Multimodal Data Set Clean-up for
Portland Oregon Metropolitan Region

Data Set Description and Dictionary

Transit Data

April 12, 2012

Portland State
UNIVERSITY

METRO   Oregon Department of Transportation   TRIMET See where it takes you.

Table of Contents

1   PORTAL Transit Data Description

This document provides the data description for the PORTAL Transit Data for the FHWA Test Data Set project submission. The transit data set provided comes from TriMet, the Portland-metro area transit agency and includes both bus and light rail data. The data includes schedule, stop event and passenger counts data for both bus and light rail and is highly detailed. The data is provided for all transit stops in the neighborhood of the I-205 corridor. Figure 1 shows a map of all transit stops in the data set. Data collection period is Sept 15, 2011 through Nov 15, 2011.

1.1   Stops

A table providing the stop name and geo-location information for all transit stops in the data set is provided.

1.2   Bus Data and Tables

Bus data includes a schedule table and a table with stop event and automatic passenger counter (APC) data. The scheduled stop time table provides schedule information for the two-month time window. To facilitate ease of use, the schedule table includes an entry for each route and each day. This structure does replicate information, but makes it very clear which route schedule applies to which date, particularly useful for holiday and weekend schedules.

The stop_event table provides information about bus activity including arrive and leave time from bus stops, passenger ons and off, dwell time, lift usage and other detailed data. A stop record is created whenever a vehicle drives past a bus stop that it is scheduled to serve. In this case, the bus records the time that it arrived and left the vicinity of the stop. The vicinity of the stop is currently defined as the area within a 30-meter (about 98 feet) radius of the recorded bus stop location. A stop record is also created whenever a vehicle door is opened. Due to the age of the system, some busses have malfunctioning Automatic Vehicle Location (AVL) units and thus a limited amount of stop data is missing.

Bus data is obtained from the TriMet Bus Dispatch System (BDS.) Bus data is obtained from the TriMet busses each evening and is then processed by a BDS data processing program. When the BDS unit fails to record times for one or two stops the BDS data processing programs generates stop event rows with estimated times. These times are estimated by interpolating between the prior and following recorded times on the basis of the distance between the stops. The Data Source column identifies stop event rows that are created by interpolation.

