

Cours	
Titre	Science des données de santé publique
Nombre de crédits	3
Sigle	MSO6624
Site StudiUM	En ligne, à l'onglet du cours sur Studium
Faculté / École / Département	École de santé publique de l'Université de Montréal
Trimestre	Hiver
Type de formation	En présentiel / En ligne

Description du cours	
Description simple	<p>Ce cours présente les aspects scientifiques et techniques autour de la donnée en santé : de la modélisation et exploration des bases de données, à l'utilisation des données pour l'apprentissage automatique</p>
Description détaillée	<p>Les données de santé sont un support essentiel à la recherche mais également à la mise en place d'un système de santé capable d'adapter automatiquement ses capacités et ses actions en fonction de nouvelles connaissances mises en évidence par ses données : un système de santé dit « apprenant ». Les outils pour créer et manipuler des données de santé, ainsi que ceux utilisés pour réaliser de l'apprentissage automatique sur ces données ne sont pas spécifiques du domaine de la santé. Cependant, ces outils demandent une certaine adaptabilité aux besoins en santé pour soutenir l'utilisabilité des données pour les soins quotidiens ainsi que leur utilisation secondaire (recherche, gestion, apprentissage automatique, etc.). Ce cours veut ainsi apporter la perspective nécessaire à l'adaptation des outils de gestion et de manipulation des données au contexte de la santé.</p> <p>Ce cours propose une approche pratique et opérationnelle pour une introduction à la science des données en santé. Il vise à donner les connaissances techniques de base pour la création et la manipulation des différentes bases de données en y associant la particularité du domaine de la santé. Ces notions seront ainsi explorées selon deux perspectives les aspects techniques de bases et les défis de leur implémentation actuelle en santé : (i) Les principes de la modélisation conceptuelle et relationnelle seront ainsi abordés ainsi que les bases du langage SQL. Une perspective sera ouverte aux autres types de bases de données (bases de données graph) ainsi qu'au langage SPARQL. (ii) Les modèles de données utilisées dans les entrepôts de données (ex., OMOP, I2B2, etc.) et de transmission des données (ex., HL7) en santé seront ensuite présentés en perspective des difficultés d'implémentation. (iii) Pour l'apprentissage machine : (a) Les grands principes de l'apprentissage automatique et de l'intelligence artificielle seront décrits pour permettre la mise en place d'algorithmes d'apprentissage sur des jeux de données ad hoc. (b) Les défis de l'implémentation de ces algorithmes sur les données en santé.</p>