**GTFSed Documentation**

**Written By:**

Hubert Yoo   
May 15, 2020 -   
[ansib1e@github.io](mailto:ansib1e@github.io)  
[ansib1e@protonmail.com](mailto:ansib1e@protonmail.com)

**Quick Overview**

**GTFS (General Transit Feed Specification)**  
 The General Transit Feed Specification (GTFS) defines a common format for public transportation schedules and associated geographic information (Wikipedia)

**GTFS Feed/Dataset** (Folder): Contains a standardized collection of comma delimited text files (CSV) that hold public transit scheduling data. The GTFS format is an internationally known standard for:

* **Journey planning**: supplies the dataset on public transit for use in multi-modal journey planner applications. When combined with a detailed model of the street/pedestrian network to allow routing to take place from point to point rather than just between spots. ArcMap Network Analyst extension incorporates GTFS for transit routing.
* **Accessibility research**: transit accessibility; estimating travel times by transit from one facility to many other points at specific times of the day.
* **Comparing service levels**: GTFS can be used to compare accessibility due to changes in transit service provision, facility location, future extensions, etc.

**Requirements:**  
The requirements bellow apply to formatting and contents of the dataset files:

* Collection of at least 6 CSV files, up to 13 files
* First line of each file must contain field names
* Preferred character encoding: **UTF-8**
* Field values may not contain tabs, carriage returns or new lines.
* Remove all spaces between fields or field names
* Each line must end with a **CRLF** or **LF** line break character
* All dataset files must be zipped together

**Field Specifications:**

* **Dates**: Service day in the YYYMMDD format. Since time within a service day can be above 24:00:00, a service day often contains information for the subsequent day(s).

*Example:* 20180913 *for September 13th, 2018.*

* **Email** - An email address.

*Example:* [***example@example.com***](mailto:example@example.com)

* **ID** - An ID field value is an internal ID, not intended to be shown to riders, and is a sequence of any UTF-8 characters. Using only printable ASCII characters is recommended. IDs defined in one .txt file are often referenced in another .txt file.

*Example: The* **stop\_id** *field in stops.txt is a ID. The* **stop\_id** *field in stop\_times.txt is an ID referencing* **stops.stop\_id**.

* **Latitude** - WGS84 latitude in decimal degrees. The value must be greater than or equal to -90.0 and less than or equal to 90.0.

*Example:* ***41.890169*** *for the Colosseum in Rome*

* **Longitude** - WGS84 longitude in decimal degrees. The value must be greater than or equal to -180.0 and less than or equal to 180.0.

*Example:* ***12.492269*** *for the Colosseum in Rome.*

* **Time** - Time in the HH:MM:SS format (H:MM:SS is also accepted). The time is measured from "noon minus 12h" of the service day (effectively midnight except for days on which daylight savings time changes occur. For more information, see the guidelines article). For times occurring after midnight, enter the time as a value greater than 24:00:00 in HH:MM:SS local time for the day on which the trip schedule begins.

*Example:* ***14:30:00*** *for 2:30PM or* ***25:35:00*** *for 1:35AM on the next day.*

* **Timezone** - TZ timezone from the https://www.iana.org/time-zones. Timezone names never contain the space character but may contain an underscore. Refer to http://en.wikipedia.org/wiki/List\_of\_tz\_zones for a list of valid values.

*Example*: **Asia/Tokyo**, **America/Los\_Angeles** *or* **Africa/Cairo**.

* **URL** - A fully qualified URL that includes http:// or https://, and any special characters in the URL must be correctly escaped. See the following http://www.w3.org/Addressing/URL/4\_URI\_Recommentations.html for a description of how to create fully qualified URL values.

**Required Dataset Files**

|  |  |
| --- | --- |
| Filename | Defines |
| agency.txt | Transit agencies with service represented in this dataset. |
| stops.txt | Stops where vehicles pick up or drop off riders. Also defines stations and station entrances. |
| routes.txt | Transit routes. A route is a group of trips that are displayed to riders as a single service |
| trips.txt | Trips for each route. A trip is a sequence of two or more stops that occur during a specific time period. |
| stop\_times.txt | Times that a vehicle arrives at and departs from stops for each trip. |
| calendar.txt | Service dates specified using a weekly schedule with start and end dates. **This file is required unless all dates of service are defined in calendar\_dates.txt** |
| calendar\_dates.txt | Exceptions for the services defined in the calendar.txt. I**f calendar.txt is omitted, then calendar\_dates.txt is required and must contain all dates of service.** |