1.3   Light Rail Data and Tables

Light rail data includes several different tables with different types of information – schedule information, data from the rail control system and data from automatic passenger counters on the trains. A schedule table is provided, which is identical in format to the schedule table for the bus data. Three data files are provided for the light rail data, each giving a different perspective on the data.
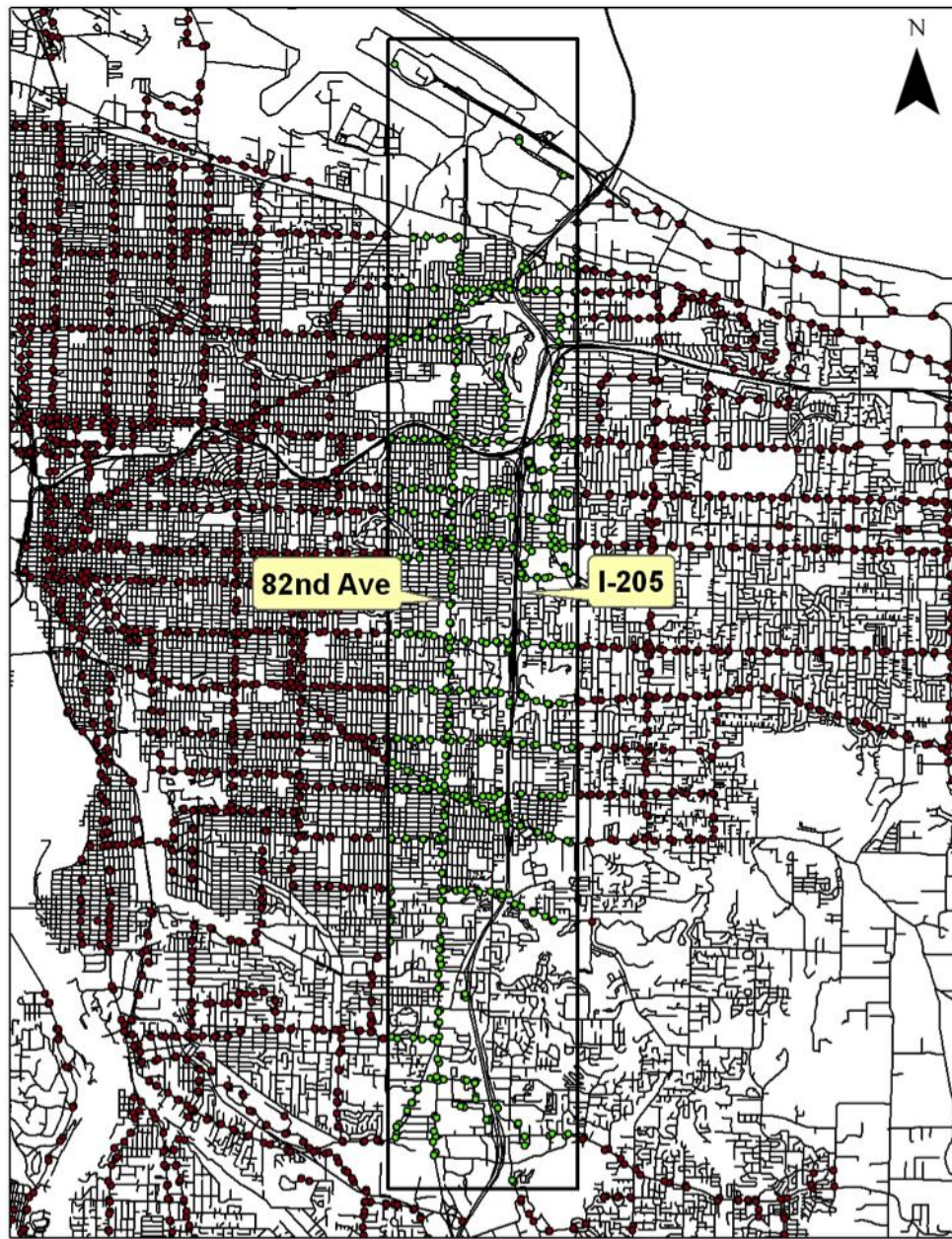
The max_stop_event table includes data from the rail control system. Loops in the rail control system are matched to the closest MAX (light rail station). The data in this table contains no schedule information, simply what was reported by the rail control system. It has the vehicle identification, route codes and times that every train was at each platform. This data is good to use if there were service disruptions, etc., because it shows all location data.

The max_stop_data_event table takes the data from the max_stop_event table and matches it to schedule; adding schedule information to the stop_event data. When trains are severely off schedule, that data may not appear in the max_stop_data_event table. The purpose of this table is to give an idea of schedule adherence and reliability. This data is quite complete, but does have some missing data due to the AIM loop data system; periodically badge numbers or car numbers are missing.

The max_stop_event_apc table includes data from the light rail automatic passenger counter system. Approximately half of the trains in the system have automatic passenger counters. This system uses GPS and motion detectors to determine stops.

1.4   GIS Files

GIS files for the TriMet bus stops and routes and light rail (MAX) lines and stops are provided. These files are part of the Oregon Metro Regional Land Information System (RLIS) data set. Additional GIS layers for the Portland Metropolitan region can be found at this link: http://www.oregonmetro.gov/rlis.

Figure 1 Transit Stops in I-205 Corridor

2 File Listing and Data Dictionaries

This document contains a listing and description of data files for the transit data and data dictionaries for the transit data.

