



**POLYTECHNIQUE
MONTRÉAL**

UNIVERSITÉ
D'INGÉNIERIE

Plan de projet

**Exo : Achalandage et ponctualité des lignes de bus 9 et 22 de la
couronne Nord de Montréal**

INF8808 - Visualisation de données

Adam Prévost - 1947205
Armelle Jézéquel - 2098157
Clara Serruau - 2164678
Jules Lefebvre - 1847158
Julien Dupuis - 1960997

24 février 2022

Table des matières

Mise en contexte	3
1.1 Contexte	3
1.2 Objectif	3
1.3 Public cible	3
Jeu de données	3
2.1 Métadonnées	3
2.2 Données	4
Questions cibles	6

1. Mise en contexte

1.1 Contexte

Le transport en commun est une activité essentielle et répandue dans la grande région de Montréal. Dans cette région, au moins une personne sur cinq utilise le transport en commun régulièrement selon *Statistique Canada*. La compagnie qui s'occupe du bon fonctionnement du réseau est *Exo*. Le domaine des transports en commun est en continuelle modernisation, où se côtoient anciennes et nouvelles infrastructures. Nous porterons notre attention sur le réseau des autobus, pour lesquels de plus en plus de données sont recueillies et analysées (achalandage, ponctualité de chaque autobus individuellement...) Elles peuvent, entre autres, servir à offrir des services de meilleure qualité à la population de Montréal et des alentours.

1.2 Objectif

L'objectif d'*Exo* est d'utiliser les données d'achalandage (nombre de personnes à bord) et de ponctualité (temps de retard ou d'avance) récoltées par GPS à chaque arrêt pour rendre compte de la qualité du service. Ces données sont usuellement étudiées séparément, mais on souhaite les étudier conjointement afin de rendre compte d'une autre réalité du réseau. Plus précisément, la qualité de service est évaluée en ciblant les lignes, les voyages, les arrêts ou les tronçons les plus problématiques. Le but de la visualisation est d'aider à prendre des décisions correctives en changeant les horaires ou les trajets par exemple. La visualisation doit donc aussi permettre d'identifier des tendances si un problème est généralisé.

1.3 Public cible

La visualisation est principalement dédiée aux employés d'*Exo* qui planifient les opérations de transport en commun. Ils utiliseront celle-ci pour avoir une idée de la qualité de service et prendre des décisions éclairées. La visualisation ne sera pas à destination des usagers du réseau de transport, donc les utilisateurs de la visualisation auront déjà de l'expertise dans le domaine.

2. Jeu de données

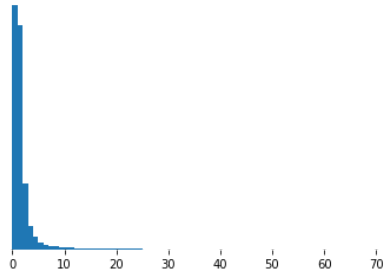
2.1 Métadonnées

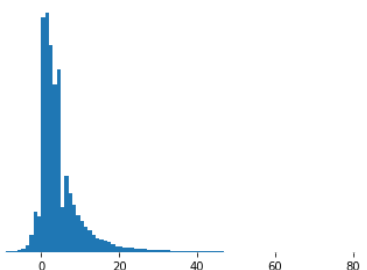
Les données sont fournies par *Exo*, et ont été récoltées de septembre à décembre 2021. Les données sont récoltées par des capteurs présents dans les bus : GPS, capteur de détection de personne montante...

Les données sont sujettes à la défaillance des capteurs et des bus. Certains bus peuvent avoir eu des problèmes durant le voyage et certains arrêts ne sont pas marqués (faute d'usager montant ou descendant).

Le fichier à traiter comprend 206 076 lignes de données.

2.2 Données

Variable	Description	Data type	Data values / range
Date	Date du voyage	date :YYYY-M M-DD	1ier septembre 2021 au 1ier décembre 2021
Secteur	Le secteur de la métropole dans lequel évolue la ligne de bus	string	une seule valeur : CITLA
Ligne	Numéro de la ligne de bus	int	{9; 22}
Voyage	Numéro du voyage effectué dans la journée (1 = 1ier bus de la journée)	int	range numérique : [1,126] range textuelles : {69_F1-01; 7_F1-01; ...} (voyages lors des jours fériés, à ignorer car non représentatifs)
Arret_code	Code de l'arrêt de bus	int	range : [80400, 84967] 231 valeurs différentes)
Arret_nom	Nom de l'arrêt de bus	string	{du Curé-Labelle / Saint-Georges; Saint-Georges / Ouimet; du Curé-Labelle / 104e Avenue Est; ...} 188 valeurs différentes
Montants	Personnes qui embarque	int	min : 0.0 max : 69.0 mean : 1.53994157495293 median : 1.0 
Etat_ponctualité	Discrétisation de <i>Minutes_ecart_planifie</i> : [-11;-1] : <i>Avance</i> [0;4] : <i>Ponctuel</i> (sauf dans le cas du dernier arrêt de la ligne : [-11;4] : <i>Ponctuel</i>) [5;80] : <i>Retard</i>	string	{Ponctuel, Retard, Avance}
Minutes_ecart_planifie	Minutes entre le moment de passage réel du bus	int	min : -11.0 max : 80.0

