Traffic Safety Facts
Research Note

April 2007 DOT HS 810 754

Differences in Driver Alcohol Involvement by Age Group and Vehicle Type

Timothy M. Pickrell*

Overview

This research note identifies significant differences in blood alcohol concentration (BAC) values between age groups and vehicle types in fatal crashes. An important finding of this study is that BAC values for motorcycle operators are distributed differently than BAC values of passenger vehicle drivers. The analysis also revealed that older drivers involved in fatal crashes tend to have lower BAC values than drivers from other age groups involved in fatal crashes. The statistical techniques used in this research note visually demonstrate the severity of the impaired-driving problem in that the majority of alcohol-involved drivers killed in fatal crashes exceed the legal per se limit of .08 grams per deciliter (g/dL). The Fatality Analysis Report-ing System (FARS) contains BAC values for drivers with positive BAC involved in fatal crashes. The data used for this research note were from the years 2000-2004, and only drivers and operators with positive BAC values were considered in the analysis.

Applications to Highway Traffic Safety Policy

Examining BAC distributions is important to highway traffic safety because these distributions demonstrate the extent of the impaired driving problem and provide a simultaneous concise summary for multiple categories like age group and vehicle type. A graphical approach to distribution analysis will quickly inform policymakers what proportion of BAC values exceed the legal limit in fatal crashes and how BACs differ by age group and vehicle type simply by looking at the shapes of the curves. Distribution analysis is just one tool to gauge the effectiveness of impaired-driving policies, providing much information in a compact form. Graphs can also be used to compare groups from different time periods in order to understand aggregate changes in alcohol consumption among fatal crash victims over time. For example, if people involved in traffic fatalities are drinking less (or more) over time, it would be evident from changes in the mean value and changes in the BAC dispersion about the mean value from one time period to the next.

Background

The National Center for Statistics and Analysis (NCSA) research note "Driver Alcohol Involvement in Fatal Crashes by Age Group and Vehicle Type" (DOT HS 810 598) showed that motorcycle operators had different degrees of alcohol involvement than did drivers of passenger vehicles for the same age groups. This research note graphically demonstrates the differences not only between motorcycle operators and passenger vehicle drivers, but differences between all vehicle types as well as differences between age groups for each vehicle type. In an attempt for a more precise understanding of the seemingly small numerical differences in BAC distribution between vehicle types and age groups, this study shows that statistically significant differences do exist between groups, even with small numerical differences.

Methodology

This study uses kernel density estimation to visually assess differences between BAC distributions by age group and vehicle type. A kernel density approximates a hypothesized probability density function from observed data, providing a method for visualizing differences between age groups and vehicle types and age groups within vehicle types. The analysis plots these distributions on the same graph so that a direct visual comparison can be made between distributions. To supplement this graphical analysis, an analysis of variance was used to determine if significant differences exist between the distributions by vehicle type and age groups. Pair-wise comparisons were performed between each of the groups to identify vehicle types and age groups with statistically significant differences. The vehicle types and age groups for this analysis are shown below:

1. **Vehicle Types:** Passenger Cars, SUVs, Pickup Trucks, Vans, and Motorcycles
2. **Age Groups:** under 20, 20-29, 30-39, 40-49, 50-59, and over 59

*Timothy M. Pickrell is a mathematical statistician with the Mathematical Analysis Division of NHTSA's NCSA.

Data Analysis

The following graphs demonstrate that BAC can have a different distribution for each vehicle type and age group. The pair-wise comparisons included at the end of the report show the comparison groups, difference between means, and the confidence limits of the difference between means. Significantly different pairs are labeled with asterisks. On each graph the vertical line represents the extent to which BAC distributions exceed BAC=.08 g/dL, the legal limit in all 50 States, the District of Columbia, and Puerto Rico. The degree of difference among dispersions indicates the difference of BAC intensity from one group to another. ANOVA shows that the variance of the distributions by vehicle type (Figure 1) is significantly different (F=104.41, p-value=<.0001).

