

Base de données sur les réseaux de transport en commun du Canada

Base de données sur les réseaux de transport en commun du Canada : Rapport de métadonnées



Date de diffusion : le 30 juillet 2025



Statistique
Canada

Statistics
Canada

Canada

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à infostats@statcan.gc.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros suivants :

- | | |
|---|----------------|
| • Service de renseignements statistiques | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur | 1-514-283-9350 |

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « Contactez-nous » > « [Normes de service à la clientèle](#) ».

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, les entreprises, les administrations et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Publication autorisée par la ministre responsable de Statistique Canada

© Sa Majesté le Roi du chef du Canada, représenté par la ministre de l'Industrie, 2025

L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

This publication is also available in English.

Table des matières

Sommaire	4
Contexte	4
La norme de données GTFS	4
Collecte et compilation des données	5
Résumé des données.....	6
Utilisation des données	6
Remerciements	7
Comment citer cet ensemble de données	7
Communiquez avec nous	7

Sommaire

La [Base de données sur les réseaux de transport en commun du Canada](#) est une collecte unifiée de données General Transit Feed Specification (GTFS) à l'échelle nationale. Les données de sociétés de transport en commun ont été recueillies à partir des sites Web des fournisseurs de transport en commun. La diffusion a pour objectif de fournir un cadre cohérent pour l'examen de la couverture et de l'accessibilité du transport en commun dans l'ensemble des secteurs de compétence en soutenant les chercheurs et les praticiens.

Les principales étapes du traitement comprenaient l'agrégation des flux GTFS de fournisseurs Canadienne dans un ensemble de données unifié, l'exécution de contrôles de validation de base pour évaluer la conformité aux normes GTFS et l'exportation des données GTFS au format géospatial.

Cette diffusion comprend les ressources suivantes :

- **fichiers de données gtfs** : fichiers .txt au format GTFS, qui comprennent notamment les arrêts, les itinéraires, les trajets, les heures d'arrêt, les calendriers et des champs supplémentaires, le cas échéant;
- **ensemble de données géospatiales nationale** : un ensemble de données géospatiales unifiée compatible avec les systèmes d'information géographique pour combiner les données sur les arrêts et les itinéraires (en format geopackage);
- **fichiers de validation des données** : fichiers résumant les avertissements et les erreurs pour chaque flux d'agence de transport.
- **métadonnées** : détails sur les sources de données et les descriptions des noms de colonnes.

Les données sont diffusées au moyen d'une [Licence du gouvernement ouvert](#), conformément à la [Directive sur le gouvernement ouvert](#).

Contexte

But

Une compilation pancanadienne de renseignements sur le transport en commun a d'abord été élaborée pour être utilisée dans un éventail d'autres produits connexes de Statistique Canada, notamment les [mesures spatiales de l'accès](#). Il est devenu évident qu'une base de données unifiée sur les infrastructures de transport en commun au Canada était nécessaire. Par conséquent, la Base de données sur les réseaux de transport en commun du Canada a été conçue comme une base de données autonome sur les transports en commun qui serait mise à la disposition des chercheurs, des décideurs politiques et du public.

L'objectif est d'offrir un aperçu régulièrement mis à jour de l'infrastructure canadienne de transport public afin de permettre aux chercheurs de faire des comparaisons au fil du temps et entre différentes municipalités ou régions de transport en commun.

La norme de données GTFS

[General Transit Feed Specification](#) (GTFS) est le format ouvert utilisé pour les données sur le transport en commun. Il est composé d'une série de fichiers .txt qui détaillent les services fournis par chaque société de transport en commun, y compris l'emplacement des arrêts, les horaires de transport en commun (notamment les variations quotidiennes, hebdomadaires ou saisonnières), les itinéraires et les trajets. Le format de données GTFS, initialement créé par les chercheurs de Google, est maintenant une norme ouverte simple et accessible dans le monde entier qui est couramment utilisée dans les logiciels et les applications de planification des déplacements.

Les données GTFS nécessitent un ensemble de six fichiers .txt de base, qui sont présentés dans le Tableau 1. En plus de ces fichiers de base essentiels, de nombreuses agences de transport publient des informations supplémentaires au format GTFS, telles que les tarifs, les informations d'accessibilité, les fichiers de formes, les informations sur l'allocation de vélos (si les véhicules sont équipés de supports à vélos), la fréquence du service et les transferts.

