# Support Information Q&A for Using MassDOT IMPACT Crash Data Files
## As of 2/5/2021

**What is the source of the data?**
Crash data are derived from the reports that Massachusetts local, State, MBTA Transit and some Campus police agencies send to the MassDOT Registry of Motor Vehicles (RMV) Division. These reports are entered into the RMV's Crash Data System (CDS). The MassDOT Highway Division's Traffic Engineering and Safety Section's IMPACT (Interactive Mapping Portal for Analysis and Crash Tracking) derives its data from that system. These systems are dependent upon the cooperation of police agencies in sending crash reports to the RMV in a complete and timely fashion.

**What crashes are included?**
The reporting threshold is any crash involving an injury or fatality, or damage to any one vehicle or other personal property in excess of $1000. Crashes not on public ways or in off-street parking lots are usually excluded, however some that were entered have been identified with the text "In Parking Lot" (or similar text) in the Landmark field, and have had crash coordinates removed, but many others have not been detected and identified in this manner.

**What years of data are available?**
The IMPACT crash data portal contains data for years from 2002 to crashes that have been entered and accepted into RMV's CDS within the past several days. Data for years prior to 2002 are in a different format and are available upon request.

**How current are the data?**
This depends on how frequently police agencies send crash report data to the RMV. Some agencies send reports every few days, others may send reports once a month or quarterly. Most reports are sent to RMV electronically, but sometimes technical issues prevent reports from being successfully received by RMV. Some agencies still send paper reports to the RMV, which must be keyed into the RMV CDS manually by clerks. Fatal crashes can require investigation and crash reconstruction, so detailed information for those crashes in IMPACT can experience significant delays.

**What are "Closed" and "Open" years of crash data?**
Crash data in IMPACT are subject to change until an entire year of data are "closed" by the RMV. Until that time additional reports may be added or deleted, and revisions may be made to the data that were previously entered. As of February 2021, the latest year of closed data is 2018.

**Are there gaps in the data?**
Yes. "Open" years of data may be incomplete. Even "closed" years of data may be incomplete if police agencies failed to submit reports, or complete data for all crashes, to the RMV. Users of crash data may want to check for gaps by using the IMPACT Crash

Data Portal (CDP) Crash Tabulation and Charting tool, or by creating pivot tables within Excel crash data files to identify potential monthly or yearly gaps in reporting by a particular police agency. The Boston Police Department does not send most crash reports to the Registry of Motor Vehicles, so the crash data files are missing a significant number of Boston reports.

**Are operator (motorist) reported crashes included?**
Not since 2010, when the RMV stopped entering operator-only reports (where there was no police report for the crash) into the CDS database. Since that time, unless a police report was submitted to and entered by RMV, it will not be included in the IMPACT data.

**Are there any duplicate crashes in the data?**
Attempts have been made to prevent duplicate crashes from appearing in the crash data, however sometimes they occur. If the crash date, time, and location are identical (or nearly identical) the crash may be a duplicate, but with a different RMV crash number. Oftentimes placeholders are added into the CDS system by the RMV for fatal crashes with minimal information, so when the full police report comes in, it appears as a duplicate until one is manually removed and often that does not happen until the end of the year. So frequently duplicates for fatal crashes do occur initially. It is possible that multiple crashes involving different vehicles occurred at the same location at the same time due to, for example, poor road surface conditions or secondary crashes resulting from an earlier crash. Duplicates can result from multiple submissions of reports of the same crash (sometimes with revisions) to RMV. These can usually be detected by checking the Crash Report IDs in the IMPACT CDP Crash Query and Visualization (Q&V) Tool (see the question, "What if I am from a police agency and want to reconcile IMPACT data with the data in my agency's records management system?" below).