Table 1 List of Transit Data Files

File Name	Primary Key	Description
transit_stops.csv	Location_id	Description of transit stops.
transit_bus_scheduled_stop_time.csv	calendar_date, route_number, , direction, trip_number, stop_time	Bus schedule information.
transit_bus_stop_event.csv	Service_date, vehicle_number, leave_time, arrive_time	Bus stop events. One record for each time a bus goes past a bus stop.
transit_max_scheduled_stop_time.csv	Calendar_date, route_number, direction, trip_number, stop_time	MAX (light rail) schedule information.
transit_max_stop_event.csv	Service_date, train, direction, location_id, loop_time, route_code	MAX (light rail) stop events. One record for each time a light rail train goes past a station.
transit_max_stop_data_event.csv	Service_date, direction, trip_number, location_id, leave_time	MAX (light rail) stop events combined with schedule data to demonstrate reliability.
transit_max_stop_event_apc.csv	service_date, route_number, direction, trip_number, leave_time	MAX (light rail) automatic passenger counter (APC) records.
buslines	N/A	Esri shapefile for the bus routes.
busstops	N/A	Esri shapefile for the bus stops.
lrt_line	N/A	Esri shapefile for the light rail (MAX) lines.
lrt_stop	N/A	Esri shapefile for the light rail (MAX) stops.

3 Data Dictionary

3.1 Stops

Table 2 Data Dictionary for PORTAL Transit Stops File

Attribute Name	Attribute Type	Description
location_id	integer	Numeric stop id. Correlates to location_id in data tables.
public_location_description	text	Text description of the location of the stop.
longitude	float	Longitude of the stop.
latitude	float	Latitude of the stop.
x_coordinate	float	X Coordinate of the stop location in the local State Plane Coordinate System. A broad overview of the State Plane Coordinate System can be found here: http://arcweb.sos.state.or.us/pages/rules/oars_700/oar_734/734_005.html Details can be found here: here: http://spatialreference.org/ref/epsg/2913/prettywkt/ (Units: feet)
y_coordinate	float	Y Coordinate of the stop location in the local State Plane Coordinate System. (Units: feet)

Sample

Attribute Name	Example 1	Example 2
location_id	20	26
public_location_description	2800 Block NE 92nd	NE 92nd & Hill
longitude	-122.569774	-122.570035
latitude	45.54267	45.54565
x_coordinate	7671863	7671825
y_coordinate	690736.9	691827

Table 3 Data Dictionary for PORTAL Transit Bus Scheduled Stop Time File and PORTAL Transit MAX Scheduled Stop Time File

Attribute Name	Attribute Type	Description
calendar_date	timestamp	Date on which this schedule was run.
base_service_key	character(1)	A designation for the types of service provided on different calendar dates. Common Service Keys, such as 'W', 'S', and 'U', specify regular Weekday, Saturday, and Sunday service. calendar_date is specified for all routes; service key information is thus not strictly necessary, but is provided as useful

		information.
service_key	character(1)	Detailed service key. Text description of the service key is provided in the service_key_description column.
service_key_description	text	Text description of the service key.
public_route_description	text	Route name.
public_direction_description	text	Direction of the bus route.
route_number	integer	The internal numeric designation of the Route.
direction	integer	A one digit numeric field indicating the direction of travel for the scheduled trip. Corresponds to full text direction in public_direction_description field. This field is a numeric representation of the text in public_direction_description. The numeric values used to have meaning (e.g. north, south, east, west), but with more complex route patterns, such simple meanings are no longer adequate. Thus the field is short hand for public_direction_description
trip_number	integer	A number that provides the most specific identification of a scheduled trip.
trip_begin_date	timestamp	Starting date for this particular schedule.
stop_time	integer	The scheduled or estimated time for the departure from a particular stop location on a particular trip. When the stop location is a timepoint the Stop Time is the scheduled time. At other stops the Stop Time is estimated by linear interpolation using the distances from the stop to the prior and next timepoint. The time is expressed in seconds after midnight, from the start of the service day. The values typically range from about 18,000 for 5:00 AM, through 86,400 for 12:00 PM, and up to roughly 97,200 for 3:00 AM of the following calendar day.
stop_distance	float	Distance from beginning of trip (feet).
schedule_status	integer	The Schedule Status identifies the relationship between the recorded stop information and the schedule of service. (See Table of Schedule Status Descriptionsbelow)
location_id	integer	The Location ID, an integer number that uniquely identifies a stop location. Correlates to location_id in Transit Stops file (transit_stops.csv).
pattern_id	integer	Routes have variations; for example every other trip route 70 goes down Tacoma or Bybee. pattern_id can be used to distinguish

