NIST Data Science Pre-Pilot Evaluation Traffic Data Codebook

NIST Data Science Team
Last Updated August 18, 2015

This document provides a codebook for the traffic detector data and the traffic event data that is given to participants for use in the NIST Data Science Pre-Pilot Evaluation. This codebook identifies the values and provides limited descriptions of the different fields that appear in each of the data sets. This codebook is organized by data set, with one Section per data set. The data sets are:

1. **Traffic Event Data.** This data is a listing of all traffic events related to traffic. These range from accidents to disabled vehicles to roadwork to bad weather.
2. **Traffic Lane Detector Inventory.** This is the inventory (list) of the different traffic lane detectors.
3. **Traffic Lane Detector Data.** This contains the different measurements of traffic volume and traffic speed from the lane detectors.

Before describing the data sets, we provide a section with general background including notation conventions of this codebook. Additionally, to better understand traffic lanes and traffic zones, we briefly describe of traffic zones in that Section.

Table of Contents

1 Background and Notation

This section provides common preliminaries that are used throughout the data set.

1.1 Notation Conventions of this Document

Here are some of the notational conventions of this codebook.

- ▶ Names of field labels are in `monospace (Consolas) font`. An example is the field `event_id`, which occurs in the Traffic Event (Incident) Data. Likewise, the format of a specific field will be in `monospace (Consolas) font`. An example is the timestamp format `[YYY]-[MM]-[DD] [hh]:[mm]:[ss][UTC]`.

- ▶ When describing each field in the data set, we will give a set of paragraphs for each field label by writing the field label in **`bold monospace (Consolas) font`**. An example is **`event_id`**, which is the first field in the Traffic Event (Incident) data.

- ▶ Emphasized words that are defined are in **bold sans-serif font**. An example is when we define a **traffic zone**.

- ▶ Example values for fields will be in `Courier New font`. An example is the timestamp `2007-02-08 14:23:56-05`. Likewise, if there is a list of field values, those fields will be in `Courier New font`. A field value used as a category heading might be in **`bold Courier New font`**.

- ▶ Throughout the document as needed, emphasized words, sometimes as headings of lists, will be in **bold sans-serif font**. The heading **Traffic Lane Detector Data** is an example.

- ▶ For titles and gentle emphasis, text in *italics* will be occasionally used.

1.2 Traffic Zones

A **traffic zone** is an aggregation of traffic lanes. To better understand a traffic zone, see Figure 1. In this figure, each dot represents a hypothetical traffic lane detector. Each box represents a traffic zone. In this figure, there are 12 traffic lane detectors, and 4 traffic zones. Notices that per road segment in each direction, there is one traffic zone detector. Each traffic zone is an aggregation of the multiple traffic lane detectors in that road segment of traffic going in that specific direction. In our data, each lane detector has both a `lane_id` and a `zone_id`. The `zone_id` of a lane detector indicates the traffic zone to which the traffic lane belongs.
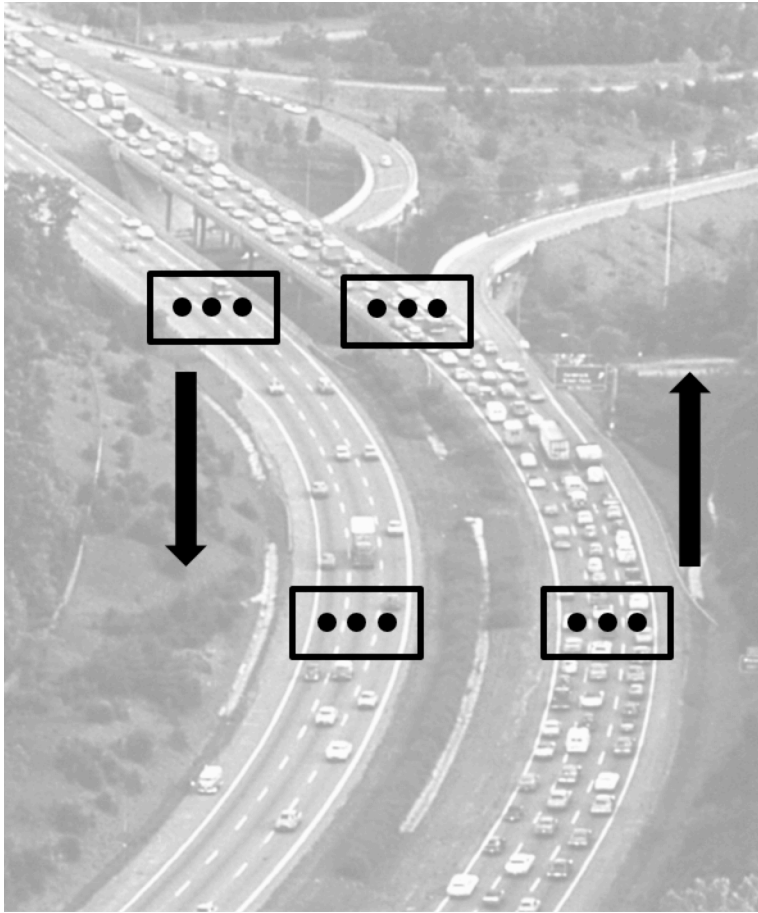
Background and Notation