	et le prévu <i>Réel - Prévu</i>		mean : 4.426027290902385 median : 3.0 
Séquence_arret	Ordonne les arrêts pour une ligne/direction	int	Ligne 9 (Métro Montmorency) : [1; 74] Ligne 9 (Lafontaine Via Gare Saint-Jérôme) : [1; 79] Ligne 22 (Gare Sainte-Thérèse) : [1; 43] Ligne 22 (Bois-des-Filion Via Georges IV) : [1; 41]
Girouette	Nom de la direction du bus	string	9 : {Lafontaine Via Gare Saint-Jérôme; Métro Montmorency} 22 : {Bois-des-Filion Via Georges IV; Gare Sainte-Thérèse}
Direction	Direction du bus	string	{sens opposé; sens}
Periode_horaire	Période de la journée durant laquelle a eu lieu le voyage	string	{AM; PM; interpointe; Soir/Nuit/Matin}
Arret_Latitude	Latitude de l'arrêt de bus	float	range : [45,55828; 45,804104]
Arret_Longitude	Longitude de l'arrêt de bus	float	range : [-74,011985; -73,720406]

Précisions sur les données :

- **Secteur** : le code CITLA correspond à la couronne nord de Montréal, plus précisément au secteur Laurentides.
- **Arret_code** , **Arret_nom** : Il y a plus de codes que de noms d'arrêt, car un arrêt peut avoir un même nom dans les 2 sens, mais avoir un code différent pour chacun.
- **Minutes_ecart_planifie**, **Etat_ponctualité** : la seconde colonne est une discrétisation de la première. Un bus est considéré ponctuel s'il a jusqu'à 4 minutes de retard. Au-delà, il est considéré comme en retard. Pour l'avance, un bus peut être considéré comme ponctuel s'il a de l'avance et qu'il s'agit du dernier arrêt, ou qu'il a moins de 4 minutes de retard et moins d'une minute d'avance. Enfin, un bus est considéré comme en avance s'il a au moins une minute d'avance.
[-11;-1] : Avance
[0;4] : Ponctuel (sauf dans le cas du dernier arrêt de la ligne : [-11;4] : Ponctuel)
[5;80] : Retard
- **Montant** : nombre fourni par un compteur de passagers dans l'entrée de l'autobus.

- **Voyage** : Certains voyages ont un code textuel. Ils correspondent à des trajets réalisés lors d'un jour férié. Ces données ne sont pas pertinentes pour l'analyse du reste de l'année, le client nous a demandé de les ignorer.
- **Séquence_arrêt** : La ligne 9 a 79 arrêts, mais l'arrêt Station Montmorency est parfois l'arrêt 1, parfois l'arrêt 79 dans la direction Lafontaine Via Gare Saint-Jérôme (est-ce une erreur, ou un cas particulier ?)

3. Questions cibles

Questions	Priorité (/3)
Aspect temporel	
Quels moments de la journée sont les plus problématiques?	☆ ☆ ☆
Quel est le retard moyen?	☆ ☆ ☆
Quel est l'achalandage moyen?	☆ ☆ ☆
Est-ce qu'un autobus peut rattraper son retard après un arrêt problématique?	☆ ☆
Quels problèmes sont récurrents et lesquels sont ponctuels?	☆ ☆
Existe-t-il des relations entre le retard et le nombre de passagers?	☆ ☆
Existe-t-il des tendances temporelles particulières dans l'échantillon donné? (impact de la rentrée scolaire, jour de la semaine problématique, Halloween, etc.)	☆
Aspect géographique	
Quel est le retard relatif gagné ou perdu à chaque arrêt?	☆ ☆ ☆
Quel est le nombre de passagers relatif gagné ou perdu à chaque arrêt?	☆ ☆ ☆
Quels sont les arrêts ou les tronçons les plus problématiques?	☆ ☆ ☆
Est-ce qu'un sens du trajet est plus ou moins problématique que l'autre?	☆ ☆
Quelle est la qualité de service sur les lignes en fonction de l'arrêt ou la zone géographique?	☆ ☆
Existe-il des contraintes géographiques particulières qui occasionnent des retards?	☆
Existe-t-il des lignes problématiques?	☆