Should You Consider An Electric Vehicle?

Miguel Rocha

12/12/21

CIST 2500-850 Applied stats for IST

Table of Contents

Introduction

The electric vehicle (EV) has changed since the first developed EV in the United State in 1890 – 1891. William Morrison of Des Moines, Iowa was the first person in the United States to develop a six-passenger wagon that its top speed was 14 mph. The EV has always been an interest in the United States since the development of the EV. EV was not very popular throughout the United States because of the high cost of the batteries, low life of the batteries, difficulty of recharging the batteries and low speed of the vehicle.

Different manufacturers throughout the world have been interested in the research and development of EV technology. The first fully EV in the United States was developed by General Motors in 1996 but not until 2008 Tesla was the first company to have an EV on the legal highways. From that time forward brands have increasingly changed the vehicle industry and exponentially grown the EV industry from the past century.

People have become concerned about the environment because of the gas-powered vehicles emitting a lot of smoke and harming the atmosphere. The EV has become adopted into cities and towns. The EV technology that manufactures have been able to produce has introduced a competitive industry.

Descriptive Statistics

Data will be converted into information to better understand. Graphs and Tables will display the information to better understand the data that the dataset has provided.

Description of Topic

The EV dataset that the project will be using, has been obtained from keggle.com (Divyanshu Gupta, 2021). With analyzing the categories that people are concern with regarding EVs such as,

- Company
- Acceleration as 0-60 mph per second
- Top speed MPH
- Battery range per mile
- Battery capacity KWH
- Battery efficiency MPGe
- Rapid charge of battery
- Price

With this data will be using descriptive and inferential statistics to attempt to get a conclusion of, should you consider an EV?

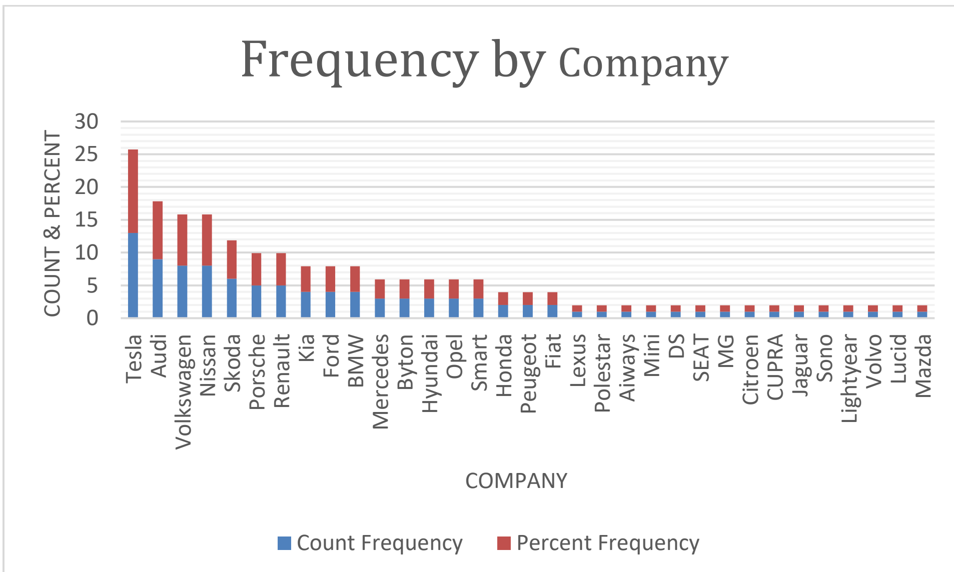
Understanding the Data

Understanding of the data will be assuming that this EV dataset is a sample of a population of vehicle brands and EVs. The dataset has different measurement of units other than USA standards. I have converted kilometers-hour to mile per hour as well as kilometers to miles. I have also converted battery efficiency watts-hour / kilometers-hour to miles per gallon equivalent.

Please note the project does not make a direct comparison of EV to an ordinary vehicle. Ordinary vehicle data that the project states is only to provide clarity and familiarity to the reader.

There are a wide range of brands and models that we will be observing. There is a lot of competition now to have the best performance EV. Figure 1 displays the frequency of the brands. There are 33 brands and 102 EVs. Fifteen brands have one EV model, that is 45.45% of all the brands, but that company that has one model can have the best performance EV. Tesla has the most models at 13 and that is 12.75% of all brands. Tesla was one of the first brands to invest a lot into the EV technology but that does not mean they have the best performance EV. The performance such as speed and battery will give us a better understanding, if an EV is something to consider purchasing.

Figure 1: EV Brand Frequency



EV Price

Price is one of the first things that a customer looks for on their purchase of an EV. The consumer believes that price is an indication of the EV value. The higher the price of the EV, the better the EV should be.

