Worst Times of the Year for Young (Aged 16-20) Driver Involvement in Traffic Crashes

Dave Brown

September 17, 2015

Published News Summary

The last three months of the year can be treacherous for all drivers, but young drivers (aged 16-20) are involved in considerably more crashes during that time than the rest of the year. During the past five years in Alabama young drivers have been involved in about 22 percent of crashes, despite their accounting for a bit less than 7 percent of all drivers. Young drivers are also reported to cause more than 16 percent of all vehicle crashes, which is more than twice expected considering their number.

Over the past five years, young drivers have been involved in an average of 79 crashes a day, out of a total 355 crashes per day (on average). During the last three months this number jumps to 84 crashes per day on average. Traffic safety professional generally conclude that 16-20 is the most vulnerable age group due to a combination of inexperience and a willingness to take risks.

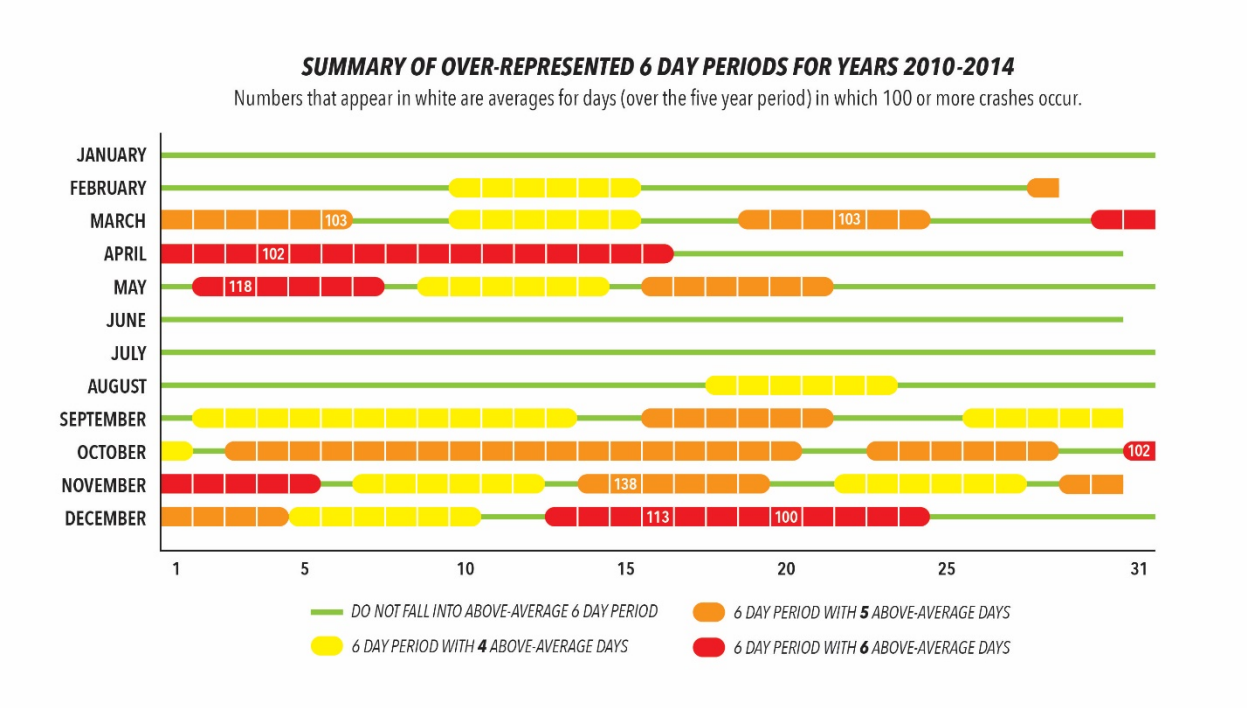
Daily crash data from 2010-2014 were added together and then grouped into six day chunks to see which of these had significantly more crashes than average. The six-day time frame tended to average various weather and outliers to smooth the data results, giving a clearer picture of which times of the year have more crashes than average. Exact 'worst day of the year' specifications are not very useful because one of the most critical factors in crash frequency is the day of the week. Fridays are generally the worst day, so the same date falls on a different day-of-the-week every year. Giving one date would not be a good way to forecast problem times.

The display below summarizes the results of the study. The green line indicates the ideal days to drive since young driver crashes are below normal. Yellow indicates slightly above, while the orange and red get progressively worse. Those days with a number are dates for which the total young driver crashes averaged over 100 per day.

Only 19 days out the 92 days in the last quarter of the year have days with less-than-average days for crashes involving young drivers. It is clear that October, November and December are collectively the worst sequences of three month over-involvement. Young drivers are not alone in this over-representation, since it was found that drivers of all ages have more crashes than average in November and December.

The worst two-week period for young drivers is the 12 days from Dec. 13 to Christmas Eve on Dec. 24.

The explanation for this probably has to do with the young drivers adding to an already high mix of people doing unusual things at this time – excess shopping, visiting, and general pre-holiday celebrations often out of their normal driving areas. While this time period is over-represented in general, but it is particularly problematic for young drivers.



During the rest of the year, January, February and the summer months of June, July and August are the months with the fewest crashes involving young drivers. The spring, March through May, have more crashes than average for young people, and especially the first half of April. This is probably attributed to spring breaks that span this time frame starting at the end of March going right through the middle of April.

This study employed the Critical Analysis Reporting Environment (CARE), a software analysis system developed by the University of Alabama Center for Advanced Public Safety (CAPS) to automatically mine information from existing databases. Crash records for the study were provided by the Alabama Law Enforcement Agency (ALEA).

Introduction

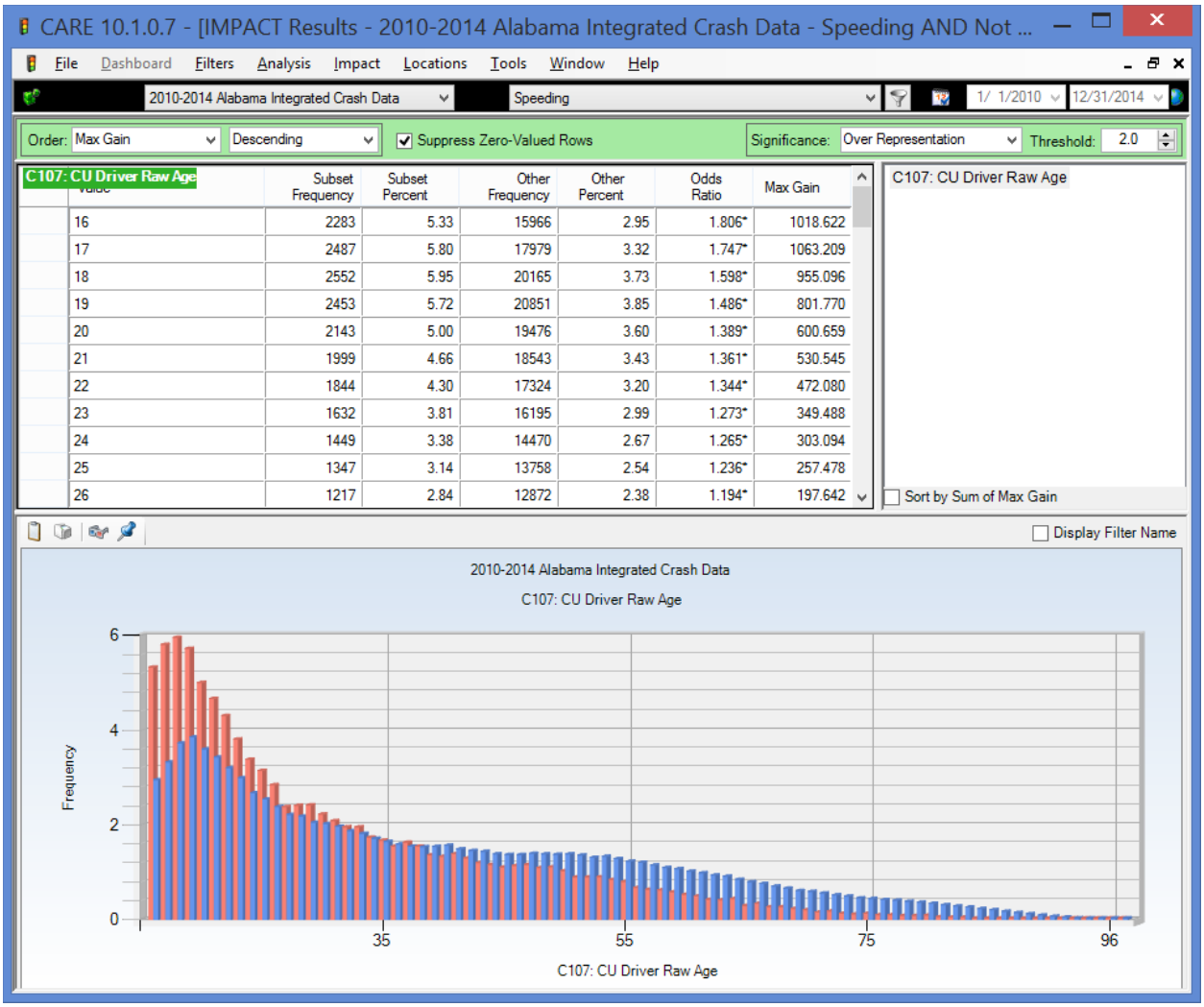
This report summarizes the findings of a research study performed by The University of Alabama Center for Advanced Public Safety (CAPS) on the subject of young driver crash involvement times. The goal of this study was to determine the worst days, weeks, and months of the year for young driver crash involvement, where "young drivers" are defined to be those of age 16-20 years, inclusive. This is important information to let young drivers and parents know when these young drivers are more apt to be in crashes so that there is additional awareness during these times; it also had the goal to raise awareness of the problems involved in young driver crashes in general.

General Issues of Younger Drivers

The study used crash data collected by investigating law enforcement officers and submitted to the Alabama Law Enforcement Agency (ALEA) central crash records database for calendar years 2010-2014 (five years). This included a total of 647,347 crashes for Alabama, of which 106,355 (16.4%) were *caused* by drivers in the young driver age range (16-20 inclusive). This age range constitutes only about 6.8% of the drivers in Alabama, indicating that they are about 2.4 times the probability of causing any given crash than the older drivers.

The overall problems and causes of younger driver crashes is indicated quite dramatically in Display 1, which compares causal driver ages in speed (red bars in the chart) and non-speed (blue bars) crashes. Even those crashes not involving speed (blue bar distribution) shows a skew to the left with the young drivers about three times the average of other drivers in causing crashes. The red bars show that this is almost doubled for speed involved crashes, which are some of the worst crashes in severity. But more importantly here, speed is an excellent proxy for all *risk taking driving behaviors*, and thus the underlying cause for most young-driver-caused crashes is a combination of inexperience and the willingness to take risks. While some young people will respond to various types of behavioral modification programs, there are others that are extremely difficult to reach. The glorification of risky driving behaviors in the media contributes heavily to this, but psychological studies have determined that the human brain is not able to fully understand the concept of risk until about the age of 25. The chart shows that the problem of risk-taking behavior continues to be over-represented through age 25, although it is greatly reduced for ages 24 and 25.

Display 1. Causal Driver Age Distributions for Speed Caused (Red) vs. Non-Speed (Blue)



C107: CU Driver Raw Age Value	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
16	2283	5.33	15966	2.95	1.806*	1018.622
17	2487	5.80	17979	3.32	1.747*	1063.209
18	2552	5.95	20165	3.73	1.598*	955.096
19	2453	5.72	20851	3.85	1.486*	801.770
20	2143	5.00	19476	3.60	1.389*	600.659
21	1999	4.66	18543	3.43	1.361*	530.545
22	1844	4.30	17324	3.20	1.344*	472.080
23	1632	3.81	16195	2.99	1.273*	349.488
24	1449	3.38	14470	2.67	1.265*	303.094
25	1347	3.14	13758	2.54	1.236*	257.478
26	1217	2.84	12872	2.38	1.194*	197.642

The major part of this study used all crash records in which young persons were *involved* (independent of crash causation). It was felt that this larger dataset would be most useful in determining when young people were more likely to be involved in crashes over the year. There were a total of 144,624 crashes in which young drivers were involved over the five year period. Of these, we saw above that 106,355 (73.5%) were caused by young drivers. This would include all single vehicle crashes in which they were involved (26,899 single-vehicle crashes). Deducting these from the totals still yields a probability of 67%, or about two-thirds of all multi-vehicle crashes in which 16-20 year old drivers are involved being cause by the younger driver.

