Tesla Death Analysis Updated

**Introduction:**

Tesla has been in the center of debate when it comes to the safety of its drivers and pedestrians. Many people tend to think the autopilot is fool-safe and is safe when it is not. Although these deaths may not be as common, it is best to try and investigate whether Tesla’s are indeed safe. The National Highway Traffic Safety Administration has reported on these many times because it is a reoccurring issue. Many other manufacturers are also cracking down on this safety aspect of electric cars as well to push their product past Tesla. The NHTSA also states that majority of the automated car crashes are involving a Tesla while other agencies and articles brag about Tesla being the safest cars on the planet.

This analysis will be investigating the question: Are Tesla cars as safe as everyone says they are? There are inconsistencies that need to be put to the test by using data analysis and visualization to see whether Tesla crashes are fatal and whether the car was assisted in driving automatically.

The dataset I chose from Kaggle.com will absolutely help because it not only shows if the deaths were in an auto-pilot driving car, but it separates each case by time, date, whether autopilot was claimed or not, whether the autopilot was confirmed, and whether the pedestrian, driver, or passenger was injured or killed. The dataset gives an abundance of detail regrading each case which I can figure out to use to the best of my ability to answer the question. Overall, the dataset will do wonders in answering the question posed due to the sheer details included.

This question is very important because as Tesla continues to hold a monopoly in the electric car industry, it is important to try and find its flaws so other companies can compete properly and fix these issues in their makes. There is a genuine misconception when it comes to Tesla cars and safety. Since not everyone owns a Tesla, it would be difficult to ride in one and get these data before jumping into the data and performing statistical inferences and data visualizations. Because this dataset is strictly being used, doing analysis on the data without physically having to go and drive Tesla cars or observe when they crash will suffice. This dataset’s detail alone is enough to conduct several statistical analyses regarding Tesla’s safety. This question is important also because it holds Tesla accountable for their vehicles and the claims they make. Just because they are the first ones in the business and hold a sort of monopoly, it does not excuse them from normal safety protocol when it comes to developing and delivering a vehicle for the public.

**Literature Review:**

Upon researching Tesla and deaths associated with its cars, a few resources that have kept a record of crashes in the past few years were found. Some of these records go back as early as 2013. Tesladeaths.com has an archive of every incident since 2013 with a source link to the news article or news segment that confirms the death along with the causes and other important details. An analysis with the data was not done to make any claims on the population of all Tesla cars and their involvements in deaths. Edward Helmore wrote about a study done that claimed that the Full Self-Driving (FSD) was linked to hitting children. This study seemed very biased as it brought in the fact that Elon hates the news agency. It felt politicized and did not seem to be an accurate representation of data. A thorough analysis to find out which variables in the dataset I found are significant and then model a visualization around those variables. Politics will not be brought into the study and will just use data analysis techniques and good ethics pertaining to data analysis.

**Exploratory Data Analysis:**

This dataset was chosen because it has many components that can be analyzed. The dataset includes critical information such as the country, and state in which the accident had occurred. These two will be the most important variables in the study because my thesis is leaning towards the fact that certain states have more relaxed driving laws which leads to more accidents in general. Tesla is somewhat to blame for improper rollout of certain safety features, but overall, states that have less severe laws for drivers, see the greatest number of accidents. By focusing on the locations of the accidents, if a certain country is involved in higher crash rates, it can be pinpointed. Then, if it’s the United States, look at everything by state and see which state has higher crash rates. If there is no true outlier or state that has significant number of crashes, move onto another variable such as the victim’s profile.

The dataset has many columns in it that are coded by using 1’s. These variables include information such as whether the victim who was hit/killed was a pedestrian or a driver. The analysis will try to investigate if there is a possibility that other drivers may have been responsible for the crashes vs the Tesla owners being at fault. Overall, the dataset gives me a plethora of analysis options to take and to really answer the question of whether Tesla cars are as safe as they claim to be or not. The dataset should have included articles linked to each case, instead, if certain information is needed, it will have to be pinpointed using the information given from the dataset. This is not something that will prevent the completion of this analysis.

**Exploratory Data Analysis (Updated):**

The analysis started off by creating a scatterplot of the data because to find out where exactly the most deaths involving Tesla cars was happening, a visualization was needed.

This is the result:

Chart, scatter chart

Description automatically generated

This was a very rudimentary plot and did not give me too many insights as to what was going on.

So the geom\_bind\_2d() function was used to create a plot that gave me a better visual.

Chart, scatter chart

Description automatically generated

This plot is much better because it is showing me where the most concentrated number of accidents involving Tesla cars is occurring. That place happens to be the United States. Then to somehow check if Auto Pilot was engaged, the fill() command was used to add that variable to a heatmap. This did not help at all because it seemed the place most suspected, the United States, was a country where Auto Pilot was not claimed.