In Figure 2 shows that the most common price of an EV is between $27,500 and $37,500. There are 33 EVs that fall between that price range. 77 out of 102 EVs are in the first 4 bins. This is 76% of all EVs. Figure 2 shows that the other classes have much fewer EVs in their price ranges. Furthermore, Figure 2 is significantly skewed to the right.

Figure 2: EV by Price

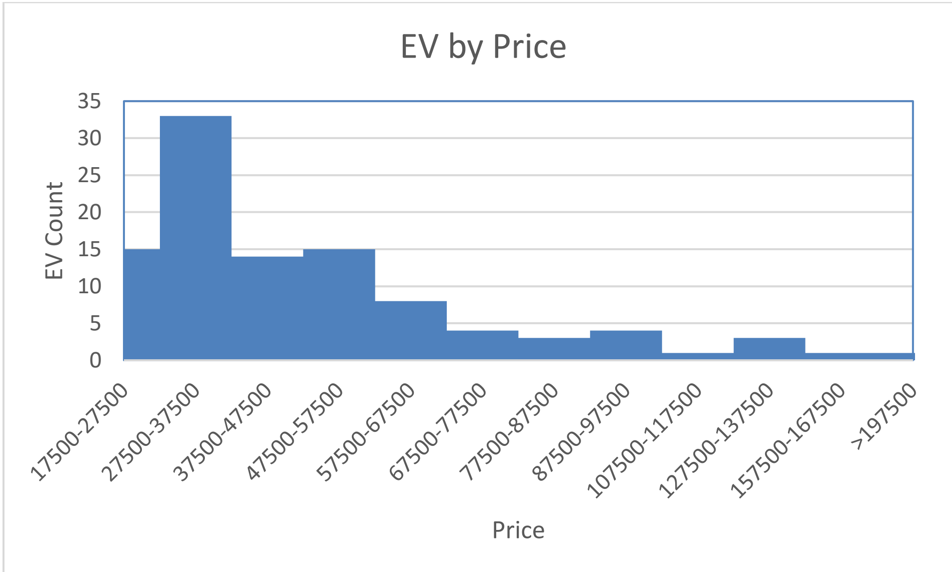


Table 1 shows that the standard deviation is high at $30,038. The average cost of an EV is $49,110 and the median cost of an EV is $39,465. 25% or less of EVs cost between $17,650 and $30,160. I see this to be the economic price point. The lower 75% of EVs cost $57,141 or less, compared with the max cost of an EV is $188,555. This large gap can be because there are some outliers on the higher side of the data. The range of 25% and 75% of the EV price is $26,981. Knowing the interquartile range of $26,981 and the upper limit is $97,613, we can see from Figure 2 that there are 6 outliers.

Table 1: EV by Price

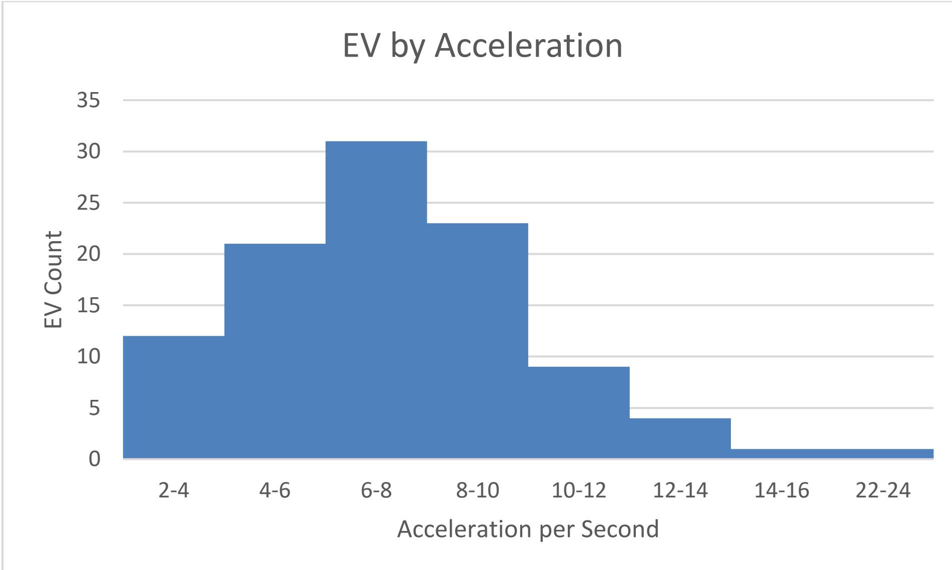
Summary Statistics - EV by Price	
Mean	49110
First Quartile	30160
Median	39465
Third Quartile	57141
Minimum	17653
Maximum	188555
Standard Deviation	30038

Acceleration

Consumers are wanting an EV that can catch up to speed with traffic or they want to get to their destination soon as possible. They want an EV that has top acceleration. The project will be specifically looking at EVs accelerating from 0 to 60 per second.

