Motor Fatality in the United States from 2000 to 2015.

Tenzin Thekchok

Abstract

The goal of this thesis is to investigate factors that affect the odds of having a fatality in a vehicle collision. For this regression research, we will focus on the motor vehicle fatality that happened during the 2000s to 2015 in the United States. To better understand what might cause a fatality we will be looking at the population of the United States, the total number of miles driven and the number of cell phone subscriptions in the United States. All of these variables might affect the odds of having a fatality. There are many more variables that could also affect the chance of fatality. Modeling will involve the use of regression. We will use Excel and R systems for data management and try to find the best fit line. By identifying the best factors, the selected model might help find a safe solution to reduce the rate of fatality.

Introduction

Throughout our history, Motor vehicles have played a huge part in our civilization. We evolved as the wheels turned. After the United States of America witnessed tremendous economic growth which has resulted in the expansion of the road network and subsequent increase in motor vehicle population. As soon as motors were introduced we started seeing rapid changes in our society and the way we transport our goods and us. It is one of mankind's biggest and best inventions. The United States of America has one of the most number of motor vehicles in the whole world and every type of motor vehicle there is. As technology advanced, we started seeing more of them, they got bigger, faster, and abundant. Because of that many people die from it every year. In the United States, the definition of road-traffic fatality that is used by the Fatality Analysis Reporting System (FARS), which is run by the NHTSA, is "a fatality in the state of Nevada is defined as any person that dies within 30 days due to a vehicle accident that occurred on a United States public road and the vehicle had an engine." The United States has the highest death from motor accidents in the world. There are many factors that could affect the death caused by a motor vehicle.

Literature Review

"In 2011, for the third year in a row, motor vehicle traffic crashes were not among the top 10 causes of death in theUnited States." (Crashstats.nhtsa). There is a plethora of research done on motor vehicle accidents because of how important it is becoming. As civilization gets better we will only see it increasing.

"When ranked by age, motor vehicle traffic crashes were the number one cause of death in 2011 among people ages 8, 13, 14 and every age 16 through 25. In contrast, motor vehicle traffic crashes were the number one leading cause of death for people ages 13 through 25 in 2010"(Crashstats) Based on their statistics most young Americans' deaths were caused by vehicles. While this study does not help to better understand what the variables are, we can conclude from their research that it is a very serious matter. Additionally, this source proves one of my hypothesis wrongs, over the years the vehicle fatality does not increase but recently since 2005 it's been decreasing.

In another study by aphapublications.org, they "study to test associations between BMI and gender and other important covariates and to show how these factors modified the association between BMI and driver motor vehicle crash fatality. Simultaneously adjusting for potential confounding factors allowed us to assess the association between BMI and fatality over and above other factors." With their research, they concluded that Male drivers who had either a high or low BMI (e.g., BMI>35 kg/m2 or <22 kg/m2) had a significantly increased risk for death compared with

those who had an intermediate BMI, but female drivers did not. Among men, the association between BMI and driver fatality was modified by the type of collision, whereas there was very little difference between either gender in the association with age, airbag deployment, or seatbelt use. The magnitude of the increased risk for a fatality at the high end of the BMI continuum among men was determined mainly by the magnitude of the change in velocity. Additionally, they found that men who had higher BMIs were at greatest increased risk for death if they sit at the front end and left side of the collisions (approximately 73% of all crashes). One of the challenges in the research was how the female's BMI doesn't have significance while made does. He concluded that "the increased risk for death due to motor vehicle crashes among obese men may have important implications for traffic safety and motor vehicle design." With that information, we can manufacture a special seat or seat belt that could help men with high BMI.

In America, drinking is a big culture and drinking can cause a person to lose their sense of distance, distorted vision and hearing. "Alcohol related crashes are the leading cause of death for young Americans, between the ages of 16 and 24 years old. For all Americans between 5 and 35 years of age, motor vehicle accidents are the number one cause of death. Over 50% of these accidents are caused by alcohol-impaired drivers."(faddintl.org). This shows how much alcohol consumption affects the rate of a fatality caused by vehicles or just the safety of people. It was ranked number one for young Americans.

Now putting all of them together, the next source tells us about the fatality rate of car accidents in the United States from the 1900s to 2018. The data he collected provides variables such as the number of deaths, vehicle miles traveled, fatalities per million population, fatality per 100,000 million Americans and the change in a capital fatality from the previous year. Even though this doesn't help us understand what might affect the rate of fatality in vehicle accidents, we at least better understand the change in the fatality rate. With the fatality rate, we can think of what happened in that year or if there is any new technology or law that helps prevent accidents.