This document will continue with an analysis of the days, weeks and months of the year that were found to be *over-represented* for young driver crashes. The term *over-represented* will be used to refer to situations that occur higher than *expected* to a statistically significant degree. The *expected* occurrence will be that given by the men or average occurrence of a comparable

control group. The Critical Analysis Reporting Environment (CARE) system was used to determine the days of the year in which a greater than expected number of young driver crashes occurred (http://www.caps.ua.edu/software/care/). Unlike some recent studies in this regard, we were not looking for just a single date, for reasons that will be explained in more detail below. However, to those who are curious, for the five year period under consideration the day of November 15th had the highest number of young driver involved (not necessarily caused) crashes, with 690 total crashes summed over the five year period (138 crashes per year).

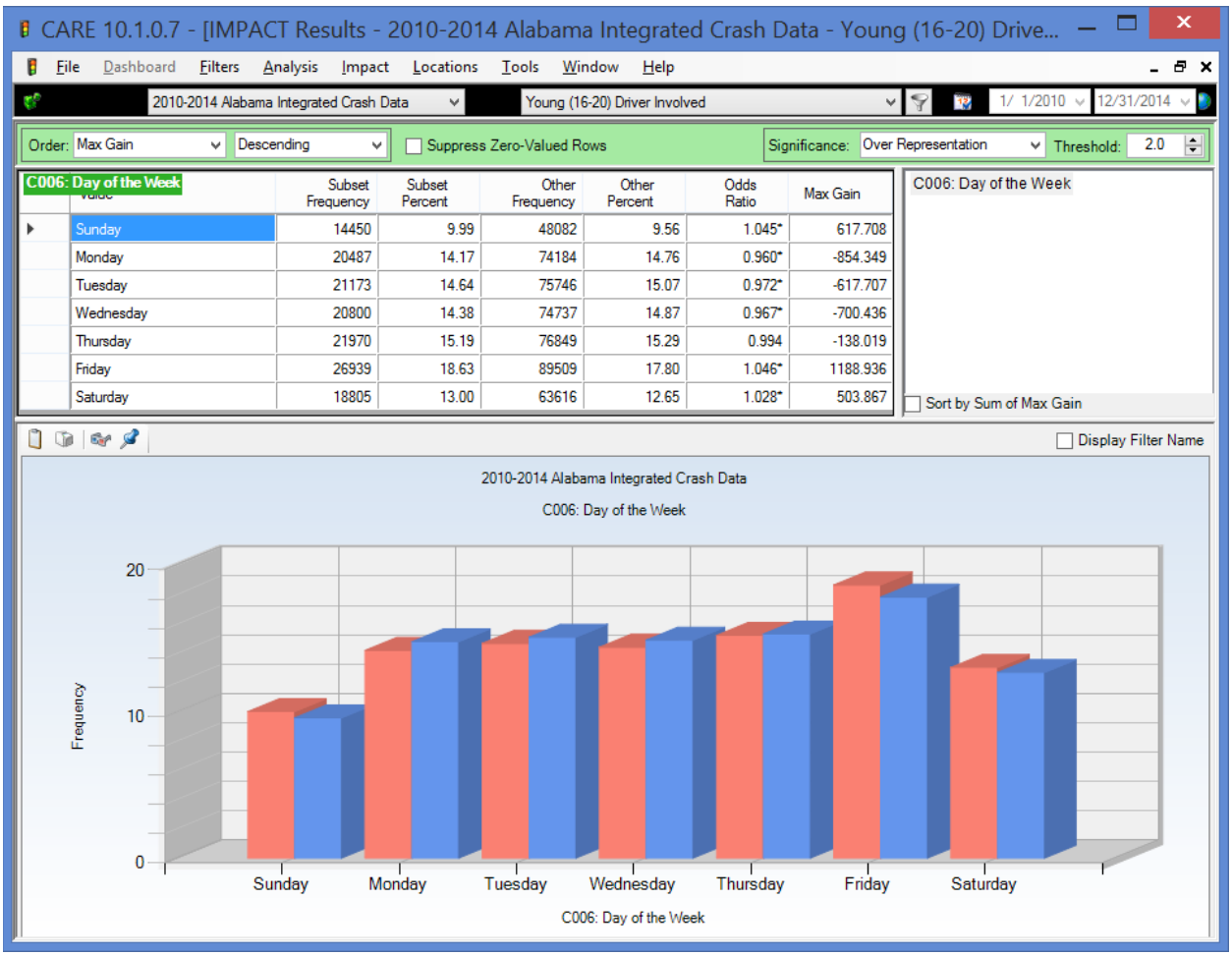
Study Approach

One problem in citing a date is that this date does not fall on the same day of the week every year – it may not again fall on the same day of the week for another seven years. Day of the week is a very important factor for young driver involved crashes, as shown in Display 2. This shows that young drivers generally follow the overall trend of Friday afternoon rush hour and night-time activity driving being the highest, and the weekends being lower. This pattern is more pronounced in the younger than in the older drivers. We can see from this that if November 15th was summed over a Tuesday-through-Saturday sequence in the five years of data, citing it for caution the following year would not be accurate if in the next year it occurred on a Sunday. Of course, this does not mean that we dismiss this data point altogether, but rather, we consolidate it into a six day period over which the special caution should be exercised, as will be illustrated below.

One day is much too low sample in any event. Things like weather and special events can have a great influence that might not carry over year to year. Because of this it was determined to consider six day periods. Six days were selected (as opposed to entire weeks) to eliminate the observed pattern of at least one day in any seven-day sequence being significantly lower than the other six. It was felt that these six-day intervals would be quite representative of the relative frequency of crashes in that general time frame.

A heuristic process was developed to determine the six-week periods that were over-represented in the five years of data (2000-2014). The overall average number of crashes per day involving young drivers was found to be 396. By the definition of an average, half of the numbers of crashes on any given day fall above this average and half below. Thus, the probability of a given reading falling above the average based on chance alone is 50% (analogous to flipping a coin). In a six day period it would be expected that half (3) of the readings would fall about the average of 396, and the other three would fall below. The probability of there being four is about 31%, and the probability of there being five of the six above average falls to less than 10%. Of course, the probability for all six to be above the average is considerably smaller, about 1.5% (which would be the probability of flipping a coin six times and getting all heads). These probabilities assume that the chances of the crashes being above or below average on any given day is strictly random. Of course, we know that is not true – these non-random days caused by a variety of factors (most notably the number of young drivers on the road) are what we are trying to surface.

Display 2. Day of the Week Comparison: Young Drivers (Red) vs Older Drivers (Blue)



C006: Day of the Week	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
Sunday	14450	9.99	48082	9.56	1.045*	617.708
Monday	20487	14.17	74184	14.76	0.960*	-854.349
Tuesday	21173	14.64	75746	15.07	0.972*	-617.707
Wednesday	20800	14.38	74737	14.87	0.967*	-700.436
Thursday	21970	15.19	76849	15.29	0.994	-138.019
Friday	26939	18.63	89509	17.80	1.046*	1188.936
Saturday	18805	13.00	63616	12.65	1.028*	503.867

The important thing to recognize at this point is that the greater the number of above-average readings in a given six day sequence (range: 4, 5 or 6), the more chance that this period of time is going to be detrimental to younger drivers.

Results of the Six Day Analyses

Displays 3a and 3b presents the results of the six day analyses, with the numbers in 3a representing the number of above-average readings (> 396) in the six-day segment. Generally the numbers in the table indicate, for any six day period that was found to be over-represented, how many of the days were above average, starting with four. The higher the number the greater the probability of this being a more dangerous time frame. In addition, the indicators were further adjusted by adding one to them if any of the days had a reading greater than or equal to 500. So, a four above-average-reading could be shown as a 5 in the table if it had one or more 500+ days.

The 500+ day cases are shown in the table by inserting the number of crashes for that day in place of the six day rating number further discussed below. No six day period had more than one 500+ day. Remember that the number of crashes is for a five year period.

Display 3a. Summary of Over-Represented Six Day Periods

Date	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1			5	7						4	7	5
2			5	7	6				4		7	5
3			5	7	592				4	5	7	5
4			5	508	6				4	5	7	5
5			517	6	6				4	5	7	4
6				6	6				4	5		4
7				6	6				4	5	4	4
8				6					4	5	4	4
9				6	4				4	5	4	4
10		4	4	6	4				4	5	4	4
11		4	4	6	4				4	5	4	
12		4	4	6	4				4	5	4	
13		4	4	6	4				4	5		6
14		4	4	6	4					5	5	6
15		4	4	6						5	690*	6
16				6	5				5	5	5	566
17					5				5	5	5	6
18					5			4	5	5	5	6
19			5		5			4	5	5	5	6
20			5		5			4	5	5		502
21			5		5			4	5			6
22			515					4			4	6
23			5					4		5	4	6
24			5							5	4	6
25										5	4	
26									4	5	4	
27									4	5	4	
28		5							4	5		
29		NA							4		5	
30		NA	7						4		5	
31		NA	7	NA		NA			NA	511	NA	

Display 3b. Summary of Over-Represented Six Day Periods

Date	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1				*							*	
2				*							*	
3				*	592						*	
4				508							*	
5			517								*	
6												
7												
8												
9												
10												
11												
12												
13												
14												
15											690	
16												566
17												
18												
19												
20												502
21												
22			515									
23												
24												
25												
26												
27												
28												
29		NA										
30		NA	*									
31		NA	*	NA		NA			NA	511	NA	

The following explains the terminology employed in Display 3a:

- Blank cell – indicates that for all six-day periods that this day was part of, there were three or less cells above the 396 average; in many of these cases all six cells were below average. The conclusion for these days is that there is *no indication that these times would be detrimental* in terms of young drivers' crashes.
- **4 (in purple)** – indicates that any four of the six days were above average. This could easily occur by chance, and so this should be regarded as the least possible indication of added danger. Consideration was given to leaving these off of the Displays 3a and 3b, but it was decided that they may play a positive role in establishing patterns.
- **5 (in orange)** – indicates that five of the six days in the sequence were above average. As indicated above, this has a much lower probability of occurring by chance than those given by the "4" indicator.
- **6 (in red)** – this indicates that all six of the days in the sequence were above average, a result that has a very low probability of occurring by chance, and thus a much greater indicator of a problem time.
- **7 (in red)** – this indicates that not only were all six days above average but that there was a reading above 500, so it was given an additional point. This is indicated by adding an asterisk (*) in Display 3b.
- **NNN** – all readings that were above 500 are entered into the table in lieu of the day's score as indicated above. This should also be interpreted that one (+1) was added to the six-day period score as indicated above.

Note that all indicators are given in six-day batches, the goal being to specify a general period of time as opposed to an exact date. Also, notice that some of the six-day periods wrap over the end of one month and the beginning of the next. Display 3b presents the cell colors as opposed to number colors, with the numbers removed to simplify the display.

Discussion of the Results

As discussed above, Displays 3a and 3b were not assembled to pinpoint a particular day to avoid driving – its goal is to provide an overall view of the patterns that are in effect. Working from the simplest to the more intricate interpretations:

- It should be clear that January, February and the summer months (June, July and August) are the months with the fewest young-driver involved crashes. Of course, this does not mean that the same concern should not be used when driving at these times – it illustrates the basic statistical reality that for every over-representation there must be compensating under-representations. Young driver crashes are not being compared against older-driver crashes; they are being compared against themselves in determining these over-represented six-day periods.
- On the other end of the scale, it is clear that October, November and December are collectively the worst sequence of three months. September might be added, but its indicators are not nearly as strong as those later on in the year.

