GTFS Studio User Guide

Background

GTFS Studio is designed to help manage, view, query and extract GTFS and GTFS-R data with a simple to use interface. GTFS data can be a difficult data source to understand and work with, especially for users new to GTFS. This system was built to make GTFS data easier to read and understand for all data consumers.

For more information about GTFS – General Transit Feed Specification you can read up via https://gtfs.org/.

GTFS

The GTFS data in GTFS Studio is from the TfNSW Open Data API endpoints. Checks for new or updated data bundles are made every 30 minutes to ensure the most recent data is available via GTFS Studio.

Data Structure

The CSV files inside the GTFS zip bundle are extracted and loaded into a relational database. The extracted GTFS tables are:

Table Name	Description	Related To
Agency	Details about the agency or brand supplying the GTFS feed.	Feed Info
Calendar	Identifying a set of dates that a service is available.	Feed Info
Calendar Dates	Used in conjunction with Calendar to activate or disable a service by date.	Calendar, Feed Info
Feed Info	Details about the GTFS data bundle.	All tables relate back to a specific feed.
Route	Specific routes for an agency.	Agency, Feed Info
Shape Path	The path of a particular trip based on the path between geospatial coordinates.	Trip, Feed Info

Table Name	Description	Related To
Stop Time	Details about each stop on a trip including arrival and departure times.	Trip, Note, Feed Info
Stop	Details about particular stops including geospatial coordinates.	Feed Info
Trip	Details about unique trips in a route.	Route, Calendar, Shape Path, Note, Feed Info

View only tables (not included in exported bundles)

Table Name	Description	Related To
Level	Describe the different levels of a station.	Feed Info
Pathway	A graph representation to describe a subway or a train, with nodes (the locations) and edges (the pathways).	Stop(from_stop), Stop(to_stop), Feed Info
Vehicle Category(Extension)	Describes the vehicles themselves.	Feed Info
Vehicle Coupling(Extension)	Describes the arrangement of vehicles in composed vehicles like trains.	Vehicle Category(parent), Vehicle Category(child), Feed Info
Vehicle Boarding(Extension)	Describes where the vehicle stops on a platform.	Vehicle Category, Stop(boarding area), Feed Info
Note	Provides descriptive information about trips and stops.	Feed Info

Page Navigation

Any of the tables in GTFS Studio can be viewed as a paginated list by clicking on the table name on the navigation bar, e.g. Feed Information, Agency, and Trip.

The pagination buttons are found at the bottom of the list. Number of items per page can be adjusted to match the user's preference.

1	1206	190		Randwick to Circular Quay	lightrail/cbdandsoutheast
1	18666	RTTA_REV			sydneytrains
0	36000	RTTA_REV	10		sydneytrains
		Rows per page:	20	1-20 of 721864 1 2 ... 36094 NEXT >	
			25		
			30		
			50		
			100		

Querying the Data

In the table view, there is a search box that allows the user to search keywords across particular fields as seen in the image below.

	Search Wynyard	Filter by FeedInfo						
	Trip Id	**Route Id**	**Service Id**	**Trip Headsign**	**Direction Id**	**Block Id**	**Shape Id**	**Trip No**
☐	1353390	2440_165X	37	City Wynyard	1		101159	
☐	1400337	2440_B1	1	City Wynyard	1		106207	
☐	1296268	2440_B1	1	Mona Vale	0		94342	
☐	1296188	2440_B1	37	Mona Vale	0		94342	
☐	1422093	2440_168X	47	City Wynyard	1		107578	
☐	1294537	2440_190X	37	Avalon	0		97384	

The specific query fields for each table are given below:

Table Name	Query Fields
Agency	Agency name, Agency ID
Calendar	Service ID
Calendar Dates	Date
Feed Info	Feed publisher name
Route	Route long name, Route short name, Route ID
Shape Path	Shape ID
Stop Times	Trip ID, Stop ID, Stop name, Trip headsign
Stops	Stop Name, Stop ID

Table Name	Query Fields
Trip	Trip ID, Trip short name, Trip long name, Trip headsign, Route ID, Shape ID

All tables can be filtered by transport type, e.g. "metro", as seen in the image below. The transport type can be found in Feed Information page, under Transport Type column, or other pages, under FeedInfo Name column.

Map Projections

Trip, Stop and Shape Path view provides coordinates projections on an interactive map.

