**Module 8: Portfolio Project, Option #1: Capstone Project—Business Intelligence Solution for U.S. Organization**

Irfan Peterson

Colorado State University – Global Campus

MIS480-1: Capstone - Business Analytics and Information Systems

Professor Bateh

August 1, 2021

**Module 8: Portfolio Project, Option #1: Capstone Project—Business Intelligence Solution for U.S. Organization**

Electric cars have been in existence as long as gas-powered cars. The first electric vehicle was developed in 1832 by Robert Anderson, however, due to the ease of manufacturing and increasing popularity of the Ford Model T, the development of electric cars came to a halt (Matulka, 2014). Fast forward to the modern age with increasing concerns of climate change and unreliable energy sources and storage, one company chose to go against the grain by offering a compelling electric vehicle with an ultimate goal to accelerate the advent of sustainable energy (Musk, 2016). Martin Eberhard, Marc Tarpenning, and their group of engineers are responsible for changing the automotive landscape and shift the industry towards more sustainable transport by bringing electric cars into the mass market. Back in 2003, they founded Tesla Motors, driven by their passion for electric cars and the desire to show the world that it is a smart alternative without sacrificing the performance of gasoline-powered cars (McFadden, 2020). In 2004, Elon Musk of SpaceX invested $6.3 million of his earnings from Paypal to become the chairman of its Board Directors where he is credit with navigating the company from bankruptcy back in 2008 to the success they are known today (McFadden, 2020). His master plan to introduce a high-end sports electric car to fund and develop a higher unit volume at lower prices was a massive success with approximately 1.5 million electric cars sold (Musk, 2016) and 16,512,273.98 tons of CO2 saved as of today (Tesla, 2021). Despite Tesla’s success and the good that they have done to help the world move toward more sustainable energy, many companies argue in bad faith and want Tesla to fail. This report seeks to disprove some of the bad faith arguments circulating Tesla.

**Tesla, Company of the Future and the New Problem Theory**

Tesla is currently one of the most important companies in the world. Their technology to transform ourselves to become more energy independent and towards a more sustainable future is admirable. Even with such a positive mission statement, Tesla encounters many negative perceptions from competitors, or any other companies that are threatened by the positive technologies that they developed. In the past, many experts did not believe Tesla will survive after more than a few years. Many investors were cautious with the fact that the company was not making a profit for many years. From 2006 to 2016, Tesla did not focus on profit as a part of its plan, upon closer look, however, it was clear that the company was focusing much of its effort reinvesting in the business upgrading their equipment, increasing their R&D to extend their product line, and expanding their presence all over the world (Mackenzie, 2020). The decision for the company to focus invest in itself paid off. Tesla is now valued with a market cap of over $400 billion at its peak, which has created a shift in the automotive industry. Many established automakers are now looking to maintain the competitive advantage by incorporating alternative energy vehicles into the lineup.

Now that the profitability arguments have been put to rest, one of the fears, uncertainty, and doubt circulating Tesla and its technology is the dangers of their vehicle and its self-driving technology. Many outlets claim that the Tesla autonomous self-driving program is a danger to everyone. It is a hazardous technology that should be outlawed or heavily regulated, which it already is. There are many perceptions that the rate for Tesla to malfunction is very common despite the claim that Tesla made that their vehicle can react to situations more effectively than humans (Dow, 2019).

**Dataset and Storage Method**

To dispel the myth surrounding the danger of Tesla’s autonomous driving technology, data compiled by Tesladeaths.com will be used. This dataset is a record of Tesla fatalities and accidents that have resulted in deaths as of today’s date and it is updated regularly for accessibility purposes. Variables such as date, country, state, number of occupants, other vehicles involved, and sources are capsulated in the dataset. The dataset is available in Excel and CSV format which makes it easy to import to data analytics programs.