Tableau 1
Définitions des fichiers d'horaire GTFS

Fichier	Présence	Description
agency.txt	Obligatoire	Informations sur l'agence de transport, comme son nom et son site web.
stops.txt	Obligatoire	Contient les coordonnées géographiques des arrêts (lieux où les véhicules prennent et déposent des passagers).
routes.txt	Obligatoire	Itinéraires suivis par les services de transport en commun. Ce fichier détaille l'identifiant, le nom et le type de chaque itinéraire.
Trips.txt	Obligatoire	Combine les données des fichiers routes, calendar et stop_times pour calculer les trajets précis qu'un passager peut emprunter.
stop_times.txt	Obligatoire	Heures d'arrivée et de départ à un arrêt donné pour un trajet spécifique. C'est généralement le fichier le plus volumineux et détaillé. Utilisé pour calculer l'heure de départ d'un service à un endroit précis.
calendar.txt and/or calendar_dates.txt	Obligatoire sous condition	calendar.txt définit les dates de service selon un calendrier hebdomadaire. calendar_dates.txt précise les exceptions (p. ex. jours fériés) aux services définis dans calendar.txt. calendar.txt est requis, sauf si toutes les dates de service figurent dans calendar_dates.txt.
fare_attributes.txt	Facultatif	Décrit les informations tarifaires des itinéraires d'une agence.
shapes.txt	Facultatif	Décrit le tracé suivi par un véhicule le long d'un itinéraire.
translations.txt	Facultatif	Traductions des éléments du jeu de données orienté vers le client.
attributions.txt	Facultatif	Attributions du jeu de données.

Note : Définitions basées sur [Référence de General Transit Feed Specification](#).

Portée

La Base de données sur les réseaux de transport en commun du Canada inclut tous les fichiers .txt de base (par exemple, itinéraires, arrêts, et trajets) pour tous les flux. La disponibilité des fichiers .txt GTFS supplémentaires (p. ex. accessibilité, tarifs, fréquences) varie en fonction des fournisseurs de services de transport en commun. Lorsque des fichiers .txt supplémentaires étaient accessibles, ils ont été agrégés et publiés.

En plus des [données GTFS statiques](#) de nombreuses sociétés de transport en commun publient un format dynamique de GTFS appelé [GTFS Realtime](#). Celui-ci fournit des mises à jour supplémentaires en temps réel sur les trajets, comme la position des véhicules et les alertes de service (stations déplacées, retards de service, interruptions inattendues, etc.). Les données GTFS Realtime nécessitent un rafraîchissement constant et dépassent donc la portée de ce premier cycle de la Base de données sur les réseaux de transport en commun du Canada. Le Laboratoire de données urbaines de Statistique Canada évalue la possibilité de mettre les données en temps réel à la disposition des utilisateurs dans les prochains cycles.

Collecte et compilation des données

Sources de données

Les sources de données potentielles des agences de transport en commun ont été compilées à partir de bases de données existantes d'informations sur le transport en commun, ainsi que d'une recherche à l'aide d'interfaces de programmation d'applications (API) à partir de deux plateformes de données de transport en commun ouvertes. La première est [Mobility Database](#) (en anglais seulement), une plateforme de données ouverte hébergée par MobilityData, une organisation à but non lucratif dédiée au partage des données GTFS, à la promotion des meilleures pratiques en matière de qualité des données et à la fourniture de documentation, de renseignements et de formations. La seconde, [Transitland](#) (en anglais seulement), est une plateforme communautaire de données ouvertes GTFS.

Un fichier de concordance a été créé pour faire correspondre les flux entre les sources et générer un identifiant unique pour chaque flux. Les flux ont été téléchargés directement à partir des sociétés de transport en commun qui disposent de données GTFS ouvertement accessibles en ligne. Dans le cas d'un petit nombre de sociétés, il était inscrit que celles-ci disposaient des données sur le transport en commun au format GTFS, mais les données n'étaient pas ouvertement accessibles sur le site Web de la société ou au moyen d'une API. Dans ces cas, les sociétés ont été contactées par courriel.

Compilation et validation des données

Pour confirmer que toutes les régions étaient adéquatement couvertes, les flux obtenus au moyen des API ont été jumelés et vérifiés manuellement pour confirmer qu'ils correspondaient à la région de couverture attendue. En cas d'erreur, les flux ont été supprimés ou, si possible, remplacés par un autre flux.

Il arrivait parfois que les flux dérivés des API incluaient des fournisseurs de services de transport en commun basés aux États-Unis. Ces cas ont fait l'objet d'une enquête individuelle. S'il était établi que le service n'effectuait aucun arrêt au Canada, ces flux étaient retirés et notés à des fins d'inventaire ultérieur.

