

CIV8760 - Gestion de données en transport

Plan de cours

Nicolas Saunier
nicolas.saunier@polymtl.ca

Automne 2022

Le responsable du cours est Nicolas Saunier, professeur titulaire au Département de génie civil, géologique et des mines (CGM) de Polytechnique Montréal. Pour plus d'informations, merci de vous référer à sa page professionnelle.

L'enseignant du cours est François Bélisle, diplômé de Polytechnique (ingénieur civil ; également titulaire d'un baccalauréat et d'une maîtrise en Mathématiques et Informatique) et est data scientist dans sa propre entreprise (Math Mobile).

Le chargé de travaux pratiques (TP) est Frédérick Chabot, diplômé de Polytechnique (ingénieur civil et maîtrise en cours en recherche en transport).

1 Description et objectif du cours

La description du cours à l'annuaire est la suivante:

Étapes principales de la collecte de données de transport: planification, gestion, validation, analyse et diffusion. Principales bases de données et microdonnées pertinentes pour l'étude des problèmes de transport. Classes et propriétés des données spatio-temporelles. Modèles de données. Méthodes et outils de traitement, analyse, modélisation et visualisation des données. Systèmes d'information géographique : systèmes de projection, concepts et intégration. Méthodes d'analyse spatiale applicables aux données de transport et outils pertinents. Nouvelles technologies de collecte de données en transport et potentialités d'analyse. Méthodes de fouille de données, d'apprentissage automatique et modèles économétriques.
(tirée du site)

Le nombre de crédits est de 3 et le triplet (3 - 1.5 - 4.5).

L'objectif du cours est de connaître et savoir appliquer des méthodes de traitement, d'analyse et de modélisation de données pour la planification, la conception et la gestion des systèmes de transport. À la fin de ce cours, vous serez en mesure de concevoir un système d'information adapté à la plupart des problèmes de transport.

Place du cours dans le programme Le cours de Systèmes de transport (CIV2710) propose une introduction au transport et quelques types de données typiques. Le présent cours explore ces idées en plus grand détail et propose une introduction à un ensemble de méthodes d'analyse et d'interprétation des données de transport. Certaines techniques de traitement de données, en particulier pour les données spatiales, vues dans le cours de

Technologies informationnelles appliquées en génie civil (CIV1120), de Programmation procédurale (INF1005D) et le Projet de caractérisation avec GPS et GIS (CIV3930) sont pertinentes pour les travaux.

Le cours de Gestion de données en transport est un cours d'ingénierie offert dans l'orientation terminale de transports avec trois autres cours: Construction et restauration de chaussées (CIV8730), Circulation (CIV8740) et Planification durable des transports (CIV8750).

2 Qualités considérées

Les huit premières qualités, ainsi que la dernière (apprentissage continu), sont approfondies dans le cours. En particulier, les étudiants apprennent à concevoir des modèles de données prenant en compte les besoins de collecte, traitement et analyse de ces données.

3 Objectifs d'apprentissage

À la fin du cours, l'étudiant sera en mesure de:

- maîtriser la nomenclature associée à la réalisation des grandes enquêtes sur la mobilité des personnes;
- sélectionner les méthodes d'acquisition de données les plus pertinentes en fonction des différents types de problèmes de transport et concevoir un processus de réalisation;
- définir les types de données selon leurs propriétés spatiales, temporelles et sémantiques;
- créer et manipuler des couches d'information géographique à l'aide de différents logiciels;
- concevoir des modèles de données de transport et développer des processus de traitement et d'analyse de ces données;
- appliquer les principales méthodes d'analyse spatiale et d'analyse spatio-temporelle aux données de transport pour modéliser différents problèmes de transport;
- concevoir des outils de consultation et visualisation analytique de données;
- appliquer différentes méthodes de fusion de bases de données en vue d'enrichir leur contenu.