**BI Tool to Perform Analytics**

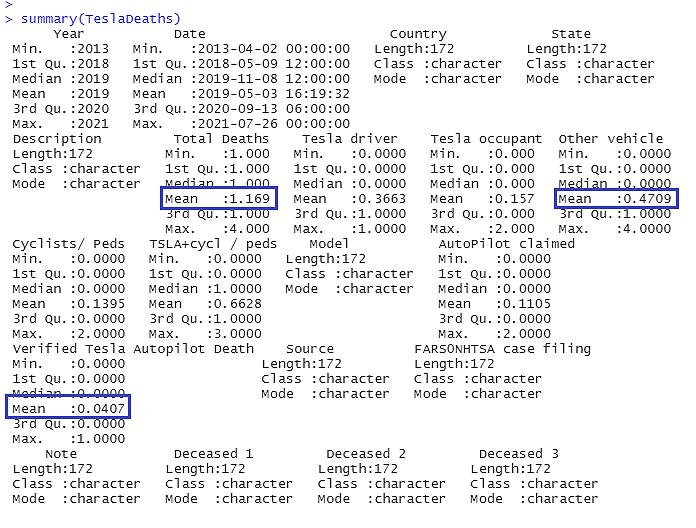
The BI tool that will be used to analyze the Tesladeath dataset is R Studio. The objective of R is for statistical computing, data mining and analytics, and as well as graphical tools (Bansal, A., & Srivastava, S, 2018). The open-sourced environment for R-Studio makes it the best program to produce diverse statistical calculations as R has different data packages to conduct various statistical tasks such as clustering, regression analysis, text mining, decision trees, time series analysis, and RHadoop supports big data processing and makes it well equipped to handle big data problems (Bansal, A., & Srivastava, S, 2018).

**Statistical Analysis**

The first task is to investigate the Tesla death data set using statistical analysis which shows the important qualities of the data that was gathered. One method that is used to gather a general view of the data is conducted through univariate analysis which consists of summary statistics, and frequency distribution analysis. According to Bhattacherjee (2012), univariate analysis refers to a set of statistical techniques that illustrates the general properties of the chosen variables, which includes the frequency distribution, central tendency, and dispersion. The central tendency on the other hand is an estimate of the center value of the distribution and is measured by calculating the mean, median, and mode (p.121). The last element of the univariate analysis is the dispersion of the data. Dispersion measures the spread around the central tendency of the data by calculating its range and standard deviation (p.122).

**Figure 1**

*Summary statistics of Tesla death*

**

The summary statistics function gives an overview of the minimum, 1st quartile, 3rd quartile, and maximum value and its central tendency which is the median and the mean value. Listed under Figure 1 above, we can see the summary statistics of the Tesla Death dataset. The analysis shows accidents that involve Tesla vehicles and resulted in deaths resulted in an average value of 1.169 deaths per event and a count of four deaths as its highest single-event tragedy. The average death rate that involves other vehicles comes in an average of 0.4709. Notice that fatalities caused by the autopilot program come in at a rate of 0.0407 which shows how relatively safe Tesla’s autonomous program is compared to human drivers.

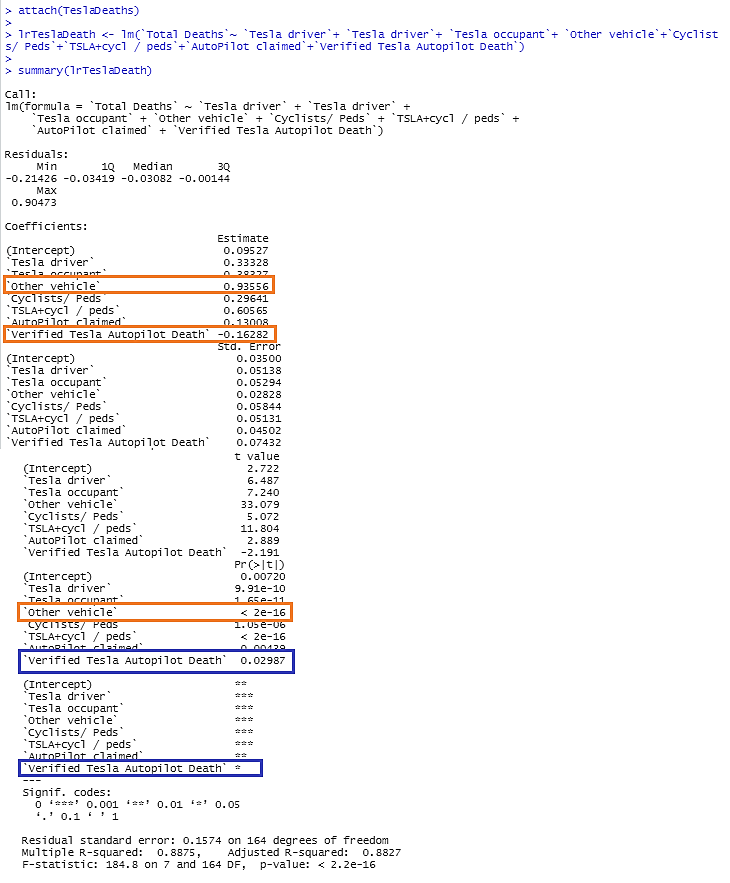
**Linear Regression Model of Tesla Death Dataset**

