CARE CRASH ON-LINE ANALYSIS README FILE
Updated August 10, 2010

As you will see, there are a number of changes that have been made with regard to the Alabama CARE on-line analysis. This file will be used to keep you in touch with the most recent changes, and we expect that this will be an ongoing process. The following are frequently asked questions with regard to the use of the on-line Dashboard and CARE.

How do I get to CARE?

The system defaults to a mode called "Dashboard." There are some tabs at the top of the screen to enable you to choose the *dataset* that you want (explained in more detail below). If you know which dataset you want, *select that first.* Then, CARE is included in the next row of buttons. Click on that and you will be presented with an interface that essentially walks you through the CARE query requirements similar to those of the CARE desktop system.

Why the Dashboard?

View it as a new interface for CARE. Try getting what you are interested in off the Dashboard first before clicking on CARE. Click on the various dropdown menus and try things out. We think this will answer many if not most of your question. It shortcuts the CARE user interface and enables you to see what the output will look like before you ask for it. We realize that this may not give you all that you need, and if not, then click the CARE button as described above. If you have suggestions as to ways we can improve the interface, please e-mail us using the contact address at the bottom of this document.

What do you mean by eCrash and pCrash?

Terminology:

- eCrash – the new *electronic crash* reporting system that should be fully implemented by January 1, 2011. The eCrash system was implemented totally by DPS and several local agencies on June 1, 2009. However many local agencies have temporarily remained with the pCrash forms for a variety of reasons.
- pCrash – this will be used to reference the old paper-forms-based reporting system that is expected to be totally phased out on January 1, 2011.

From the above it should be clear that the 2009 data available for on-line processing were obtained partially from eCrash and partially from pCrash. And, for that matter, any datasets made in the future that will include either or both of these years will also have data from both the eCrash and the pCrash systems.

Does this mean that you will need to run two different CARE runs to produce information for these years?

No. We have created what we are calling "Integrated" datasets that include both the eCrash and the pCrash data in single datasets.

How was this integration accomplished? – are there not some variables and codes that are strictly eCrash and others that are strictly pCrash?

The answer to the second question is yes – DPS wisely chose to go to a MMUCC-compatible set of variables and codes when they made the modifications in the eCrash report. Almost everything changed in some way, shape, or form (with the obvious exception of age, day of the week, month, etc.).

To answer the first question, the following rules were observed when there were discrepancies:

- Variables/attributes that are now in eCrash but were not in pCrash – these are prefixed with an E.
- Variables/attributes that are were in pCrash but are no longer in eCrash – these are prefixed with a P.
- Codes within variables/attributes that are now in eCrash but were not in pCrash – these are prefixed with an E.
- Codes within variables/attributes that were in pCrash but are no longer in eCrash – these are prefixed with a P.
- Variables and/or codes that are common to both systems do not have a prefix.
- * – an asterisk on a code indicates that there may be another variable that addresses this same code value.

This means that in some cases where things were similar *but not identical* that it may be necessary for you to combine some of the p and e code values so that you produce the results that you are seeking.

Does this mean that many of the variable counts will not be comparable between 2008 and 2009?

Yes, many of them will ***not*** be comparable. The MMUCC-compliant codes are in many cases quite different from those in the pCrash system, and an entirely new set of rules exists for officers entering crash reports that did not even exist for the pCrash form. In fact, there is no report form now – it is all paperless except for the hard copy produced for the public.

Please do not attempt to compare 2009 CARE results with existing 2008 CARE results unless it is for those variables that are obviously comparable (such as ages, dates, time of day, day of the week, etc.). While we have lost the ability to make some of these comparisons, we have gained a tremendous amount in terms of federal standardization, accuracy, completeness, consistency and timeliness. In a few years the comparability problem will be negligible. Progress has a price.

Does this mean that the results from the pCrash system were wrong?