Figure 1: **Kernel Density Plots of BAC by Vehicle Type**

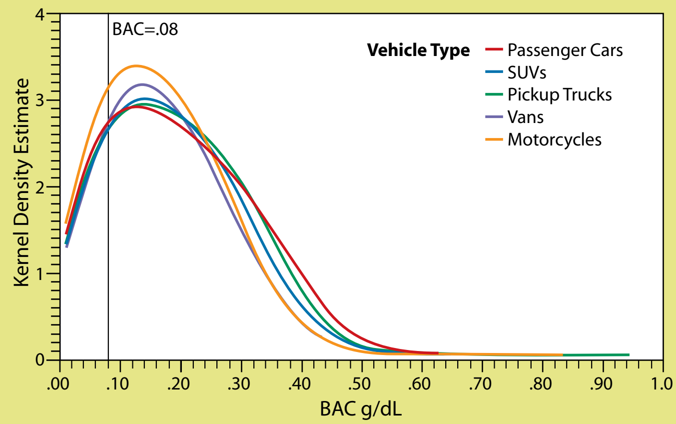


Figure 1 provides insight into the extent of alcohol involvement by vehicle type for drivers with positive BAC involved in fatal crashes and demonstrates that the majority of drivers with positive BAC values exceed the legal limit of .08. The variance of the distributions by vehicle type is significantly different (F=104.41, p-value=<.0001). Since BAC distributions show differences by vehicle type it should be interesting to look for difference by age group across all vehicle types.

Figure 2: **Kernel Density Plots of All Driver BAC by Age Group**

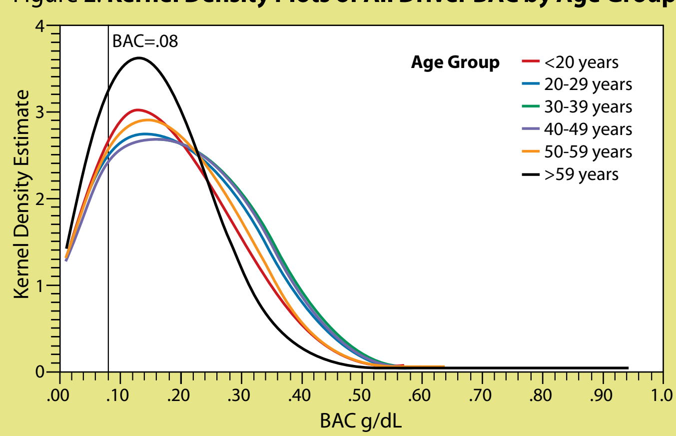


Figure 2 examines the BAC distribution from the perspective of age, and the differences are more pronounced here than for vehicle type. Drivers over age 59 have a distribution clustered more narrowly around their average value. Drivers with positive BAC values in other age groups have more widely dispersed BAC distributions. The variance of the distributions by age group is significantly different between age groups (F=186.56, p-value=<.0001).