**Is more data available for each crash than is contained in the Extract file?**
Yes. The Data Extract standard report file contains 25 columns of data for each crash. By using the IMPACT CDP Q&V Advanced Search query type, users may obtain over 160 data fields for each crash, including linked roadway attribute data, at either the crash level, vehicle level or person level. Or, the Open Data Portal can be used to export all the data for a specific year and data type (crash, vehicle, person) – see the next section. The standard extract file contains only crash level data (each crash is a separate row regardless of how many vehicles or people were involved in the crash). No personally identifiable information (PII) are available from the IMPACT crash data portal.
The link to the IMPACT Crash Query and Visualization Tool is:
https://apps.impact.dot.state.ma.us/cdv/
and use "Enter as Guest."
A partial listing of available data fields and values available in IMPACT can be viewed at: https://apps.impact.dot.state.ma.us/fdl/

**Can I obtain Statewide crash data for an entire year?**
Yes. Use the IMPACT Open Data Portal at this link:
https://massdot-impact-crashes-vhb.opendata.arcgis.com/

Complete data for each crash, including linked roadway data where available, are available from this site so the data files are quite large.

**How do I obtain a copy of a police report for a crash?**
Police reports of individual crashes may be obtained for a fee from RMV (not from IMPACT). Reports may be requested online from the RMV's website at:
https://www.mass.gov/how-to/request-a-crash-report

**What if I am from a police agency and want to reconcile IMPACT data with the data in my agency's records management system?**
Use the Crash Q&V Tool (see link above) and select "Crash Report IDs" from the checkbox list of crash data level fields in the data fields you select. This is your agency's report ID number. For agencies that send paper instead of electronic reports to RMV, there may be issues with using this data field that may require some manual reconciliation. "Police Agency Type" (e.g., Local vs. State Police) can also be selected from the checkbox list or included in the search criteria, as well as "State Police Troops" for State Police reported crashes.

**Is the crash location data for each crash in IMPACT identical to what was submitted in the police agency report?**
Not necessarily. As part of the review and geocoding process (see below), the MassDOT Highway Division Traffic Engineering and Safety Section tries to locate (geocode or attach geographic coordinates to) each crash using computer maps to enable spatial analysis of the data. This is done through a combination of automated and manual review processes. Crashes that require manual review by traffic engineers and analysts for geocoding often have the crash location data in IMPACT edited based on information in the diagrams and narratives in each report, and usually use the roadway names from the MassDOT Office of Transportation Planning Statewide road inventory file. Often intersecting roadway names, address numbers or landmarks that appear in diagrams or narratives can be added to supplement or clarify the police provided location data as part of the geocoding process, and to improve the quality of crash location data for users. When there are differences between data in the Location section of the police report and the diagram and narrative, the location data are often edited to enable the crash to be geocoded. In cases of major inconsistencies between these different sections of a police report, the narrative/diagram is often used to override the police-provided location section information. This can even include a change to the reported city/town where the crash occurred. However, if anyone obtains a copy of the original police report from the RMV, that report contains the location information as originally submitted by the police agency to RMV.

**How are crashes geocoded in IMPACT?**
It is very important to traffic engineers at MassDOT as well as for public users of crash data to have as accurate a description of the crash locations as possible in order to develop plans and projects to implement safety improvement programs. About 80% of all reported crashes can be automatically geocoded (located) to a point using Geographic Information System (GIS) tools to match the police provided location data to the MassDOT Planning road inventory file using city/town, roadway names, route numbers,

distance and direction from intersection, route and milemarker, and route and exit number.  Crashes with locations referenced by an address can also be automatically geocoded using street address databases, however addresses can often be somewhat imprecise.  The remaining approximately 20% of records require manual review to try to geocode as many crashes as possible.  Crashes that are included in CDS and IMPACT that are found during manual review to have occurred in parking lots or private ways may have their coordinates removed.  Not all crashes are able to be geocoded due to incomplete or ambiguous crash location data in the police reports, but overall, about 95% of all Statewide crashes are able to be geocoded to a point using a combination of automated and manual review processes.  This is a Statewide percentage and may vary by police agency or type of roadway.  The IMPACT database includes all reported crashes whether geocoded or not, but obviously crashes without coordinates will not appear on the maps and cannot be selected using spatial query tools.  There are delays between the time crashes are entered in the IMPACT database and when they are geocoded (if possible) if geocoding must be done manually.

