Support Information for Using Year 2002, 2003 and 2004 MassHighway Crash Data Files

As of 12/28/2005

Note: This document pertains only to year 2002, 2003 and 2004 crash data.
See the file named Support_Information_10-02a for crash data for earlier years.

The year 2002, 2003 and 2004 crash data files from Mass. Highway Dept. are significantly different than data files for previous years. Crash data beginning with year 2002 are derived from a new Registry of Motor Vehicles (RMV) computer system called the Crash Data System (CDS). Entirely new forms were used to collect crash reports from police and operators, replacing the old forms that had been in use with only minor changes over many decades. CDS is an entirely new RMV Oracle database, which replaced the former COBOL Accident Records System (ARS).

This document summarizes some of the many differences between crash data for 2002-2004 and prior years and contains some caveats in the use of the data.

- CDS contains many new and redefined data elements, developed in accordance with NHTSA (National Highway Traffic Safety Administration) guidelines called the Model Minimum Uniform Crash Criteria (MMUCC).
- There are fewer crashes in the 2002-2004 statewide data than in files for prior years. The Statewide 2002 file contains a total of 139,036 reported crashes; year 2003 contains 141,676 crashes; and year 2004 contains 138,635 crashes. In comparison, the 2001 file contained 181,247 records, and the 2000 file contained 189,046 records. So, users should be aware that comparing raw crash count numbers for 2002-2004 with prior years may not produce a valid comparison of crash frequency between years. An Excel file named Total Crashes by Town and Year1990-2004.xls is available to show the differences in total crash reports by city/town from year to year.
- The data fields used to collect crash location data were significantly changed. Prior to 2002, only five fields were used to collect location data (roadway, nearest intersecting roadway or landmark, distance and direction from nearest intersecting roadway, and milemarker). The new CDS data contain a total of 30 crash location fields. Now there is a separate field for each type of discrete data (roadway names, route number, address number, milemarker, exit number, landmark, etc.) so that each can be more easily validated and edited as necessary. This was done in an effort to improve the quality of the crash location data by standardizing the spelling of roadway names, and by attempting to collect and capture more complete and accurate crash location data, which is of special interest for traffic engineering purposes. The 30 data fields have been consolidated (concatenated) into a smaller number of crash location fields in the crash reports. These consolidated fields are grouped around the different location referencing methods used for locating crashes (roadway/nearest intersecting roadway; route/milemarker; route/exit number; roadway/street address number; and roadway/landmark.
- If the crash date, time and location are identical (or nearly identical), the crash may be a duplicate, but with a different crash number.

- The format of the Mass. Highway standard crash report is very different than in prior years, due to the changes in data elements and values. Consequently, the user may encounter some formatting challenges when trying to view the report on the screen, and when trying to print it.
- The distinction between AT INTERSECTION and NOT AT INTERSECTION crashes is often not a consistent one. This is because persons filling out the report may have had different interpretations of the definition of an intersection. Crashes at interchanges, ramps, rotaries, and driveways could have been reported either at an intersection or not at an intersection. Sometimes the NOT AT INTERSECTION side of the form was used for the convenience of reporting crashes referenced by a street address number, milemarker, ramp, or landmark. Some crashes contain data for both AT INTERSECTION and NOT AT INTERSECTION.

Sorting/Searching Information Regarding MassHighway Crash Data

The data MassHighway has supplied is in a Microsoft Excel 2000 format. Due to the different format of the 2002-2004 data from prior years, sorting the data by location will be difficult because of the five different columns containing location data. Crash data are not completely standardized. Several different variations of a street name (or other field) may exist. Due to the format of the year 2002-2004 data, searching may be more useful than sorting. Search all five crash location columns for each occurrence of a street name.

When selected records are found, they can be copied and pasted into another Sheet in the same Workbook.

Explanation of columns and abbreviations in Excel Spreadsheets

A. **Crash Number** – Unique number used by Registry of Motor Vehicles to identify each crash. Each crash could have several reports (police, operator(s), so this is the master record ID number. There is no relationship between this number and local police department incident numbers.
B. **City/Town Name** – The city or town in which the crash occurred. If the crash was reported as occurring in a locality (neighborhood name) within the city or town, this is shown in parentheses after the proper city/town name. However if the crash was just reported as occurring in the city/town (rather than in a locality/neighborhood), then the locality name is not shown. (Therefore, searching/sorting by locality name will not identify all crashes that actually occurred in that locality.)
C. **Crash Date** – Date of the Crash
D. **Crash Time** - Time of Crash
E. **Crash Severity** – Type of Crash
    - Fatal injury
    - Non-fatal injury
    - Property damage only (none injured)
    - Not Reported
    - Unknown