No. *Different* does not mean wrong. The variable structure in most cases, the particular codes, and the procedures that are being applied now are different, and so the numbers in many very similar variables will not be exactly the same as they would had pCrash remained as the sole data collection method for the state.

Does this not affect the filter definitions as well?

Yes. The integrated datasets have required that all new filters be created. Most of this has been taken care of in the standardized filters that are available on the dropdowns. However, if you created special filters in the past, they will not be able to be "converted" for use in the integrated datasets. It will be necessary to recreate them.

Why are there three different datasets for any given time period?

All three of these are essential, depending on what you want to count. Recognize that all computer numerical outputs are counts of something. It is essential that you as a user understand exactly what it is that is being counted. What do the outputs mean? The following answers this questions for each of the output types:

- **Crash.** These named datasets count crashes – they do not count injuries, fatalities, occupants or anything else. For example, the day-of-the-week frequency output tells how many crashes happened on those respective days. This is no doubt the most used dataset. The driver information in the crash dataset is generally for the causal driver of the crash, although there are some second driver variables for the crashes. However, there is no way to get information from all drivers or all vehicles from the Crash datasets.
- **Driver-Vehicle.** These datasets produce information by counting "units," and since there is usually one driver and one vehicle associated with each of these units, information on them has been consolidated into a single dataset. For example, from these datasets it is possible to obtain age or gender distributions for all drivers involved in all crashes (not just the causal driver), or the make and year of all vehicles involved in crashes. This information cannot be obtained from Crash datasets.
- **Person.** These datasets count the number of persons with various characteristics, including all drivers, vehicle occupants, pedestrians and other non-motorists involved in crashes. For example, we can get the severity of injury of all persons involved in crashes. This is the only dataset that can produce this information.

Are there any exception to these definitions? For example, can't we get the number of fatalities from the Crash datasets?

That is not really an exception. There are several Crash variables (C051-C060) that count "numbers of" things. They do not provide any information on them other than the raw counts. For example C060 counts the number killed. How does it do this if all the Crash datasets do is count crashes? Answer: it counts the number of crashes in which there was one fatality, the number in which there were two fatalities, the number in which there were three fatalities, etc. If you want to know the total number of fatalities (within the constraint of the current filter), it will be a two step process:

- Run a frequency on C060 (Number Killed); then
- Multiply out the number of crashes for each the categories.

The above recognizes that the Crash dataset is still counting crashes, and only indirectly enabling the calculation of other metrics, most of which can be obtained directly using the other dataset types.

What is FARS data and why is it there?

FARS is the Fatal Analysis Reporting System, which is a federal (NHTSA) system for collecting and providing common data on fatal crashes throughout the country. We had put it under CARE as a service to those who might be familiar with CARE and wish to use it to process FARS data. By using IMPACT it is possible to compare Alabama with National results or Alabama with any other state.

Why is the on-line system limited to a single year and being able to do only one query at a time?

There are some features of the on-line system that make it superior to the desktop system, like the ease of getting results on the dashboard. However, we realize the advantages of the desktop version especially for intensive research work. We will seek over time to incorporate as many of these capabilities into the web version. We limited the on-line system to a single year so that there would not be confusion in interpreting what the outputs mean – all results for the Alabama Crash data are for 2009. The current system has been limited to one query at a time so that one user will not overwhelm its bandwidth requirements by asking for "everything."

What are some things that the desktop (downloadable) version of CARE is able to do that I cannot do with the on-line version.

The following is a partial list:

- Multiple year analyses;
- Location analyses (authorized users only);
- Filter generation directly from frequency or crosstab outputs;
- Filter combinations, conversions and imports;
- Multiple (e.g., all) frequencies in a given run;
- Information mining (variables ranked by significance) over the entire dataset; and
- Data and output exports (e.g., in Excel).

These restrictions, however, do not nullify the handy nature of the on-line version for the majority of simple CARE queries. The CARE desktop system is available on this website under the DOWNLOADS button.

Where can I get additional information?

Email care@cs.ua.edu or call 205-348-7920.