- In particular, the 12 days prior to Christmas seem like they could be a candidate for the worst two weeks period, but we have not yet looked at the spring. Also note the 7's at the end of October and the beginning of November. The explanation for this has to do with the young drivers adding to an already high mix of people doing unusual things at this time – excess shopping, visiting, and general pre-holiday celebrations. This time period is over-represented in general and not just for young drivers.
- Within the last three months is a particularly bad sequence at the end of October (Halloween) and the first few days of November (note the 7's). Relatively speaking the rest of November is not nearly as bad, although the "worst day" is embedded in a six day period in which only four of the days are above average – the time frame got a penalty point for the 690 on November 15th. While there is nothing sacred about this day, it is embedded in an above-average six-day sequence, and the pre-Thanksgiving holiday should certainly be one for special caution. While the other two times indicated are only slight indicators, the pattern shows an overall problem when taken collectively.
- Tuning to the spring, it is clear that the first half of April is a major problem, and this is easily attributed to spring breaks that span this time frame starting at the end of March going right through the middle of April.
- March and May are quite comparable, and while not as concentrated as April or December, these months should not be discounted.

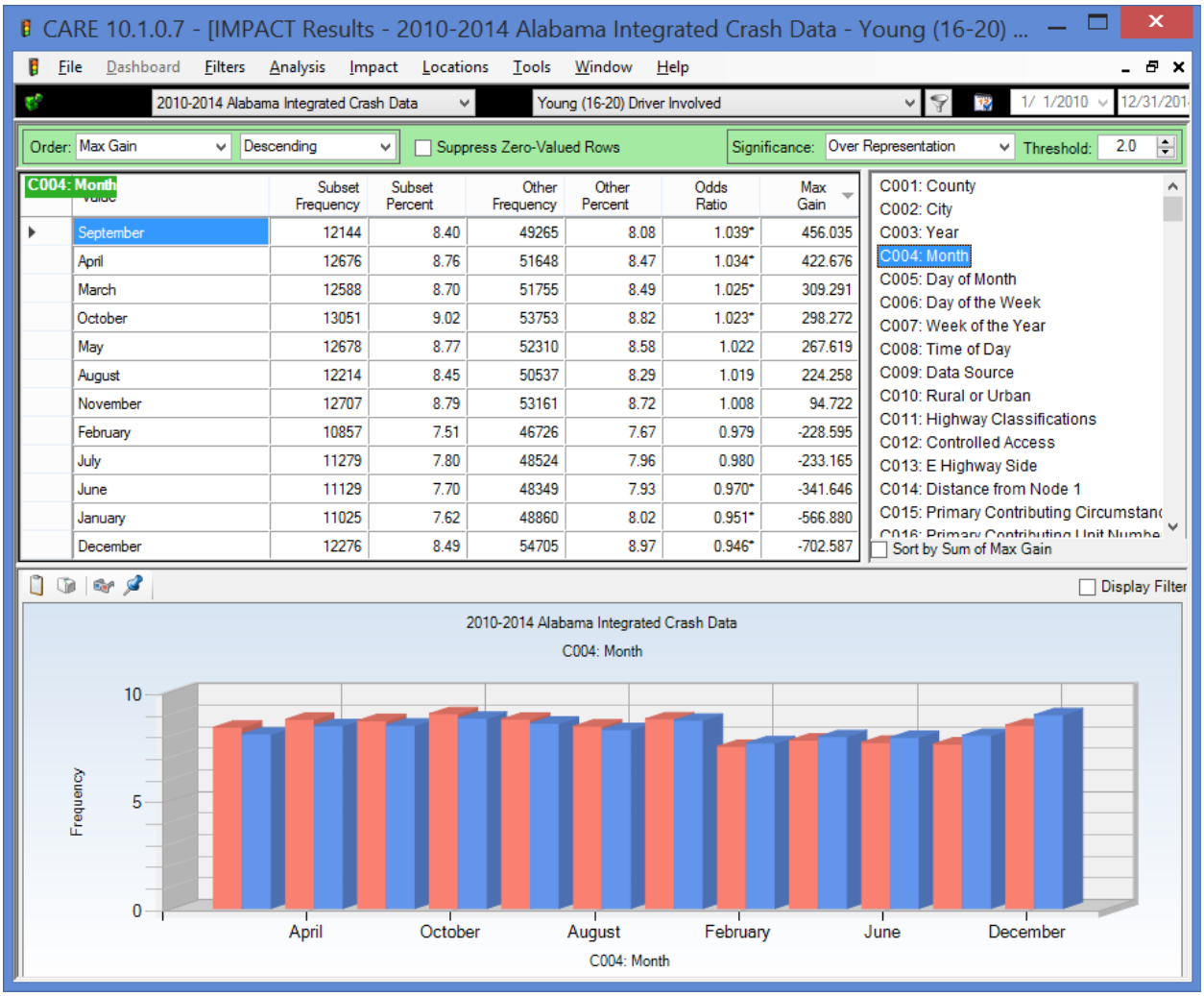
Additional consideration can be given to the ordering of the months once they are compared with the crashes that do not involve younger drivers. This is done in Display 4. Note that the comparison here is between *younger (16-20) drivers and older drivers*, so some of the results may seem inconsistent with those given above. But recognize that Display 4 discounts those times when both younger and older drivers are over-represented.

This display is arranged in a "worst-first" ordering so that the most over-represented months are shown at the top of the list when compared to crashes that do not involve younger drivers. Let us consider each of these in turn in adding this to the information that we already have in Displays 3a and 3b:

- September – The fact that September was not considered to be one of the largest problem times for young drivers is not contradictory to September having the highest over-representation. The differences in the odds ratios for the significantly over-represented months is not large, and neither are the odds ratios themselves. This shows a very close mapping of the young driver problem times with *heavy crash times in general*, with September being one of the months that had the highest over-representations of the younger drivers when compared to the older. The obvious reason would be the increase in younger drivers going back to school and being involved in other school and dating activities – things that would not have as much impact on older drivers.
- April and March – these were expected to be over-represented consistent with the findings above. Spring break is a particular problem time for young drivers that does not map to crashes for older drivers.

- October – We see this being over-represented because of the increase school activities during this month, which is again more of a young driver than a general problem.
- May, August and November – Display 4 shows that the proportion, while above its expected value, is not significantly different. This indicates that the over-representations for young drivers during these months are shared by older drivers.
- February and July – there are under-represented for younger drivers as well as older drivers, to a quite comparable degree as indicated by a lack of significant differences and the results of Displays 3a and 3b.
- January and June – the results here are consistent with Displays 3a and 3b in that there are no problem times for young drivers, and the indication is that the older drivers have significantly more than their expected proportion of crashes occurring during these times (i.e., young drivers are significantly *under-represented*).
- December – the under-representation is also seen to be significant in December, which has both the greatest under-representation for the younger drivers coupled with one of the greatest problems indicated by Displays 3a and 3b. All this is saying is that the pre-Christmas problem time is not only shared by the older drivers, but seems to be even a great problem for the older drivers than it is for the younger drivers. This in no way should be seen to be a mitigating factor; in fact, it is the increased older-driver traffic at this time that is a major factor in causing the increase in crashes by younger drivers. This is no reason to marginalize the danger to younger drivers at this time.

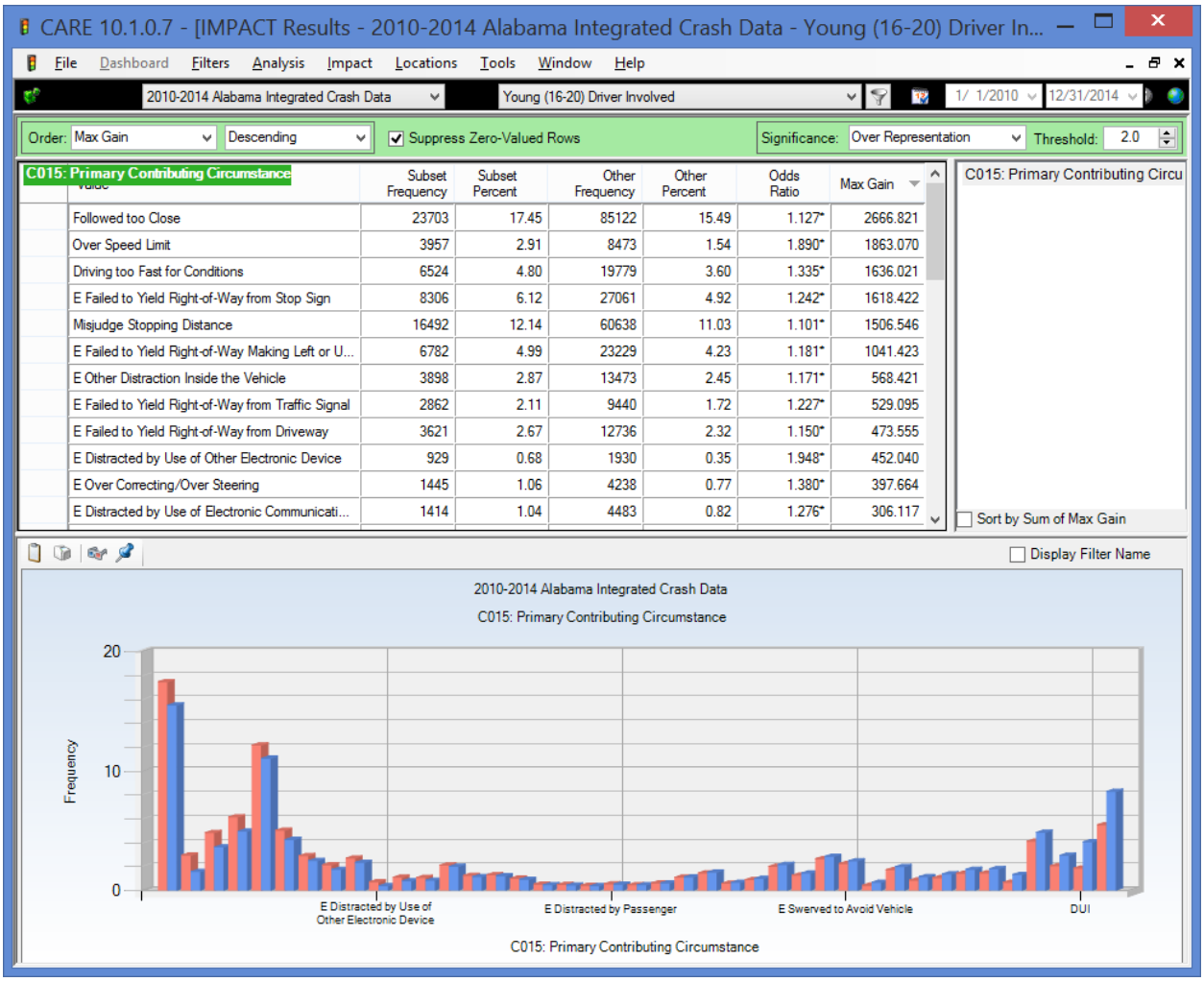
Display 4. IMPACT for Month comparing Young Driver Involved vs. Compliment



