



# City of Albuquerque Information Technology Services Division Data Management

## Transit Scheduled Route Metadata Requirements

### Contact Information

*Who is the contact for this dataset? The contact will be the City employee who is accountable for the data provided in this dataset and can act as front-line support in the event of any questions about the data.*

<b>Name</b>	Joe Saraphon
<b>Department/Division</b>	Transit ("ABQ RIDE")
<b>Phone</b>	505-724-3113
<b>Email</b>	saraphon@cabq.gov

### What Does this Dataset Describe?

*What is the name of this dataset? How should a user identify this dataset in any communication with contact above? Provide a shorter description of the Dataset that can act as a one-line summary of the dataset when dealing with stakeholders. Provide a longer description of the data that can be readily understood by non-technical users.*

<b>Dataset Title</b>	Transit Scheduled Route Data
<b>Short Description</b>	Bus Stops, Stop Times, Routes, etc that change about once a quarter
<b>Full Non-Technical Description</b>	
Data that changes about once a quarter and includes detailed route and schedule information - Bus Stops, Stop Times, Routes, Agency, Calendar, Calendar Dates, Shapes, Trips	

### How Should this Dataset be Cited?

*How should external sources refer to this dataset in publications or documentation? Often this will simply be the URL and the date retrieved.*

[http://data.cabq.gov/transit/gtfs/google\\_transit.zip](http://data.cabq.gov/transit/gtfs/google_transit.zip)

## Does the Dataset Reflect a Particular Time Period?

Provide any date restrictions that may affect the validity of the data. The table fields are defined as follows:

	Definition
<i>Start Date</i>	<i>Start date of the time period within which this data falls. Format: MM/DD/YYYY HH:MM:SS.</i>
<i>End Date</i>	<i>End date of the time period within which this data falls. Format: MM/DD/YYYY HH:MM:SS.</i>
<i>Dataset Refresh Interval</i>	<i>Time period between Dataset refreshes. Format: “nn [seconds/minutes/hours/days/weeks/months/years]” or the word “Static” if never refreshed.</i>
<i>Data Expiration Date</i>	<i>Date after which the data must be considered stale and no longer of sufficient utility (fit-for-purpose). Format: MM/DD/YYYY HH:MM:SS.</i>
<i>Dataset Review Date</i>	<i>Date after which this dataset will be reviewed by the City for utility (fit-for-purpose) and usage. Format: MM/DD/YYYY HH:MM:SS.</i>
<i>Comments</i>	<i>Specific comments related to any time-specific features of this dataset.</i>

<b>Start Date</b>	05/19/2012
<b>End Date</b>	12/18/2012
<b>Dataset Refresh Interval</b>	Quarterly or when data changes, usually associated with schedule changes
<b>Dataset Expiration Date</b>	N/A – ABQ RIDE does not intend to discontinue this data feed in the foreseeable future.
<b>Dataset Review Date</b>	ABQ RIDE reviews the data before the end date above to ensure that current data is available.
<b>Comments</b>	
The files are comma delimited with field names in the first row; the format is dictated by Google at	

<https://developers.google.com/transit/gtfs/reference>. Please consult that site for more details on data definitions. Abbreviated descriptions are provided below.

## Dataset Definition/Format

Provide a field-by-field breakdown and definition of each record. This section acts as the formal data dictionary for an individual record.

Field Name	Format	Description
google_transit.zip		
agency.txt		
agency_phone	xxx-xxx-xxxx	Contact phone number for the agency
agency_url	http://	The website for the agency
agency_id	numeric	The identification number for the agency
agency_name	alpha	The name of the agency
agency_timezone	Country/City	The time zone the agency is in
agency_lang	Char(2)	The language that will be used for the agency
calendar.txt		
service_id	Alphanumeric	The identification number for the service type (e.g. weekday schedules versus Saturday, etc). May include unique services that run only once, as defined in the calendar_dates.txt table. Multiple service id's may be active on the same date.
start_date	Date (yyyymmdd)	The start date of the service
end_date	Date (yyyymmdd)	The end date of the service.
monday	0/1	0 indicates that the service normally does not run on that day; 1 indicates that the service normally runs on that day
tuesday	0/1	0 indicates that the service normally does not run on that day; 1 indicates that the service normally does run on that day.
wednesday	0/1	0 indicates that the service normally does not run

