

TP1 - Méthodologies de collecte de données : étude de cas, conception et analyse

CIV8760 - Gestion de données en transport
Frédéric Chabot & Nicolas Saunier

8 Septembre 2023

Ce premier travail pratique s'intéresse aux méthodes de collecte de données dans le domaine du transport et vise à vous familiariser avec les divers paramètres les caractérisant. Ce travail s'articule autour de deux mandats :

1. Étude de cas dans un cadre existant;
2. Développement d'une étude de cas.

1 Étude de cas dans un cadre existant

Pour ce premier mandat, vous devrez effectuer certaines analyses à l'aide d'une base de données portant sur une étude récente faite au Parc Jean-Drapeau. L'étude porte sur l'usage de certains sous-espaces au Parc Jean-Drapeau. La Société du Parc Jean-Drapeau se demande notamment si l'animation et le mobilier mis en place dans ces espaces favorisent ou non leur usage (ou si les gens ne font que passer par là) et de quelle(s) façon(s) exactement (quelle(s) activité(s) y sont faites). Les documents suivant (disponible sur Moodle) vous aideront à mieux comprendre le contexte et la méthodologie afin de produire vos analyses :

- [Guide du bénévole](#) ;
- [Présentation du plan directeur du Parc Jean-Drapeau](#)
- [Présentation de la méthodologie des études de la vie dans les espaces publics \(ÉVEP\) par le Centre d'écologie urbaine de Montréal](#) ;
- [Présentation concernant l'usage de l'application mobile pour l'observation de la vie dans les espaces publics \(OVEP\)](#).

Cette dernière présentation vous aidera non seulement à comprendre certaines limites des analyses que vous pouvez faire avec les résultats collectés par cet outil, mais aussi à pouvoir utiliser l'outil pour la seconde partie de ce travail pratique.

À noter que toutes manipulations des données pour cette partie du travail doit être faites avec SQL (obligatoire). Chaque requête utilisées pour extraire l'information nécessaire doit être précisées dans le rapport. Veuillez vous référer à l'exemple ci-dessous pour le format de présentation attendu des requêtes :

```
SELECT * FROM thisTable GROUP BY uneColonne
```

Vous ne pouvez pas mettre de simples captures d'écran de vos requêtes. Les cartographies devront être faites à partir du logiciel QGIS (ou équivalent). Enfin, vous êtes libres d'utiliser l'outil que vous voulez pour faire les graphiques/tableaux (Excel, Python, etc.).

1.1 Analyse des données

Pour l'analyse, vous devez proposer sous forme graphique une analyse temporelle, une spatiale et une multivariée (plus de 2 variables). Vous devez fournir un minimum de 2 graphiques par type d'analyse (donc 6 au total). Commentez chacun d'eux et faites des liens avec la méthodologie et/ou la matière du cours.

De plus, vous devez faire une analyse de la variable "Type de groupe". C'est-à-dire que vous devez proposer une méthode pour transformer cette variable catégorielle en variable numérique. En utilisant la variable "Nombre exacte", faite une comparaison des résultats obtenus avec votre méthode. Justifiez les différences perçues.

1.2 Modèle de données

Après avoir pris connaissance de l'étude, veuillez préciser si les tableaux (feuilles) de [ce chiffrier](#) pour l'entrée des données respectent les trois normes. Indiquez si la base de données suit ou ne suit pas les trois normes. Veuillez finalement présenter schématiquement quel serait le modèle approprié selon vous en prenant le soin de noter les différentes associations, dimensions, cardinalités et fonctionnalités. De plus, résumez dans un (ou plusieurs) tableau à quelle entité correspond chaque variable de la base de données.

2 Conception d'une étude de cas

Comme vous l'avez vu en cours, plusieurs outils existent pour procéder à une collecte de données. Le choix du bon outil est ainsi une étape importante de la collecte de données. Ce choix peut notamment dépendre des limites en place (budget, cadre temporel, etc.) et des caractéristiques à collecter. Pour ce mandat, vous devez ainsi concevoir une étude de cas dans laquelle vous devrez faire de la collecte par observation directe. Vous serez donc en mesure de vivre vous-même les conséquences des choix qui sont faits pour concevoir une telle étude de cas. Le type d'étude se colle à celui de la première partie de ce travail pratique.