Les flux GTFS ont été validés à l'aide du [MobilityData Canonical GTFS Schedule Validator](#) (en anglais seulement). Cet outil de validation vérifie la conformité avec la [référence de General Transit Feed Specification](#) et les [bonnes pratiques de GTFS Schedule](#) et génère une liste des violations de la conformité ou des éléments qui compromettent autrement la qualité des données. Ces éléments sont classés par degré de gravité (erreurs, avertissements et informations). Un résumé des erreurs, des avertissements et des avis d'information relevés pour chaque flux est inclus dans un fichier CSV de validation des données, publié en plus de l'ensemble de données primaires de GTFS Canada. Le fichier CSV de validation des données sert à aider les chercheurs à évaluer la qualité ou à signaler toute erreur qui pourrait se trouver dans les flux inclus dans GTFS Canada.

Le progiciel [gtfstools](#) (en anglais seulement) dans R a été utilisé pour valider, manipuler et fusionner les flux GTFS dans l'ensemble de données final de GTFS Canada. Un problème de compilation important s'est présenté lors de l'identification et du traitement des doublons (c.-à-d. des flux qui couvrent la même région). Les flux qui contenaient des renseignements identiques ont été relevés et signalés à l'aide de [gtfstools](#). D'autres cas (p. ex. un flux couvert par un autre flux plus volumineux) ont été relevés à l'aide de [QGIS](#) (en anglais seulement). Tous les cas en double ont fait l'objet d'une enquête individuelle afin de s'assurer que le flux à jour, actif et complet avait été sélectionné et qu'aucune donnée n'avait été omise par erreur lors de la suppression des doublons.

Résumé des données

La compilation finale des données contient un total de 139 flux qui couvrent les 10 provinces et 2 territoires. Ce total comprend également quelques services d'autobus, de train ou de traversier interrégionaux qui opèrent entre des municipalités ou des régions.

Toutes les agences ont inclus les six fichiers GTFS requis. Cependant, certains comportaient des avertissements d'erreur qui ont été détectés lors de l'étape de validation des données. Plusieurs fichiers GTFS optionnels ont été collectés lorsqu'ils étaient disponibles, mais n'étaient pas présents auprès de tous les fournisseurs de données. Les renseignements sur les fichiers fournis par chaque organisme sont inclus dans le fichier récapitulatif de validation inclus dans le dossier de téléchargement.

Le fichier Geopackage final contient une couche d'arrêts (points) et une couche de formes (lignes). Afin de réduire la taille du fichier, la géométrie de la couche des formes a été simplifiée en utilisant une tolérance de 250m.

Utilisation des données

Limites

Les données de la Base de données sur les réseaux de transport en commun du Canada ont été compilées à partir de sources de données ouvertes. Exception faite d'une vérification de la validation de la conformité aux normes de l'industrie de GTFS, la qualité des données a été acceptée « telle quelle ». Le fichier CSV de validation des données qui les accompagne est destiné à fournir aux utilisateurs une idée de la qualité des flux. À ce stade, aucune tentative n'a été faite pour corriger les erreurs ou les avertissements dans les flux.

Amélioration future

Pour les prochains cycles de données de la Base de données sur les réseaux de transport en commun du Canada, le Laboratoire de données urbaines vise à poursuivre des partenariats et à discuter des possibilités de partage de données avec les fournisseurs de services de transport en commun qui ne téléversent pas de données

GTFS à l'heure actuelle. L'objectif est d'améliorer la précision et la couverture de l'ensemble de données en veillant à ce que toutes les municipalités ou régions qui disposent de données sur le transport en commun soient représentées.

Le Laboratoire de données urbaines étudie également la possibilité de publier des données GTFS en temps réel dans les futures versions.

Remerciements

La Base de données sur les réseaux de transport en commun du Canada est le fruit des efforts soutenus d'organisations qui produisent et tiennent à jour des données ouvertes, et de la collaboration d'organismes qui ont autorisé l'intégration de leurs données accessibles au public, ou directement fourni leurs données à des fins de publication sous la forme de données ouvertes. Nous remercions ces organisations de leur soutien et de leur contribution. Ces données ont été soutenues par Logement, infrastructures et collectivités Canada (LICC).

Comment citer cet ensemble de données

Statistique Canada. (2025). La Base de données sur les réseaux de transport en commun du Canada. <https://www150.statcan.gc.ca/n1/pub/23-26-0003/232600032025001-fra.htm>.

Communiquez avec nous

Pour toute question, correction ou omission, veuillez communiquer avec nous à l'adresse statcan.lode-ecdo@statcan.gc.ca. Veuillez inclure le titre de la base de données ouverte dans l'objet du courriel.