Methodology and Data

For this data, my goal is to understand what might cause or affect the rate of fatality by motor accidents from 2000 to 2015. I used time series analysis and put the years in order to get the number of deaths. With that, I can further my research and find if there are any laws or new technology that might reduce the number of fatalities. So for that, I used multiple regression to get a better understanding of which variables best explain the data I collected. I collected data online from different websites from the year 2000 to 2015 and I used three variables that will help me understand the cause of death by a vehicle accident. I choose the number of vehicles in the U.S, the number of cell phone subscriptions and the population as my independent variables and fatality as the dependent variable. With the data collected, I ran the regression model on excel.

Result

The following is the regression summary after I ran the regression in Excel with the data I collected.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.973080569
R Square	0.946885793
Adjusted R Square	0.933607241
Standard Error	1186.754104
Observations	16

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	301293408.1	100431136	71.30941787	6.43278E-08
Residual	12	16900623.63	1408385.302		
Total	15	318194031.8			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	99137.53789	76280.02186	1.29965272	0.218132445	-67062.35239	265337.4282	-67062.35239	265337.4282
X Variable 1	0.03538604	0.007420574	4.768639288	0.000457332	0.019217998	0.051554081	0.019217998	0.051554081
X Variable 2	4.16828E-06	4.014E-05	0.103843553	0.919008864	-8.32893E-05	9.16258E-05	-8.32893E-05	9.16258E-05
X Variable 3	-0.000550556	0.000237171	-2.321341964	0.038672999	-0.001067307	-3.38038E-05	-0.001067307	-3.38038E-05

If I were to put it in the formula it will be

Fatality=999137.5+0.035numberofvehicles+4.17cellphonesubscription-0.001population

- The Y- intercept is 99137.54 and that means that there are 99137 deaths when the independent variables are 0.
- The value of B1 is 0.0354. That means, holding everything else fixed when there is one more vehicle on the road in the U.S the rate of fatality increases by 0.0354.

- The value of B2 is -0.000004 or -0.00001. That means, holding everything else fixed when there is one marginal cell phone subscription the rate of fatality decreases by 0.000001.
- The value of B3 is -0.00055. That means, holding everything else fixed when the population increases by one the fatality rate decreases by 0.0006.

Based on the data with R square, Multiple R and P-value. The value we got from Multiple R is 0.97 which is 97%. The Multiple R represents how strong the linear relationship is, with that we can understand that the relationship between the variables is 97% which is very high. The R square also is known as the Coefficient of Determination, It tells you how many points fall on the regression line. Similar to Multiple R it helps us understand the regression line. The value we got from the model is 0.95 or 95%. In other words, it means that 95% of the variables fall on the regression line and it is very good. And lastly, P-value represents " hypothesis testing to help you support or reject the null hypothesis. The p-value is evidence against the null hypothesis. The smaller the p-value, the stronger the evidence that you should reject the null hypothesis."(statisticshowto.datasciencecentral.com). From the data, we can say that from the three variables, the number of cars in the United States is the most significant when it comes to explaining the fatality rate because it has the lowest P-value of 0.00045 while the number of cellphone subscriptions is very high with .92 or 92%. Which is extremely high.

Conclusion

It makes sense that the more cars we have and the bigger the population, the higher the rate of fatality. But it can be otherwise because of technology and laws that were passed. For example, we have seat belts, airbags and drinking age limits. We used fatality as a dependent variable and the number of cars, cellphone subscriptions, and population. Of course, there are many more variables that could affect the fatality rate. After running the regression we can conclude that the data have a very positive relationship because it's close to 1. Additionally, from the three variables numbers of the car on road best explain the fatality rate so the best solution for the government is to regulate car production more. Maybe increase the tax on oil and cars or provide more public transportation like trains, subways, and busses.

References

- U.S. Vehicle-Miles. (n.d.). Retrieved December 12, 2019, from https://www.bts.gov/content/us-vehicle-miles.
- US Population by Year. (n.d.). Retrieved December 12, 2019, from https://www.multpl.com/united-states-population/table/by-year.
- US Census Bureau. (2016, November 30). State Intercensal Tables: 2000-2010. Retrieved December 12, 2019, from https://www.census.gov/data/tables/time-series/demo/popest/intercensal-2000-2010-state.html.
- Holst, A. (2019, September 4). US mobile cellular subscriptions 2000-2018. Retrieved December 12, 2019, from https://www.statista.com/statistics/186122/number-of-mobile-cellular-subscriptions-in-the-united-states-since-2000/.
- Drunk Driving Fatality Statistics. (n.d.). Retrieved December 12, 2019, from https://www.responsibility.org/alcohol-statistics/drunk-driving-statistics/drunk-driving-fatality-statistics/.
- Motor vehicle fatality rate in U.S. by year. (2019, December 7). Retrieved December 12, 2019, from https://en.wikipedia.org/wiki/Motor_vehicle_fatality_rate_in_U.S._by_year.
- 2019 Drunk Driving Statistics: The Zebra. (n.d.). Retrieved December 12, 2019, from https://www.thezebra.com/drunk-driving-statistics/.