While summary statistics give a good overview of the dataset, multiple linear regression must be performed to help identify which factors and variables make the most impact on the fatalities involving Tesla and the autonomous program Multiple linear regression is a methodology that allows researchers to answer questions that consider the role that multiple independent variables play in accounting for variance in a single dependent variable (Nathans et al, 2012, p.1). The dependent variable, also known as the response variable, is being modeled and predicted, while the independent variable, known as the predictor variable, is the variable used to predict the response (zyBookz, 2018, Section 6.1).

Since we want to understand the mechanics behind the total fatality and contributing factors, The total deaths will be the response variable, while the remaining seven variables will be the explanatory or predictor variable. The first step to conduct this analysis is to use the linear model function notated as lm(). Once the lm() function is configured, we can display the results using the summary of the linear regression model. The summary will display values such as standard error, coefficient estimates, and p-value. The p-value determines which variables have the most impact on the total death variable.

**Figure 2**

*Multiple linear regression of Tesla Death dataset*



The result generated from the multiple linear regression model as shown in Figure 2 above further supports our hypothesis. Although all the predictor variables resulted as statistically significant, most likely from the very limited sample size due to Tesla being a newer brand or how relatively safe it is as a vehicle, we can see that fatalities verified as a result from Tesla Autopilot have the highest p-value at 0.02987. Note that the lower the p-value, the stronger of a predictor it is. Additionally, when looking at the coefficient estimate value, the verified Tesla autopilot death variable comes in at a negative value. A negative value indicates that as the independent variable increases, the dependent variable decreases, meaning that the Autopilot software decreases the fatality rate (Frost, 2021).

**Data Visualization**

For data visualization needs SAS is the preferred software. SAS is statistical software that enables us to dissect data for advanced analytics, multivariate analysis, data management, and data presentation. However, compared to R studio, SAS does a better job in data visualization.

**Figure 3**

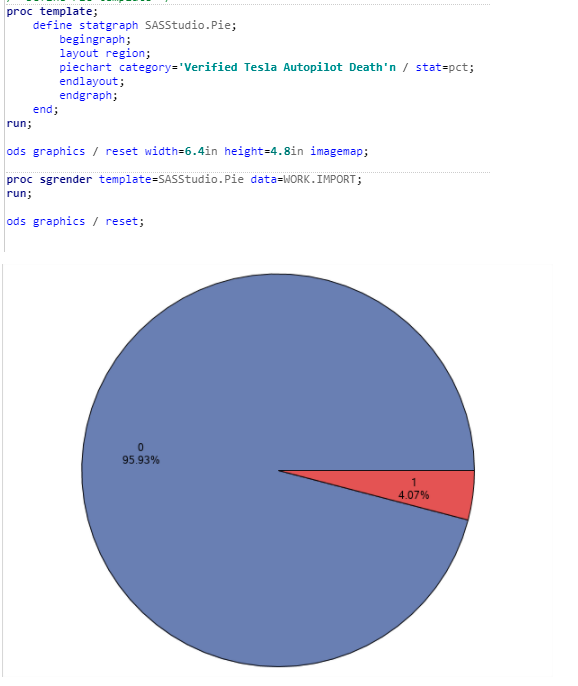
*Stacked bar chart of a fatality caused by Tesla Autopilot versus other vehicles*

****

As a continuation of statistical analysis, the frequency distribution is a summary of measure for a variable’s frequency of individual values or ranges for that variable, a numerical value or percentage may be used to describe the variable (Bhattacherjee p.121). A frequency distribution can be depicted using a bar chart. A bar chart is a great way to show the characteristic of a categorical variable (zyBookz, 2018, Section 8.3.3). Displayed in Figure 3 above, is a variation of a bar chart called the stacked bar chart. We can see the distribution of fatalities; the graphs show an overwhelming amount caused by other vehicles compared to the Tesla Autopilot system. Additionally, single event fatalities in the amount of two or more do not involve the autopilot system.

