Analysis of the Most Critical Factors in Teen (15-19) Driver Causes Vehicle Crashes

Jesse Norris and David Brown

November 17, 2011

This report was done at the request of William D. King, RPh, MPH, DrPH, Professor of Pediatrics, Division Director, Southeast Child Safety Institute.

The purpose of this report is to document and further explain the results of an analysis of data from 2008 through the most current data that were available within CARE as of the date of this report. The summary of the results of this analysis, which considered all attributes in the CARE database, is given in Table 1. This is also available in an Excel spreadsheet file which accompanies this report. Note that while there were a very large number of variables/values that were found to be significant, these seemed to be the ones that could best be controlled in some way (education, enforcement, etc.).

Please observe the following aspects of Table 1:

- Attributes are given by Variable and Value (highlighted in yellow). All attributes within a given variable are grouped together and then ordered by the Max Gain (explained below).
- Values prefixed by an E are strictly from the eCrash system; while those prefixed by a P are from the paper forms based system; if they have no prefix then the value is common to both the E and P systems.
- The "Subset" for this analysis was created by a filter that only allowed teen driver caused injury (including fatal) crashes. This provided the basis for the analysis even though the original request was to deal with fatalities. There are not enough teen-driver caused fatal crashes to produce statistically significant results for most attributes. This justifies the use of all injury (including fatal) crashes, coupled with the fact that injury crashes predict the attributes of fatality crashes as well, if not better, than fatality crashes alone. Note that in the remainder of this report we will use the term "injury" to refer to include fatal and non-fatal injury.
- The "Other" columns provide a control to which the "Subset" columns were compared. In this case the "Other" columns represent the subset formed from all injury crashes that were not caused by teen drivers. Thus, we are comparing injury crashes caused by teen drivers to injury crashes that were not caused by teen drivers.
- The rationale for this comparison is that it highlights where teen drivers are doing things differently from their older driver counterparts. We further reason that injury crashes and fatalities can be reduced if we are able to get the teen drivers to behave like older drivers.

- The "Significant" column indicates whether there was a statistically significant difference found between the Subset Percentage and the Other Percentage. This was found to be true for all of the attributes listed. This indicates that teen drives are behaving in a different way when it comes to all of these crash attributes, i.e., the differences observed are not just due to chance.

Table 1. Most Critical Teen Driver Caused Crash Factors

Highlighting Results	2008-2011 (to current) Data	Subset	Subset	Other	Other		Odds	Injury
Variable	Value	Frequency	Percentage	Frequency	Percentage	Significant	Ratio	Max Gain
C015: Primary Contributing Circ	Over Speed Limit	1514	10.07%	8969	2.05%	TRUE	4.91	1205.39
C015: Primary Contributing Circ	Driving too Fast for Conditions	794	5.28%	14377	3.29%	TRUE	1.61	299.31
C015: Primary Contributing Circ	E Failed to Yield Right-of-Way from Stop Sign	635	4.22%	9895	2.27%	TRUE	1.87	294.53
C015: Primary Contributing Circ	E Failed to Yield Right-of-Way Making Left or U-Turn	415	2.76%	8172	1.87%	TRUE	1.48	133.82
C015: Primary Contributing Circ	E Fatigued/Asleep	221	1.47%	3472	0.79%	TRUE	1.85	101.53
C015: Primary Contributing Circ	P Vehicle Left in Road	181	1.20%	2916	0.67%	TRUE	1.80	80.67
C015: Primary Contributing Circ	E Aggressive Operation	145	0.96%	2127	0.49%	TRUE	1.98	71.81
C015: Primary Contributing Circ	E Failed to Yield Right-of-Way from Traffic Signal	183	1.22%	3469	0.79%	TRUE	1.53	63.64
C015: Primary Contributing Circ	E Distracted by Use of Electronic Commun Device	117	0.78%	1796	0.41%	TRUE	1.89	55.20
C015: Primary Contributing Circ	E Distracted by Use of Other Electronic Device	74	0.49%	728	0.17%	TRUE	2.95	48.95
C017: First Harmful Event	E Ran Off Road Right	483	3.22%	7506	1.72%	TRUE	1.87	224.47
C017: First Harmful Event	E Ran Off Road Left	227	1.51%	4096	0.94%	TRUE	1.61	85.92
C022: E Manner of Crash	Single Vehicle Crash (all types)	2989	19.89%	50945	11.66%	TRUE	1.71	1236.08
C129: CU Vehicle Maneuvers	E Negotiating a Curve	820	5.47%	8833	2.02%	TRUE	2.70	516.23
C129: CU Vehicle Maneuvers	P Avoid Object in Road	168	1.12%	2936	0.67%	TRUE	1.66	67.03
C202: CU Contributing Circumstance	E Over Correcting/Over Steering	256	1.70%	2478	0.57%	TRUE	3.00	170.68
C202: CU Contributing Circumstance	E Distracted by Passenger	103	0.69%	1610	0.37%	TRUE	1.86	47.57