Figure 3: **Kernel Density Plots of Passenger Car Driver BAC by Age Group**

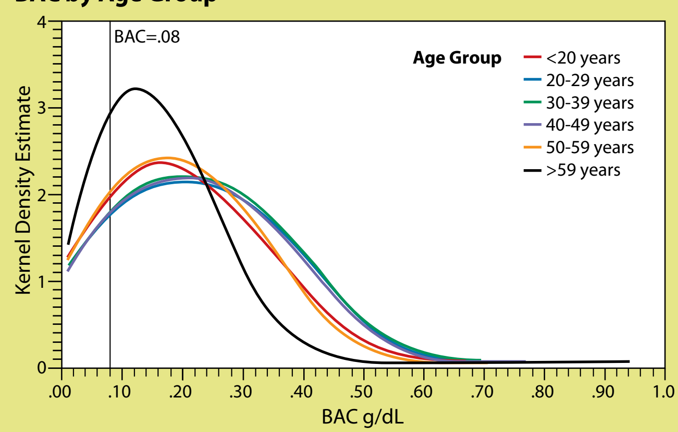


Figure 3 shows that passenger car drivers in the age ranges of 20-29, 30-39, and 40-49 had essentially the same distribution of BAC values (overlapping blue, purple, and green curves). Drivers over age 59 had less variation in BAC distribution than passenger car drivers of other age groups. The variance of the distributions by age group is significantly different between age groups (F=155.08, p-value=<.0001).

Figure 4: **Kernel Density Plots of Sport Utility Vehicle Driver BAC by Age Group**

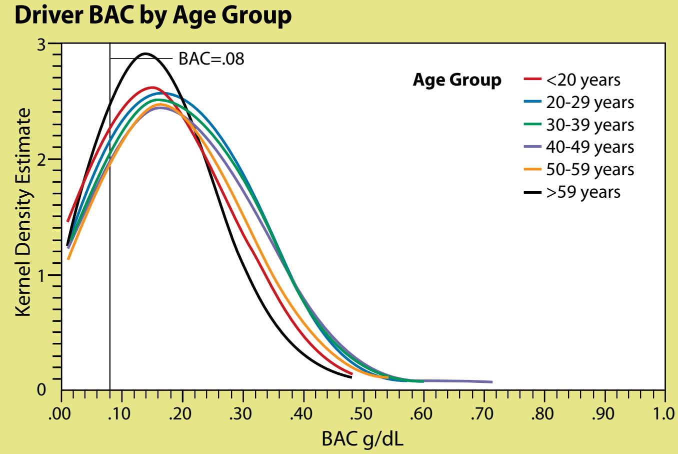


Figure 4 shows that for SUVs the age groups 20-29, 30-39, and 40-49 had nearly identical BAC distributions while drivers in the age groups under 20 and 50-59 had BAC distributions similar to each other. Drivers over age 59 display a BAC distribution with less variance than the other age groups. The variance of the distributions by age group is significantly different between age groups (F=26.70, p-value=<.0001).

Figure 5: **Kernel Density Plots of Pickup Truck Driver BAC by Age Group**

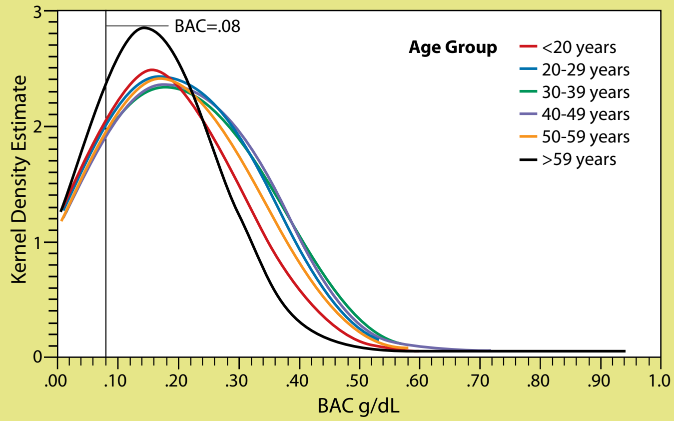


Figure 5 reveals that drivers over age 59 had a very different BAC distribution from all other pickup truck drivers. Again, the age groups 20-29, 30-39, and 40-49 were nearly identical in their BAC distribution. The variance of the distributions by age group is significantly different between age groups (F=55.17, p-value=<.0001).