Figure 1: Traffic zones are an aggregation of traffic lanes. The dots are traffic lane detectors and the boxes are the traffic zones. In this figure, there are 12 traffic lane detectors: 4 traffic zones with 3 detectors in each zone. There is one traffic zone detector per road segment in each direction. In the data, there are different numbers of lane detectors in each traffic zone. This photo was obtained from National Archives: https://catalog.archives.gov/id/546711 and then modified.

2   Traffic Event Data

The traffic event data has the following fields: `event_id`, `event_description`, `start_tstamp`, `confirmed_tstamp`, `created_tstamp`, `closed_tstamp`, `event_type`, `event_subtype`, `location`, `latitude`, `longitude`, `number_of_responders`, and `lanes_affected`.

This file is a comma-separated file, where there is one traffic event (record) per row (line), and each row contains the fields in the order they are described here. There is a header row with the field labels.

For this data set, timestamps (times) for each of the fields is in the following format, where identifiers such as the year (YYYY) are enclosed in brackets `[]`:
`[YYY]-[MM]-[DD] [hh]:[mm]:[ss][UTC]`. The format of a timestamp is as follows:

- ▶ `[YYYY]` is the year, written as 4 digits,
- ▶ `[MM]` is the month, written as 2 digits,
- ▶ `[DD]` is the day, written as 2 digits,
- ▶ `[HH]` is the hour in 24-hour time, written as 2 digits,
- ▶ `[mm]` is the minute, written as 2 digits,
- ▶ `[ss]` is the second, written as 2 digits
- ▶ `[UTC]` is the UTC (Coordinated Universal Time) timezone; notice that the timezone has a + or – before the number.

In this format, the hyphens, spaces, and colons are as in the format string. An example time is `2007-02-08 14:23:56-05`, which is 14:23:56 on February 8, 2007 at the UTC -5 timezone, or 2:23:56pm on February 8, 2007 at EST without Daylight Savings Time.

event_id is the unique identifier of the event or incident. Each id is a text string with the following format:

- ▶ `EventSource_EventLabel`

Where `EventSource` is a tag identifying which source it is from. Notice that some `EventSource` strings contain an underscore "_" and notice that the `EventLabel` is separated from the `EventSource` by an underscore "_" The different `EventSource` fields are:

1. `MDOT_CHART`,
2. `DDOT_CAPTOP`,
3. `DDOT_CONSTRUCTION`,
4. `VaTraffic`,
5. `VDOT_NOVA`,

Due to noise in the data the `EventSource` may not always be an exact match.

event_description is the textual description of the event, which is a text string.

start_tstamp is the time at which the event began (moment of impact for collisions, scheduled start time for road construction, etc.).

confirmed_tstamp is the time at which the traffic management center confirmed the event. One scenario for this would be if a traffic accident was called in by a traveler, the event would not be confirmed until a highway service patrol or other responder arrived on scene and informed the operator that there was indeed an accident.

created_tstamp is the time at which the event is created within the traffic management system. In the above scenario, this would be the time at which the operator received the initial accident report from the traveler.

closed_tstamp is the time at which the event is considered over. For accidents, this typically indicates when all lanes have been reopened, damaged vehicles have been removed, and all responders have left the scene.

event_type is a label that identifies the overall category or incident type. The 37 different event_type values are:
- ▶ accidentsAndIncidents:
- ▶ adviceInstructionsMandatory:
- ▶ adviceInstructionsRecommendation:
- ▶ AlternateRoute:
- ▶ closures:
- ▶ DelayStatusCancellation:
- ▶ deviceStatus:
- ▶ disasters:
- ▶ disturbances:
- ▶ incidentResponseEquipment:
- ▶ incidentResponseStatus:
- ▶ laneRoadway:
- ▶ mobileSituation:
- ▶ obstruction:
- ▶ parkingInformation:
- ▶ pavementConditions:
- ▶ precipitation:
- ▶ responderGroupAffected:
- ▶ restrictionClass:
- ▶ roadwork:
- ▶ specialEvents:
- ▶ sportingEvents:
- ▶ suggestionAdvice:
- ▶ systemInformation:
- ▶ temperature:
- ▶ trafficConditions:
- ▶ transitMode:
- ▶ transitOperations:

Traffic Event Data

- ▶ travelerGroupAffected:
- ▶ unusualDriving:
- ▶ vehicleGroupAffected:
- ▶ visibilityAndAirQuality:
- ▶ warningAdvice:
- ▶ weatherConditions:
- ▶ winds:
- ▶ winterDrivingIndex:
- ▶ winterDrivingRestrictions:

event_subtype is the subcategory of each event. For each **event_type**, here are the possible **event_subtype** values. Note that any misspellings are copied from the data and are as found in the data.

- ▶ **accidentsAndIncidents**: incident, disabled vehicle, multi vehicle accident, numerous accidents, accident involving a truck, accident, vehicle on fire, minor accident, hazardous material spill, abandoned vehicle, accident involving a pedestrian, accident involving a semi trailer, serious accident, injury accident, accident involving a bicycle, accident involving a bus, accident involving a motorcycle, accident involving a train, accident involving hazardous materials, earlier accident, medical emergency, secondary accident, accident investigation work, stalled vehicle, disabled truck, disabled semi trailer, disabled bus, disabled train, vehicle spun out, vehicle in water, vehicles slowing to look at accident, jackknifed semi trailer, jackknifed trailer home, jackknifed trailer, spillage occuring from moving vehicle, acid spill, chemical spill, fuel spill, oil spill, spilled load, toxic spill, overturned vehicle, overturned truck, overturned semi trailer, overturned bus, derailed train, stuck vehicle, truck stuck under bridge, bus stuck under bridge, accident cleared, incident cleared
- ▶ **adviceInstructionsMandatory**: keep to the right, keep to the left, use right lane, use left lane, heavy vehicles use right lane, heavy vehicles use left lane, observe signals, observe signs, no passing, no smoking, no open flames, use shoulder as lane, do not drive on the shoulder, allow emergency vehicles to pass, clear a lane for emergency vehicles, pull over to the edge of the roadway, wait for escort vehicle, in emergency wait for police patrol, reduce your speed, observe speed limits, check point, entry requirements, insurance requirements, firearms requirements, pet requirements, shoulder travel no longer allowed
- ▶ **adviceInstructionsRecommendation**: drive carefully, drive with extreme caution, approach with care, keep your distance, increase normal following distance, test your

brakes, cross intersection with care, pass with care, use low beam headlights, use fog lights, use hazard warning lights, do not leave your vehicle, leave your vehicle and proceed to next safe place, turn off engine, close all windows turn off heater air conditioner and vents, turn off air conditioner to prevent engine overheating, turn off mobile phones and two way radios, prepare to stop, stop at next rest area, stop at next safe place, only travel if absolutely necessary, drive to another service area, use through traffic lanes, use local traffic lanes, use left hand parallel roadway, use right hand parallel roadway, use heavy vehicle lane, observe recommended speed

▸ **AlternateRoute:** detour, no detour available, detour where possible, follow signs, follow detour signs, follow special detour markers, do not follow detour signs, detour in operation, follow local detour, compulsory detour in operation, no suitable detour available, detour is no longer recommended, local drivers are recommended to avoid the area, trucks are recommended to avoid the area, consider alternate route, consider alternate parking, consider alternate destination, consider alternate area

▸ **closures:** closed to traffic, closed, closed ahead, closed intermittently, closed for repairs, closed for the season, blocked, blocked ahead, reduced to one lane, reduced to two lanes, reduced to three lanes, collapse, out, open to traffic, open, reopened to traffic, clearing, cleared from road

▸ **DelayStatusCancellation:** delay and disruption, delays, short delays, long delays, very long delays, delays of uncertain duration, delayed until further notice, busy, very busy, crowded, overcrowded, cancellations, route cancelled and no replacement, service cancelled, service suspended, service withdrawn, service fully booked, all services fully booked, next departure, next arrival, very frequent service, frequent service, fairly frequent service, regular service, irregular service, not operating, system busy, system very busy, system crowded, system overcrowded, headway, extra services in operation, delays clearing, delays cleared, nromal services resumed, operating