- The Odds Ratio column indicates the extent of the difference found. It is just the Subset Percentage divided by the Other Percentage. As an example, the 4.91 in the "Over Speed Limit" row indicates that teen driver caused crashes are nearly five times more likely to be speeding relates than those caused by older drivers. This is a remarkable finding, and the most significant one found. Since speed is a proxy for risk taking in general, it shows the inclination of young drivers to take risks of all types.
- Injury Max Gain. This column indicates the number of crashes that could be saved over the approximate 3.75 years of the study had teen drivers had the same percentage of crashes of the corresponding attribute value that older drivers had. In the "Over the Speed Limit" example, this indicates that about 1,205 injury crashes could have been reduced if teen drivers had behaved like older drivers. The ordering of the output is based on this column.
- Output ordering. Note that the Injury Max Gain is ordered within variables. Values within variables are mutually exclusive. Values between attributes are not. For example, a crash could be caused by speed, be a run-off-the-road crash, and be single vehicle all at the same time. Thus, it is best to consider Max Gain comparisons within variables as opposed to those that might be between variables. The Max Gain results are not additive, at least not among variables; they are generally additive within a given variable.

- Fatal crash estimates. To estimate the number of fatal crashes from these data, recognized that there were a total of 15,031 injury crashes observed in the study, of which 293 were fatal. Thus the injury to fatality ratio is 15,031 to 293, or about 51 to 1. Thus, any of the teen-caused crash count number given in Table 1 can be divided by 51 to obtain an estimate of the number of fatal crashes. This would include the Max Gain column, since it is in units of teen-caused injury crashes reduced.

Considering the Max Gain column, it is clear that the most significant factor is speed. The first two items in the table relate to speed, and taken together they account for over 1500 crashes that could have been reduced by behavioral change. It should be clear that this would also reduce a disproportionately larger number of fatalities, since speed is both a cause of the crash (lack of control, especially for novice drivers) and an increase in the severity of the crash. Studies based on Alabama data have shown an approximate doubling of the probability of any crash being fatal for each ten miles per hour of impact speed. For a further analysis of speed and risk taking in young drivers, see: http://www.safehomealabama.gov/InfoTraining/YoungDriverIssues.aspx -- especially the short article at the bottom of the page.

Going down the list, virtually all of the factors involve risk taking of some sort. The only exception might be "Avoid Object in Road," which could be out of the control of the driver. While "Single Vehicle Crash" might not on the surface imply risk-taking, it does show the very high proportion of teen-caused crashes that are (in tennis language) un-forced. That is, they were not due to interactions with other vehicles, and thus they were almost totally within the control of the teen driver.