Figure 6: **Kernel Density Plots of Van Driver BAC by Age Group**

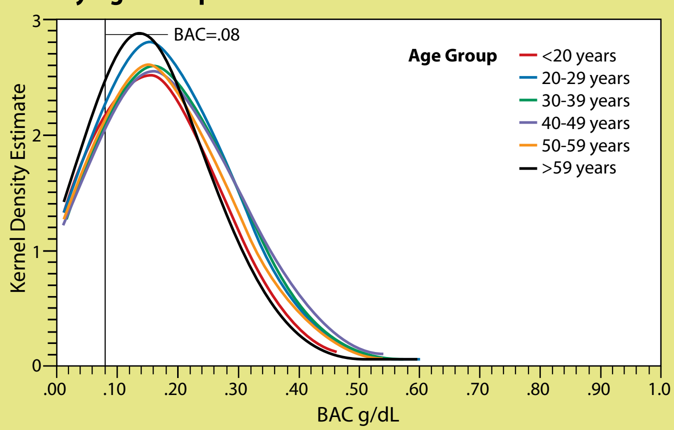


Figure 6 shows a repeating pattern for the age groups 20-29, 30-39, and 40-49 as those BAC distributions are very similar. Drivers under 20 had a BAC distribution distinct from the other age groups, in that it had less dispersion. The variance of the distributions by age group is significantly different between age groups (F=10.41, p-value=<.0001).

Figure 7: **Kernel Density Plots of Motorcycle Driver BAC by Age Group**

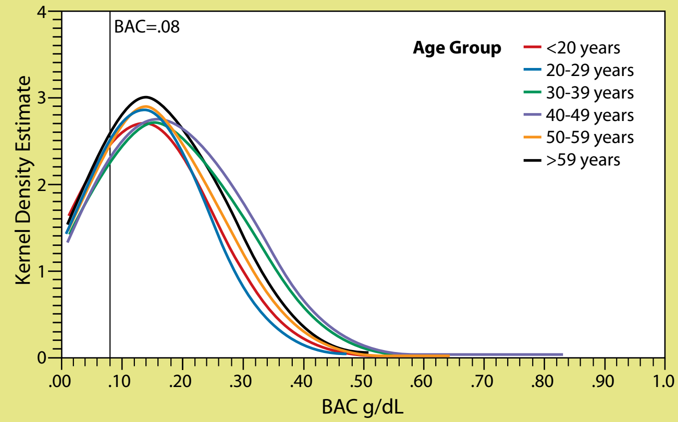


Figure 7 demonstrates that motorcycle operators have different BAC distributional relationships compared to Passenger Vehicles, Sport Utility Vehicles, Pickup Trucks, or Vans. Motorcycle operators in the under-20 and over-59 age groups had nearly identical BAC distributions. Also motorcycle operators in the 20-29 and 50-59 age groups had very similar distributions. The variance of the distributions by age group is significantly different between age groups (F=23.93, p-value=<.0001).

Pair-wise comparisons are included in the appendix to demonstrate which vehicle types, age groups, and age groups within vehicle types are significantly different.

Findings

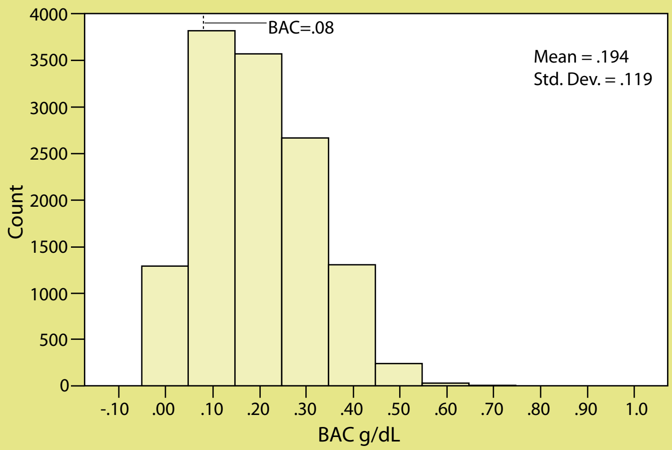
- The majority of drivers or operators with alcohol involved in fatal crashes had a BAC value exceeding the legal limit of BAC=.08 g/dL.
- Motorcycle operators with positive BAC values tend to have similar BAC distributions regardless of age group.
- Older driver BAC distributions tend to display smaller mean values with less variation than BAC distributions from all other driver age groups.
- Different age groups have different levels of alcohol involvement leading to fatal crashes. Not all drivers or vehicle types can be regarded in the same way.
- Small differences exist between group BAC mean values, but the distributions of BAC values are significantly different between vehicles types and also between age groups within each vehicle type.

