



École d'été interdisciplinaire en numérique de la santé du 26 au 30 mai 2025

Introduction à R avec un cas d'étude Plan d'activité pédagogique

Type de l'activité : ☒ Atelier ☐ Présentation

Objectif : Le présent document décrit le contenu scientifique et l'organisation de l'activité pédagogique « Introduction à R avec un cas d'étude ».

1 Introduction

L'activité pédagogique « Introduction à R avec un cas d'étude » se déroulera dans le cadre de l'école d'été interdisciplinaire en numérique de la santé (EINS) et s'inscrit dans le thème « Analyse de données en santé ».

1.1 Renseignements sur le présentateur

Prénom, nom : Jean-Philippe Morissette

Affiliation principale :

Chargé de cours au Département de mathématiques, Faculté des sciences, Université de Sherbrooke

Courriel : jean-philippe.morissette@usherbrooke.ca

Biographie

Enseignant et assistant de recherche spécialisé en statistique, Jean-Philippe Morissette estime que l'apprentissage et le développement des connaissances scientifiques passent avant tout par le plaisir et l'intérêt des personnes impliquées. Curieux, il tente de rendre l'expérience d'apprentissage motivante et unique, et est prêt à utiliser de nouveaux outils afin de favoriser la compréhension et le processus d'intégration des savoirs.

Après avoir donné divers cours de statistiques offerts dans des programmes d'informatique, d'intelligence d'affaires, de marketing, d'administration générale et, évidemment, de mathématiques, il se lance maintenant dans un atelier appliqué en santé.

2 Description

Cette section présente le contenu de l'activité et les principaux objectifs.

2.1 Contenu

Devant la quantité de données qui sont produites et disponibles, il devient nécessaire d'utiliser un logiciel si l'on souhaite pouvoir communiquer des résultats. Que ce soit pour le nettoyage, l'importation, la transformation, la visualisation ou la modélisation de données, tout doit passer par un logiciel d'analyse de données. Dans cet atelier, il sera proposé aux personnes étudiantes d'utiliser le langage de programmation R ainsi que l'environnement de développement RStudio comme outils d'analyse.

R est un logiciel libre qui permet l'utilisation de différentes bibliothèques spécialisées en ligne. Au cours de la dernière décennie, une série de bibliothèques ont été introduites afin de faciliter et de standardiser la manipulation de données (voir *tidyverse*) ainsi que la visualisation de données (voir *ggplot2*). Ces bibliothèques ont été développées afin que la majorité des étapes de l'analyse de données puisse être faite de façon plus *humaine* : la lecture et la rédaction de l'analyse suivent une approche grammaticale semblable au langage usuel. Lors de l'atelier, les personnes étudiantes pourront apprivoiser le langage R via l'exploration d'un jeu de données inspiré d'un cas réel en santé en tentant de répondre à une question de recherche qui concerne ce jeu de données.

2.2 Objectifs de formation

Cette activité permettra à une personne étudiante de :

- O1.s'initier au langage de programmation R, de connaître certaines de ses particularités et d'utiliser l'environnement de développement RStudio ;
- O2.se familiariser avec les principes fondamentaux ainsi que les bonnes pratiques en matière d'organisation des données ;
- O3.développer ses capacités à manipuler efficacement des données ;
- O4.construire et analyser un modèle linéaire afin de répondre à une question de recherche.

3 Références

Cette section présente les principales références documentaires utilisées pour construire l'activité et les références pour approfondir des concepts présentés.

3.1 Références essentielles

Statistics Team. (s.d.). *Introduction to R for Health Data Science*. Repéré en ligne le 15 avril 2025, à partir de https://bookdown.org/m_p_sperrin/introduction_to_r/https://r4ds.had.co.nz/index.html

Wickham, H., Çetinkaya-Rundel, M., & Golemund, G. (2023). *R for Data science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media.

3.2 Références complémentaires

Wickham, H. (2016). *ggplot2*. In Use R! <https://doi.org/10.1007/978-3-319-24277-4>

Boehmke, B. C., PhD. (2016). *Data Wrangling with R*. Springer.

Peng, R. D. (2012). *R Programming for data science*.