MTA Railroad Ticket Sales

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 you are MTA is the largest transportation network in North America.

The Long Island Rail Road (LIRR) is the busiest commuter railroad in North America, serving customers from Manhattan to the eastern tip of Suffolk County on Long Island.

Metro-North Railroad (MNR) is the second-busiest commuter railroad in North America, connecting Manhattan with the Bronx, the Hudson Valley, and Connecticut. Service east of the Hudson River consists of the Hudson, Harlem, and New Haven lines, which originate at Grand Central Terminal. Service west of the Hudson River consists of the Port Jervis and Pascack Valley lines, which originate at Hoboken Terminal in New Jersey, and are operated under contract with New Jersey Transit.

This dataset presents daily ticket sales for the Long Island Rail Road and Metro-North Railroad, aggregated by sales channel, ticket and fare attributes, and origin, destination, via, and device station details. Each row represents the total number of tickets sold for a unique combination of these dimensions on a given day.

This dataset was published through MTA's Open Data Program, which is committed to increasing transparency through high-quality open data and accompanying resources. 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

This dataset aggregates ticket sales records from four source systems that collectively cover all paper and mobile ticket sales for LIRR and MNR. FareGo captures paper ticket sales for both LIRR and MNR from newer ticket vending machines located in the railroad stations. Central Support System (CSS) is

used exclusively by MNR and captures paper ticket sales from older ticket vending machines as well as over-the-counter sales at staffed station ticket windows. Ticket Selling System (TSS) is used exclusively by LIRR and captures paper ticket sales from older ticket vending machines, over-the-counter sales at staffed ticket windows, and on-board sales. Masabi is utilized by both LIRR and MNR and captures mobile ticket purchases made through the MTA's TrainTime app. Each source provides transaction-level ticket sales data.

Records from each channel are mapped to a common set of columns and combined into a single table. Station identifiers for origin, destination, via, and device stations, which vary in format across source systems, are mapped to a unified station reference dataset to produce consistent station codes, names, zones, and branch assignments. The 'Sold As' field use a standardized set of values across all channels, providing a consistent classification of the ticket purchase type. The 'Via' (code, name, zone, and branch) fields include categorical information about the station that the ticketed passenger went through if the origin and destinations are not connected by one route. Given the nature of the types of trips that use 'Via', this is not a common field populated in the dataset. In addition, it is important to note that MNR ticketing does not use the 'Via' fields. Device station fields (code, name, zone, and branch) identify the physical location of the ticket vending or selling machine where the purchase occurred. These fields are not applicable to Masabi eTickets, which are purchased via mobile device. Other fields, including 'Sale Packet' and 'Ticket Type', retain the values provided by each source system and may differ in naming conventions across channels.

Records are aggregated daily by railroad, sales channel, sales packet, article number, ticket type, sold as type, fare type, and the full set of station attributes (origin, destination, via, and device). The measure column 'Number of Tickets' reports the count of tickets sold for each unique combination of these dimensions on a given date. For all non-commutation tickets, the 'Number of Tickets' sold field already incorporates the ridership impact for those ticket types, e.g., a 'Number of Tickets' value of 70 in a row for a Ten Trip ticket implies that seven physical tickets were sold valid for a total of 70 rides.

Statistical and Analytic Issues

Data currently is not available for West of Hudson service since it is operated under contract with New Jersey Transit.

Limitations of Data Use

There are no limitations on the data at this time.

Release Notes

Version 1.0.0 initial release (07/23/2026)