Appendix:

Kernel Density Estimation Background

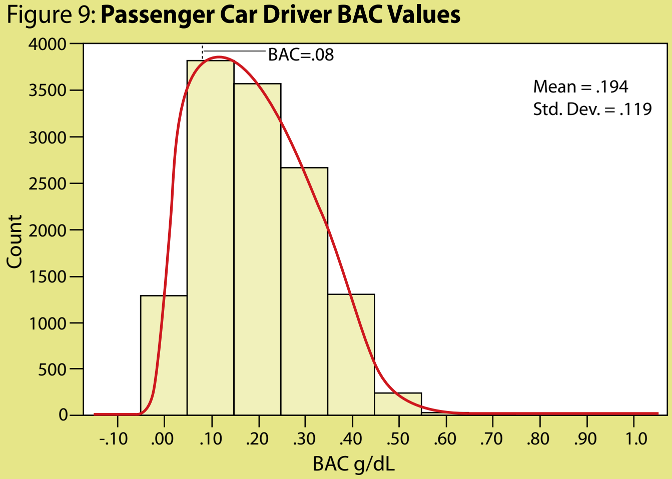
The statistical technique of kernel density estimation is an extension of constructing a simple histogram. A histogram shows the number of observations in a data set that fall within a particular range, commonly known as a frequency distribution. To illustrate this example, a histogram of BAC values for passenger vehicle drivers (2000-2004) is shown below.

Figure 8: **Passenger Car Driver BAC Values**



The histogram in Figure 8 shows how many BAC values occur in the ranges specified on the horizontal axis. For example, the range with midpoint BAC=.020 has 3,500 observations, the height of the rectangle is directly proportional to the number of observations in the range. A histogram shows not only how many values fall into each range but gives an idea about the dispersion of the data about a mean value, in this case .194 g/dL. A kernel density is an extension of the histogram, much like an outline smoothing out the rough edges of the histogram. In Figure 9, the same histogram is displayed with a kernel density superimposed.

Figure 9: **Passenger Car Driver BAC Values**

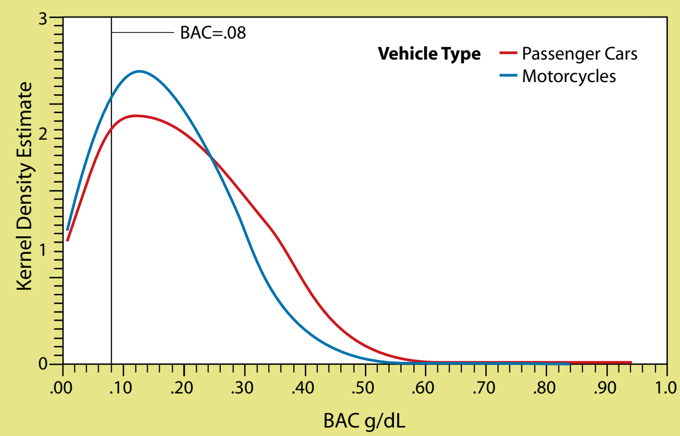


The kernel density conveys similar information about the density of the data. However, the kernel density is a plot of a mathematical function where each BAC value (x) corresponds to a kernel density estimate $f(x)$, as opposed to displaying the frequency of values as a histogram does. The kernel density estimate for each BAC value is expressed as:

$$\hat{f}(x) = \frac{1}{na_n}\sum_{i=1}^{n} K\left(\frac{x - x_i}{a_n}\right), \text{ with the normal distribution } K = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