between different trip patterns.

Sample

Attribute Name	Example 1	Example 2
calendar_date	15SEP2011:00:00:00	15SEP2011:00:00:00
base_service_key	W	W
service_key	A	A
service_key_description	Weekday MAX	Weekday MAX
public_route_description	MAX Red Line	MAX Red Line
public_direction_description	To Portland International Airport	To Portland International Airport
route_number	90	90
direction	0	0
trip_number	1010	1020
trip_begin_date	04SEP2011:00:00:00	04SEP2011:00:00:00
stop_time	15900	18120
stop_distance	84127.708	104560.018
schedule_status	4	4
location_id	8346	10572
pattern_id	36	46

Schedule Status Definitions

Value	Description
0	Off-Trip Stop: Vehicle on an unidentified trip or Deadhead.
1	Secondary Stop: Vehicle on a scheduled trip but stop appears to provide secondary service for a location where there is another recorded stop that provides a better match to the scheduled stop.
2	Primary Stop: Vehicle on a scheduled trip and stop provides the primary service at the indicated stop location. When multiple vehicle stops are recorded near a scheduled stop location the primary stop is assigned on the basis of being the nearest to the indicated stop location or having the nearest door opening within the close vicinity of the scheduled stop location. There can be only one primary stop for any scheduled stop.
3	Pseudo-Timepoint Stop: The stop is the Primary Stop at a location that corresponds to an unofficial (or "pseudo") timepoint.
4	Timepoint Stop: The stop is the Primary Stop at a location that corresponds to an official scheduled timepoint.
5	First Stop of Trip: The stop is the Primary Stop for the first stop and timepoint of a scheduled trip.
6	Last Stop of Trip: The stop is the Primary Stop for the last stop and timepoint of a scheduled trip. In these rows the value of the Leave Time column has been set to the value of the recorded arrive time.

3.2 Bus Stop Event Data

Table 4 Data Dictionary for PORTAL Transit Bus Stop Event File

Attribute Name	Attribute Type	Description
service_date	timestamp	The calendar date associated with the service. Typically this is the date the vehicle leaves the garage. When the vehicle is on the road at midnight, the service provided after midnight is associated with the previous day. Such late service is usually completed by 3:00 AM.
vehicle_number	integer	The Vehicle Number of the bus recording the data. This is the number that is painted on both the interior and exterior of the bus. In the data, the Vehicle Number is stored as a five-character field with leading zeros. For example, the Vehicle Number for bus 512 is represented as '00512'.
leave_time	integer	When the bus is no longer within 30 meters of the bus stop location a departure time is recorded. The departure time is recorded at all stops even if the bus doesn't stop to serve passengers. (See Figure 2) (Units: seconds post midnight)
train	integer	The Train or Block number stored as a number. Scheduled trips are blocked together into trains for assignment to vehicles.
route_number	integer	The internal numeric designation of the Route. For Example, Route 1 has the Route Number of 1 for the Greeley Line and 101 for the Vermont Line.
direction	integer	A one digit numeric field indicating the direction of travel for the scheduled trip. The field contains either the character Zero or One, where 0 specifies outbound and 1 specifies inbound. On cross-town routes 0 often specifies Northbound and 1 often specifies Southbound.
service_key	character(1)	A designation for the types of service provided on different calendar dates. Common Service Keys, such as 'W', 'S', and 'U', specify regular Weekday, Saturday, and Sunday service.
trip_number	integer	A number that provides the most specific identification of a scheduled trip.
stop_time	integer	The scheduled or estimated time for the departure from a particular stop location on a particular trip. When the stop location is a timepoint the Stop Time is the scheduled time. At other stops the Stop Time is estimated by linear interpolation using the distances from the stop to the prior and next timepoint. The time is expressed in seconds after midnight, from