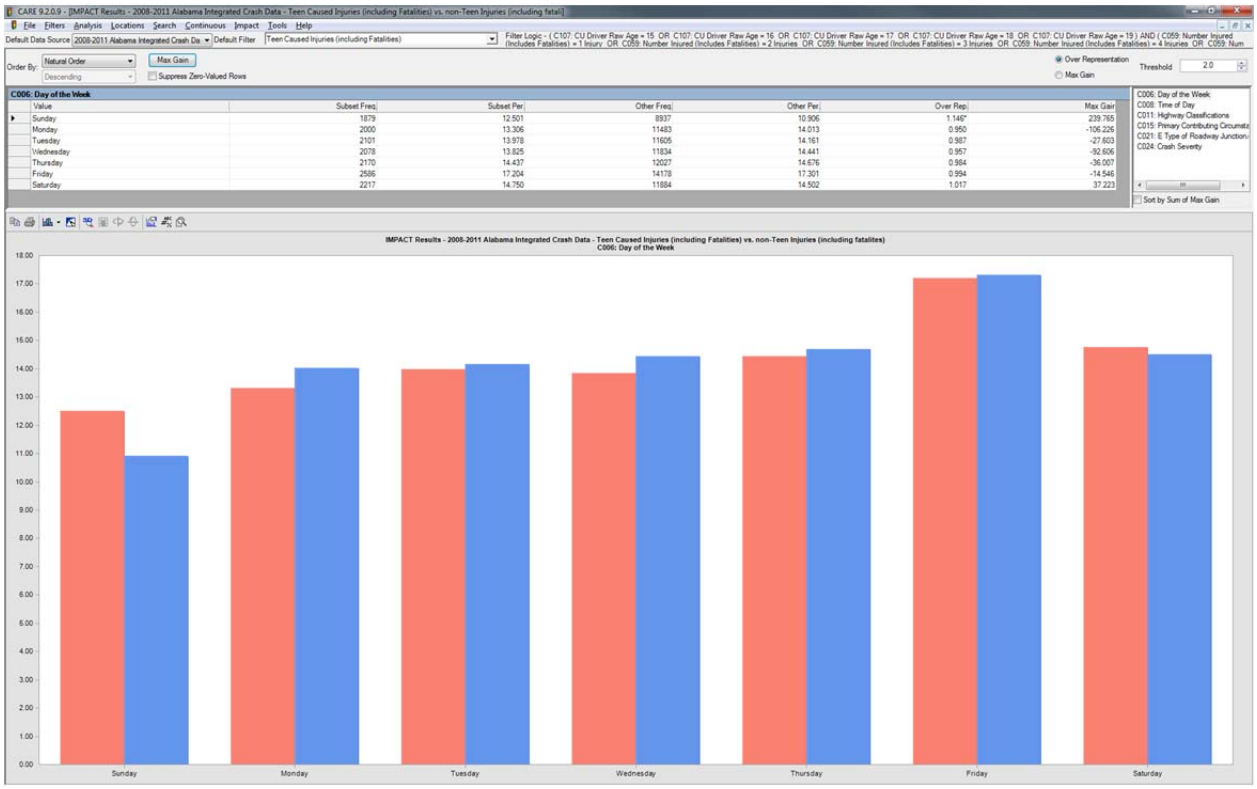
Distracted driving misinformation. For purposes of keeping this in front of us, we left the distracted driving categories in, although their numbers indicate them to be less than critical. However, these frequency numbers are misinformation. The only thing that is valid in these rows is the Odds Ratio, since the same errors are being made on both sides of the comparison. (IMPACT can salvage useful information from deficient data as long as we can assume that the reporting is consistent on both sides of the comparison.) The raw numbers display a gross (albeit explainable) under-reporting on the part of law enforcement (not their fault). The problem is with the competing values within the PCC variable. As an example, other competing values are the various "Failure to Yield …" values. Now, if you are an officer filling out the form and you know they failed to yield but were not sure if they were on a cell phone or even if the cell phone caused the crash, which one would you put in – you cannot put both of them in. We are working on a solution to this problem, which will be to create an entirely different variable which will be something like "Officer's opinion as to the electronic device distraction." Like the similar thing for DUI, it will not be something they will have to prove in court – just an opinion, but we could not ask for any better data except if it came from the cell phone companies. National studies

have generally shown that one out of nine crashes are caused by some form of distraction; we feel that this would be a conservative estimate.

