



Projet de Probabilités et Statistiques : Etude de la santé

Cédric Travelletti et Desmons Florian

Rappel

L'objectif des projets est de vous familiariser avec les outils mathématiques que vous avez ou allez étudier en cours de Probabilités et Statistiques en utilisant le langage de programmation **Julia**. Vous serez amenés à explorer un jeu de données réel, à en extraire des informations pertinentes, et à formuler des conclusions appuyées par des méthodes statistiques rigoureuses.

L'évaluation du projet se déroulera en trois étapes distinctes. Vous devrez tout d'abord soumettre un cahier des charges avant le 24 octobre 2025, dans lequel vous présenterez votre sujet, les données choisies, vos objectifs, votre répartition du travail et les premières pistes d'analyse. Ensuite, une présentation orale individuelle aura lieu le 28 novembre 2025 afin de faire le point sur l'avancement du projet, les difficultés rencontrées et les ajustements envisagés. Enfin, une présentation finale en groupe est prévue le 9 janvier 2026, au cours de laquelle vous exposerez l'ensemble du travail réalisé, les résultats obtenus et leur interprétation.

Avant de commencer l'analyse, il est essentiel de bien **comprendre** les enjeux de votre projet. Pour cela, nous vous encourageons à consulter différentes ressources documentaires telles que des livres, des articles scientifiques, des sites spécialisés ou des vidéos. Cette phase de documentation constitue une **étape fondamentale** du projet, car elle vous permettra de poser des bases solides pour votre étude et de mieux orienter vos choix en fonction de vos préférences. Les projets proposés offrent une grande diversité d'approches et de thématiques, et il est vivement recommandé de **personnaliser** votre travail afin de le rendre **plus stimulant et pertinent**.

Explication des données

Les données fournis par l'observatoire Suisse de la santé (<https://ind.obsan.admin.ch/fr>) sont principalement organisées en trois fichiers CSV distincts et ils décrivent quelques indicateurs de santé. Plus en détail, l'espérance de vie, le taux de mortalité et les causes de décès. Ils couvrent chaque année de 1998 à 2023 et sont calculées pour 100000 habitants. On va décrire plus dans le détail chaque fichier:

- Espérance de vie (esp_vie.csv): ce jeu de données fournit l'espérance de vie moyenne à la naissance (var1), accompagnée de certaines variables statistiques, répartis par canton et par sexe.
- Taux de mortalité : ce jeu de données comprend à la fois le taux de mortalité standardisé et le taux de mortalité brut, correspondant respectivement aux variables var1 et var2, également répartis par canton et par sexe. Il est possible de fournir un jeu de données détaillant le taux de mortalité par tranche d'âge.
- Causes de décès : ce jeu de données recense les causes de décès pour l'ensemble de la Suisse, sans subdivision par canton mais répartis par sexe. Var1 contient la proportion de mort pour une cause en %. Un jeu de données supplémentaire peut toutefois être fourni pour répartir les décès par tranche d'âge.

Ces données permettent d'analyser l'évolution de l'espérance de vie et des taux de mortalité au fil du temps, de comparer les différences entre sexes et cantons, et d'étudier l'impact des différentes causes de décès sur la population et de réaliser des analyses comparatives du taux de mortalité et de l'espérance de vie entre cantons, sexes ou périodes.

Enfin, il est possible d'ajouter un ensemble de données permettant d'observer les particules fines présentes dans différentes villes suisses afin de voir s'il existe un lien avec l'évolution du taux de mortalité ou de l'espérance de vie.

Comment lire les données

Afin de lire et étudier les données voici un module intéressant dans **Julia** :

- CSV (<https://csv.julidata.org/stable/>)