire du code, ce qui valide pleinement l'apprentissage visé.

Compte rendu personnel DIOUF Mariama– SAÉ Analyse de données, reporting et datavisualisation

Apprentissage critique validé : Comprendre les intérêts de la data visualisation et de l'infographie.

Durant cette SAÉ, j'ai participé activement à la réalisation d'un tableau de bord interactif destiné à l'Observatoire MaVie, en me concentrant particulièrement sur la création des graphiques et visuels permettant de valoriser les données collectées. L'objectif était de représenter de manière claire, lisible et engageante les profils des répondants, leurs caractéristiques socio-professionnelles et les circonstances des accidents de la vie courante.

Pour cela, j'ai produit quatre pages principales de visualisation que vous pouvez voir sur le tableau de bord.

Ce travail m'a permis de comprendre concrètement l'intérêt de la data visualisation dans un projet de traitement de données :

Synthétiser des informations complexes en un coup d'œil (ex. : taux d'accidents selon l'heure de la journée). Faciliter la comparaison entre catégories (ex. : genre et lieu d'accident). Mettre en évidence des tendances invisibles dans les données brutes (ex. : majorité de cadres parmi les répondants). Améliorer la communication des résultats auprès d'un public non spécialiste.

J'ai également travaillé sur l'ergonomie du tableau de bord (couleurs cohérentes, choix des types de graphiques adaptés à chaque type de variable, structuration par thème), afin que l'ensemble soit à la fois esthétique et fonctionnel.

En conclusion, cette SAÉ m'a permis de mettre en application les principes fondamentaux de la data visualisation. J'ai appris à choisir les bons graphiques selon les types de données, à rendre les résultats compréhensibles et exploitables, et à mesurer l'impact visuel d'une infographie bien construite. Ces compétences valident pleinement l'apprentissage critique « Comprendre les intérêts de la data visualisation et de l'infographie ».

Compte rendu personnel GASSE Armand-Valentin– SAÉ Analyse de données, reporting et datavisualisation

Apprentissage critique validé : Comprendre qu'une analyse correcte ne peut émaner que de données propres et préparées

Grâce à ce projet, j'ai pu me rendre compte que pour réaliser des graphiques pertinents, il est nécessaire et primordial d'avoir des données nettoyées, car dans le cas contraire, il est difficile de faire de bonnes analyses.

En effet, il y avait par exemple dans notre jeu de données, une colonne où des informations sur les personnes habitants le foyer des répondants étaient rassemblées dans une liste. Ainsi, pour traiter et modéliser efficacement ces données, il a fallu que l'on s'occupe de créer plusieurs colonnes distinctes avec Nathaël, Mariama et Paul, pour mettre une caractéristique d'un membre du foyer par colonne, et ainsi pouvoir réaliser un graphique de qualité.

Par ailleurs, le nettoyage des données m'a permis de comprendre les données à traiter, et aussi de pouvoir les interpréter plus facilement et plus clairement que si elles n'étaient pas préparées.

Pour finir, le nettoyage des données est non seulement nécessaire pour les destinataires qui reçoivent nos graphiques et nos conclusions, mais il est aussi facilitateur pour le data analyst qui est en charge de la manipulation des données, avec notamment l'ajout essentiel de nouvelles colonnes dans cet exemple, mais il y a aussi la suppression de variables inutiles car ces dernières ne sont pas utiles à exploiter ou car elles contiennent trop de valeurs vides. Ainsi, si la préparation des données est correctement effectuée, le data analyst, qui était une des principales fonctions que j'occupais durant ce projet, pourra visualiser clairement les données importantes à sa disposition, lui permettant de réaliser des graphiques que l'on peut qualifier « d'optimiser », mettant donc en évidence les informations les plus pertinentes.