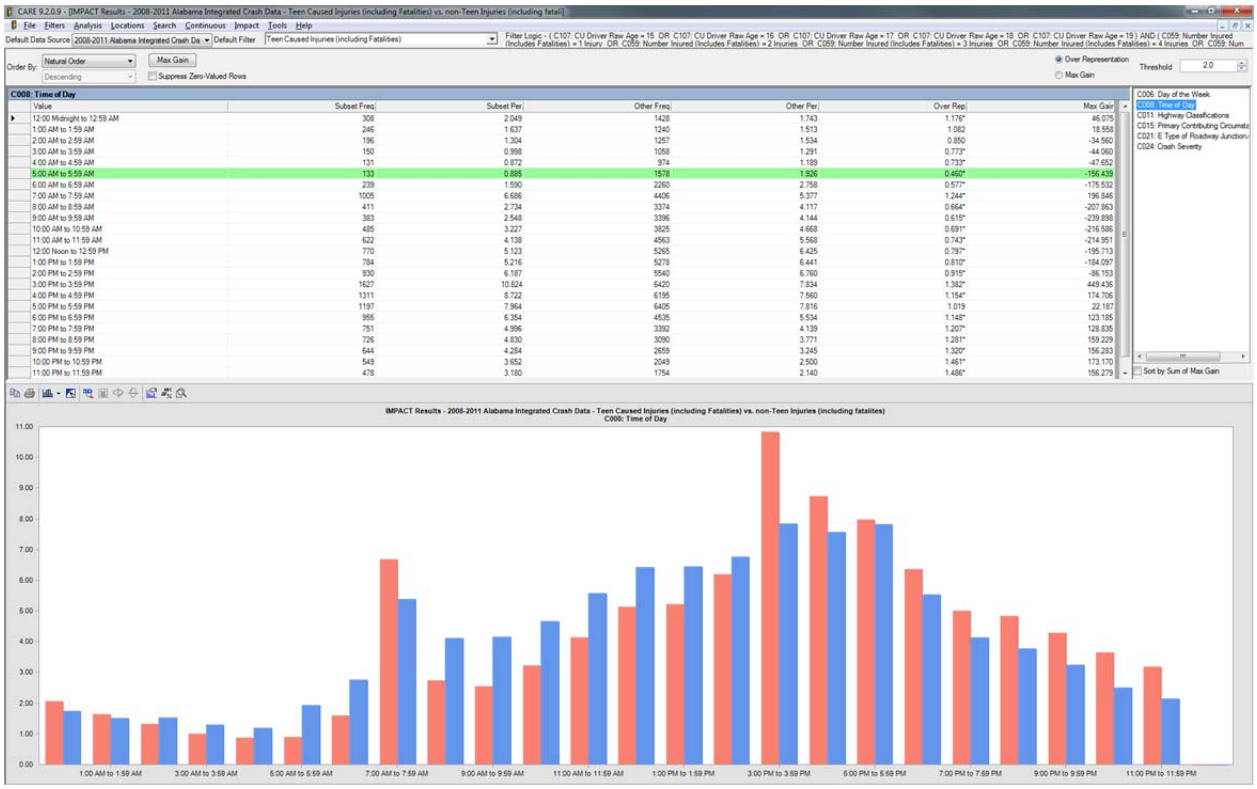
The following IMPACT outputs (red bars = teen caused; blue bars = older driver caused) may be of interest to those working with young drivers:

- Display 1. Day of the Week. There is not a large deviation from the general population. Typically crashes are down on weekends, but these are injury/fatal crashes that are being compared, so there would be a fairly large number of DUI caused crashes on the part of the older drivers on weekends and Friday nights.
- Display 2. Time of Day. Before and after school are over-represented; the later afternoon and night over-representations are probably due to teen driving on the weekends. If exact days and times were important for selective enforcement, they could be resolved with a cross-tab of time of day by day of the week.
- Display 3. Highway Classification. The over-representation on county roads is very important since these are our worst roads from both engineering and EMS points of view.
- Display 4. Primary Contributing Circumstances. These were largely captures in Table 1; this is a more complete summary.
- Display 5. Crash Severity for Non-PDO Crashes. Young people themselves tend to be survivors of crashes, and thus we see the fatal injury category quite under-represented. This variable is for the worst injury that occurred in the crash.
- Display 6. Driver Restraint Use Comparison. This compares restraint use of injured drivers who are 16-19 years old against injured drivers who are older. This is comparing number of persons who are restrained (or any of the other values) – it is not a count of crashes, as was the case in all of the IMPACT comparisons above. The top line is the most telling, with the younger drivers who are not using their restraints in a proportion that is about 32.5% higher than the older drivers. Also, unlike the comparisons able, this result is for *all* 15-19 year old drivers involved in crashes, and not just those who are causing crashes. This, however, has very little effect on the conclusions that can be drawn.