Shape Id

94703

Line
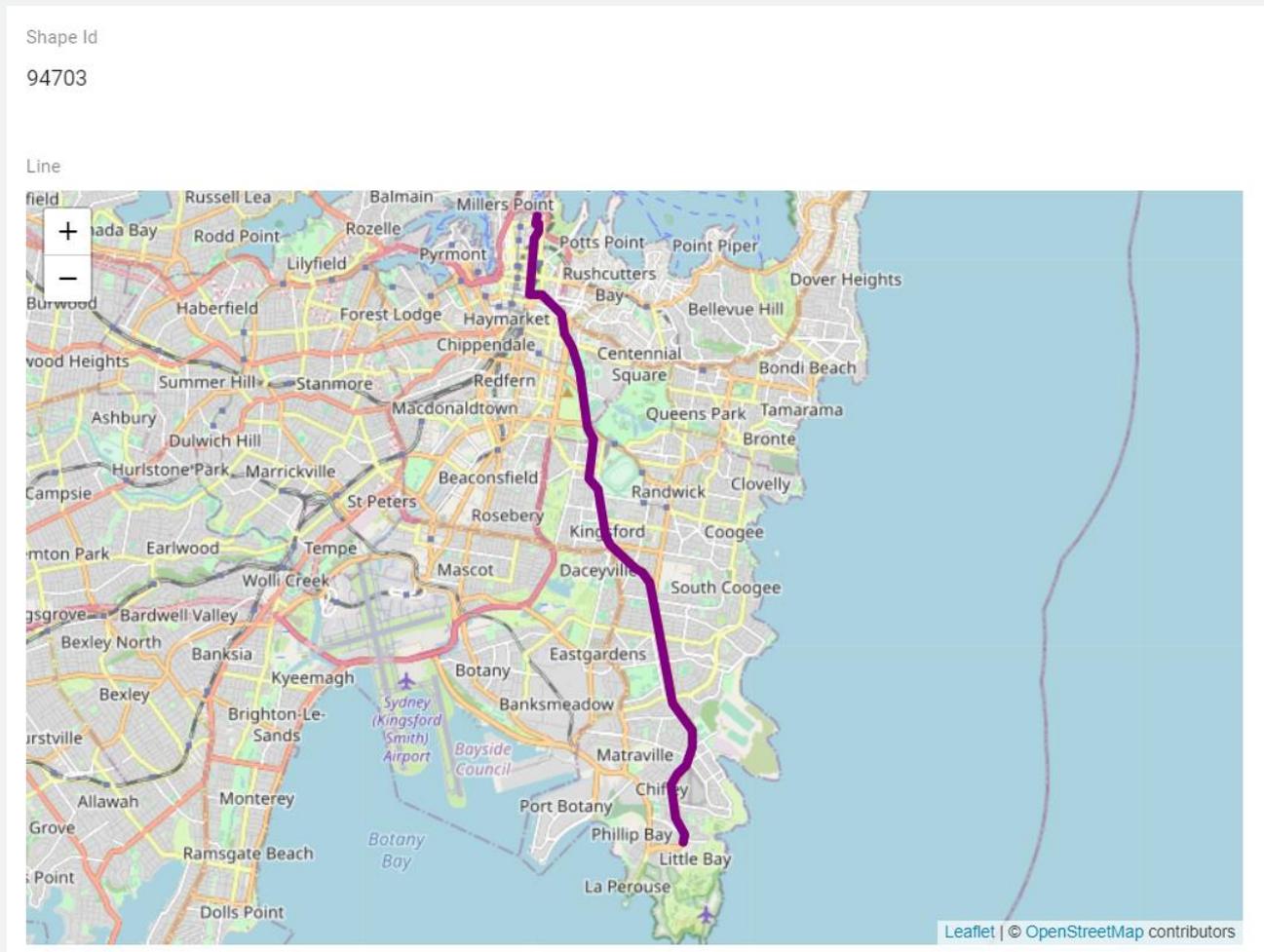
Exporting Data Subsets

GTFS Studio allows users to query the data and select a smaller subset for export as a GTFS bundle. Each user has the ability to generate <u>one</u> custom data subset. The user needs to delete the generated custom data subset if the user wishes to generate another data subset.

A data subset bundle contains all CSV files from the original data bundle but each data file is filtered by the user's selection. The bundling process determines all trips that relate to the selection and then extracts information from all the other tables that are related to those selected trips providing the user with a generated custom data subset.

Exporting data subsets can take place on any of the table lists. The steps to create custom data subset are as follows:

1. Choose a data table and either provide a search query or select a specific page to refine the search query.

2. Use the checkboxes to select the records to include in the bundle.

3. From the selected items toolbar click "EXPORT" to generate a bundle.

4. The request for a data subset will be added to the queue and the screen below may appear while the task is waiting to start.

5. Once the bundle has been generated, details of data collected and link for download will be displayed.

- The user can leave the page and come back later by selecting "Your Custom Dataset" at the bottom of navigation bar. Once the data bundle is ready, the user can click on the link to download the generated data subset.

6. If the user wants to build a different data bundle, the user will need to return to this page and click the "DELETE" button to remove the current bundle.

GTFS Realtime

GTFS Realtime (GTFS-R) data is collected from the TfNSW Open Data API endpoints. This is a periodic process that constantly updates realtime data every minute. The most recent realtime data will be available as well as a period of historic data, up to 30 days.

Data Structure

Data from the realtime stream is collated and converted into a number of data tables. These tables are:

Table Name	Description
Vehicle Positions	Realtime updates of vehicle positions
Trip Updates	Realtime updates of the trip delay compared with scheduled time.

Filtering Realtime Data

The Vehicle Positions and Trip Updates realtime data tables can be filtered for particular agencies, trips, routes, services and shape ids as well as a specific time period. The user can add many filters and the export function will search for any trip ids that match the filters.

1. Pick the data, e.g. trip updates, vehicle positions or both

2. Select time range to filter GTFS-R feeds

3. Use search fields to filter trips

4. Click Export CSV to generate the bundle. Exporting to CSV works the same way as exporting GTFS data. The task will be added to a queue and once it is completed, the user will be able to click the link to download the generated data subset.

5. If the user want to build a different data bundle, the user will need to return to this page and click the "DELETE" button to remove the current bundle.