Where n is the sample size, K is the normal distribution function, and a_n is the *smoothing constant* such that as a_n becomes larger, the shape of the curve becomes more smooth. The key component of kernel density estimation is its use of an underlying probability distribution function, in this case the normal or Gaussian distribution function (commonly known as a bell curve). The underlying bell-shaped curve is used as a pattern which is altered by the actual data to arrive at the final shape of the curve. The real advantage of kernel density estimation is clear when two distributions are plotted on the same axis. It is impossible to produce simultaneous histograms on the same plot as shown below with kernel density estimation; however, multiple kernel density plots can easily be displayed on the same set of axes. Figure 10 plots the kernel density function for passenger car and motorcycle operator BAC values on the same axis.

Figure 10: **Kernel Density Plots: Passenger Car Driver and Motorcycle Operator BAC Values**



This graph summarizes the BAC distributional behavior of two categories (passenger cars and motorcycles) on the same graph and makes a direct comparison much easier than with traditional histograms. Furthermore it allows each distribution to be displayed in contrast to the critical value of BAC=.08. As the graph above shows the majority of each distribution exceeds the legal limit of BAC=.08, demonstrating the importance of this technique with regard to highway traffic safety policy.

Additional Data

Pair-wise comparisons between each of the categories (vehicle types and age groups) show which contrasts are statistically significant (at the .05 level) by the asterisk notation (***). The analysis of variance method used shows where two groups have statistically significant different mean values. For those contrasts not statistically significantly different, we cannot say that those BAC distributions are significantly different from each other.

Vehicle Type Pair-Wise Mean Comparisons

Comparisons significant at the 0.05 level are indicated by ***.

Vehicle Type Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
Passenger Cars - Pickups	0.004096	-0.000292	0.008485	
Passenger Cars - Motorcycles	0.029014	0.02393	0.034099	***
Passenger Cars - SUVs	0.012568	0.00785	0.017285	***
Passenger Cars - Vans	0.021657	0.016554	0.026761	***
Pickups - Motorcycles	0.024918	0.019648	0.030188	***
Pickups - SUVs	0.008471	0.003554	0.013388	***
Pickups - Vans	0.017561	0.012273	0.022849	***
SUVs - Motorcycles	0.016447	0.0109	0.021994	***
SUVs - Vans	0.00909	0.003526	0.014654	***
Vans - Motorcycles	0.007357	0.001479	0.013235	***

Age Group Pair-Wise Mean Comparisons – Across All Vehicle Types

Comparisons significant at the 0.05 level are indicated by ***.

Age Group Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
<20 years - >59 years	0.015693	0.008356	0.02303	***
20-29 years - <20 years	0.020381	0.013003	0.02776	***
20-29 years - >59 years	0.036074	0.030574	0.041574	***
20-29 years - 50-59 years	0.009111	0.003297	0.014926	***
30-39 years - <20 years	0.025605	0.018269	0.032942	***
30-39 years - >59 years	0.041298	0.035854	0.046742	***
30-39 years - 20-29 years	0.005224	-0.000276	0.010724	
30-39 years - 40-49 years	0.00026	-0.005211	0.00573	
30-39 years - 50-59 years	0.014335	0.008574	0.020097	***
40-49 years - <20 years	0.025346	0.017989	0.032702	***
40-49 years - >59 years	0.041039	0.035568	0.04651	***
40-49 years - 20-29 years	0.004965	-0.000562	0.010491	
40-49 years - 50-59 years	0.014076	0.008289	0.019863	***
50-59 years - <20 years	0.01127	0.003695	0.018845	***
50-59 years - >59 years	0.026963	0.021201	0.032724	***

Age Group Pair-Wise Mean Comparisons – Passenger Car Drivers

Comparisons significant at the 0.05 level are indicated by ***.