Après avoir fait les TP en laboratoire informatique, l'étudiant sera en mesure de:

- créer des couches d'information spatio-temporelles et structurer un système d'information géographique intégrant différents types d'objets;
- appliquer des techniques d'analyse spatiale à l'aide de différents outils informatiques pour modéliser des problèmes de transport;
- concevoir des outils interactifs de visualisation analytique de données à l'aide d'outils informatiques (tableur, etc.);

- appliquer des techniques d'estimation des modèles statistiques avec composantes spatiales;
- extraire et interpréter des informations de bases de données de transport, à l'aide de systèmes de gestion de bases de données et d'outils de fouille de données.

4 Méthodes d'enseignement

Les méthodes d'enseignement sont traditionnelles. Les séances de cours reposent sur des exemples (incluant différents types de contenu comme des vidéos, animations et logiciels) pour introduire les concepts, et des exercices pour les mettre en œuvre et tester la compréhension. Le cours se déroule en salle informatique de sorte que chacun puisse utiliser l'ordinateur pour faire certains exemples et exercices.

La participation en cours est très fortement encouragée. S'ajoutent quelques activités, quiz et recherches documentaires entre les cours pour vérifier la compréhension.

Les séances de laboratoire commencent par une courte introduction du sujet, puis les étudiants travaillent par eux-mêmes. D'autres explications sont données à toute la classe selon les questions et les besoins.

5 Modalités de fonctionnement

François Bélisle est joignable en tout temps par courriel (<mailto:belisle.francois@gmail.com>); on peut également se recontrer, virtuellement ou en personne, en prenant rendez-vous.

6 Mécanismes d'évaluation

L'évaluation de l'apprentissage se fera selon quatre catégories d'évaluation dont la pondération est établie comme suit:

Contrôles	Pondération
Activités et quiz	10 %
TP	30 %
Contrôle périodique	30 %
Examen final	30 %

Pour les évaluations de certaines questions ouvertes dans les examens ou TP, le barème de correction est explicité en annexe ??.

7 Documentation pertinente

Le matériel du cours est disponible sur le site moodle du cours. Un ensemble de calepins Jupyter propose des exemples et des activités en Python illustrant le cours.

Des ressources pour installer et apprendre le langage Python sont présentées sur le site GitHub du cours. Des ressources complémentaires sur la programmation et la science des données sont disponibles sur le Wiki des Transports Intelligents de Polytechnique Montréal (en anglais).

D'autres ressources sont listées en annexe ??.

8 Calendrier des rencontres et des activités d'évaluation

Le cours est donné le mardi de 9h30 à 12h20 dans la salle L6624. Les séances de TP sont données le vendredi de 8h30 à 11h20 les semaines B2 (commençant donc le 8 septembre) dans la salle L6655.

À titre indicatif, le calendrier des cours et des sujets abordés est le suivant:

Date	Sujet
30 août	Présentation du cours et introduction
6 septembre	Méthode de collecte de données
13 septembre	Traitement de données
20 septembre	Bases de données (modèles)
27 septembre	Bases de données (langage SQL)
18 octobre	Contrôle périodique
25 octobre	Données spatiales
1 novembre	Analyse statistique
8 novembre	Modèles statistiques
15 novembre	Visualisation de données
22 novembre	Fouille de données 1
29 novembre	Fouille de données 2
7 décembre	Synthèse

Le contrôle périodique est prévu pour le 18 octobre, après la semaine de relâche.

À titre indicatif, le calendrier et les sujets des TP sont les suivants:

Date	Sujet
9 septembre	Initiation à Python et à un logiciel d'enquête
23 septembre	Réalisation et analyse d'une enquête transport
7 octobre	Modèles et bases de données
28 octobre	Données spatiales et statistiques
11 novembre	Analyses statistiques
25 novembre	Fouille de données

Les sujets des TP peuvent être appelés à changer.

Les rapports des TP seront rendus généralement une semaine après la séance (les dates de rendu seront confirmées dans chaque énoncé). Les rapports de TP de ce cours devraient suivre le format suggéré par le guide de rédaction préparé pour le programme de génie civil (disponible sur le site Moodle du cours).