**Optional Dataset Files**

|  |  |
| --- | --- |
| Filename | Defines |
| fare\_attributes.txt | Fare information for a transit agency’s routes. |
| fare\_rules.txt | Rules to apply fares for itineraries |
| shapes.txt | Rules for mapping vehicle travel paths, sometimes referred to as route alignments |
| frequencies.txt | Headway (time between trips) for headway-based service or a compressed representation of fixed-schedule service |
| transfers.txt | Rules for making connections at transfer points between routes. |
| pathways.txt | Pathways linking together locations within stations. |
| levels.txt | Levels within stations. |
| translations.txt | Translations of customer-facing dataset values |
| feed\_info.txt | Dataset metadata, including publisher, version, and expiration information |
| attributions.txt | Dataset attributions |

**Field Definitions / Required Tables (Required)**

**agency.txt**

The agency table provides information about the transit agency as such, including name, website and contact information.

|  |  |  |  |
| --- | --- | --- | --- |
| **Field Name** | **Type** | **Required** | **Description** |
| agency\_id | ID | **Conditionally Required** | Identifies a transit brand which is often synonymous with a transit agency. Note that in some cases, such as when a single agency operates multiple separate services, agencies and brands are distinct. This document uses the term "agency" in place of "brand". A dataset may contain data from multiple agencies. This field is required when the dataset contains data for multiple transit agencies, otherwise it is optional. |
| agency\_name | Text | **Required** | Full name of the transit agency |
| agency\_url | URL | **Required** | URL of the transit agency |
| agency\_timezone | Timezone | **Required** | Timezone where the transit agency is located. If multiple agencies are specified in the dataset, each must have the **same** agency\_timezone |

**routes.txt**

The routes table identifies distinct routes. This is to be distinguished from distinct routings (or paths), several of which may belong to a single route.

|  |  |  |  |
| --- | --- | --- | --- |
| **Field Name** | **Type** | **Required** | **Description** |
| route\_id (primary key) | ID referencing stops.stop\_id | **Required** | Identifies the serviced stop. All stops serviced during a trip must have a record in stop\_times.txt. Referenced locations must be stops, not stations or station entrances. A stop may be serviced multiple times in the same trip, and multiple trips and routes may service the same stop. |
| route\_short\_name | Text | **Required** | Short name of a route |
| route\_long\_name | Text | **Conditional Required** | Full name of a route |
| route\_type | Enum | **Required** | Indicates the type of transportation used on a route.  Valid options are:  **0** - Tram, Streetcar, Light rail. Any light rail or street level system within a metropolitan area.  **1** - Subway, Metro. Any underground rail system within a metropolitan area.  **2** - Rail. Used for intercity or long-distance travel.  **3** - Bus. Used for short- and long-distance bus routes.  **4** - Ferry. Used for short- and long-distance boat service.  **5** - Cable tram. Used for street-level rail cars where the cable runs beneath the vehicle, e.g., cable car in San Francisco.  **6** - Aerial lift, suspended cable car (e.g., gondola lift, aerial tramway). Cable transport where cabins, cars, gondolas or open chairs are suspended by means of one or more cables.  **7** - Funicular. Any rail system designed for steep inclines.  **11** - Trolleybus. Electric buses that draw power from overhead wires using poles.  **12** - Monorail. Railway in which the track consists of a single rail or a beam. |

**trips.txt: (Constant)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Field Name** | **Type** | **Required** | **Description** |
| trip\_id (primary key\_ | ID referencing stops.stop\_id | **Required** | Identifies the serviced stop. All stops serviced during a trip must have a record in stop\_times.txt. Referenced locations must be stops, not stations or station entrances. A stop may be serviced multiple times in the same trip, and multiple trips and routes may service the same stop. |
| route\_id (foreign key) | ID referencing routes.rout\_id | **Required** | Identifies a route. |
| service\_id (foreign key) | ID referencing calendar.service\_id or calendar\_dates.service\_id | **Required** | Identifies a set of dates when a service is available for one or more routes. |

**stops.txt**

The stops table defines the geographic locations of each and every actual stop or station in the transit system as well ass, and optionally, some of the amenities associated with those stops.