		on that day; 1 indicates that the service normally does run on that day.
thursday	0/1	0 indicates that the service normally does not run on that day; 1 indicates that the service normally does run on that day.
friday	0/1	0 indicates that the service normally does not run on that day; 1 indicates that the service normally does run on that day.
saturday	0/1	0 indicates that the service normally does not run on that day; 1 indicates that the service normally does run on that day.
sunday	0/1	0 indicates that the service normally does not run on that day; 1 indicates that the service normally does run on that day.
calendar_date.txt		
service_id	Alphanumeric	The identification number for the service.
date	Date (yyyymmdd)	Date that the service exception occurs.
exception_type	Numeric	2 indicates that service does not operate that date; 1 indicates that service does operate that date.
route.txt		
route_long_name	Alphanumeric	The long name of the bus route
route_type	Numeric	A number indicating whether the route is a bus route, light rail, commuter train, etc. All ABQ RIDE routes are bus routes, corresponding to route_type 3.
route_text_color	Alphanumeric	The color the route text will be displayed in on the map
route_color	Alphanumeric	The color the route will be displayed in on the map
agency_id	Numeric	The identification number for the agency
route_id	Numeric	The identification number of the bus route
route_url	Nulls	Not currently used
route_desc	Nulls	Not currently used

route_short-name	Numeric	The bus route number used to publicly identify the route.
shapes.txt		
shape_id	Numeric	The identification number of the route shape to be displayed on the map
shape_pt_lat	Alphanumeric (-99.999999)	The latitude where the shape should be displayed
shape_pt_lon	Alphanumeric (-99.999999)	The longitude where the shape should be displayed
shape_pt_sequence	Numeric	The sequence the shapes should be displayed in on the map
shape_dist_traveled	Numeric (9.9999)	The distance between the first shape and the current shape
stop_times.txt		
trip_id	Numeric	The identification number of the trip
arrival_time	hh:mm:ss	The time that the bus is scheduled to arrive at the stop
departure_time	hh:mm:ss	The time that the bus is scheduled to depart from the stop
stop_id	Numeric	The internal system identification number of the bus stop
stop_sequence	Numeric	The sequence number of the bus stop on the route for that trip.
stop_headsign	Nulls	Not currently used
pickup_type	Nulls	Not currently used
drop_off_type	Nulls	Not currently used
shape_dist_traveled	Numeric (9.9999)	The distance between the first shape and the current shape where the stop occurs.
stops.txt		
stop_lat	Alphanumeric (-99.999999)	The latitude where the stop is located

stop_code	Numeric	The bus stop number ABQ Ride uses to identify the stop to the public.
stop_lon	Alphanumeric (-99.999999)	The longitude where the stop is located
stop_id	Numeric	The internal system identification number of the bus stop
stop_url	Nulls	Not currently used
parent_station	Nulls	stop_id of the parent station (e.g. the one point that would be displayed if displaying all the individual stops would create a cluttered map, like at a major transfer center.)
stop_desc	Alphanumeric	The description of the stop to be displayed, often a nearby address. ABQ RIDE does not currently use this system-generated data.
stop_name	Alphanumeric	The name of the stop to be displayed
location_type	Numeric	Indicates whether a stop is a parent station or not. 1 indicates a parent station; 0 for all others.
zone_id	Nulls	Not currently used
trips.txt		
block_id	Alphanumeric	Identifies the block a trip is associated with; a block is a series of trips performed by the same bus.
route_id	Numeric	The identification number of the bus route
direction_id	0/1	Indicates the (relative) direction of travel (e.g. eastbound versus westbound or inbound versus outbound). There is no pattern to which direction is 0 or 1.
trip_headsign	Nulls	Not currently used
shape_id	Numeric	The identification number of the shape to be displayed on the map
service_id	Alphanumeric	The identification number for the service that the trip occurs in.
trip_id	Numeric	The identification number of the trip

## **Dataset Technical Description**

*Provide a technical description of the dataset. This should be a complete technical description aimed at developers and expert users who need to understand the scope, strengths and limitations of the dataset.*