9 Fraude et plagiat

Les règlements des études du baccalauréat de Polytechnique Montréal s'appliquent à ce cours (en particulier concernant la fraude, le plagiat, les absences, etc.). Veuillez les consulter sur le site du secrétariat général et sur le site de prévention du plagiat.

En outre, j'aimerais insister sur la rédaction de vos travaux dans un français de bonne qualité, car vous serez aussi jugés sur vos qualités de présentation dans votre vie professionnelle. Les rapports doivent suivre le guide de rédaction de génie civil. Les retards dans le rendu des travaux entraîneront automatiquement des pertes de points (10 % de la note par jour de retard).

A Évaluation des questions ouvertes

A	B	C	D
Modèle de données E/A			
Toutes les associations entre entités sont correctes (fonctionnalités et cardinalités), avec plusieurs exemples de attributs, y compris l'identifiant. Les choix sont justifiés.	Il y a quelques erreurs dans les fonctionnalités/cardinalités ou des attributs peuvent manquer.	Il y a plusieurs erreurs d'associations entre les entités et des attributs manquent. Certains choix ne sont pas justifiés.	Le modèle proposé ne permet pas de représenter les données, la plupart des associations sont erronées, plusieurs entités n'ont aucun attribut.
Modèle de données relationnel			
Toutes les entités sont représentées par des tables, avec une clef primaire clairement identifiée et des attributs (et leur types), et les clefs externes représentent correctement les liens entre les tables (les associations 1-n et n-m du modèle E/A). Les choix sont justifiés.	Il y a quelques erreurs dans les clefs primaires ou externes, ou des attributs peuvent manquer.	Il y a plusieurs erreurs de clefs primaires et externes et des attributs manquent. Certains choix ne sont pas justifiés.	Le modèle proposé ne permet pas de représenter les données, il y a des erreurs de compréhension des clefs primaires et externes.

B Autres ressources

Il y a plusieurs associations professionnelles et organismes gouvernementaux dans le domaine du transport qui publient des manuels, recommandations, font de la formation:

- AASHTO: American Association of State Highway and Transportation Officials
- AQTr: Association Québécoise des Transports
- ATC: Association du Transport du Canada
- CITE: Canadian Institute of Transportation Engineers
- FHWA: Federal Highway Administration
- IEEE: Institute of Electrical and Electronics Engineers
- ITE: Institute of Transportation Engineers (gratuit pour les étudiants)
- MTQ: Ministère des Transports du Québec
- NHTSA: National Highway Traffic Safety Administration
- TRB: Transportation Research Board

Parmi les institutions universitaires dans le domaine du transport, on peut noter:

- CIRRELT: Centre Interuniversitaire de Recherche sur les Réseaux d'Entreprise, la Logistique et le Transport
- HEC-Montréal: École des hautes études commerciales de Montréal
- INRS: Institut National de la Recherche Scientifique
- IFSTTAR: Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux (anciennement Institut National de la Recherche sur les Transports et leur Sécurité, INRETS)

Parmi les revues avec comité de lecture dans le domaine des transports, on peut noter:

- Transportation Research Record: Journal of the Transportation Research Board
- Transportation Research Part A, B, C, D, E, F
- Accident Analysis and Prevention
- Journal of Intelligent Transportation Systems
- Journal of Advanced Transportation
- Computer-Aided Civil and Infrastructure Engineering
- IEEE Transactions on Intelligent Transportation Systems
- IEEE Transactions on Vehicular Technology
- ITE Journal
- Recherche Transports Sécurité

- Routes et transports (de l'AQTr)

Je vous recommande les moteurs de recherche suivants pour trouver des articles et constituer une bibliographie:

- Google Web et Scholar
- Les sites des éditeurs (dont le contenu est accessible lorsque vous vous connectez à Internet de l'École): Elsevier, Springer, IEEE explore, ACM, Scopus, Taylor and Francis, Public Library of Science (PLOS), etc.
- Les dépôts publics comme ArXiv et engrXiv
- Transportation Research Information Services (TRIS) database