|  |  |  |  |
| --- | --- | --- | --- |
| **Field Name** | **Type** | **Required** | **Description** |
| stop\_id (primary key) | ID referencing stops.stop\_id | **Required** | Identifies the serviced stop. All stops serviced during a trip must have a record in stop\_times.txt. Referenced locations must be stops, not stations or station entrances. A stop may be serviced multiple times in the same trip, and multiple trips and routes may service the same stop. |
| stop\_name | Text | **Required** | Name of the location. Use a name that people will understand in the local and tourist vernacular.  Required for locations which are stops (location\_type=0), stations (location\_type=1) or entrances/exits (location\_type=2).  • Optional for locations which are generic nodes (location\_type=3) or boarding areas (location\_type=4). |
| stop\_lon | Longitude | **Required** | Longitude of the location. |
| stop\_lat | Latitude | **Required** | Latitude of the location. |

**stop\_times.txt**

|  |  |  |  |
| --- | --- | --- | --- |
| **Field Name** | **Type** | **Required** | **Description** |
| stop\_id (primary key) | ID referencing stops.stop\_id | **Required** | Identifies the serviced stop. All stops serviced during a trip must have a record in stop\_times.txt. Referenced locations must be stops, not stations or station entrances. A stop may be serviced multiple times in the same trip, and multiple trips and routes may service the same stop. |
| trip\_id (foreign key) | ID referencing trips.trip\_id | **Required** | Identifies a trip. |
| arrival\_time | Time | **Required** | Arrival time at a specific stop for a specific trip on a route. If there are not separate times for arrival and departure at a stop, enter the same value for arrival\_time and departure\_time. For times occurring after midnight on the service day, enter the time as a value greater than 24:00:00 in HH:MM:SS local time for the day on which the trip schedule begins. |
| departure\_time | Time | **Required** | Departure time from a specific stop for a specific trip on a route. For times occurring after midnight on the service day, enter the time as a value greater than 24:00:00 in HH:MM:SS local time for the day on which the trip schedule begins. If there are not separate times for arrival and departure at a stop, enter the same value for arrival\_time and departure\_time. See the arrival\_time description for more details about using timepoints correctly. |
| stop\_sequence | Non-negative integer | **Required** | Order of stops for a particular trip. The values must increase along the trip but do not need to be consecutive.  Example: The first location on the trip could have a *stop\_sequence*=*1*, the second location on the trip could have a *stop\_sequence*=*23*, the third location could have a *stop\_sequence*=*40*, and so on. |

**calendar.txt**

The calendar table defines service patterns that operate recurrently such as, for example, every weekday. Service patterns that do not repeat such as for a one-time special event will be defined in the calendar\_dates table.

|  |  |  |  |
| --- | --- | --- | --- |
| **Field Name** | **Type** | **Required** | **Description** |
| service\_id | ID | **Required** | Uniquely identifies a set of dates when service is available for one or more routes. Each service\_id value can appear at most once in a calendar.txt file. |
| monday | Enum | **Required** | Indicates whether the service operates on all Mondays in the date range specified by the start\_date and end\_date fields.  1 – service is available for all Mondays in the range  0 – Service is not available for Mondays in the date range |
| tuesday | Enum | **Required** | Function is the same was as monday except applies to Tuesday |
| wednesday | Enum | **Required** | Function is the same was as monday except applies to Wednesday |
| thursday | Enum | **Required** | Function is the same was as monday except applies to Thursday |
| friday | Enum | **Required** | Function is the same was as monday except applies to Friday |
| saturday | Enum | **Required** | Function is the same was as monday except applies to Saturday |
| sunday | Enum | **Required** | Function is the same was as monday except applies to Sunday |
| start\_date | Date | **Required** | Start service day for the service interval. |
| end\_date | Date | **Required** | End service day for the service interval. This service day is included in the interval. |

**References**

1. W. (2019, December 16). General Transit Feed Specification. Retrieved May 15, 2020, from https://en.wikipedia.org/wiki/General\_Transit\_Feed\_Specification
2. Reference &nbsp;|&nbsp; Static Transit &nbsp;|&nbsp; Google Developers. (n.d.). Retrieved May 15, 2020, from https://developers.google.com/transit/gtfs/reference/
3. Using GTFS Data in ArcGIS. (n.d.). Retrieved May 15, 2020, from http://esri.github.io/public-transit-tools/
4. Google, T. M. (n.d.). Google/transit. Retrieved May 15, 2020, from https://github.com/google/transit/blob/master/gtfs/spec/en/reference.md

**Hyperlinks**

<https://en.wikipedia.org/wiki/General_Transit_Feed_Specification>

<http://esri.github.io/public-transit-tools/>