		the start of the service day.
		The values typically range from about 18,000 for 5:00 AM, through 86,400 for 12:00 PM, and up to roughly 97,200 for 3:00 AM of the following calendar day. (Units: seconds past midnight)
arrive_time	integer	When the vehicle passes a stop without opening a door, this field contains the time the vehicle arrived within the vicinity of the stop. The vicinity of a stop is the area within a 30-meter (98 feet) radius of the recorded stop location. Otherwise, this field contains the time the door was first opened. The time is expressed in seconds after midnight, from the start of the service day. (See Figure 2) (Units: seconds past midnight)
dwell	integer	The number of seconds the door is open. (See Figure 2.)
location_id	integer	The Location ID, an integer number that uniquely identifies a stop location. Correlates to location_id in Transit Stops file (transit_stops.csv).
door	integer	The number of times the door was opened at the stop.
lift	integer	The number of times the vehicle lift was used at the stop.
ons	integer	The raw APC count of the number of persons boarding the vehicle.
offs	integer	The raw APC count of the number of persons leaving the vehicle. See Ons for more information.
estimated_load	integer	Estimated load when leaving the stop from adjusted APC data. This field will be zero if the APC counts were identified as invalid.
maximum_speed	integer	Maximum speed in mph since prior stop record. This value is not reliable in areas where GPS signals are reflected, such as areas with tall building and on some bridges. Such reflected signals create multi-path interference that may cause GPS units to generate false coordinates.
train_mileage	float	Cumulative distance in miles from the start of the train's recorded service. This distance is the sum of the tenth mile odometer ticks in the current and all prior stop records generated during the service day for the block.
pattern_distance	float	An estimate of the linear distance, measured in feet, from the beginning of the route's pattern to the vehicle's current location. The estimate is derived from the position of the vehicle, the distance of the vehicle from nearby stops, and the distance of those stops from the beginning of the route pattern. The measure identifies the vehicle's position in relation to

			the scheduled route rather than the distance travelled by the bus. Since the estimate is generated independently from the position of each stop, it is non-cumulative and will not necessarily increase monotonically. The measure may decrease if the vehicle backtracks along the route or in rare situations where the estimate is unstable because of unusual route geometry. The Pattern Distance column provides a convenient way to relate the recorded stop information to other route features, such as shape points, that are maintained within a route based linear referencing system.
location_distance	float		The distance between the vehicle position recorded by BDS and the location of the scheduled stop. The unit of the measure is feet and the number is stored as a floating-point value.
x_coordinate	float		This column contains the X Coordinate of the stop location in the local State Plane Coordinate System. This coordinate value is also called the Easting. The coordinate system is specified as the SPCS83-Oregon North Zone. The unit of the measure is International Feet.
y_coordinate	float		This column contains the Y Coordinate of the stop location in the local State Plane Coordinate System. This coordinate value is also called the Northing.
data_source	integer		A numeric field indicating the source of the row data with the following codes. 0 Row added during post-processing. 1 Stop, but not Location ID, recorded by vehicle. 2 Stop and Location ID recorded by vehicle.
schedule_status	integer		The Schedule Status identifies the relationship between the recorded stop information and the schedule of service. (See Table of Schedule Status Descriptions below.)