▸ **deviceStatus:** traffic lights not working, sign down, lane control signs not working, lane control signs working incorrectly, lane control signs operating, variable message signs not working, variable message signs working incorrectly, variable message signs operating, emergency telephones not working, emergency telephone number not working, traffic lights working incorrectly, ramp control signals not working, ramp control signals working incorrectly, temporary traffic lights not working,

temporary traffic lights working incorrectly, traffic signal control computer not working, traffic signal timings changed, overheight warning system triggered, equipment failure, railroad crossing equipment failure, tunnel ventilation not working, power failure, widespread power outages, technical problems, lines in road faded, damaged light standard, traffic signal stuck on flash, guide rail, fencing, light standard hanging by wires, call box, signal cabinet, detector, improper use of State vehicle or equipment, bulb out, not yet operational, not yet installed, electronic signs repaired, emergency call facilities restored, traffic signals repaired, railroad crossing equipment now working normally

▶ **disasters:** flash flood, brush fire, earthquake damage, wildfire, major flood, reservoir failure, levee failure, tsunami, tidal wave, volcanic eruption, ash fall, lava flow, serious fire, forest fire, building fire, grass fire, fire danger extreme, fire danger very high, fire danger high, fire danger medium, fire danger low, air crash, rail crash, toxic release, toxic leak, radioactive release, radioactive hazard, reactor leakage, explosion, major hazardous materials fire, major hazardous materials release, disaster cleared

▶ **disturbances:** security incident, march, security alert, bomb alert, terrorist incident, assault, crime, robbery, fare dispute, shooting, gunfire on roadway, suicide, fight, gang fight, person harassment, person injured, sick customer, unruly passenger, person intoxicated, crowd control problem, demonstration, public disturbance, riot, civil unrest, civil emergency, strike, public transit strike, stampede, teargas used, checkpoint, high velocity shell fire, explosives in use, air raid, weapons of mass destruction threat, military operations, security problem cleared, traffic disturbance cleared

▶ **incidentResponseEquipment:** ground fire suppression, heavy ground equipment, aircraft, marine equipment, support equipment, medical rescue unit, other, ground fire suppression other, engine, truck or aerial, quint, tanker pumper combination, brush truck, aircraft rescue firefighting, heavy ground equipment other, dozer or plow, tractor, tanker or tender, aircraft other, aircraft fixed wing tanker, helitanker, helicopter, marine equipment other, fire boat with pump, boat no pump, support apparatus other, breathing apparatus support, light and air unit, medical rescue unit other, rescue unit, urban search rescue unit, high angle rescue, crash fire rescue, bls unit, als unit, mobile command post, chief officer car, hazmat unit, type i hand crew, type ii hand crew, privately owned vehicle, other apparatus resource, ambulance, bomb squad

van, combine harvester, construction vehicle, farm tractor, grass cutting machines, hazmat containment tow, heavy tow, light tow, flatbed tow, hedge cutting machines, mobile crane, refuse collection vehicle, resurfacing vehicle, road sweeper, roadside litter collection crews, salvage vehicle, sand truck, snowplow, steam roller, swat team vantrack laying vehicle unknown vehicle white lining vehicle dump truck supervisor vehicle snow blower rotary snow blower road grader steam truck

▶ **incidentResponseStatus:** unconfirmed report, initial response en route, follow up response on scene, initial response on scene, follow up response en route, confirmed report, scene is unsecured at this time, response scene secured, rescue and recovery work in progress, extraction in progress, clearance work in progress, body removal operations, fire containment contained, fire containment not contained, event cleared, traffic clearing, incident closed

▶ **laneRoadway:** all roadways, through lanes, left lane, right lane, center lane, middle lanes, middle two lanes, right turning lanes, left turning lanes, upper deck lanes, lower deck lanes, reversible lanes, right exit lanes, left exit lanes, right merging lanes, left merging lanes, right exit ramp, right second exit ramp, right entrance ramp, right second entrance ramp, left exit ramp, left second exit ramp, left entrance ramp, left second entrance ramp, escape ramp, hard shoulder, soft shoulder, right shoulder, left shoulder, highways, right hand parallel lanes, left hand parallel lanes, connecting lanes, express lanes, local lanes, toll lanes, electronic toll lanes, toll plaza, inspection lane, hov lanes, bus lanes, carpool lanes, truck lanes, emergency lanes, passing lanes, climbing lanes, slow lane, service road, cycle lane, bridge, overpass, elevated lanes, underpass, tunnel

▶ **mobileSituation:** motorcade, abnormal load, wide load, long load, slow vehicle, farm equipment, horse drawn vehicles, overheight load, overweight load, tracked vehicle, vehicle carrying hazardous materials, slow moving maintenance vehicle, convoy, military convoy, refugee convoy, mobile situation repositioning, winter maintenance vehicles, snowplows, slow moving maintenance vehicle warning cleared, exceptional load warning cleared, hazardous load warning cleared, convoy cleared