Additional IMPACT Analyses

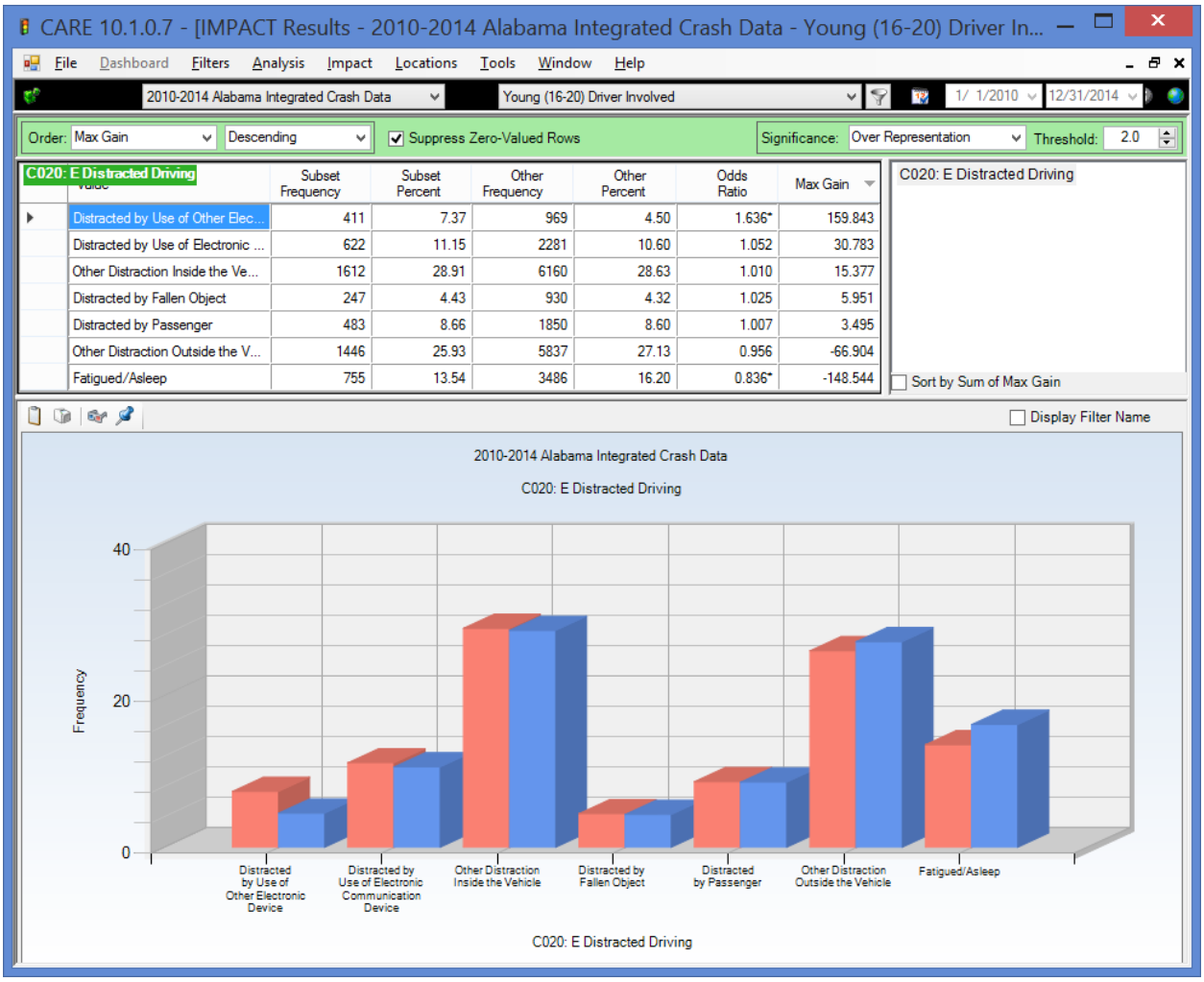
This section presents a number of IMPACT runs that surface some of the major characteristics of crashes in which younger drivers are involved as compared to older drivers (aged 21 and above). For information regarding the interpretation of IMPACT outputs,
see: http://www.caps.ua.edu/software/care/
And scroll down to the bottom of the page for the IMPACT tutorial.

C015: Primary Contributing Circumstances



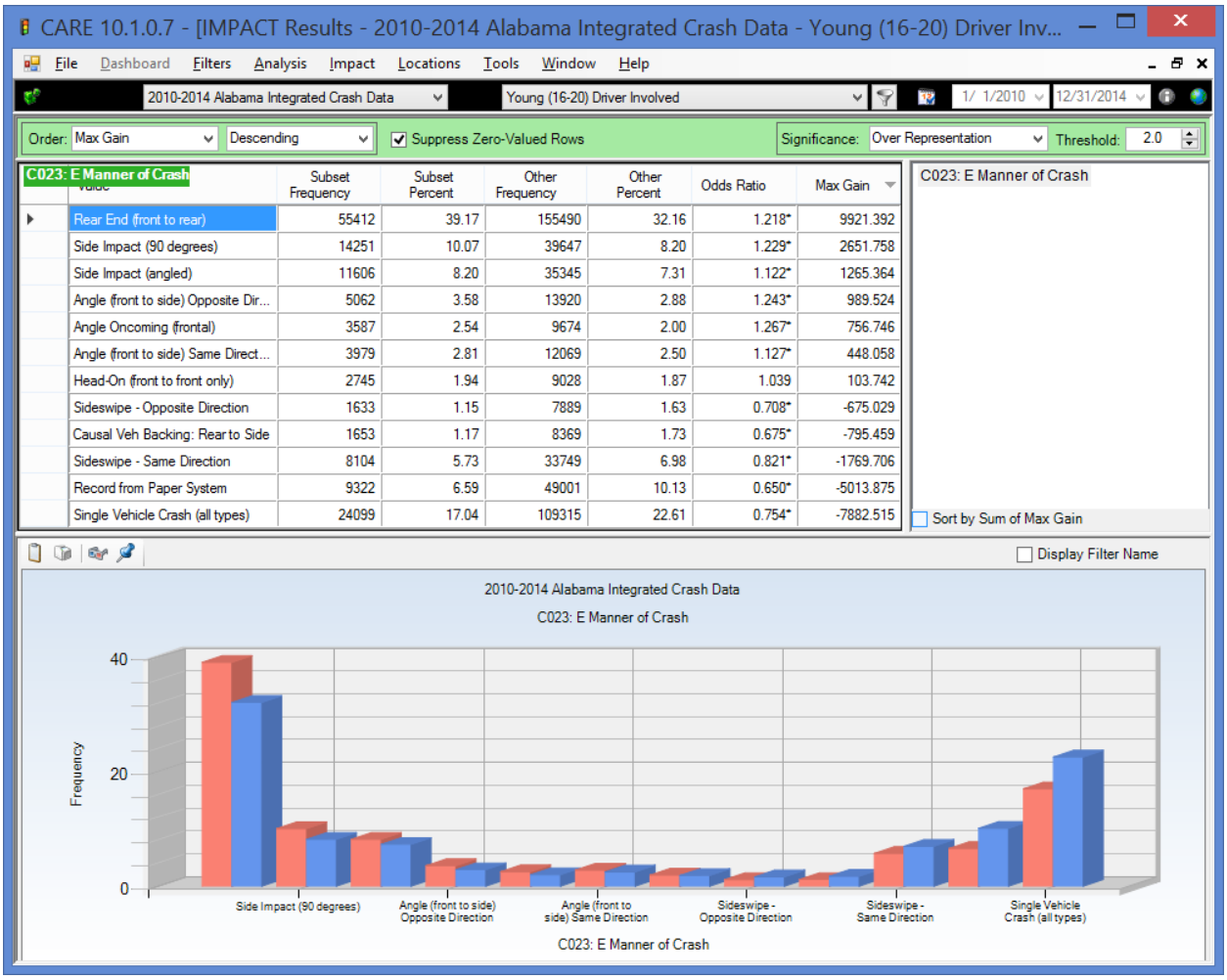
C015: Primary Contributing Circumstance	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
Followed too Close	23703	17.45	85122	15.49	1.127*	2666.821
Over Speed Limit	3957	2.91	8473	1.54	1.890*	1863.070
Driving too Fast for Conditions	6524	4.80	19779	3.60	1.335*	1636.021
E Failed to Yield Right-of-Way from Stop Sign	8306	6.12	27061	4.92	1.242*	1618.422
Misjudge Stopping Distance	16492	12.14	60638	11.03	1.101*	1506.546
E Failed to Yield Right-of-Way Making Left or U...	6782	4.99	23229	4.23	1.181*	1041.423
E Other Distraction Inside the Vehicle	3898	2.87	13473	2.45	1.171*	568.421
E Failed to Yield Right-of-Way from Traffic Signal	2862	2.11	9440	1.72	1.227*	529.095
E Failed to Yield Right-of-Way from Driveway	3621	2.67	12736	2.32	1.150*	473.555
E Distracted by Use of Other Electronic Device	929	0.68	1930	0.35	1.948*	452.040
E Over Correcting/Over Steering	1445	1.06	4238	0.77	1.380*	397.664
E Distracted by Use of Electronic Communicati...	1414	1.04	4483	0.82	1.276*	306.117

C020: Distracted Driving



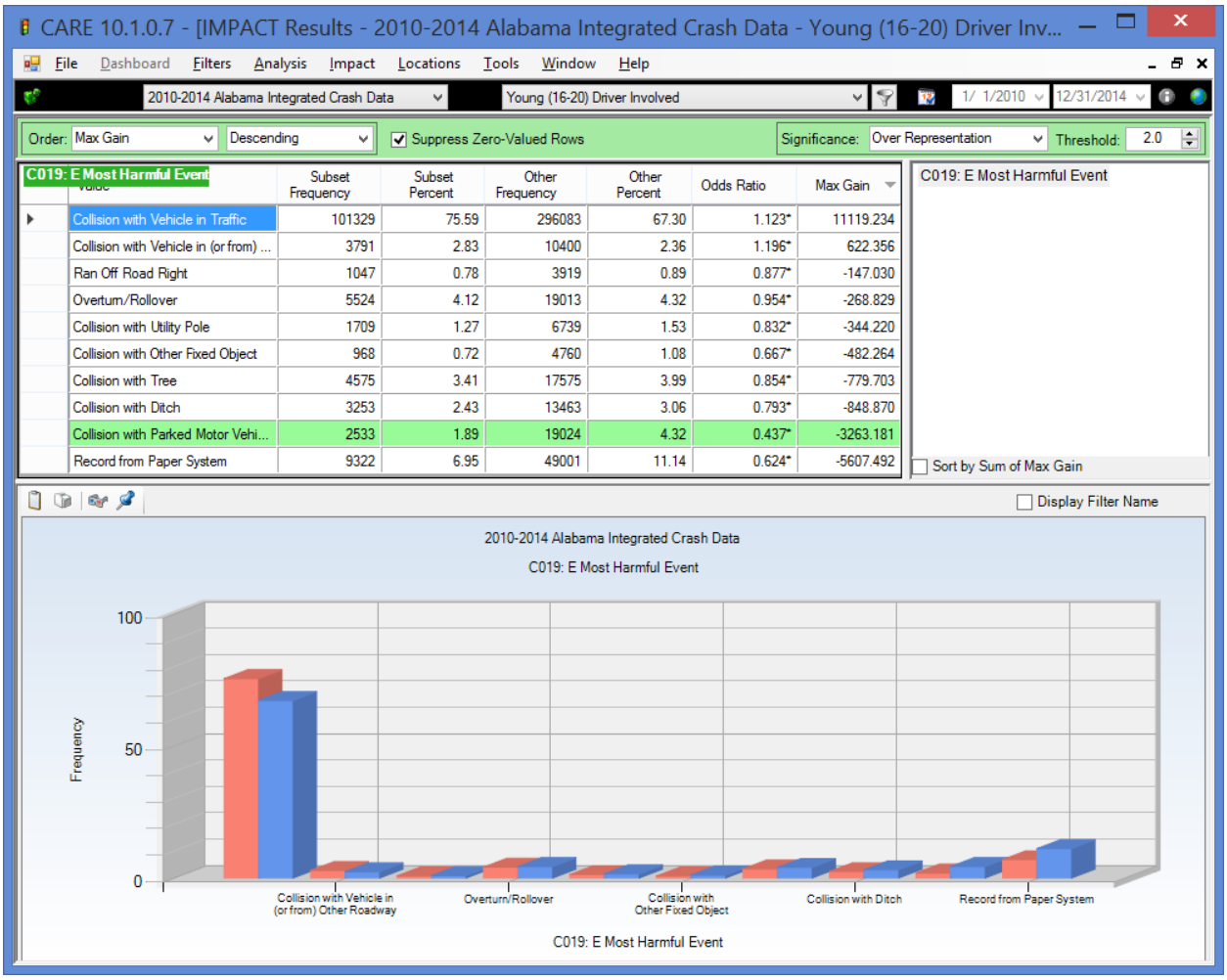
C020: E Distracted Driving	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
Distracted by Use of Other Elec...	411	7.37	969	4.50	1.636*	159.843
Distracted by Use of Electronic ...	622	11.15	2281	10.60	1.052	30.783
Other Distraction Inside the Ve...	1612	28.91	6160	28.63	1.010	15.377
Distracted by Fallen Object	247	4.43	930	4.32	1.025	5.951
Distracted by Passenger	483	8.66	1850	8.60	1.007	3.495
Other Distraction Outside the V...	1446	25.93	5837	27.13	0.956	-66.904
Fatigued/Asleep	755	13.54	3486	16.20	0.836*	-148.544