### What are the location coordinates?
In the Extract files, geocoded crashes have values in the "X Coordinate" and "Y Coordinate" columns (the last two columns in the file).  Coordinates are shown only for crashes that were successfully geocoded to a point or to an approximate point based on available crash location data.  Users should be aware that for many crashes (especially ones located at or near an exit, highway interchange, rotary, etc.) these coordinates are only approximate, depending on the quality of the source crash location data.  Crashes referenced by an exit number or interchange may in fact actually have occurred some distance from that exit or interchange.  Coordinates are in Massachusetts Mainland State Plane NAD 83 meters.  Users of IMPACT data may obtain coordinates in the form of Latitude/Longitude by using the Crash Query and Visualization Tool (see link above) and using the "Latitude" and "Longitude" data elements (these are pre-selected as defaults) from the checkbox list of crash data level fields in the data fields you select.

### How can I see the location coordinates using the IMPACT mapping tool?
Use the Crash Q&V Tool (see link above) to first select data fields of interest, then select the Query Type, then Visualize Results.  The selected crashes that meet the search criteria and that have been geocoded will appear on the map.  If the Query Type is selected as Spatial Search the user can use the drawing tools to define the spatial area of the search such as to a specific intersection or section of roadway.  This will limit the search results to the geocoded crashes within the geographic area defined by the drawing tool.

### Are there special issues involving crashes reported at milemarkers and exit numbers?
Yes.  In some cases, crash locations that were described by exit numbers or milemarkers were found to be in a different city/town from that reported by police.  Corrections/edits were made by MassDOT Highway staff when crash location data could be reconciled from the information contained in police narratives and diagrams.  In some cases, conflicting or ambiguous information could not be reconciled.  In many cases, crash location information was not complete enough in police reports to enable geocoding to a

point (for example, the roadway or route was provided but not a specific location along that roadway or route).

Often, crash locations on Interstate and other limited access highways as provided by police were limited to route and nearest exit number, without sufficient information as to distance and direction from the nearest exit. In some cases, the exit or milemarker reported was in a different city/town than contained in the police report. In such cases, the crash location data have often been edited in the Landmark box to indicate "VICINITY OF", "NORTH OF", "PRIOR TO", etc. using the limited (often ambiguous) information that could be gleaned from police diagrams and narratives, to provide at least some information to data users as to crash location, even if geocoding to a point was not possible. "VICINITY OF MM __" (Milemarker) in the Landmark field usually indicates that the police-provided route/milemarker and city/town combination could not be validated with the milemarker data in the MassDOT Planning road inventory file linear referencing system. When the police-described location was in a different city/town than the Planning road inventory file indicated, the geocoded point could not be determined with sufficient confidence. In some cases, the police report provided a whole milemarker (rounded to the nearest mile) when if the tenths-of-a-mile decimal number had been provided the milemarker location would be in a different city/town.

In November 2020, MassDOT Highway began the process of renumbering exits on Interstate and major limited access highways based on milemarkers instead of consecutively numbered exits along each route. As route signs are replaced, they will contain both new and old exit numbers to facilitate this transition. Although the stated intention of State Police command is to have troopers continue to use old exit numbers, this has not been universally applied with crash reporting. As of February 2021, the sign renumbering work is continuing. IMPACT automated geocoding tools are still based on using the old exit numbers, so as of this time the intention is to continue to include the OLD exit numbers in the IMPACT crash database, with perhaps occasional references to new exit numbers in the Landmark data field. So, during this transition, both new and old exit numbers will appear in the crash data.

**<u>Who do I contact if I have questions about the data or have issues using the IMPACT portal?</u>**
If you have questions, please contact [DOTCrashDataRequest@dot.state.ma.us](mailto:DOTCrashDataRequest@dot.state.ma.us)
As of February 2021, due to staff temporarily working from home this email address is not checked every day so please have patience.