▶ **obstruction:** debris on roadway, fallen trees, subsidence, mudslide, rockfall, drawbridge open, obstruction on roadway, animal struck, downed cables, landslide, object on roadway, objects falling from moving vehicle, storm damage, people on roadway, bicyclists on roadway, sightseers obstructing access, large numbers of visitors, animal on

roadway, large animal on roadway, herd of animals on roadway, advertising signs, over turned trees, tree limbs, utility pole down, downed power lines, road surface collapse, frost jacking, frost heave, pavement buckled, pothole, flooding, broken water main, collapsed sewer, wash out, washboard, sewer overflow, gas leak, snowmelt, avalanche, clearance work, obstruction cleared

▶ **parkingInformation**: normal parking restrictions lifted, parking meter restrictions lifted, special parking restrictions in force, full parking lot, full parking garage, all parking lots full, no parking spaces available, only a few spaces available, spaces available, no parking, parking on one side of street only, parking on both sides of street, parallel parking only, parking meters not available, use of parking meters restricted, event parking, handicapped parking, long term parking, overnight parking, short term parking, parking by permit only, special parking restrictions lifted, no parking information available

▶ **pavementConditions**: ice, surface water hazard, impassable, almost impassable, passable with care, passable, danger of hydroplaning, wet pavement, treated pavement, slippery, mud on roadway, leaves on roadway, loose sand on roadway, loose gravel, fuel on roadway, oil on roadway, road surface in poor condition, melting tar, icy patches, black ice, ice pellets on roadway, ice build up, freezing rain, wet and icy roads, melting snow, slush, frozen slush, snow on roadway, packed snow, packed snow patches, plowed snow, wet snow, fresh snow, powder snow, granular snow, frozen snow, crusted snow, deep snow, snow drifts, drifting snow, expected snow accumulation, current snow accumulation, dry pavement, snow cleared, pavement conditions improved, skid hazard reduced, pavement conditions cleared

▶ **precipitation**: snow, severe weather, blizzard, heavy snow, light snow, snow showers, winter storm, ice glaze, heavy frost, frost, ice storm, sleet, rain and snow mixed, rain changing to snow, damaging hail, hail, thunderstorms, thundershowers, extremely heavy downpour, heavy rain, rain, light rain, drizzle, showers, dew, precipitation cleared

▶ **responderGroupAffected**: emergency vehicle units, federal law enforcement units, state police units, county police units, local police units, ambulance units, rescue units, fire units, hazmat units, light tow unit, heavy tow unit, freeway service patrols, transportation response units, private contractor response units

▶ **restrictionClass**: restrictions, ramp restrictions, truck restriction, speed restriction, noise restriction, traffic regulations have been changed, local access only, no trailers, no high profile vehicles, hazardous materials truck restriction, no through traffic, no motor vehicles,

width limit, height limit, length limit, axle load limit, gross weight limit, axle count limit, carpool lane available, carpool restrictions changed, hov 2 no single occupant vehicles, hov 3 no vehicles with less than three occupants, bus lane available for all vehicles, truck lane available for all vehicles, permits call in basis, permits temporarily closed, permits closed, permits open, restrictions for high profile vehicles lifted, width limit lifted, height limit lifted, length limit lifted, axle load limit lifted, weight limit lifted, axle count limit lifted, carpool restrictions lifted, lane restrictions lifted, ramp restrictions lifted, motor vehicle restrictions lifted, restrictions lifted

▶ **roadwork:** bridge maintenance operations, storm drain, emergency maintenance, road maintenance operations, overgrown grass, road marking operations, paving operations, overgrown brush/shrubs, work on underground services, water main work, bridge construction, road construction, construction work, road widening, work in the median, overgrown trees, major road construction, long term road construction, road reconstruction, opposing traffic, narrow lanes, construction traffic merging, single line traffic alternating directions, cracks, bumps, drop off, bridge demolition work, seismic retrofit, blasting, avalanche control activities, gas main work, work on underground cables, new road construction layout, new road layout, temporary lane markings, temporary traffic lights, road maintenance cleared, normal road layout restored, road work clearance in progress, road construction cleared, normal traffic lanes restored, road work cleared

▶ **specialEvents:** major event, parade, concert, fair, airshow, hot air ballooning, state occasion, vip visit, show, festival, exhibition, performing arts, outdoor market, carnival, fireworks display, trade expo, movie filming, presidential visit, procession, funeral procession, crowd, holiday traffic crowds, event ended