Chart, waterfall chart

Description automatically generated

This plot was created to see whether there were even accidents involving Auto Pilot were happening in the United States. So, the result was that there were accidents in the United States involving Auto Pilot.

Scatter chart

Description automatically generated with medium confidence

Chart, waterfall chart

Description automatically generatedThis leads me the last plot for the Exploratory Data Analysis.

This plot shows the concentration of accidents in each state while also showing how many deaths occurred in each case. These plots will help analyze the ‘Auto Pilot Confirmed’ column of the data so see the true results.

Second Exploratory Data Analysis

The second dataset was also found on Kaggle.com. It contained a similar dataset which recorded deaths and their locations. The locations included country and states just like the previous dataset. After doing a similar analysis of the data using methods that were used on the previous dataset, very similar results were found between both datasets. The most accidents involving Tesla cars happened in the United States.

In the comparison of the number of deaths per incident between the two datasets, the United States is still the country with the highest number of deaths. The only difference is that Germany and China also have accidents that occurred in which three people died. To see if this is because of a few cases or if this was more common in those countries, another visualization was done to see if those two other countries have more frequent accidents involving Tesla cars.

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated**Initial Dataset Supplemental Dataset**

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated**Initial Dataset Supplemental Dataset**

This visualization proves that both the datasets have a trend. Majority of the accidents involving Tesla cars do happen in the United States. The most deaths and number of people who died per accident are highest in the United States. A new method that may need to be investigated is to see whether Tesla deaths have gone up over the years. This would require a breakdown of each year and to compare the accident to death ratios for that year to see if that value has increased or decreased. As Tesla claims to have safer cars now than before, this value would have to have a steady increase over the years.

**Validation**

To validate this prototype, a continual analysis will be conducted with the dataset initially found to prove that Tesla cars are indeed unsafe when it comes to Full Self-Driving systems. A daily article search will need to be conducted as well because news is coming in daily about new Tesla failures. There is a new article from Reuters.com released on February 16, 2023, that revealed a Tesla recall. 362,000 Tesla cars were recalled because they were deemed unfit by the NHTSA (Shepardson, 2023). These cars were going through intersections in an unlawful way and would drive well past speed limits on highways. These are all the types of issues that need to be considered when doing the analysis of Tesla cars. To validate the prototype, a constant check on news and articles followed by new techniques to visualize the data to prove that Tesla cars are unsafe.

**Project Failures**

Some failures of the project include the dataset itself, the fact that news is coming out frequently regarding the status of Tesla cars, and the lack of true data on the topic at hand. The dataset from Kaggle.com is very limited in terms of information. It does not include any articles pertaining to each incident that occurred. This makes it hard to confirm information without having to do extensive research for each case. The fact that news comes out frequently and Tesla is on the forefront of investigations and recalls, it makes it difficult to come to any conclusions because data is still being processed and being released to this day. This leads to a lack of data or lack of foreseeable data. It could also make data currently being used become obsolete because it has either been disproven or the issue has already been fixed, which would require more pending data from the newer models to confirm if the issue has really been resolved.

**Web Application Framework**

The Shiny package in RStudio could be used to create an interactive drop-down web application that could allow users to select a country, state, tine, model of Tesla, and auto-pilot features to see what the accident to death ratios are and if they have increased or decreased over the years to help visualize this study.

# Bibliography

Devastator, T. (2023, February 1). *Tesla Deaths Updated (2023)*. Retrieved February 2023, from Kaggle.com: https://www.kaggle.com/datasets/thedevastator/tesla-accident-fatalities-analysis-and-statistic

Helmore, E. (2022, August 2022). *Tesla’s self-driving technology fails to detect children in the road, group claims*. Retrieved February 2023, from theguardian.com: https://www.theguardian.com/technology/2022/aug/09/tesla-self-driving-technology-safety-children

Shepardson, D. (2023, February 16). *Tesla recalls 362,000 U.S. vehicles over Full Self-Driving software*. Retrieved 2023 February, from Reuters.com: https://www.reuters.com/business/autos-transportation/tesla-recalls-362000-us-vehicles-over-full-self-driving-software-2023-02-16/

Srinivasan, S. (n.d.). *Tesla Autonomous Deaths data*. Retrieved February 2023, from Kaggle.com: https://www.kaggle.com/datasets/sripaadsrinivasan/tesla-death-data

tesladeaths.com. (n.d.). *Tesla Deaths*. Retrieved February 2023, from tesladeaths.com: https://www.tesladeaths.com