F. ***Number of Vehicles*** – Total number of vehicles involved in the crash
G. ***Total Nonfatal Injuries*** - Number of persons injured in the crash excluding fatalities
H. ***Total Fatal Injuries*** - Number of persons killed in the crash
I. ***Manner of Collision*** - Manner of Collision or Collision Type
- Angle
- Head-on
- Rear-end
- Rear-to-Rear
- Sideswipe, opposite direction
- Sideswipe, same direction
- Single vehicle crash
- Unknown
- Not reported

J. ***Vehicle Travel Directions*** – Direction that each vehicle was traveling at time of the crash; V1 = Vehicle 1, V2 = Vehicle 2, etc.
K. ***Most Harmful Events*** – Most harmful event for each vehicle
- Collision with motor vehicle in traffic
- Collision with parked motor vehicle
- Collision with pedestrian
- Collision with cyclist (bicycle, tricycle, unicycle)
- Collision with animal – deer
- Collision with animal – other
- Collision with moped
- Collision with workzone maintenance equipment
- Collision with railway vehicle (train, engine)
- Collision with other movable object
- Collision with curb
- Collision with tree
- Collision with utility pole
- Collision with light pole or other post/support
- Collision with guardrail
- Collision with median barrier
- Collision with ditch
- Collision with embankment
- Collision with highway traffic sign post
- Collision with overhead sign support
- Collision with fence
- Collision with mailbox
- Collision with impact attenuator/crash cushion
- Collision with bridge
- Collision with bridge overhead structure
- Collision with other fixed object (wall, building, tunnel)
- Collision with unknown fixed object
- Overturn/rollover

- Fire/explosion
- Immersion
- Jackknife
- Cargo/equipment loss or shift
- Other
- Other non-collision
- Unknown non-collision
- Unknown
- Reported but invalid

L. Road Surface Condition –The condition of the road's surface at the time of the crash
- Dry
- Wet
- Snow
- Ice
- Sand, mud, dirt, oil, gravel
- Water (standing, moving)
- Slush
- Other
- Unknown

M. Ambient Light – Light conditions
- Daylight
- Dawn
- Dusk
- Dark – lighted roadway
- Dark – roadway not lighted
- Dark – unknown roadway lighting
- Other
- Unknown

N. Weather Condition – A maximum of two weather conditions may be reported
- Clear
- Cloudy
- Rain
- Snow
- Sleet, hail, freezing rain
- Fog, smog, smoke
- Severe crosswinds
- Blowing sand, snow
- Other
- Unknown

O. At Roadway Intersection – If crash location information was entered in the AT INTERSECTION side of the report, the route numbers and/or roadway names will appear in this column. The route/roadway where the crash occurred will appear first, followed by a slash (/), followed by up to two combinations of routes and/or roadway names.

P. *Distance from Nearest RoadwayIntersection* – If crash location information was entered in the NOT AT INTERSECTION side of the report, and if the crash was referenced as occurring at some distance and direction from the nearest intersecting street, the crash location information will appear in this column. However, sometimes only a route and/or roadway name will appear, or other information such address numbers may appear in this column.

Q. *Distance from Nearest Milemarker* - If crash location information was entered in the NOT AT INTERSECTION side of the report, and if the crash was referenced as occurring at some distance and direction from the nearest milemarker, the crash location information will appear in this column. However, sometimes only a route and/or roadway name will appear, or other information may appear in this column.

R. *Distance from Nearest Exit* - If crash location information was entered in the NOT AT INTERSECTION side of the report, and if the crash was referenced as occurring at some distance and direction from the exit or interchange, the crash location information will appear in this column. However, sometimes only a route number or other information may appear in this column.

S. *Address/Distance from Nearest Landmark* - If crash location information was entered in the NOT AT INTERSECTION side of the report, and if the crash was referenced as occurring at a street address or at a landmark, or at some distance and direction from them, the crash location information will appear in this column. However, sometimes only a roadway name, route number, or other information may appear in this column. There may be some data in this column that duplicates data in other crash location columns. Landmark text is limited to a maximum of 32 characters (the portion enclosed by parentheses).

Note: Beginning with year 2004 data, column "S" has been divided into two columns, "S" and "T." Column "S" contains landmark data, and column "T" contains street address number and/or street name data.

Crash location data in columns O through T as described above will only be shown in the format described above if it was correctly entered by police or operators and/or RMV. Offset and/or direction of offset may be missing, or the nearest intersecting street/milemarker/exit number may be missing. Redundant location data often appears in multiple columns due to the format of the report and the desire to attempt to show all crash location data regardless of the type of referencing method used.