▶ **sportingEvents:** sports event, game, tournament, track and field event, baseball game, basketball game, boxing match, football game, soccer game, golf tournament, hockey game, tennis tournament, wrestling match, road race, automobile race, bicycle race, race event, marathon, horse show, rodeo, water sports event, winter sports event, skating event, dog sled race, sporting event ended

▶ **suggestionAdvice:** why not ride share, is this your no ride day, why not use transit, why not park and ride, why not try paratransit, why not travel by rail, why not use the subway, why not try rapid transit, why not travel by bus, your parking ticket covers the return ride, avoid the rush hour, do not allow unnecessary gaps, follow the vehicle in

front smoothly, do not slow down unnecessarily, watch for muggers, watch for pickpockets, watch for thieves, sorry for any delay, thank you for your understanding, we appreciate your patience, we are grateful for your cooperation

▶ **systemInformation:** test message, information available on radio, information available on tv, call to get information, information available via internet, no information available, null description, information service is being suspended, information service resumed, message canceled

▶ **temperature:** extreme heat, maximum temperature, temperature, minimum temperature, current temperature, heat index, hot, hotter, heat, warmer, warm, mild, cool, cooler, cold, colder, very cold, extreme cold, wind chill, dewpoint, relative humidity, temperatures close to the seasonal norm, less extreme temperatures

▶ **trafficConditions:** traffic congestion, stopped traffic, stop and go traffic, slow traffic, heavy traffic, traffic building, long queues, traffic lighter than normal, traffic heavier than normal, traffic much heavier than normal, current speed, speed limit, travel time, traffic flowing freely, traffic easing, traffic returned to normal, no problems to report, traffic congestion cleared

▶ **transitMode:** travel, transit, bus, trolleybus, rail, commuter rail, subway, rapid transit, light rail, streetcar, dial a ride, park and ride, shuttle, free shuttle, airport shuttle, taxies, ferry, passenger ferry, vehicle ferry, aerial tramway, automated guideway, cable cares, monorail, air travel

▶ **transitOperations:** unknown transit problem, sleeping customer, assault on passenger, assault on employee, broken seat, bus alarm, crime or drug deal, eating on board, equipment problem with air conditioning, equipment problem with air system, equipment problem with brakes, equipment problem with chassis or suspension, equipment problem with cooling system, equipment problem with doors, equipment problem with electrical, equipment problem with engine, equipment problem with exterior or body, equipment problem with fare collection, equipment problem with fuel or exhaust, equipment problem with horn, equipment problem with interior, equipment problem with liftkneeling, equipment problem with lights, equipment problem with lubrication, equipment problem with radio or communication, equipment problem with signs, equipment problem with steering, equipment problem with tires, equipment problem with transmission, equipment problem with unknown alarm, equipment problem with wipers, fare dispute expired pass, fare dispute expired transfer, fare dispute expired

upgrade, fare dispute other, fare dispute refuses to pay, lift passenger cycle completed, lift passenger ready to alight, lift passenger ready to board, lost article, objects thrown, passenger accident alighting, passenger accident boarding, passenger accident fallen on board, passenger load, passenger accident other, passenger sick or injured, right of way, theft, theft of service, waiting to get relief for schedule break, waiting to get relief after run is finished, waiting to provide relief

▶ **travelerGroupAffected:** through traffic, holiday traffic, residents, visitors, long distance traffic, local traffic, regional traffic, arrivals, departures, airline travelers, commuter airline travelers, domestic airline travelers, international airline travelers, pedestrians, bicyclists

▶ **unusualDriving:** vehicle traveling wrong way, reckless driver, prohibited vehicle on roadway, emergency vehicles on roadway, high speed emergency vehicles, high speed chase, dangerous vehicle warning cleared, emergency vehicle warning cleared

▶ **vehicleGroupAffected:** all vehicles, bicycles, motorcycles, cars, light vehicles, cars and light vehicles, cars with trailers, cars with recreational trailers, vehicles with trailers, heavy vehicles, trucks, buses, articulated buses, school buses, vehicles with semi trailers, vehicles with double trailers, high profile vehicles, wide vehicles, long vehicles, hazardous loads, exceptional loads, abnormal loads, convoys, maintenance vehicles, delivery vehicles, vehicles with even numbered license plates, vehicles with odd numbered license plates, vehicles with parking permits, vehicles with catalytic converters, vehicles without catalytic converters, gas powered vehicles, diesel powered vehicles, ipg vehicles, military convoys, military vehicles