Sample

Attribute Name	Example 1	Example 2
service_date	15SEP2011:00:00:00	15SEP2011:00:00:00
vehicle_number	2927	2927
leave_time	21198	21444
train	1267	1267

|		
badge	1149	1149
route_number	4	4
direction	0	0
service_key	W	W
trip_number	1020	1020
stop_time	21181	21417
arrive_time	21188	21402
dwell	0	15
location_id	1486	13297
door	0	2
lift	0	0
ons	0	0
offs	0	1
estimated_load	3	3
maximum_speed	33	24
train_mileage	34.6	35.7
pattern_distance	26054.15129	32002.65884
location_distance	176.6057792	44.76730189
x_coordinate	7666298.96	7672220.638
y_coordinate	677239.7283	676713.8646
data_source	2	2
schedule_status	2	2
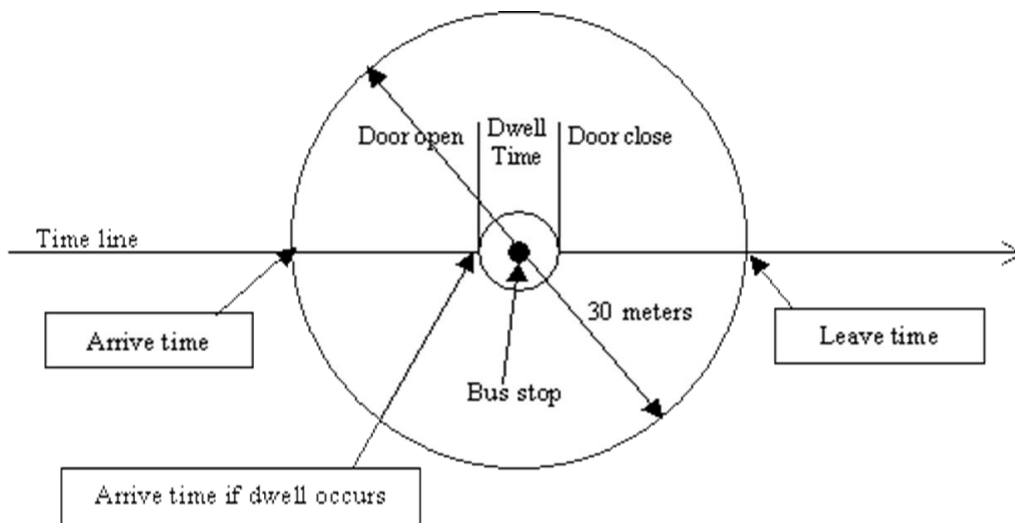
Figure 2 Diagram with representation of Arrive Time, Leave Time and Dwell Time

3.3   Light Rail Event Data

Table 5 Data Dictionary for PORTAL Transit MAX Stop Event File

Attribute Name	Attribute Type	Description
service_date	Timestamp	Date of Service

train	Integer	Scheduled Vehicle ID
direction	Smallint	Direction of travel (0 – Eastbound/Northbound) – (1 – Westbound/Southbound)
location_id	integer	MAX Station ID. Correlates to location_id in Transit Stops file (transit_stops.csv).
tratkh	text	MAX Station Name (corresponds to location_id)
loop_time	Integer	Time vehicle crossed loop. (Units: seconds past midnight.)
time_difference	Integer	Calculated offset time from loop to MAX station. The loop detectors that are the source of this stop event data are not located directly at MAX station; this field is the calculated travel time (offset) between the MAX station and the loop detector. (Units: seconds)
route_code	Integer	Numeric destination code
destination	Character (100)	Text description of destination.
car1	Integer	Actual Light Rail Vehicle (LRV) number
car2	Integer	Actual Light Rail Vehcile (LRV) number Light Rail trains are always scheduled with 2 cars, but one car can be dropped for various reasons, such as mechanical issues.