The google\_transit.zip folder contains the following files in a zipped format:

File name	File description
agency.txt	The agency file contains all of the companies, government entities and other establishments that have transit service described in the other files.
calendar.txt	The calendar file provides the overall dates that the other files apply to.
calendar_dates.txt	The calendar dates file provides the dates that a bus or entire service is on an exception schedule.
routes.txt	The routes file provides information about how the route should be displayed on the map.
shapes.txt	The shapes file provides the location where a specified route shape should be displayed on the map.
stop_times.txt	The stop times file provides the details of when each trip that a bus takes on a given route will be at each stop.
stops.txt	The stops file provides details about each stop.
trips.txt	The trips file provides the information to connect the stop times, shapes, trips, route and service together.

## **Dataset Assumptions**

*What technical and business assumptions are implied in the creation of this dataset? Examples could include the way in which a salary figure was calculated or data that was omitted for a specific reason.*

The data in the google\_transit.zip folder is updated as needed, generally on a quarterly basis. The format and content of the files are dictated by Google at <https://developers.google.com/transit/gtfs/reference>.

In general, the data are generated by ABQ RIDE's scheduling software. Scheduled times for many stops are estimated based on assuming buses travel a constant speed between scheduled "timepoints." ABQ RIDE's printed schedules (available here: <http://data.cabq.gov/transit/Maps/>) show only the timepoints.

## **Who Produced the Dataset?**

*Which department in the City produced this dataset? Note that this might not always be the data owner. An example of this could be a dataset that ITSD produced on behalf of EHD who owned the*

*data.*

The Transit Department, Planning Division is responsible for establishing the bus schedules, routes and trips. They use the Trapeze software application to generate the data in the Google-dictated format. Transit IT is responsible for maintaining the data on data.cabq.gov.

### **Who Owns the Data?**

*Where did this data originate? Who owns the data used in this dataset? Note that this might not always be the dataset producer. An example of this could be a dataset that ITSD produced on behalf of EHD who owned the data.*

The Transit Department is the data owner.

### **Why was the Dataset Created?**

*All datasets should have an explicit reason for existence and should, somehow, have value to someone. What is the perceived value that this dataset will bring?*

The dataset was originally created to allow Google to map the bus route information on Google Earth, Google Transit, and other map applications.

### **How was the Dataset Created?**

*How was this dataset produced? Was it a manual process? An automated process? What were the main IT systems involved in producing this dataset?*

The dataset is generated by the Trapeze Transit software suite used by the Transit department. The process is mostly but not entirely automated.

### **What Similar or Related Data Should the User be Aware of?**

*Are there any other datasets available that may contain related or similar information? Might there be situations in which these other datasets might be a better alternative?*

Following a different Google standard (<https://developers.google.com/transit/gtfs-realtime/>) there is now data on the real-time location and direction of each bus which is updated every 12 seconds. See the <http://data.cabq.gov/transit/realtime/MetaData.pdf> for more details on how the dataset folders work together.

<http://data.cabq.gov/transit/realtime/>

### **How Reliable are the Data?**

*Are there any concerns about overall data reliability? Are there any data problems that the user*

*needs to be aware of? Are there any constraints with data accuracy? What levels of confidence with this dataset could the user reasonably assume?*

The data is very reliable.

### **How Well Have the Observations Been Checked?**

*What quality assurance steps have been performed? Sometimes, a third-party verification/audit process may also be required. If so, provide the name of the third-party who performed the verification.*

This data is currently being used by Google Earth and Google Transit and has been verified by the data owner. ABQ RIDE uses tools provided by Google to assist in quality assurance, including the feedvalidator and schedule\_viewer tools available here:

<https://developers.google.com/transit/tools>.

### **Are there Legal Restrictions on the Access or Use of the Data?**

*Are there any specific legal or compliance restrictions for this data? How might this affect the way in which end users might access and use this data?*

Please see the City's general data disclaimer here: <http://www.cabq.gov/abq-data/abq-data-disclaimer-1>

### **Legal Disclaimer**

The City's standard copyright, disclaimers and legal statements may be found at <http://www.cabq.gov/about/legal>. The City data policy governing data.cabq.gov may be found at <http://www.cabq.gov/abq-data/>.