▶ **visibilityAndAirQuality:** fog, visibility reduced, dense fog, patchy fog, freezing fog, ice fog, mist, haze, visibility blocked, white out, blowing snow, smoke hazard, spray hazard, low sun glare, snow glare, blowing dust, blowing sand, dust storms, sandstorms, air quality good, air quality fair, air quality poor, air quality very poor, severe exhaust pollution, smog alert, pollen count high, pollen count medium, pollen count low, swarms of insects, fog clearing, visibility improved, fog forecast withdrawn, pollution alert ended, air quality improved

▶ **warningAdvice:** police at scene, alert, risk, watch, warning, danger, danger of explosion, danger of fire, extra police patrols in operation, look out for vehicles stopped under bridges, increased risk of accident, rescue and recovery work in progress at scene, emergency vehicles at scene, traffic being directed around accident area, police directing traffic, rescue workers directing traffic,

> repairs in progress, pilot car in operation, look out for flagger, police checks in operation, truck check point, lockdown, security check point, single occupancy vehicle check point, mandatory speed limit in force, speed limit in force for heavy vehicles, warning canceled, watch canceled, alert canceled, ended, cleared, canceled

▶ **weatherConditions:** overcast, cloudy, mostly cloudy, partly cloudy, partly sunny, mostly sunny, sunny, fair skies, clear skies, mostly clear, mostly dry, dry, uv index very high, uv index high, uv index moderate, uv index low, uv index very low, barometric pressure, ozone alert, lighting unknown, artificial exterior light, artificial interior light, darkness, dusk, dawn, moonlight, daylight

▶ **winds:** strong winds, hurricane, tornado, hurricane force winds, tropical storm, gale force winds, storm force winds, moderate winds, light winds, calm, gusty winds, crosswinds, windy, strong winds have eased, strong wind forecast withdrawn

▶ **winterDrivingIndex:** driving conditions good, driving conditions fair, difficult driving conditions, very difficult driving conditions, hazardous driving conditions, extremely hazardous driving conditions

▶ **winterDrivingRestrictions:** winter equipment recommended, winter equipment required, snow chains recommended, snow chains required, snow chains prohibited, studded tires prohibited, snow tires recommended, snow tires required, four wheel drive recommended, four wheel drive required, snow tires or chains recommended, snow tires or chains required, winter driving requirements lifted

The subtype `incident` of `accidentsAndIncidents` is sometimes a generic type that classifies events that are not given a clear label.

location is a textual description of the location of the event.

latitude is the latitude in degrees of the event location.

longitude is the longitude in degrees of the event location.

number_of_respondents is a field whose meaning is unknown.

lanes_affected is the number of lanes whose traffic was influenced by the event or incident. This field is sometimes missing.

3   Traffic Lane Detector Inventory

The traffic lane detector inventory has the following fields: `lane_id`, `zone_id`, `lane_number`, `name`, `state`, `road`, `direction`, `location_description`, `lane_type`, `organization`, `detector_type`, `latitude`, `longitude`, `bearing`, `default_speed`, `interval`.

The files with this data contain one lane detector per row, and are comma-separated. There is a header row with the field labels, which are in the order that they are described in this codebook.

`lane_id` is the unique identifier of the lane.

`zone_id` is the foreign key that identifies which traffic zone this lane is in

`lane_number` is the number of that lane.

`name` is the text name of that specific lane. It will often make sense to pair the name with the `location_description` field in order to better understand the lane. Sometimes the `name` will contain text that has information contained in other fields, such as the `direction` and the `lane_number` (although the text could be spelled differently than the values in the other fields).

`state` is the name of the state that this lane is in. The different values for `state` are:
- ▶ D.C.
- ▶ Maryland
- ▶ Northern Virginia

`road` is the name of the road or route that this lane is on.

`direction` is direction of the traffic in that lane. The different possible values for `direction` are:
- ▶ East,
- ▶ Inner Loop,
- ▶ None,
- ▶ North,
- ▶ Other,
- ▶ Outer Loop,
- ▶ Reversible HOV,
- ▶ South,
- ▶ West,

`location_description` is the textual description of the lane. This description includes the Route number and sometimes a mile marker near that lane. The abbreviation MM inside this field stands for "Mile Marker." One will note that the text in the

`location_description` sometimes contains information found in other fields (although sometimes spelled differently), such as the `road` field.

