Analysis Project 2 – Driving Accidents

**Brief Introduction**

Car accidents are way too common and can be fatal, but at the least expensive and cumbersome. Driving is such a common thing in the United States due to its distances that a 16-year-old and a 95-year-old person can obtain a valid driver's license. In addition to this vast variability introduced by all the individual variability that the American country beholds there are additional behavioral and vehicle-related variability added to avoid accidents (i.e.: wearing a seat belt, vehicle age, having functional airbags, driver distractions). Some of these factors are included in a data set that will be used in this analysis to determine what factors play significant roles in accident fatalities.

Rationale:

I do not think that the age of the car matters as much because I believe that cars nowadays are made of cheaper and weaker materials than older, more robust vehicles. I also do not think that frontal impacts would be as fatal as non-frontal impacts unless the passengers did not wear a seatbelt. I consider the seatbelt the most important factor in accident fatality prevention. Additionally, my encounters with car accidents remind me that the most protection of the passengers are placed in front of them, because the engine tends to absorb most of the frontal impact. On the other hand, if a vehicle is hit from the side, much less layers of protection is available, and most vehicles are easier to flip from a side impact. Also, while wearing a seatbelt is important at any point to protect the passengers, seatbelts would offer more protection for frontal impacts to avoid flying through the windshield. Seatbelts may not as important or protective when the vehicle has a side impact with the exception if the vehicle flips. Finally, the age of the occupants in the vehicle matters because older adults may have a higher chance of passing from serious injuries compared to young adults.

Hypotheses:

The state proposed that vehicle age is one of them when controlling for the speed of the car, therefore, I will reanalyze that data to ensure their result accuracy using this data set. Additionally, I laid out the following alternative hypotheses:

1.) I hypothesize that the fatality of car accidents will be significantly associated with impact area and seatbelt usage when controlling for car speed and vehicle age.

2.) I hypothesize that the fatality of a car accident will be significantly higher when not wearing a seatbelt in a frontal impact compared to a non-frontal impact when controlling for the speed and vehicle age.

3.) I hypothesize that the fatality of a car accident will be significantly associated with the age of the occupant(s) when controlling for speed, impact area, and seatbelt usage.

**Methods**
A logistic regression was used to model the data with a binomial random component and a logit link function. Accident fatality (0 = alive, 1 = dead), impact site (0 = frontal, 1 = not frontal), airbag deployment (0 = deployed, 1 = not deployed), and seatbelt usage (none = 0, belted = 1) were dummy coded. McFadden's pseudo R2 was used to estimate effect size. The Wald test was used to determine the significance of the results since $N = 20,974$.

Model 1: Odds(died | alive) = $b_0 + b_1$(Speed) + $b_2$(Vehicle Age) + e

Model 2: Odds(died | alive) = $b_0 + b_1$(Speed) + $b_2$(Vehicle Age) + $b_3$(Frontal) + $b_4$(Seatbelt) + e

Model 3: Odds(died | alive) = $b_0 + b_1$(Speed) + $b_2$(Vehicle Age) + $b_3$(Frontal) + $b_4$(Seatbelt) + $b_5$(Frontal*Seatbelt) + e

Model 4: Odds(died | alive) = $b_0 + b_1$(Speed) + $b_2$(Frontal) + $b_3$(Seatbelt) + $b_4$(Age of Occupant) + e

**Results**
Vehicle age had no significant association with accident fatality when controlling for vehicle speed (Model 1, $z = 1.67, p = .095$) and this effect was moderate (*McFadden* = .19). At the mean level of vehicle age ($M = 6.74, SD = 5.31, Range = 0$-49) the odds of dying from a car accident were .37. One additional year in vehicle age multiplied the odds of dying from a car accident by 1.01 (CI[.99 – 1.02]). Vehicle age by itself showed a significant association with accident fatality ($z = 6.35, p < .001$), which indicates that vehicle age was no longer significant when vehicle speed was included in Model 1.

For Model 2 the results showed that wearing a seatbelt ($z = 12.32, p < .001$) and a frontal impact ($z = 14.20, p < .001$) were significantly associated with accident fatality when controlling for vehicle speed and vehicle age. This model had a moderate effect (*McFadden* = .23). The odds of dying when not wearing a seatbelt multiplied by 2.51 comparing to wearing a seatbelt (CI[2.17 – 2.91]). The odds of dying when the impact was non-frontal multiplied by 2.89 compared to frontal accidents (CI[2.49 – 3.35]).

