Drafted Data Dictionary for VT\_Crashes

2015:

* **X** corresponds to coordinates for mapping. This variable is the same as Easting
* **Y** corresponds to coordinates for mapping. This variable is the same as Northing
* **ObjectID** is a unique identifier for each entry.
* **REPORTINGA** and **REPORTINGAGENCY** – REPORTINGA is a for the reporting agency, and REPORTINGAGENCY is the name of the reporting agency.
* **REPORTNUMB** – This is the number associated with each reported crash. It appears that the REPORTNUMB can be the same for two different entries, which suggests that they were two different vehicles in the same crash. Rows #2 and #11 share REPORTNUMB and other identical information for example.
* **LOCALID** – Unknown, categorical
* **ACCIDENTTI** – Date and time of accident.
* **ACCIDENTDA** – Date of accident. A less descriptive version of ACCIDENTTI
* **STREETADDR** – The address where the accident occurred.
* **NUMBER** – Associated to the stretch of road on which the accident occurred.
* **NUMBER1**
* **NUMBER2**
* **NUMBER3**
* **INTERSECTI** – Indicates intersecting street, if crash occurred at an intersection. This seems to be related to RoadCharacteristics, which provides a description of intersections.
* **DISTANCE\_1** – Unknown
* **POSTEDSPEE** – Appears to be the posted speed limit, although the value is unavailable in many cases. Additionally, 99 is present many times, despite 65mph being the highest speed limit in Vermont.
* **DIRFROMNEA** – Indicates cardinal directions. It is unclear what the significance of this is, and this value is missing for many rows.
* **DIROFCOLLI –** Provides a description of the collision direction.
* **MILEMARKER1** - Unknown
* **MILEMARKER\_1**
* **MILEMARKER\_2**
* **AOTROADWAYGROUPid** – The id for which Vermont Agency of Transportation road group the road where the accident occurred belongs to.
* **AOTROADWAY** **–** The full name of the road group where the accident occurred.
* **AOTACTUALM** - Unknown
* **CrashType** – Brief description of the category of crash which the accident belongs to.
* **RoadCharacteristics** – Describes if the accident occurred at an intersection or not.
* **NonReportableAddress** - Unknown
* **VCSG\_CITYORTOWNid -**  Vermont Crash System Group (VCSG). Essentially a numeric ID assigned to a city or town.
* **CITYORTOWN** – The name of the city or town.
* **EASTING –** The horizontal mapping coordinate.
* **NORTHING –** The vertical mapping coordinate
* **VCSG\_AOTROUTEid** - Vermont Crash System Group (VCSG) route ID
* **AOTROUTE** – Vermont Agency of Transportation route name. The name for the route of a given ID.
* **LRSNUMBER** – Linear Reference System (LRS) number.
* **VCSG\_LATITUDE -**Latitude
* **LAT\_DD** -Latitude
* **VCSG\_LONGITUDE** - Longitude
* **LONG\_DD –** Longitude
* **HOWMAPPED** – The method of mapping an accident. If it uses XY, it appears to uses the VCSG coordinates. If it uses LRS, it appears to be related to the LRSNUMBER, if it uses \_DD values, it goes by GEOCODE.
* **Animal** – Was an animal involved in the accident? If yes, animal type is typically listed.
* **Impairment –** Was the driver impaired? Typically alcohol or drugs.
* **Involving –** Were bystanders involved? Pedestrians or Bicycles for example.
* **Weather –** Description of weather at the time of the accident.
* **DayNight -**  Day/night conditions for the accident.
* **InjuryType –** Indication of a physical injury, or if there was property damage.
* **LOC\_ERROR –** This occurs if the LRSNUMBER is not present.

Source Item: <https://vtrans.vermont.gov/planning/maps/route-logs>

<https://geodata.vermont.gov/datasets/f493ca1c91454baaa75934a6a3931651_53>

Combined:

* **ACCIDENTDATE** – Date of accident. A less descriptive version of ACCIDENTTI
* **ACCIDENTTI** – Time and date of accident
* **AOTACTUALMILEPOINT -** Unknown
* **AOTROADWAY -** The full name of the road group where the accident occurred
* **AOTROADWAYGROUPid** - The id for which Vermont Agency of Transportation road group the road where the accident occurred belongs to.
* **AOTROUTE –** Route name where accident occurred
* **AOTROUTEid –** Code for route where accident occurred.
* **Animal** - Was there an animal involved in the accident? Ex: Deer, domestic
* **CITYORTOWN** – Name of town where accident occurred
* **CITYORTOWNid –** ID for town where accident occurred.
* **CrashType**  - Brief description of the category of crash which the accident belongs to.
* **DIRFROMNEA –** Cardinal direction of travel for the accident
* **DIROFCOLLI –** Description of the collision direction for year 2015
* **DISTANCE\_1** – Unit of distance, unknown use
* **DayNight –** If day/night
* **DirOfCollision** – The equivalent of DIROFCOLLI for years 2016-2018
* **EASTING –** Horizontal coordinate (equivalent to X.
* **GIS\_LATITUDE –** Essentially the same as LAT\_DD. The coordinates used for HOWMAPPED-GEOCODE
* **GIS\_LONGITUDE -** Essentially the same as LAONG\_DD. The coordinates used for HOWMAPPED-GEOCODE
* **HOWMAPPED** – Describes the mapping method used
* **INTERSECTIONWITH –** Describes the intersection the accident occurred at if there was an accident.
* **Impairment –** Was the driver impaired? Ex: Drugs/alcohol
* **InjuryType –** Was there an injury, or just property damage?
* **Involving -**  Were bystanders such as cyclists or pedestrians involved?
* **LATITUDE** – Used in the 2018 set, this appears to act as the only latitude value for the set
* **LAT\_DD** – This is a general latitude value used for HOWMAPPED: GEOCODE
* **LOCALID –** Unknown, categorical
* **LOC\_ERROR -**  This indicates if the LRSNUMBER is not present.
* **LONGITUDE –** A generic longitude value used in the 2018 set.
* **LONG\_DD –** A general latitude value used for HOWMAPPED: GEOCODE
* **LRSNUMBER** - Linear Reference System (LRS) number. Referred to as VCSG\_LRSNUMBER in 2017
* **MILEMARKER1** - Unknown
* **MILEMARKER\_1**
* **MILEMARKER\_2**
* **NORTHING –** Provides vertical coordinate (equivalent to Y)
* **NUMBER -** Associated to the stretch of road on which the accident occurred.
* **NUMBER1**
* **NUMBER2**
* **NUMBER3**
* **NonReportableAddress -** Unknown
* **OBJECTID –** ID within each year
* **POSTEDSPEE –** Present in 2015. Appears to be the posted speed limit, although the value is unavailable in many cases. Additionally, 99 is present many times, despite 65mph being the highest speed limit in Vermont.
* **RDFLNAME** – Name of stretch road the accident occurred on, not present in 2015.
* **REPORTINGA –** Code for reporting agency
* **REPORTINGAGENCY** – Name of the reporting agency
* **REPORTINGAGENCYid –** Code for reporting agency
* **REPORTNUMBER –** Code for report
* **ReportingAgency –** Name of the reporting agency
* **RoadCharacteristics** – Description of the kind of junction (if present) the accident occurred on.
* **RoadCondition** - Were the things such as obstructions or conditions due to weather?
* **RoadGroup** – Name of which road group the road belong to, similar to AOTROADWAYGROUPID (possible merger needed.)
* **Route** – Name of road stretch
* **STREETADDRESS –** Street address for accident location
* **SurfaceCondition -**  Describes things like ICE, WET, DRY
* **VCSG\_AOTROUTE** -
* **VCSG\_AOTROUTEid** -
* **VCSG\_CITYORTOWN** -
* **VCSG\_CITYORTOWNid** -
* **VCSG\_EASTING** -
* **VCSG\_LATITUDE** -
* **VCSG\_LONGITUDE** -
* **VCSG\_LRSNUMBER** – Effectively the same as LRSNUMBER
* **VCSG\_NORTHING** -
* **Weather** – Describes weather conditions at the time of the accident.
* **X** – Horizontal coordinates equivalent to Easting
* **Y –** Vertical coordinates equivalent to Northing
* **Year** – Year for row

Notes:

Dropping:

Dropped Columns:

* ACCIDENTTI – Somewhat redundant with ACCIDENTDATE, and only present in 2015
* AOTROADWAYGROUPid
* AOTROUTEid
* CITYORTOWNid
* MILEMARKER1
* MILEMARKER\_1
* MILEMARKER\_2
* NUMBER
* NUMBER1
* NUMBER2
* NUMBER3
* OBJECTID
* RDFLNAME – Somewhat redundant with STREETADDRESS, and missing for 2015
* REPORTINGAGENCYid
* VCSG\_AOTROUTEid
* VCSG\_CITYORTOWNid
* NORTHING -Redundant with Y but with more nulls
* EASTING – Redundant with X but with more nulls.
* DIRFROMNEA – Missing too many values
* CrashType – Redundant with InjuryType
* DISTANCE\_1 – Only in 2015, and use unknown
* POSTEDSPEE – Only present in 2015
* VCSG\_LATITUDE - Redundant with LAT\_DD
* VCSG\_LONGITUDE – Redundant with LONG\_DD

Dropping Notes:

* ID columns are dropped in favor of names.
* X & Y dropped in favor of Easting and Northing.
* Columns which are not useful such as Numbers need dropped.
* Columns with excess null values that cannot be amended will need to be dropped.

Merging:

* DIROFCOLLI + DirofCollision
* GIS\_LATITUDE + LAT\_DD
* GIS\_LONGITUDE + LONG\_DD
* VCSG\_LRSNUMBER + LRSNUMBER
* REPORTINGA + REPORTINGAGENCYid
* REPORTINGAGENCY + ReportingAgency
* VCSG\_EASTING + EASTING
* VCSG\_NORTHING + NORTHING
* VCSG\_AOTROUTE + AOTROUTE
* VCSG\_CITYORTOWN + CITYORTOWN
* AOTROADWAY+RoadGroup

Merging Notes:

* Geographic data primarily needs consolidated to LAT\_DD and LONG\_DD & Northing and Easting