C023: Manner of Crash



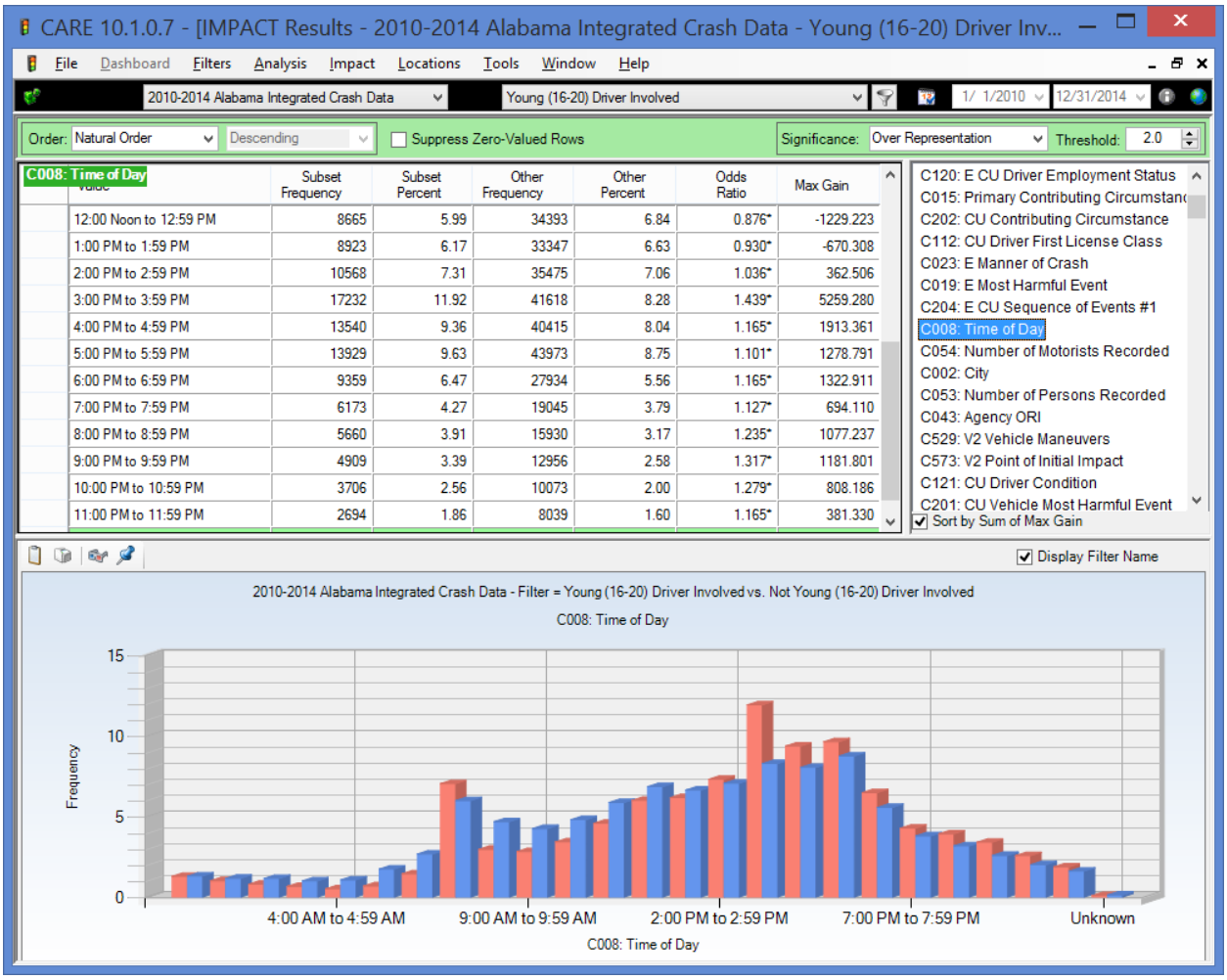
C023: E Manner of Crash	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
Rear End (front to rear)	55412	39.17	155490	32.16	1.218*	9921.392
Side Impact (90 degrees)	14251	10.07	39647	8.20	1.229*	2651.758
Side Impact (angled)	11606	8.20	35345	7.31	1.122*	1265.364
Angle (front to side) Opposite Dir...	5062	3.58	13920	2.88	1.243*	989.524
Angle Oncoming (frontal)	3587	2.54	9674	2.00	1.267*	756.746
Angle (front to side) Same Direct...	3979	2.81	12069	2.50	1.127*	448.058
Head-On (front to front only)	2745	1.94	9028	1.87	1.039	103.742
Sideswipe - Opposite Direction	1633	1.15	7889	1.63	0.708*	-675.029
Causal Veh Backing: Rear to Side	1653	1.17	8369	1.73	0.675*	-795.459
Sideswipe - Same Direction	8104	5.73	33749	6.98	0.821*	-1769.706
Record from Paper System	9322	6.59	49001	10.13	0.650*	-5013.875
Single Vehicle Crash (all types)	24099	17.04	109315	22.61	0.754*	-7882.515

C019: Most Harmful Event (Suppressing Values < 900 Occurrences)



C019: E Most Harmful Event	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
Collision with Vehicle in Traffic	101329	75.59	296083	67.30	1.123*	11119.234
Collision with Vehicle in (or from) ...	3791	2.83	10400	2.36	1.196*	622.356
Ran Off Road Right	1047	0.78	3919	0.89	0.877*	-147.030
Overturn/Rollover	5524	4.12	19013	4.32	0.954*	-268.829
Collision with Utility Pole	1709	1.27	6739	1.53	0.832*	-344.220
Collision with Other Fixed Object	968	0.72	4760	1.08	0.667*	-482.264
Collision with Tree	4575	3.41	17575	3.99	0.854*	-779.703
Collision with Ditch	3253	2.43	13463	3.06	0.793*	-848.870
Collision with Parked Motor Vehi...	2533	1.89	19024	4.32	0.437*	-3263.181
Record from Paper System	9322	6.95	49001	11.14	0.624*	-5607.492

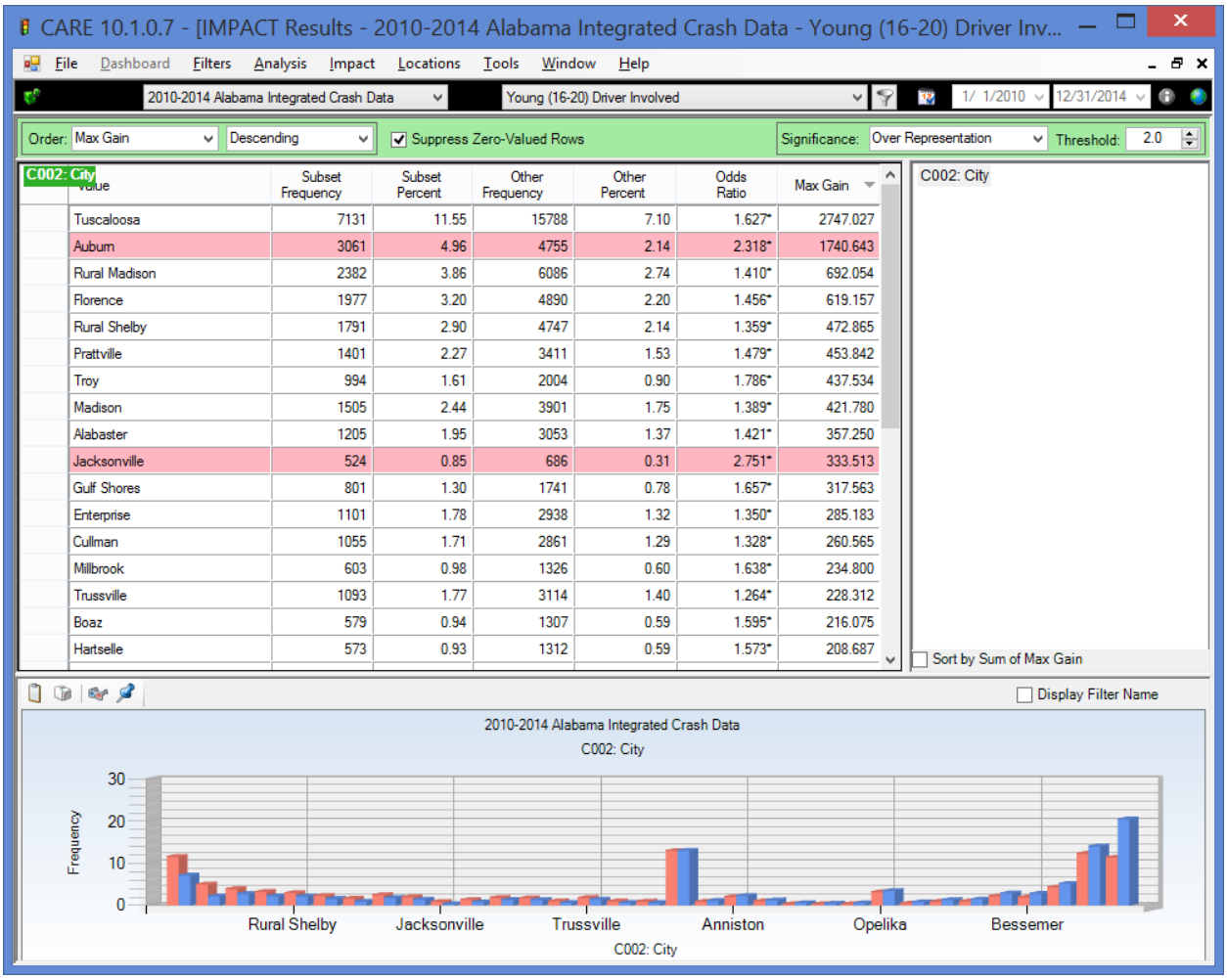
C008: Time of Day



C008: Time of Day Value	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
12:00 Noon to 12:59 PM	8665	5.99	34393	6.84	0.876*	-1229.223
1:00 PM to 1:59 PM	8923	6.17	33347	6.63	0.930*	-670.308
2:00 PM to 2:59 PM	10568	7.31	35475	7.06	1.036*	362.506
3:00 PM to 3:59 PM	17232	11.92	41618	8.28	1.439*	5259.280
4:00 PM to 4:59 PM	13540	9.36	40415	8.04	1.165*	1913.361
5:00 PM to 5:59 PM	13929	9.63	43973	8.75	1.101*	1278.791
6:00 PM to 6:59 PM	9359	6.47	27934	5.56	1.165*	1322.911
7:00 PM to 7:59 PM	6173	4.27	19045	3.79	1.127*	694.110
8:00 PM to 8:59 PM	5660	3.91	15930	3.17	1.235*	1077.237
9:00 PM to 9:59 PM	4909	3.39	12956	2.58	1.317*	1181.801
10:00 PM to 10:59 PM	3706	2.56	10073	2.00	1.279*	808.186
11:00 PM to 11:59 PM	2694	1.86	8039	1.60	1.165*	381.330