For Model 3 the interaction was not significant ($z = -1.27, p = .205$) when controlling for vehicle age and speed. This model had a moderate effect size (*McFadden* = .23). Simple slope analysis revealed that for frontal accidents the odds of dying when wearing a seat belt were multiplied by .83 compared to not wearing a seatbelt (CI[.62 – 1.12], see Figure 1). The odds of dying for frontal accidents when wearing a seatbelt were .15, and when not wearing a seatbelt were .13. For non-frontal accidents the odds of dying when wearing a seatbelt compared to not wearing a seatbelt were multiplied by 1.21 (CI[.90 – 1.61]). The odds of dying for non-frontal accidents when wearing a seatbelt were .49, and when not wearing a seatbelt the odds of dying were .59.

The results for Model 4 showed that age of occupant was a significant predictor of accident fatality when controlling for vehicle speed, accident impact, and seatbelt usage ($z = 17.59, p < .001$) Model 4 had a large effect size (*McFadden* = .27). At the mean level of occupant age ($M = 37.34, SD = 17.98, Range = 16$-$97$) the odds of dying from a car accident were .17. One additional year in the occupant's age multiplied the odds of dying by 1.03 (CI[1.03 – 1.04]). The odds of dying from a car accident for 10 years of change was multiplied by 1.40 (CI[1.35 – 1.48]).
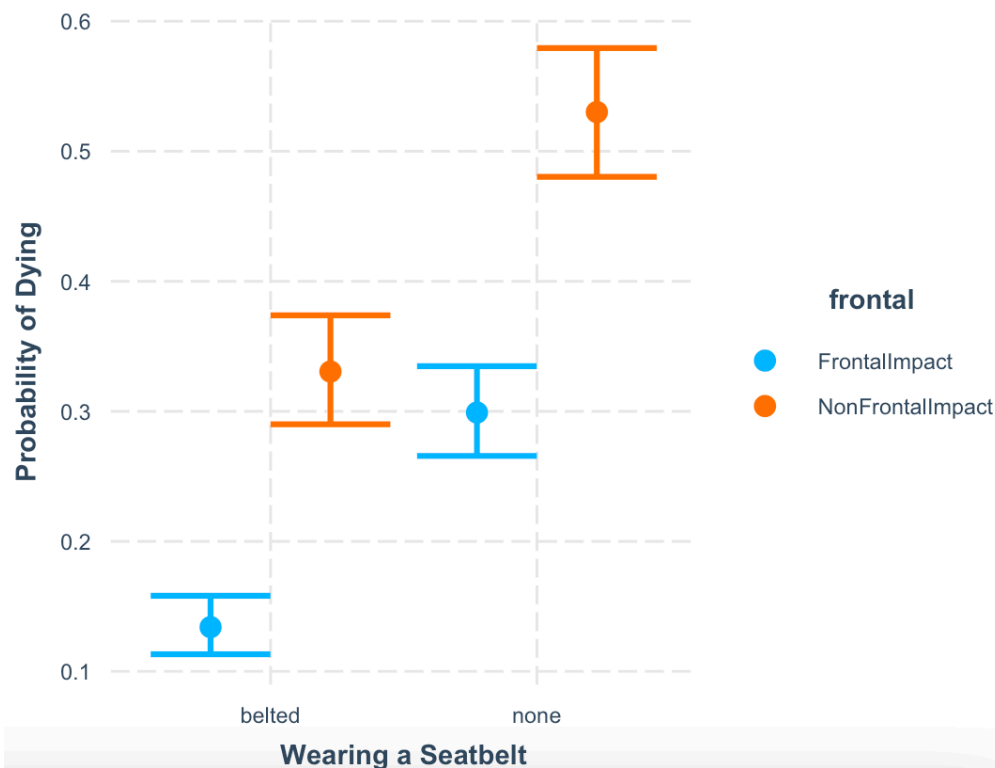


*Figure 1.* Model 3 interaction plot.

**Brief Discussion**
The results from Model 1 indicate that after controlling for vehicle speed, the age of the vehicle was not a significant predictor of accident fatality. The state's assumption of older vehicles contributing to accident fatality was no longer true once the speed of the vehicle was accounted for. This suggests that vehicle speed at the time of the crash has a stronger association with accident fatality than vehicle age.
The second analysis suggests that impact contact and seatbelt usage were significantly associated with the odds of dying in a car accident when controlling for speed and vehicle age. This means that wearing a seatbelt and having a frontal versus non-frontal car accident significantly impact car accident survival.
The interaction between seatbelt usage and impact site was not significant, suggesting that there is no significant difference in the odds of survival between wearing a seatbelt or not wearing a seatbelt in a frontal accident.
The last analysis looked found that the age of occupants had a significant association with accident fatality when controlling for vehicle speed, seatbelt usage, and impact site. This suggests that the age of the occupant also matters when considering dying from a car accident.


Limitations:
It is possible that other variables such as airbag availability or deployment or the status of the occupant (driver or passenger) may contribute to accident fatality, however, the scope of this research focused on the above-mentioned variables based on the outlined rationale and theory. Future research should address other variables and observe their relationship with car accident fatality.