Age Group Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		

<20 years - >59 years	0.039062	0.025804	0.052319	***
20-29 years - <20 years	0.029289	0.014912	0.043665	***
20-29 years - >59 years	0.06835	0.057915	0.078786	***
20-29 years - 30-39 years	0.001982	-0.009882	0.013847	
20-29 years - 40-49 years	0.00465	-0.00729	0.016589	
20-29 years - 50-59 years	0.026892	0.014536	0.039248	***
30-39 years - <20 years	0.027306	0.012897	0.041716	***
30-39 years - >59 years	0.066368	0.055887	0.076848	***
30-39 years - 40-49 years	0.002667	-0.009311	0.014646	
30-39 years - 50-59 years	0.02491	0.012516	0.037304	***
40-49 years - <20 years	0.024639	0.010168	0.03911	***
40-49 years - >59 years	0.0637	0.053135	0.074266	***
40-49 years - 50-59 years	0.022242	0.009777	0.034708	***
50-59 years - <20 years	0.002397	-0.01242	0.017213	
50-59 years - >59 years	0.041458	0.030424	0.052492	***

Age Group Pair-Wise Comparisons – Sport Utility Vehicle Drivers

Comparisons significant at the 0.05 level are indicated by ***.

Age Group Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
<20 years - >59 years	0.005855	-0.010626	0.022337	
20-29 years - <20 years	0.02769	0.012447	0.042933	***
20-29 years - >59 years	0.033545	0.020241	0.04685	***
20-29 years - 50-59 years	0.005718	-0.006926	0.018362	
30-39 years - <20 years	0.03064	0.015384	0.045896	***
30-39 years - >59 years	0.036495	0.023176	0.049815	***
30-39 years - 20-29 years	0.00295	-0.008802	0.014703	
30-39 years - 40-49 years	0.002479	-0.009454	0.014413	
30-39 years - 50-59 years	0.008668	-0.003992	0.021328	
40-49 years - <20 years	0.028161	0.012778	0.043544	***
40-49 years - >59 years	0.034016	0.020551	0.04748	***
40-49 years - 20-29 years	0.000471	-0.011446	0.012387	
40-49 years - 50-59 years	0.006189	-0.006623	0.019001	
50-59 years - <20 years	0.021972	0.006019	0.037925	***
50-59 years - >59 years	0.027827	0.013715	0.04194	***

Age Group Pair-Wise Comparisons – Pickup Truck Drivers

Comparisons significant at the 0.05 level are indicated by ***.

Age Group Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
<20 years - >59 years	0.005242	-0.009738	0.020222	
20-29 years - <20 years	0.027787	0.012298	0.043276	***
20-29 years - >59 years	0.033029	0.021649	0.044409	***
20-29 years - 50-59 years	0.000991	-0.011264	0.013246	

30-39 years - <20 years	0.039447	0.024076	0.054818	***
30-39 years - >59 years	0.044689	0.03347	0.055908	***
30-39 years - 20-29 years	0.01166	-0.00023	0.02355	
30-39 years - 40-49 years	0.001164	-0.010592	0.01292	
30-39 years - 50-59 years	0.012651	0.000545	0.024757	***
40-49 years - <20 years	0.038283	0.022897	0.053669	***
40-49 years - >59 years	0.043525	0.032285	0.054764	***
40-49 years - 20-29 years	0.010496	-0.001414	0.022406	
40-49 years - 50-59 years	0.011487	-0.000638	0.023612	
50-59 years - <20 years	0.026796	0.011141	0.042451	***
50-59 years - >59 years	0.032038	0.020433	0.043643	***

Age Group Pair-Wise Comparisons – Van Drivers

Comparisons significant at the 0.05 level are indicated by ***.