2.1 Site, problématique, hypothèse(s) et objectif

Vous devez ensuite choisir et décrire un site, une problématique (inventée) et un objectif qui défend la pertinence de faire votre collecte.

Par exemple, vous pourriez décider de prendre la place Shamrock à Montréal se trouvant près du Marché Jean-Talon et vérifier l'usage des bancs/tables publics, car selon votre expérience personnelle, il semble manqué de places assises à certains moments de la journée.

Il serait donc intéressant de savoir si c'est le cas, à quel moment, si un groupe d'âge ou un genre est davantage affecté, etc.

2.2 Sélection de l'outil

Selon vos préférences et ce à quoi vous avez accès, vous devez tout d'abord choisir entre la méthode papier-crayon et l'application mobile (voir les différents documents affichés sur Moodle dans la section TP1).

2.3 Sélection des variables/attributs

Vous devez choisir au moins 3 caractéristiques/variables à observer en plus d'une 4e qui est la géolocalisation de l'observation faite (*Mapping*). À noter que pour la géolocalisation, vous pouvez faire de l'observation par "sous-espace" dans votre espace afin d'en faire une collecte prenant en compte l'aspect spatial. Par exemple, pensez à un échiquier comme représentation de votre espace et qu'une personne observée pourrait donc avoir certains attributs (genre, âge, etc.), en plus de faire une certaine activité dans la zone (case) D5 de cet échiquier.

2.4 Présentation de la méthodologie

Vous devez présenter votre ou vos grilles sous forme de tableau, image ou schéma (selon l'outil utilisé). De plus, vous devez faire un choix sur le nombre de grilles, la période de temps pour observer par grille, s'il y a répétition de la collecte sur plusieurs jours ou semaine, etc. Veuillez justifier tout vos choix en fonction des paramètres mentionnés précédemment (variables, site, problématique, objectifs, etc.). Il est important dans ce mandat d'avoir au moins collecté pour 30 minutes au total dans un espace ayant plus de 2 personnes par minute environ (60 personnes observées minimum).

Cette étape vous permettra notamment de bien analyser vos résultats (faire des liens et exposer des limites).

2.5 Analyse sommaire

Si vous avez procédé à un traitement/nettoyage des données, veuillez en parler et donner des justifications claires.

Vous devez présenter les résultats sous forme de statistiques descriptives agrégées et inclure des figures (analyses bivariées et multivariées) et une cartographie aidant à la compréhension des résultats. Vos résultats doivent permettre de comprendre le type d'espace et son usage.

Au vue des résultats obtenus, analysez si les paramètres de votre collecte étaient appropriés au contexte d'étude (achalandage de l'espace *vs* le format de collecte choisi, caractéristiques analysées *vs* les attentes selon l'espace, etc.). Soyez transparents quant aux problèmes rencontrés durant et après votre collecte et proposez des solutions (autre outil, taille de l'espace, etc.). Décrivez quels seraient les avantages et désavantages de vos solutions par rapport à ce qui a été fait ici.

3 Modalités

Pour le second mandat, vous devez avoir récolté au moins le 2/3 de vos observations avant le second laboratoire (22 septembre) pour commencer vos analyses avec les observations que vous aurez.

Ce travail se fait par groupe de deux ou trois. Vous pouvez accéder à [ce fichier](#) pour indiquer votre numéro d'équipe et votre information personnelle. Un **rapport** au format PDF ou Word, ne dépassant pas **30 pages**, doit faire état des mandats de ce travail pratique. La date de rendu est le **5 octobre 2023 à 23h59**. Le fichier doit être déposé en format électronique sur Moodle.

Le nom du fichier doit porter la nomenclature suivante :
EQ{numéro d'équipe}_TP{numéro du TP}_{semestre d'étude (A, H ou E)}{année d'étude}.
Par exemple, "EQ01_TP1_A23".

Une attention particulière sera portée à la rédaction (les fautes de français seront sanctionnées tout comme une mauvaise organisation générale du travail), comptant pour 5% sur la note finale.

Veillez consulter le [Guide de rédaction pour ingénieur civil](#) disponible sur Moodle à la section Ressources pour vous assurez de remettre un rapport en bonne et due forme.