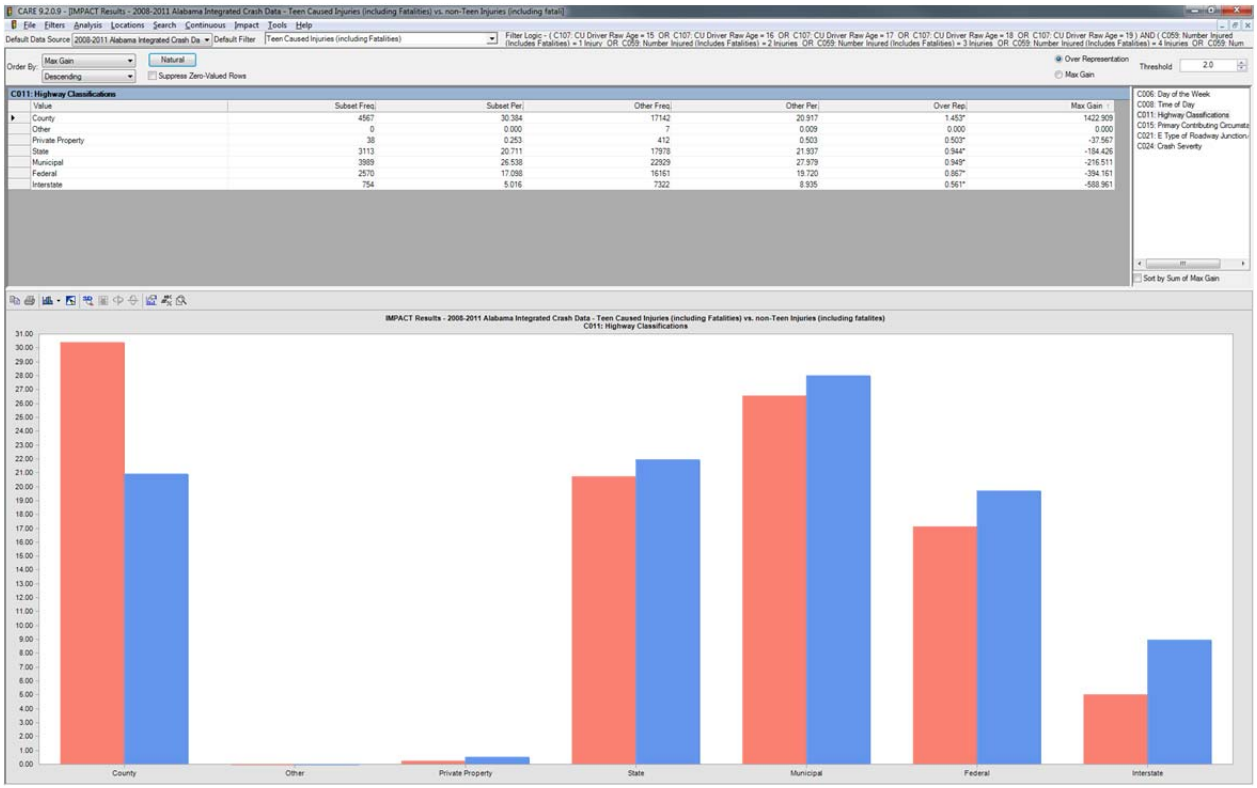
Display 1. Day of the Week



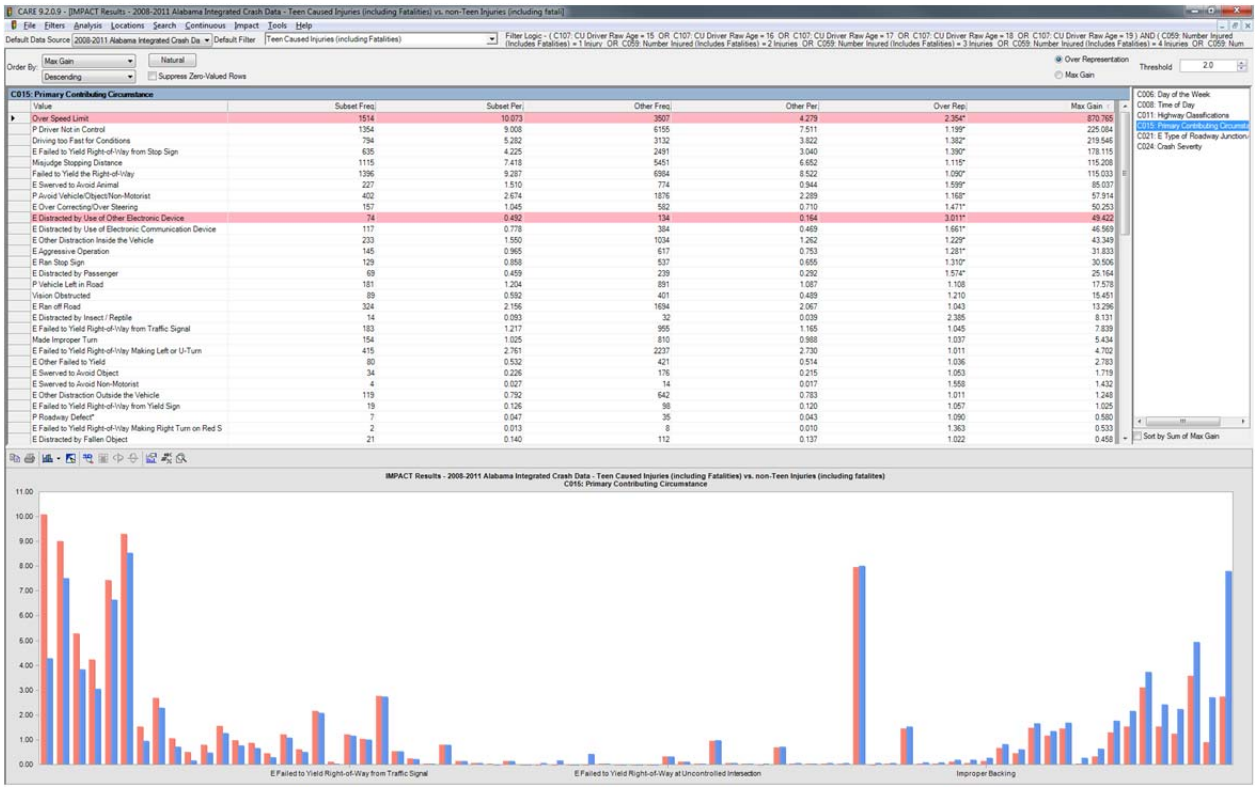
Display 2. Time of Day



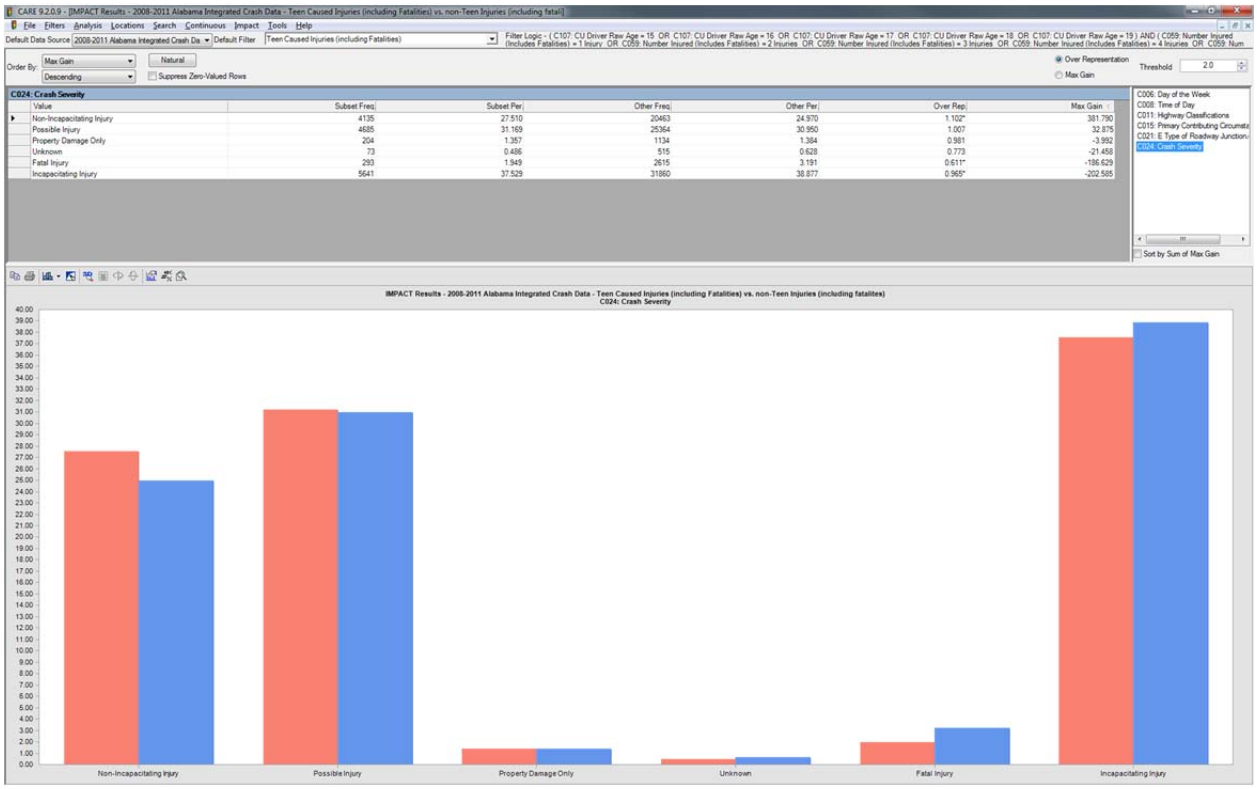
Display 3. Highway Classification



Display 4. Primary Contributing Circumstances



Display 5. Crash Severity for Non-PDO Crashes



Display 6. Driver Restraint Use