Age Group Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
<20 years - >59 years	0.005255	-0.013775	0.024285	
20-29 years - <20 years	0.013743	-0.004565	0.032052	
20-29 years - >59 years	0.018998	0.005225	0.032772	***
20-29 years - 50-59 years	0.002091	-0.011334	0.015516	
30-39 years - <20 years	0.017867	-0.000234	0.035969	
30-39 years - >59 years	0.023122	0.009625	0.036619	***
30-39 years - 20-29 years	0.004124	-0.008335	0.016584	
30-39 years - 50-59 years	0.006215	-0.006926	0.019356	
40-49 years - <20 years	0.020134	0.001988	0.03828	***
40-49 years - >59 years	0.025389	0.011832	0.038946	***
40-49 years - 20-29 years	0.006391	-0.006134	0.018915	
40-49 years - 30-39 years	0.002267	-0.009953	0.014487	
40-49 years - 50-59 years	0.008482	-0.004721	0.021684	
50-59 years - <20 years	0.011653	-0.007127	0.030432	
50-59 years - >59 years	0.016907	0.002514	0.0313	***

Age Group Pair-Wise Comparisons – Motorcycle Operators

Comparisons significant at the 0.05 level are indicated by ***.

Age Group Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
<20 years - >59 years	0.001309	-0.022371	0.024988	
20-29 years - <20 years	0.017715	-0.00307	0.038499	
20-29 years - >59 years	0.019024	0.00236	0.035687	***
20-29 years - 50-59 years	0.005925	-0.006902	0.018753	
30-39 years - <20 years	0.033006	0.012407	0.053606	***
30-39 years - >59 years	0.034315	0.017883	0.050747	***

30-39 years - 20-29 years	0.015291	0.003405	0.027178	***
30-39 years - 50-59 years	0.021217	0.008691	0.033743	***
40-49 years - <20 years	0.037438	0.016853	0.058023	***
40-49 years - >59 years	0.038747	0.022332	0.055161	***
40-49 years - 20-29 years	0.019723	0.007861	0.031585	***
40-49 years - 30-39 years	0.004432	-0.007103	0.015966	
40-49 years - 50-59 years	0.025648	0.013146	0.03815	***
50-59 years - <20 years	0.011789	-0.009367	0.032946	
50-59 years - >59 years	0.013098	-0.004027	0.030223	

References

Barnes, George R., et al. "The Visualization of Continuous Data Using PROC KDE and PROC CAPABILITY," SAS Users Group International 26, Long Beach, CA: Paper 176-26.

Ferrell, Jennifer. "A Comparison of the Mixed Procedure and the Glimmix Procedure," SAS Users Group International 31, San Francisco, CA: Paper 189-31.

Pickrell, Timothy. "Driver Alcohol Involvement in Fatal Crashes by Age Group and Vehicle Type," NHTSA Research Note, DOT HS 810 598, July 2006. Washington, DC: National Highway Traffic Safety Administration.

Shankar, Umesh. "Alcohol Involvement in Fatal Motorcycle Crashes," NHTSA Research Note, DOT HS 809576, April 2003. Washington, DC: National Highway Traffic Safety Administration. Available at: www-nrd.nhtsa.dot.gov/pdf/nrd-30/NCSA/RNotes/2003/809-576.pdf.

Subramanian, Rajesh. "Alcohol Involvement in Fatal Motor Vehicle Traffic Crashes, 2003," NHTSA Research Note, DOT HS 809 822, August 2005. Washington, DC: National Highway Traffic Safety Administration. Available at: www-nrd.nhtsa.dot.gov/pdf/nrd-30/NCSA/Rpts/2005/809822.pdf

For additional copies of this research note, please call 800-934-8517 or fax your request to 202-366-3189. This research note and other general information on highway traffic safety may be accessed by Internet users at: www-nrd.nhtsa.dot.gov/departments/nrd-30/ncsa/AvailInf.html.

www.nhtsa.gov

U.S. Department of Transportation
National Highway Traffic Safety Administration