**Figure 4**

*Pie chart of fatality caused by Tesla Autopilot versus other vehicles*



Pie charts show the relative frequency of the distribution by dividing the circle into each category slice in an appropriate size (zyBookz, 2018, Section 8.6). Figure 4 shows the ratio of fatalities caused by the Tesla Autopilot system (in red) versus situations that arise from other vehicles. This shows that there is a four percent likelihood of the autopilot system causing a fatality, which is much lower, thus further supports the argument.

**Benefits of Business Intelligence/Business Analytics**

The ever-increasingly competitive business environment puts a premium on the way information is perceived. Millions and often billions of dollars in investments made into this sector have made it clear that the benefits are worth the cost. Upon the many benefits, businesses seek BI to make their firm more efficient. Goals to increase productivity can be met using BI to gain insight into underperforming processes, detect ways to decrease costs, uncovers insights efficient processes (Synoptek, 2020). Accuracy is another avenue to be improved with business intelligence. Real-time information from a wide range of sources of data is used to produce accurate reporting that delivers faster decision-making speed (Predictiveanalyticstoday,2020).

**Conclusion**

Electric cars are changing the world for the better. Not only does it provide a more sustainable way to travel, but companies such as Tesla are creating technology that makes their vehicle safer, despite the bad faith news circulating. Statistical analysis as shown above supports the argument. The p-value from the multiple regression model states that the autonomous software decrease fatality events and the summary supports that notion as well.

**References**

*6 Amazing Benefits of Business Intelligence*. Synoptek. (2020, September 7). <https://synoptek.com/insights/it-blogs/6-amazing-benefits-business-intelligence/>.

*Top 15 Benefits of Business Intelligence Software in 2021 - Reviews, Features, Pricing, Comparison*. PAT RESEARCH: B2B Reviews, Buying Guides & Best Practices. (2021, May 7). <https://www.predictiveanalyticstoday.com/top-benefits-of-business-intelligence-software/>.

Bachman, E. (2021, July 1). *Digital record of Tesla crashes resulting in death*. TeslaDeaths.com. <https://www.tesladeaths.com/>.

Bhattacherjee, A. (2012). *Social science research: Principles, methods, and practices*. Retrieved from <https://scholarcommons.usf.edu/oa_textbooks/3>

Bansal, A., & Srivastava, S. (2018). [Tools used in data analysis: A comparative study](https://csuglobal.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=129311324&site=eds-live). *International Journal of Recent Research Aspects, 5(*1), 15–18.

*Carbon Impact: Tesla.* Tesla, Inc. (2021). https://www.tesla.com/carbonimpact.

Dennis, A., Wixom, B., & Roth, R. (2019). *Systems analysis & design* (7th ed.). Wiley. ISBN: 9781119496489, 1119496489

Dow, J., & Jameson Dow Jameson has been driving electric vehicles since 2009. (2019, October 23). *Tesla Autopilot safety almost 9x safer than average driving*. Electrek. https://electrek.co/2019/10/23/tesla-autopilot-safety-9x-safer-than-average-driving/.

Frost, J. (2021, June 8). *How to INTERPRET p-values and coefficients in regression analysis*. Statistics By Jim. https://statisticsbyjim.com/regression/interpret-coefficients-p-values-regression/.

Mackenzie, W. (2020, January 20). When will tesla make a profit? Retrieved April 04, 2021, from <https://www.forbes.com/sites/woodmackenzie/2020/01/20/when-will-tesla-make-a-profit/?sh=30b08aed1ba9>

Matulka, R. (2014, September 14). Timeline: History of the electric car. Retrieved April 19,

2021, from <https://www.energy.gov/timeline/timeline-history-electric-car>

McFadden, C. (2020, July 6). *The Short but Fascinating History of Tesla*. Interesting

Engineering. https://interestingengineering.com/the-short-but-fascinating-history-of-tesla.

Musk, E. (2016, July 21). *Master Plan, Part Deux.* Tesla, Inc.

https://www.tesla.com/blog/master-plan-part-deux.

Nathans, Laura L., Oswald, Frederick L., & Nimon, Kim (2012). Interpreting Multiple Linear Regression: A Guidebook of Variable Importance. Practical Assessment, Research & Evaluation, 17(9). Available online: http://pareonline.net/getvn.asp?v=17&n=9

zyBooks. (2018). Statistics for data analytics. Retrieved from https://www.zybooks.com/catalog/statistics-for-data-analytics/