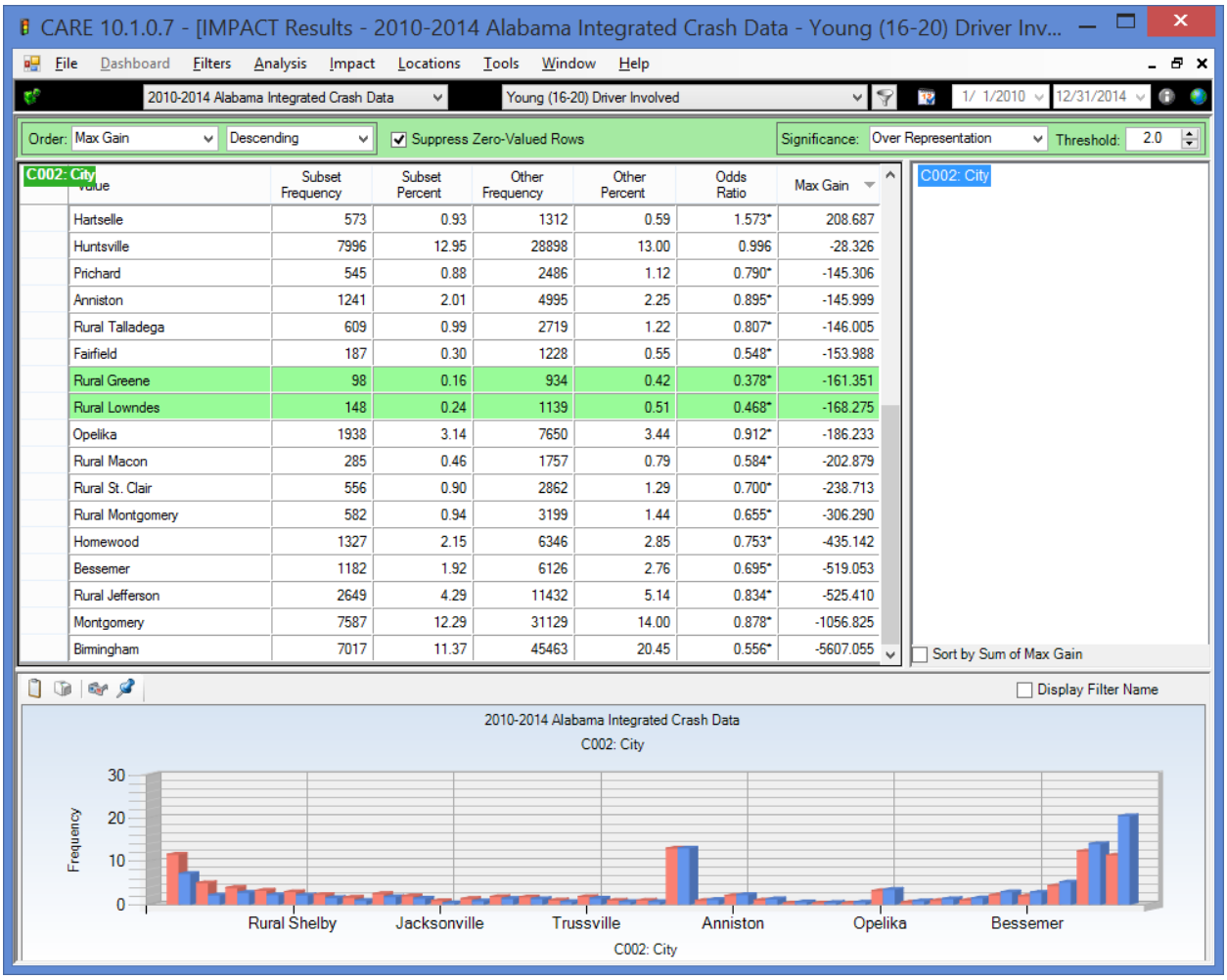
C120: E CU Driver Employment Status
C015: Primary Contributing Circumstanc
C202: CU Contributing Circumstance
C112: CU Driver First License Class
C023: E Manner of Crash
C019: E Most Harmful Event
C204: E CU Sequence of Events #1
C008: Time of Day
C054: Number of Motorists Recorded
C002: City
C053: Number of Persons Recorded
C043: Agency ORI
C529: V2 Vehicle Maneuvers
C573: V2 Point of Initial Impact
C121: CU Driver Condition
C201: CU Vehicle Most Harmful Event
Sort by Sum of Max Gain

C002: City (Worst Cities with regard to Young Driver Proportions)



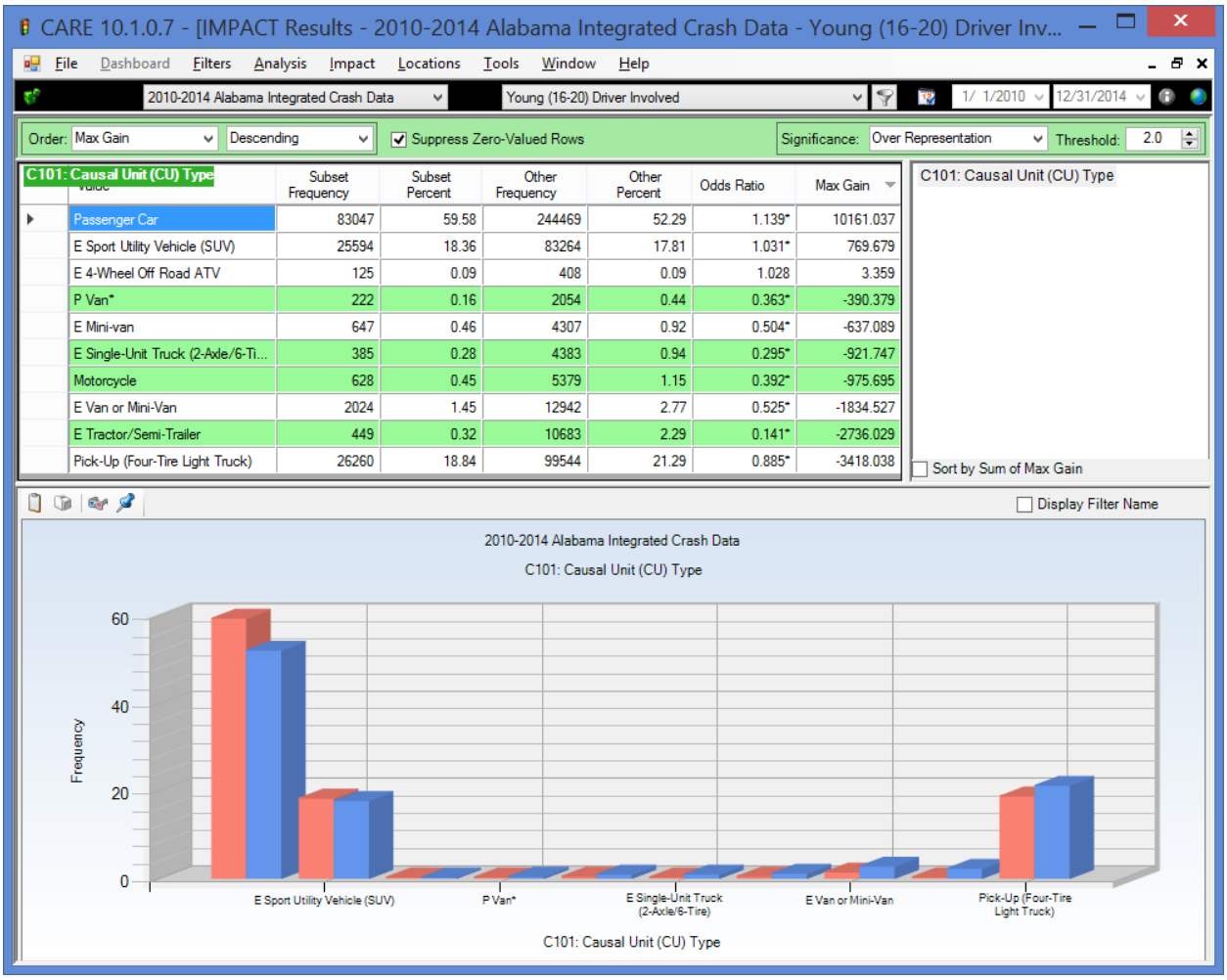
C002: City Value	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
Tuscaloosa	7131	11.55	15788	7.10	1.627*	2747.027
Auburn	3061	4.96	4755	2.14	2.318*	1740.643
Rural Madison	2382	3.86	6086	2.74	1.410*	692.054
Florence	1977	3.20	4890	2.20	1.456*	619.157
Rural Shelby	1791	2.90	4747	2.14	1.359*	472.865
Prattville	1401	2.27	3411	1.53	1.479*	453.842
Troy	994	1.61	2004	0.90	1.786*	437.534
Madison	1505	2.44	3901	1.75	1.389*	421.780
Alabaster	1205	1.95	3053	1.37	1.421*	357.250
Jacksonville	524	0.85	686	0.31	2.751*	333.513
Gulf Shores	801	1.30	1741	0.78	1.657*	317.563
Enterprise	1101	1.78	2938	1.32	1.350*	285.183
Cullman	1055	1.71	2861	1.29	1.328*	260.565
Millbrook	603	0.98	1326	0.60	1.638*	234.800
Trussville	1093	1.77	3114	1.40	1.264*	228.312
Boaz	579	0.94	1307	0.59	1.595*	216.075
Hartselle	573	0.93	1312	0.59	1.573*	208.687

C002: City (Best Cities with regard to Young Driver Proportions)



