|  |
| --- |
|  |
| Automobile Crashes Involving Deer |
| Jan 2003 – Dec 2007 |

|  |
| --- |
| Sarah Rawls, Phillip Domschke, Michael Davidoff, Alex Lentz, Jim Box  26-Aug-13 |

# Summary of Findings

Based on accident data from the NC Department of Transportation, there is a marked increase in automobile accidents involving deer in the fourth quarter of the year. In the years 2003 – 2007, there was an average of 929 deer-related accidents per month in the months of January – September. In those same years, there were 2796 deer-related accidents per month on average in October – December.

|  |  |
| --- | --- |
| Month | Avg Deer Related Accidents |
| Jan – Sep | 969 |
| October | 2321 |
| November | 3760 |
| December | 2306 |

Based on these findings, we strongly recommend an outreach program to your customers via radio advertising and billing inserts to remind your drivers of the added risks of deer related accidents in these months.

Contents

[Summary of Findings 1](#_Toc364975846)

[Analytical Methods 2](#_Toc364975847)

[Plots of Deer Related Accidents Over Time 2](#_Toc364975848)

[Regression Analysis – Base SAS 3](#_Toc364975849)

[Regression Analysis – Enterprise Miner 5](#_Toc364975850)

[Discussion 7](#_Toc364975851)

# Analytical Methods

## Plots of Deer Related Accidents Over Time

|  |
| --- |
|  |
| *Fig 1: Number of Deer Related Accidents per Month, 2003 - 2007* |

|  |
| --- |
|  |
| *Fig 2: Percent of Automobile Accidents Involving Deer per Month, 2003 - 2007* |

Figure 1 shows the monthly number of deer related accidents for 203 – 2007. We clearly see that there is a peak in November each year for several years. There seems to be an increasing trend of the peak each year since 2005. Figure 2 shows the percentage of accidents that involve deer. The fact that the patterns match indicates that cyclical increase in deer related accidents is not due to more accidents occurring; it is due to more interactions between drivers and deer in the fourth quarter of each year.

## Regression Analysis – Base SAS

The first analysis was a regression of Deer on the time variable and the seasonal dummy variables Mon1 – Mon12. This model returned the parameter estimates in Table 1 below:

|  |
| --- |
|  |
| *Table 1: Parameter Estimates of Linear Regression Model #1* |

There is a problem with mon12, because it is a perfect linear combination of all the other month variables in the model. This causes the model to not be full rank, which means that in PROC REG it will choose nonzero solutions for all the variables that are linearly independent of the previous variables, but will put zero as the solutions for the variables that are not linearly independent. In this case SAS produced nonzero estimates for all the variables except mon12. Another problem with not having a full rank model is that the solution given by the PROC REG output is not unique. The estimates that SAS produces are biased as well, and in the output SAS does put a B, for biased, beside all the variables that are biased.

We reran the model, taking Mon12 out of the model statement. The beta estimates remained the same for each regression. This is due to the fact that when the first model was coded mon12 was the last variable and SAS saw that it was a linear combination of the previous variables and set it equal to zero. Thus the beta estimates will be the same for the second regression that doesn’t include mon12, since all the previous variables are not linear combinations of the other variables.

Table 2 shows the ANOVA outputs for this second model (these were identical to the first).

|  |
| --- |
|  |
| *Table 2: ANOVA outputs for Linear Regression Model #2* |

The F test indicates that our model has statistically significant results. From the value of R-Square, we see that the model explains 98% of the variation of deer related crashes over time. The parameter estimates shown in Table 1 show the predicted number of deer related crashes each month relative to the number of crashes in December (indicated by the Intercept estimate), once you add in the month indicator t.

Based off of the beta estimates there are time trends in the number of accidents - fewer in the Spring and Summer and an increase in the late Fall and early Winter. This corresponds to the graphs that were shown in figures 1 and 2.

## Regression Analysis – Enterprise Miner

We ran the same models in Enterprise Miner. Figure 3 shows the same plot produced in Figure 1. Other than formatting, these figures are identical.

|  |
| --- |
|  |
| *Fig 1: Number of Deer Related Accidents per Month, 2003 – 2007(Enterprise Miner)* |

We reran Linear Regression Model #1, which included all seasonal dummies for mon1 – mon12. Table 3 shows the parameter estimates.

|  |
| --- |
|  |
| *Table 3: Parameter Estimates of Linear Regression Model #1(Enterprise Miner)* |

Note that the parameter estimates are different, and there is a different month with a 0 estimate. There is still the problem with specifying all twelve months, as one will be a perfect linear combination of the other 11. With the SAS output, Mon12 was the last variable we specified it the model, so that was the one that SAS identified as a linear combination of the others. With Enterprise Miner, the variables were added into the model by alphanumeric order, so Mon9 was the last specified and thus the one identified as a linear combination of the others. The values of the parameter estimates are different from Table 1 because they are using the Mon9 as the reference month.

We manually went in to the variable selection panel and set Mon12 as rejected and got the parameter estimates shown in Table 4.

|  |
| --- |
|  |
| *Table 4: Parameter Estimates of Linear Regression Model #2 (Enterprise Miner)* |

These parameter estimates are equal to the ones provided in Table 1, with small differences due to rounding.

To compute the predicted number of deer crashes in December 2004, we need to use the parameters for the Intercept (since Mon12 is not in the model) and for the time variable t. December 2004 is the 24th month in the data, so the prediction is the following:

PVDec2004 = 2062.0 + 6.8007(24) = 2225.2. The observed value for December 2004 was 2147, so our prediction was 3.6% higher than the observed value.

# Discussion

The current estimated white-tailed deer population in North Carolina is 1.35 million. There was a rapid population increase of deer in the state up until the early 1990s, when either-sex hunting season restrictions were loosened. This means that hunters were now able to hunt antlerless deer. The relaxation of these laws saw stabilization and in some cases a decrease in deer population for parts of the state. There are several areas that are continuing to see an increase in deer population which are the urban and suburban areas of North Carolina. The urbanization and suburban increase in North Carolina has limited the space deer have and there are stricter hunting laws in place in towns and suburban areas. We can see this in the county level laws in which most prohibit firing of a gun from the right-of-way of a highway and most urban areas have restricted archery seasons. November happens to be the middle of deer hunting season in North Carolina, which causes deer to be more skittish and move to safer areas.

The sharp increase in accidents in November could also be due to the rut (deer mating season) that occurs on average from the last week of October to the first week of December, depending on where you are located in the state. During this time male deer (bucks) increase their “home turf” to look for possible mates. This causes deer to come out and roam in more areas and during the day which is very rare during non-hunting or mating seasons.

Based off the results of the estimates of the regression models and the deer population information I would recommend either a paper ad or radio ad. In conclusion November is a pressure filled time for deer with mating and hunting season all at once, which causes many to roam where they normally wouldn’t and more frequently as well. If you wish to make your clients more aware and possible decrease the number of accidents a radio or paper ad in September or October would make the clients aware and hopefully safer drivers during that time of the year.