Sample

Attribute Name	Example 1	Example 2
service_date	2011-09-2100:00:00	2011-09-2100:00:00
train	9001	9001
direction	0	0
location_id	8356	8357
loop_time	14280	14315
time_difference	69	5
route_code	0	0
destination	ClevelandOpenTrack	ClevelandOpenTrack
badge	20	20
car1	240	240
car2	115	115

Table 6 Data Dictionary for PORTAL Transit MAX Stop Data Event File

Attribute Name	Attribute Type	Description
service_date	timestamp	Date of Service
base_service_key	character (1)	W – weekday, S – Saturday – U – Sunday, X - Holiday
service_key	character (1)	Scheduling service key (A, "Weekday MAX") (B,"Saturday MAX" (c,"Streetcar Construction MAX") (C,"Sunday MAX") (d,"Race for the Cure") (p,"Monday - Thursday MAX") (P,"Late hours Friday MAX")

		Lower case designation indicates alternate service for various reasons.
route_number	integer	90 red, 100 blue, 190 yellow, 200 green
direction	smallint	Direction of travel (0 – Eastbound/Northbound) – (1 – Westbound/Southbound)
trip_number	integer	Scheduling trip number
train	integer	Scheduled Vehicle ID
location_id	integer	MAX Station ID
stop_time	integer	Scheduled time. (Units: seconds past midnight)
schedule_status	smallint	2,3 regular stops, 4 intermediate time point, 5 beginning of route, 6 end of route
leave_time	integer	Time vehicle crossed loop (Units: seconds past midnight)
time_difference	integer	Calculated offset time from loop to MAX station. Calculated offset time from loop to MAX station. The loop detectors that are the source of this stop event data are not located directly at MAX station; this field is the calculated travel time (offset) between the MAX station and the loop detector. (Units: seconds)

Sample

Attribute Name	Example 1	Example 2
service_date	2011-09-2100:00:00	2011-09-2100:00:00
base_service_key	W	W
service_key	A	A
route_number	100	100
direction	0	0
trip_number	1010	1010
train	9001	9001
location_id	8356	8357
stop_time	14220	14360
schedule_status	5	2
leave_time	14280	14315
time_difference	69	5

Table 7 Data Dictionary for PORTAL Transit MAX Stop Event APC File

Attribute Name	Attribute Type	Description
service_date	timestamp	Date of Service
service_key	character (1)	scheduling service key (A, "Weekday MAX") (B,"Saturday MAX" (c,"Streetcar Construction MAX") (C,"Sunday MAX") (d,"Race for the Cure") (p,"Monday - Thursday MAX")

		(P,"Late hours Friday MAX")
		Lower case designation indicates alternate service for various reasons.
route_number	integer	90 red, 100 blue, 190 yellow, 200 green
direction	smallint	Direction of travel (0 – Eastbound/Northbound) – (1 – Westbound/Southbound)
trip_number	integer	scheduling trip number
location_id	integer	MAX Station ID
stop_time	integer	Scheduled time. (Units: seconds past midnight)
distance	double precision	Distance from beginning of trip (Units: feet)
ons	smallint	Total boardings of vehicle(s)
offs	smallint	Total alightings of vehicle(s)
estimated_load	double precision	On board Load
arrive_time	integer	Arrive time at station (Units: seconds past midnight)
leave_time	integer	Leave time at station (Units: seconds past midnight)
lrv	smallint	Number of Light Rail Vehicles (LRV). Number of cars in the train.
train	integer	Scheduled Vehicle ID

Sample

Attribute Name	Example 1	Example 2
service_date	2011-09-2100:00:00	2011-09-2100:00:00
service_key	A	A
route_number	100	100
direction	1	1
trip_number	1350	1350
location_id	8363	8364
stop_time	21540	21603
distance	1933.272	2320.351
ons	5	6
offs	0	0
estimated_load	5	11
arrive_time	21620	21701
leave_time	21640	21730
lrv	2	2
train	9008	9008