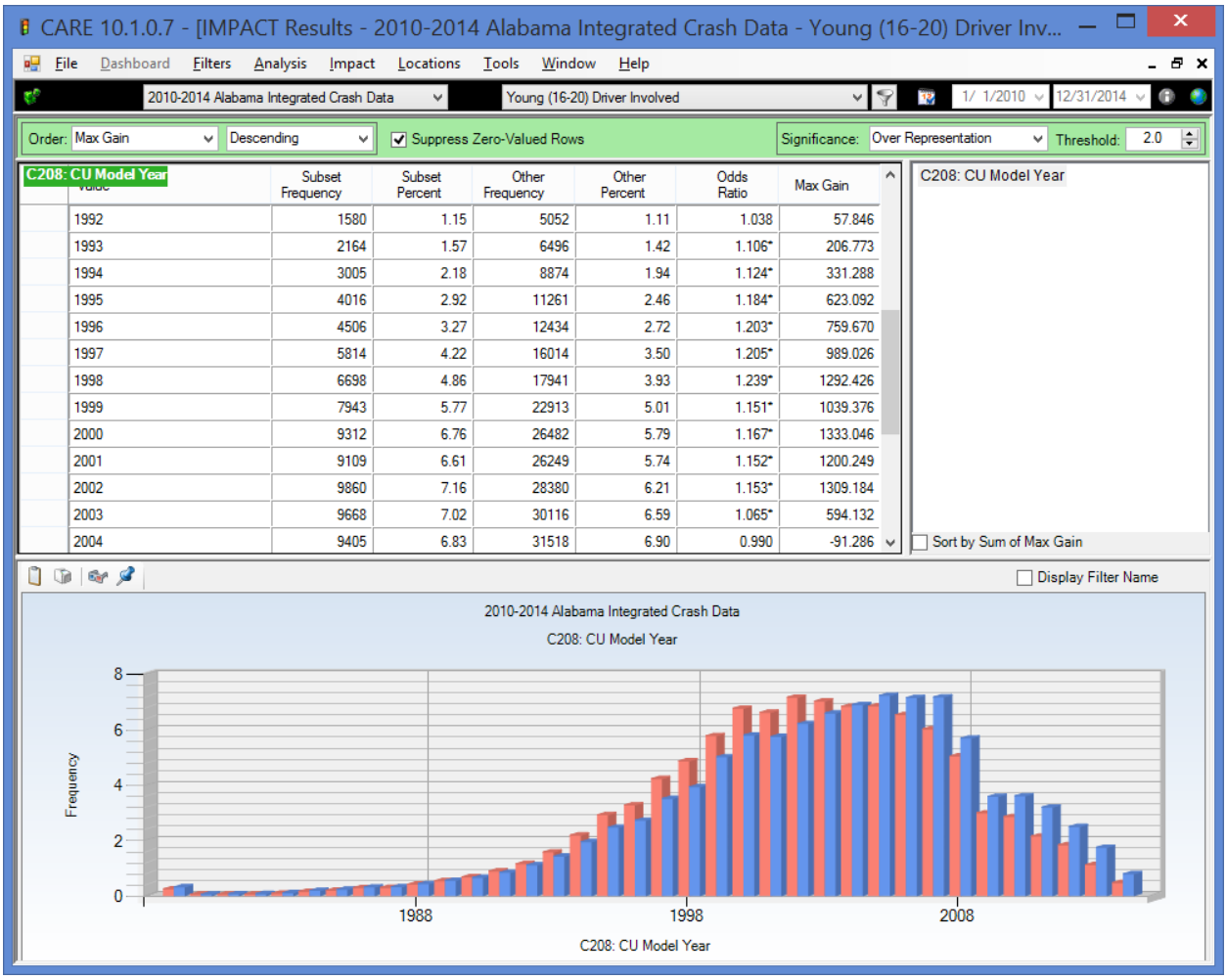
C002: City Value	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
Hartselle	573	0.93	1312	0.59	1.573*	208.687
Huntsville	7996	12.95	28898	13.00	0.996	-28.326
Prichard	545	0.88	2486	1.12	0.790*	-145.306
Anniston	1241	2.01	4995	2.25	0.895*	-145.999
Rural Talladega	609	0.99	2719	1.22	0.807*	-146.005
Fairfield	187	0.30	1228	0.55	0.548*	-153.988
Rural Greene	98	0.16	934	0.42	0.378*	-161.351
Rural Lowndes	148	0.24	1139	0.51	0.468*	-168.275
Opelika	1938	3.14	7650	3.44	0.912*	-186.233
Rural Macon	285	0.46	1757	0.79	0.584*	-202.879
Rural St. Clair	556	0.90	2862	1.29	0.700*	-238.713
Rural Montgomery	582	0.94	3199	1.44	0.655*	-306.290
Homewood	1327	2.15	6346	2.85	0.753*	-435.142
Bessemer	1182	1.92	6126	2.76	0.695*	-519.053
Rural Jefferson	2649	4.29	11432	5.14	0.834*	-525.410
Montgomery	7587	12.29	31129	14.00	0.878*	-1056.825
Birmingham	7017	11.37	45463	20.45	0.556*	-5607.055

C101: Causal Unit (CU) Type



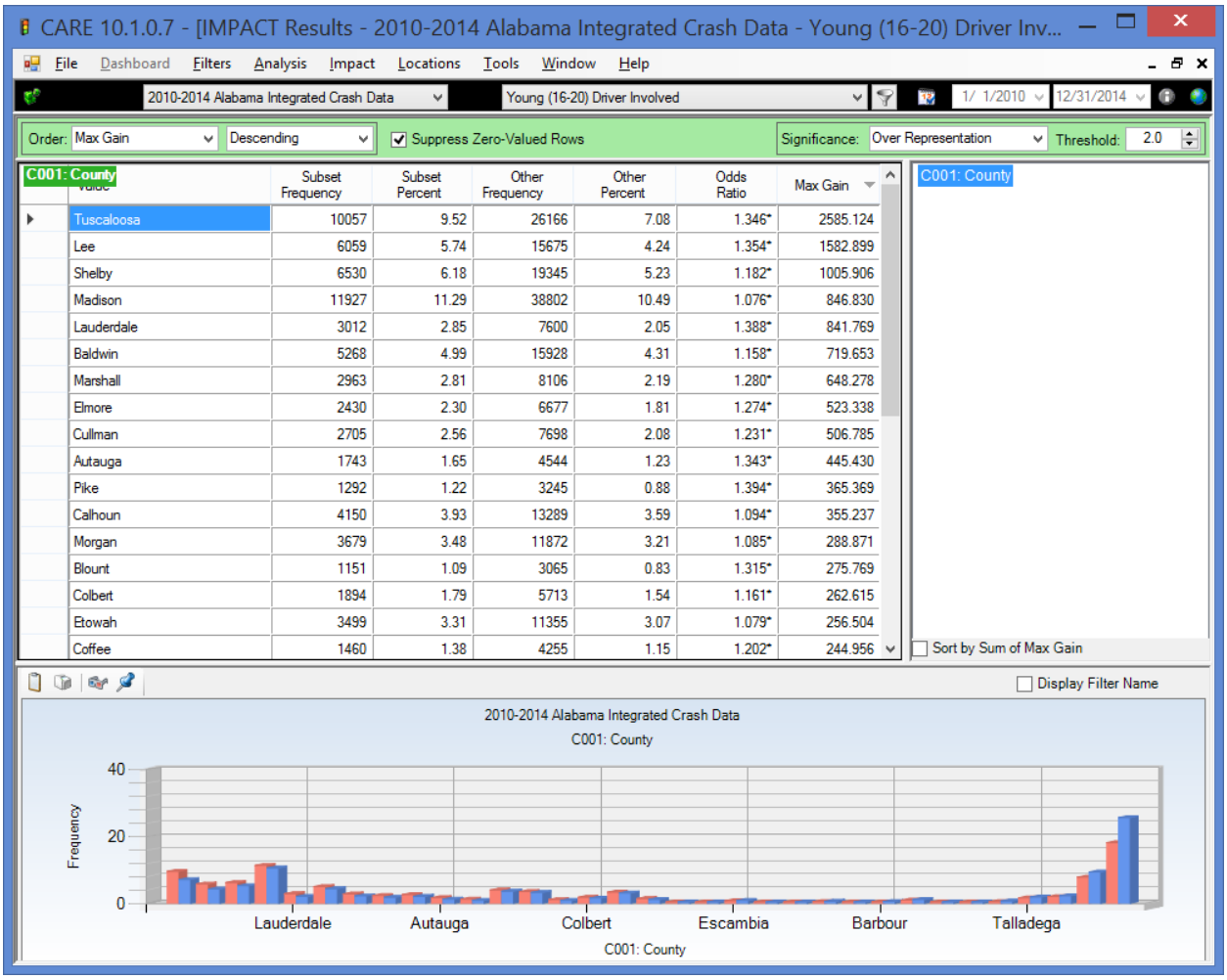
C101: Causal Unit (CU) Type	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
Passenger Car	83047	59.58	244469	52.29	1.139*	10161.037
E Sport Utility Vehicle (SUV)	25594	18.36	83264	17.81	1.031*	769.679
E 4-Wheel Off Road ATV	125	0.09	408	0.09	1.028	3.359
P Van*	222	0.16	2054	0.44	0.363*	-390.379
E Mini-van	647	0.46	4307	0.92	0.504*	-637.089
E Single-Unit Truck (2-Axle/6-Ti...	385	0.28	4383	0.94	0.295*	-921.747
Motorcycle	628	0.45	5379	1.15	0.392*	-975.695
E Van or Mini-Van	2024	1.45	12942	2.77	0.525*	-1834.527
E Tractor/Semi-Trailer	449	0.32	10683	2.29	0.141*	-2736.029
Pick-Up (Four-Tire Light Truck)	26260	18.84	99544	21.29	0.885*	-3418.038

C208 CU Model Year



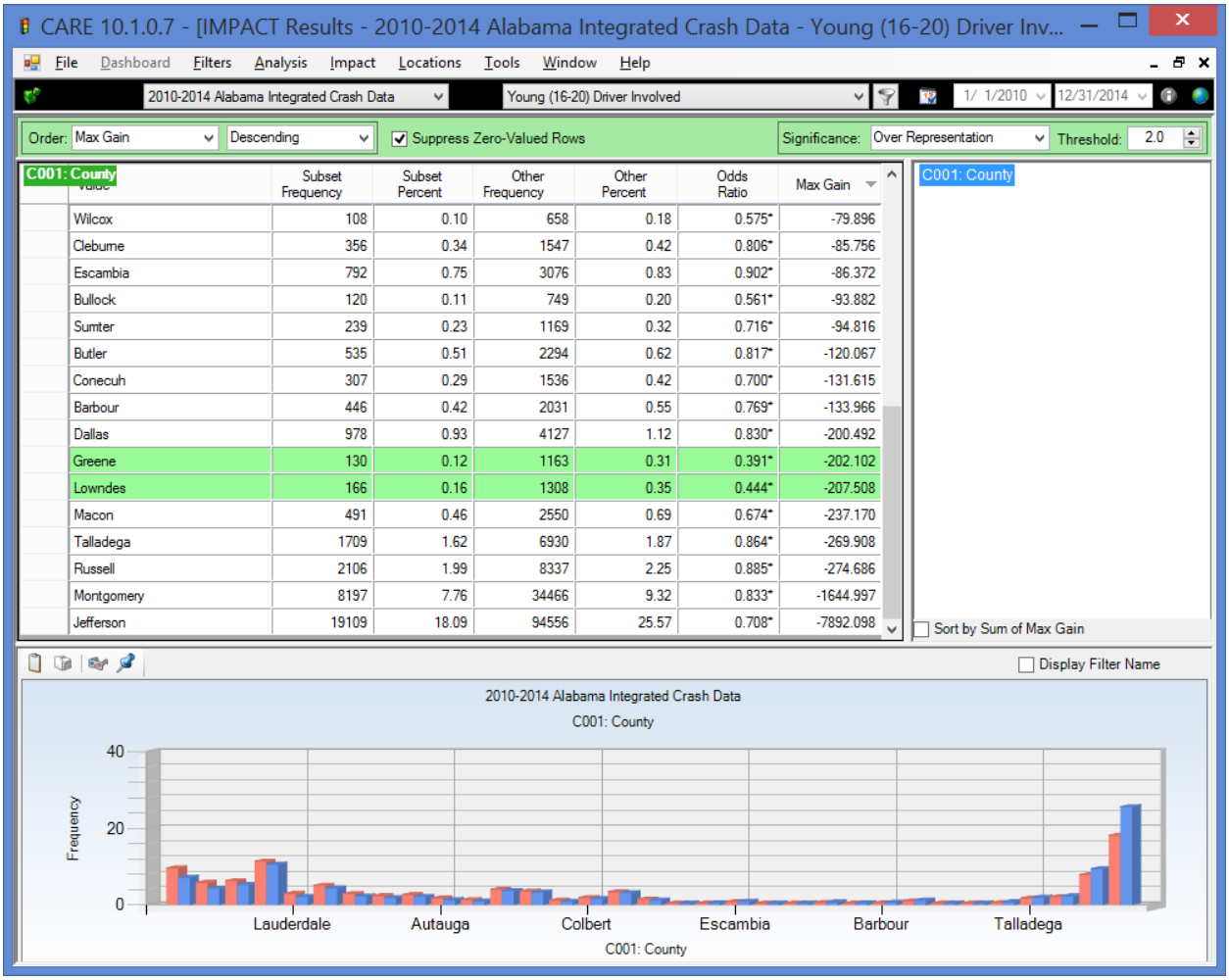
C208: CU Model Year Value	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
1992	1580	1.15	5052	1.11	1.038	57.846
1993	2164	1.57	6496	1.42	1.106*	206.773
1994	3005	2.18	8874	1.94	1.124*	331.288
1995	4016	2.92	11261	2.46	1.184*	623.092
1996	4506	3.27	12434	2.72	1.203*	759.670
1997	5814	4.22	16014	3.50	1.205*	989.026
1998	6698	4.86	17941	3.93	1.239*	1292.426
1999	7943	5.77	22913	5.01	1.151*	1039.376
2000	9312	6.76	26482	5.79	1.167*	1333.046
2001	9109	6.61	26249	5.74	1.152*	1200.249
2002	9860	7.16	28380	6.21	1.153*	1309.184
2003	9668	7.02	30116	6.59	1.065*	594.132
2004	9405	6.83	31518	6.90	0.990	-91.286