`lane_type` is the label for the kind of lane. The typical `lane_type` value is Normal, and the list of all the `lane_type` values is:

- ▶ Bus,
- ▶ CD,
- ▶ HOV,
- ▶ HOV Off Ramp,
- ▶ HOV On Ramp,
- ▶ Normal,
- ▶ Off Ramp,
- ▶ On Ramp,
- ▶ RHOV,
- ▶ Shoulder,
- ▶ Toll,
- ▶ Unknown,

`organization` is the source that provided the inventory data. The different `organization` values are:

- ▶ HERE, (CATT Lab)
- ▶ MDOT_CHART,
- ▶ VDOT_NOVA,

`detector_type` is the kind of detector for that lane at that location. The possible values for `detector_type` are:

- ▶ 170E,
- ▶ Acoustic,
- ▶ Magnetometer (Sensys),
- ▶ Microwave,
- ▶ Radar,
- ▶ RTMS,
- ▶ Single Loop,
- ▶ Unknown,

`latitude` is the latitude in degrees of the lane location.

`longitude` is the longitude in degrees of the lane location.

`bearing` is a field whose meaning is unknown.

`default_speed` is usually the speed limit, in miles per hour, of that lane. This field is often missing, and this field is occasionally given something else by the data provider.

`interval` is the typical number of seconds between detector measurements.

4   Traffic Lane Detector Data

The traffic lane detector inventory data has the following fields: `lane_id`, `measurement_start`, `speed`, `flow`, `occupancy`, `quality`.

The fields are in left-to-right order, separated by commas "," in between each field. There is one measurement per row (line). There is a header row with the field labels, which are in the order that they are described in this codebook.

For this data set, timestamps (times) for each of the fields is in one of the two the following formats, where identifiers such as the year (YYYY) are enclosed in brackets `[]`: `[YYY]-[MM]-[DD] [hh]:[mm]:[ss][UTC]` and `[YYY]-[MM]-[DD] [hh]:[mm]:[ss.ddd][UTC]`. The format of a timestamp is as follows:

- ▶ [YYYY] is the year, written as 4 digits,
- ▶ [MM] is the month, written as 2 digits,
- ▶ [DD] is the day, written as 2 digits,
- ▶ [HH] is the hour in 24-hour time, written as 2 digits,
- ▶ [mm] is the minute, written as 2 digits,
- ▶ [ss] is the second, written as 2 digits
- ▶ `[ss.ddd]` is the second measurement to three decimal places, taking 6 digits: `ss` for the whole number of seconds, and `ddd` for the fractional part of the second measurement.
- ▶ `[UTC]` is the UTC (Coordinated Universal Time) timezone; notice that the timezone has a + or – before the number.

In this format, the hyphens, spaces, and colons are as in the format string. An example time is `2014-11-01 00:26:40.427-04`, which is 00:26:40.207 on November 1, 2014 at the UTC -4 timezone (EST with Daylight Savings Time) and is parsed using the second format. Another example is `2014-11-30 23:54:09-05`, which is 11:54:09pm on November 30, 2011 at timezone UTC -5 (EST without Daylight Savings Time).

For these measurements, the `speed` is the average traffic speed from the previous time, and the `volume` is the total number of vehicles that passed through since the previous time. For instance, for these two measurements, assuming that these are two consecutive measurements from the same detector [labels added for clarity]:

- ▶ `2008-03-13 06:40:46-04,` `speed=46.6,` `vol=137`
- ▶ `2008-03-13 06:45:46-04,` `speed=46.3,` `vol=124`

This means that during the 5 minutes from 06:40:46 to 06:45:46, the average (mean) vehicle speed was 46.3 miles per hour and 124 vehicles pass through this detector.

lane_id is the foreign key, which is the lane identifier for the detector from which this reading came.

measurement_start is the time at which the reading was collected.

speed is the average (arithmetic mean) speed, in miles per hour, of all vehicles observed passing through the detector during this reading.

flow is the number of vehicles observed passing through the detector during this reading. Each reading happens every `interval` seconds (the `interval` field is in the traffic lane inventory). For instance, if the value of interval value is `60`, that means that the flow value is the number of vehicles to have passed through this detector in the last `60` seconds. This is true even if the previous measurement is missing, meaning that if the previous measurement is missing, the flow value is still only for the last `interval` seconds.

occupancy is the percentage of time a vehicle was observed within the detector during this reading. From examining sample data, this occupancy can be very small (0.3) with a high traffic speed, and can even be negative and over 100. For example, an occupancy value of `9.3` means that the detector was occupied 9.3% of the time from the previous measurement until the current measurement.

quality is the quality indicator of the reading. The possible values of the `quality` field are:
- ▶ `0`, (valid)
- ▶ `1`, (invalid)
- ▶ `2`, (incomplete)
- ▶ `3`, (unknown).

Not all providers include a `quality` flag with their feed (most don't) and this field defaults to `0` (valid), so there may be readings that don't make sense but are still flagged as `0` (valid).

5 Acknowledgements