MTA Bus Route Segment Speeds

Overview

General Description

The Metropolitan Transportation Authority (MTA) is a public-benefit corporation responsible for public transportation in the state of New York serving 12 counties in southeastern New York, along with two counties in southwestern Connecticut under contract to the Connecticut Department of Transportation (CDOT). The MTA is the largest transportation network in North America.

Bus service within New York City is operated by MTA agencies New York City Transit (NYCT) and MTA Bus Company (MTABC).

This dataset provides data on how fast buses are traveling between pairs of subsequent timepoints (the major stops on a bus route) for every bus route in the system. It will provide the average speed in miles per hour between pairs of timepoints, the average travel time in minutes, the road distance in miles, and the number of bus trips for each bus route aggregated by month, day of the week, and hour of day. This allows for the calculation of weighted averages across multiple routes. The dataset also provides the type of bus trip (Local, Limited, SBS, Express, School), the borough of the bus route, the names of the timepoints, their coordinates, and the stop sequence order of the first two timepoints in a pair.

While timepoints may generally be the major bus stops on a route, this is not always the case, with the selection of timepoints also determined by factors internal to the process of writing bus schedules. Timepoints are the only stops on bus routes for which NYCT formally schedules arrival times. Times at other locations are approximations determined by interpolation. The selection of timepoints for each bus route can change from schedule to schedule. This dataset calculates bus speeds between timepoints as opposed to between all bus stops to keep the data to a workable size, as many routes can have well over 50 stops.

This dataset was published during the first phase of the MTA's commitment to increasing transparency. We continually examine all our published and publishable data with a view to both providing datasets that can be effectively utilized by our customers and the public at large, and to providing regular, automated updates to these datasets efficiently and sustainably. Consequently, this dataset may be restructured and/or combined with other similar datasets in the future.

Data Collection Methodology

The GPS system on buses is used to determine when each bus is at its major stops along a route. The MTA's route-matching algorithm (internally known as "Bus Matching 2.0 (BM2)" is used to match pings to a bus route.

BM2 uses GPS pings from BusTime, which consist of location data, timestamp, bus ID and inferred route, and turns them into a clean table with arrival times for each bus at each stop on the correct route. It does so by using modern geospatial analytics to identify stretches of pings that are geospatially similar to route shapes in General Transit Feed Specification, an open standard used to display public transit schedule data. Stops are inferred along the trips by identifying when the bus first crossed a radius around each stop. BM2 properly accounts for buses with the wrong route or destination given, detours, deadheads along route, layovers upstream of terminals, etc., and removes impossible pings and duplicates. Non-revenue trips in the data are removed. BM2 is a newly developed method for tracking performance on our bus fleet.

Bus stop arrival times are then combined with data for the road distance between each stop to calculate how fast a bus travels between timepoint stops.

Statistical and Analytic Issues

The speed calculations include everything a bus will encounter on its route, including the time the bus dwells at stops to pick up passengers, traffic lights, instances where buses change operators mid-route (reliefs), road closures, delivery vehicles, and traffic slowdowns, all of which are important for representing the true travel time of our riders.

Data users should be mindful that bus speeds can vary significantly if there are holidays, and that bus speeds are different for different holidays (since some employers allow their workers to take different days off). These days include MLK Day, Presidents Day, Columbus Day, Veterans Day, Thanksgiving, and Jewish holidays like Yom Kippur. Data is presented as it was collected by the MTA's GPS network: users should be mindful that these represent coarse estimates of bus speeds, which due to variety of factors (sensor misalignment, sensor interference, etc.) may differ slightly from true bus speeds.

This dataset includes shuttle bus routes like the Bx90 and the Q107.

Limitations of Data Use

There are no limitations on the data at this time.

Release Notes

Version 1.0.0 initial release (08/28/2024)