C001: County (Worst Counties for Young Driver Proportions)



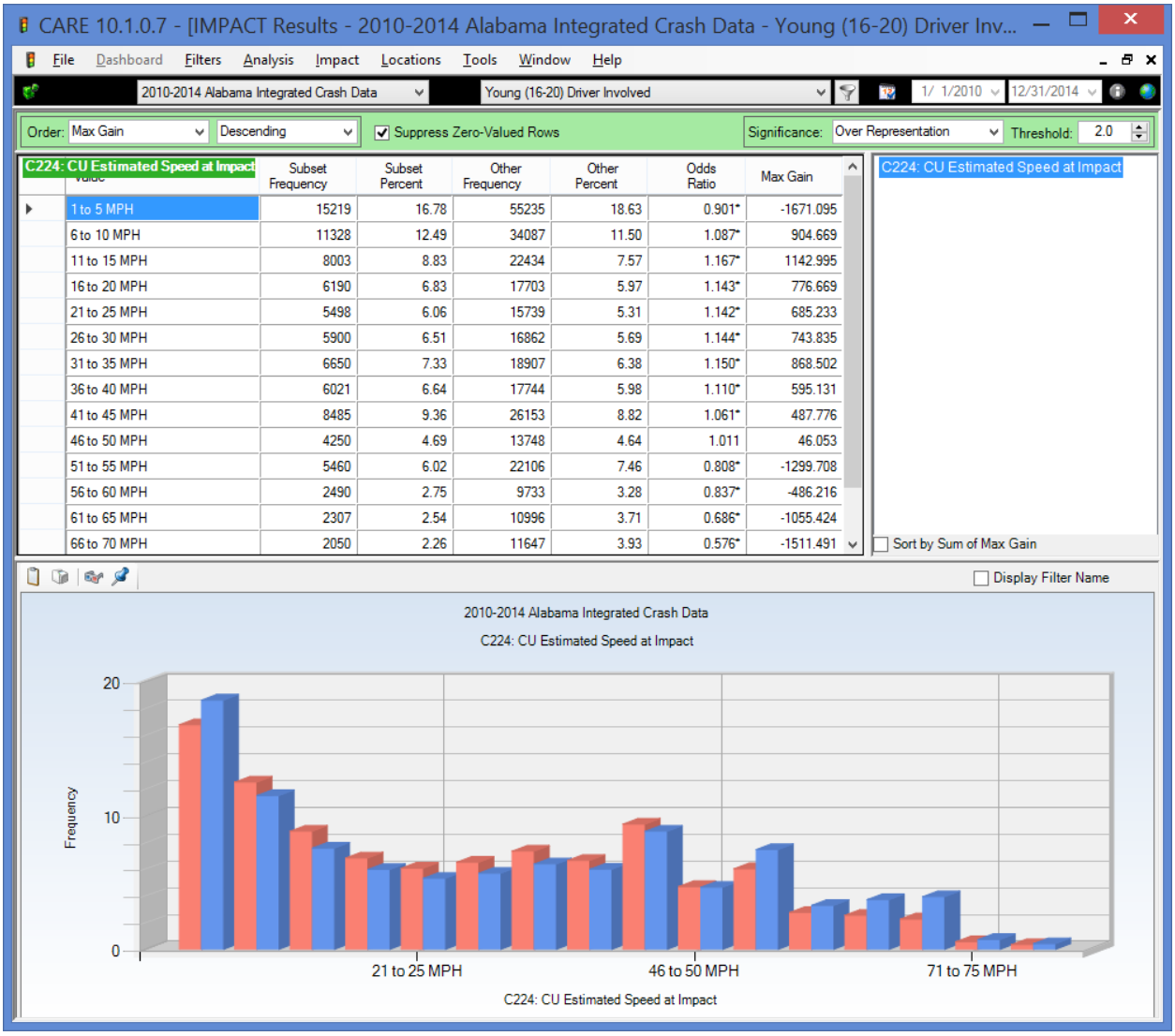
C001: County	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
Tuscaloosa	10057	9.52	26166	7.08	1.346*	2585.124
Lee	6059	5.74	15675	4.24	1.354*	1582.899
Shelby	6530	6.18	19345	5.23	1.182*	1005.906
Madison	11927	11.29	38802	10.49	1.076*	846.830
Lauderdale	3012	2.85	7600	2.05	1.388*	841.769
Baldwin	5268	4.99	15928	4.31	1.158*	719.653
Marshall	2963	2.81	8106	2.19	1.280*	648.278
Elmore	2430	2.30	6677	1.81	1.274*	523.338
Cullman	2705	2.56	7698	2.08	1.231*	506.785
Autauga	1743	1.65	4544	1.23	1.343*	445.430
Pike	1292	1.22	3245	0.88	1.394*	365.369
Calhoun	4150	3.93	13289	3.59	1.094*	355.237
Morgan	3679	3.48	11872	3.21	1.085*	288.871
Blount	1151	1.09	3065	0.83	1.315*	275.769
Colbert	1894	1.79	5713	1.54	1.161*	262.615
Etowah	3499	3.31	11355	3.07	1.079*	256.504
Coffee	1460	1.38	4255	1.15	1.202*	244.956

C001: County (Best Counties from Young Driver Proportion POV)



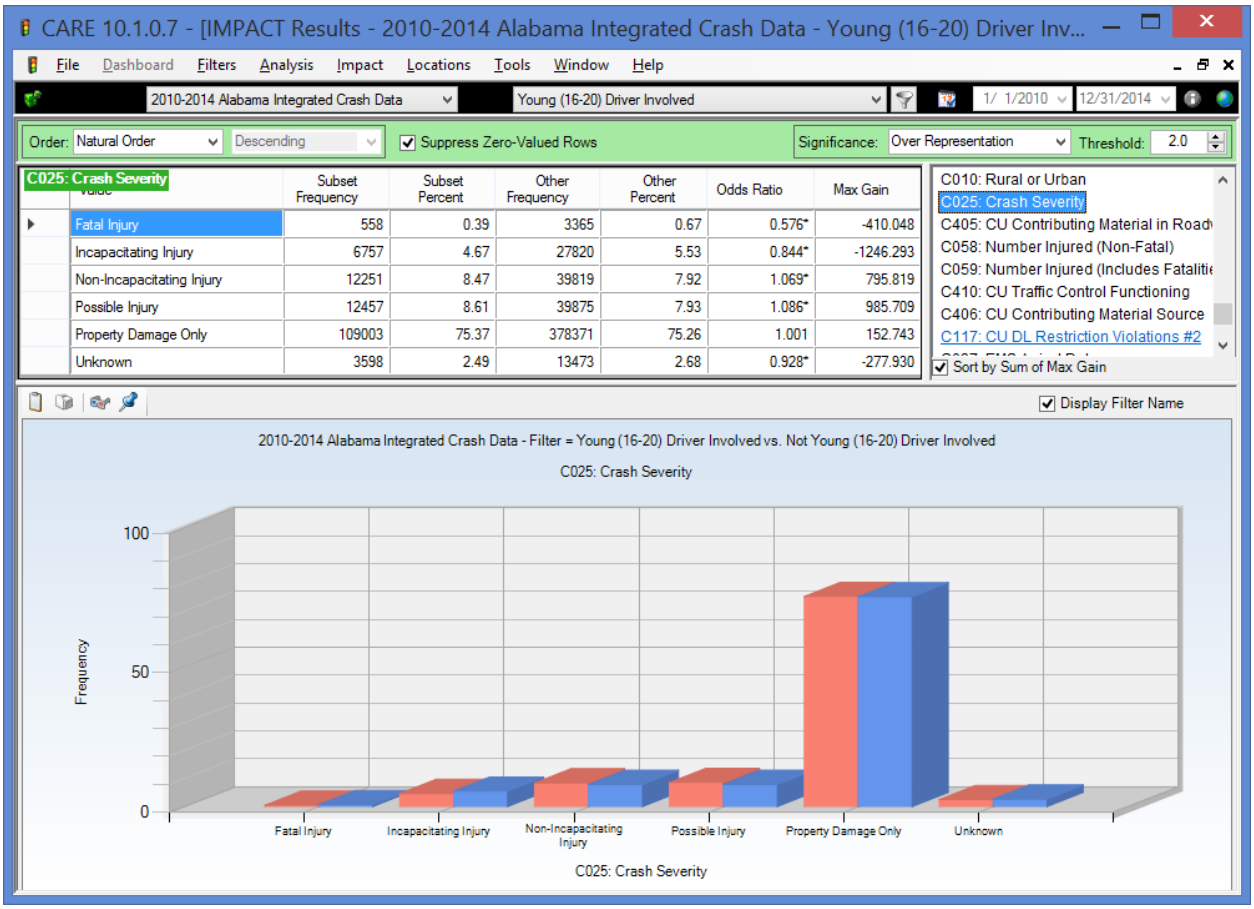
C001: County	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
Wilcox	108	0.10	658	0.18	0.575*	-79.896
Clebume	356	0.34	1547	0.42	0.806*	-85.756
Escambia	792	0.75	3076	0.83	0.902*	-86.372
Bullock	120	0.11	749	0.20	0.561*	-93.882
Sumter	239	0.23	1169	0.32	0.716*	-94.816
Butler	535	0.51	2294	0.62	0.817*	-120.067
Conecuh	307	0.29	1536	0.42	0.700*	-131.615
Barbour	446	0.42	2031	0.55	0.769*	-133.966
Dallas	978	0.93	4127	1.12	0.830*	-200.492
Greene	130	0.12	1163	0.31	0.391*	-202.102
Lowndes	166	0.16	1308	0.35	0.444*	-207.508
Macon	491	0.46	2550	0.69	0.674*	-237.170
Talladega	1709	1.62	6930	1.87	0.864*	-269.908
Russell	2106	1.99	8337	2.25	0.885*	-274.686
Montgomery	8197	7.76	34466	9.32	0.833*	-1644.997
Jefferson	19109	18.09	94556	25.57	0.708*	-7892.098

C224: Causal Unit (CU) Estimated Speed at Impact



C224: CU Estimated Speed at Impact	Subset Frequency	Subset Percent	Other Frequency	Other Percent	Odds Ratio	Max Gain
1 to 5 MPH	15219	16.78	55235	18.63	0.901*	-1671.095
6 to 10 MPH	11328	12.49	34087	11.50	1.087*	904.669
11 to 15 MPH	8003	8.83	22434	7.57	1.167*	1142.995
16 to 20 MPH	6190	6.83	17703	5.97	1.143*	776.669
21 to 25 MPH	5498	6.06	15739	5.31	1.142*	685.233
26 to 30 MPH	5900	6.51	16862	5.69	1.144*	743.835
31 to 35 MPH	6650	7.33	18907	6.38	1.150*	868.502
36 to 40 MPH	6021	6.64	17744	5.98	1.110*	595.131
41 to 45 MPH	8485	9.36	26153	8.82	1.061*	487.776
46 to 50 MPH	4250	4.69	13748	4.64	1.011	46.053
51 to 55 MPH	5460	6.02	22106	7.46	0.808*	-1299.708
56 to 60 MPH	2490	2.75	9733	3.28	0.837*	-486.216
61 to 65 MPH	2307	2.54	10996	3.71	0.686*	-1055.424
66 to 70 MPH	2050	2.26	11647	3.93	0.576*	-1511.491

Higher Speeds:

76 to 80 MPH	311
81 to 85 MPH	101
86 to 95 MPH	87
91 to 85 MPH	18
96 to 100 MPH	72
Over 100 MPH	58

C025: Crash Severity