Figure 3 shows that 31 out of 102 EVs are in the 6-8 second range. Also shows that the data is slightly skewed to the right because there are few EVs that have much slower acceleration times.

Figure 3 EV by Acceleration



Referring to Table 2 half of the EVs go 0–60 in 7.3 seconds, and the average EV goes 0–60 in 7.39 seconds. 75% of all EVs are at 9 seconds or below and 25% are at 5.1 seconds or below. The slowest EV is at 22.4 seconds that is very slow. The fast EV is 2.1 seconds and that is compatible with a fast acceleration ordinary vehicle. The standard deviation is at 3 seconds and clearly there is some variation. The interquartile range is 3.9 seconds. The upper limit is 14.85 seconds, and the lower limit is -0.75. Looking at Figure 3 we can see that there are 2 outliers on the upper tail.

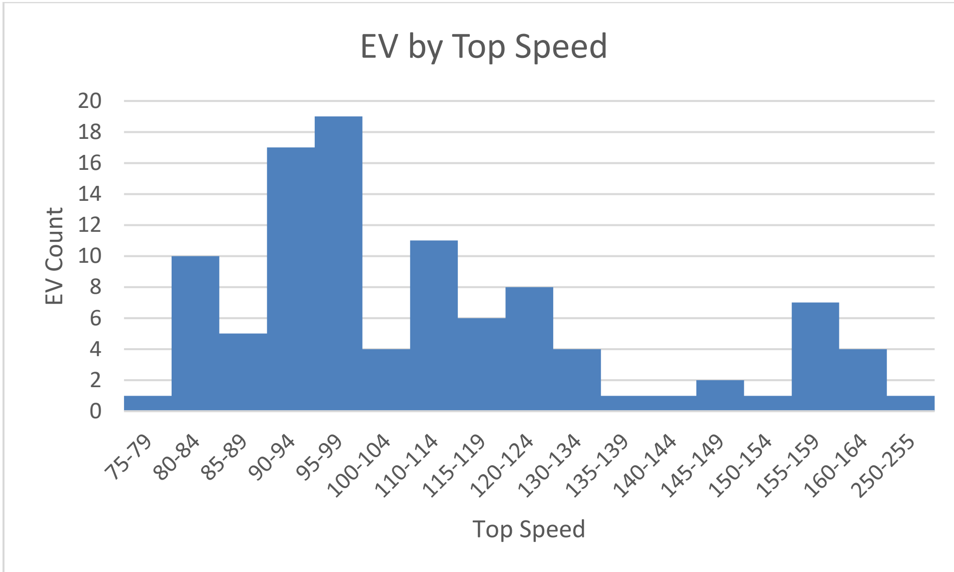
Table 2: EV by Acceleration

Summary Statistics - Vehicle by Acceleration	
Mean	7.39
First Quartile	5.1
Median	7.3
Third Quartile	9
Smallest Value	2.1
Largest Value	22.4
Standard Deviation	3.03

Top Speed

Even though there are limits to how fast you can drive. Knowing the top speed of an EV is important because there are speed limits that you need to maintain. Most EVs do not need to have a very high-top speed because the highest speed limit in the USA is 85 mph. Figure 4 shows that the most common top speed is 95-99 mph that is 19 or 19% of EVs. Figure 4 shows that there is a right tail skewness that could be because EVs companies are not so concerned of making the fastest car but there are few companies that are wanting a supercar.

Figure 4: EV by Top Speed



In Table 3 shows that the average is 111 mph and half of EV go 99 mph. The standard deviation is high at 27.2 mph. The fastest EV goes 255 mph but 75% EVs go up to 124 mph, there is a good size gap there. The slowest EV is 76 mph. This EV will have limitations regarding where it can be driven. 25% of EVs have a top speed of 93 mph. The interquartile range is 31 mph. The upper limit is 171 mph, and the lower limit is 47 mph. Figure 4 shows that there is one outlier at 255 mph.

Table 3: EV by Top Speed

Summary Statistics - EV by Top Speed	
Mean	111
First Quantile	93
Median	99
Third Quantile	124
Minimum Value	76
Maximum Value	255
Standard Deviation	27.2

Battery Range

At the beginning of the manufacturing of EVs, battery range was one of the performance factors why people did not want to purchase EVs. Battery range determines how many miles you can drive before charging the battery.

Figure 5 shows that the number of EVs that have a battery range of 324 miles or more drastically decreases. From the range of 325- 603 miles there are only 4 EVs. There are 29 or 28% of EVs that have the most common battery range of 190 – 234 miles. Figure 5 shows that the data is close to normal.

Figure 5: EV by Battery Range Per Mile

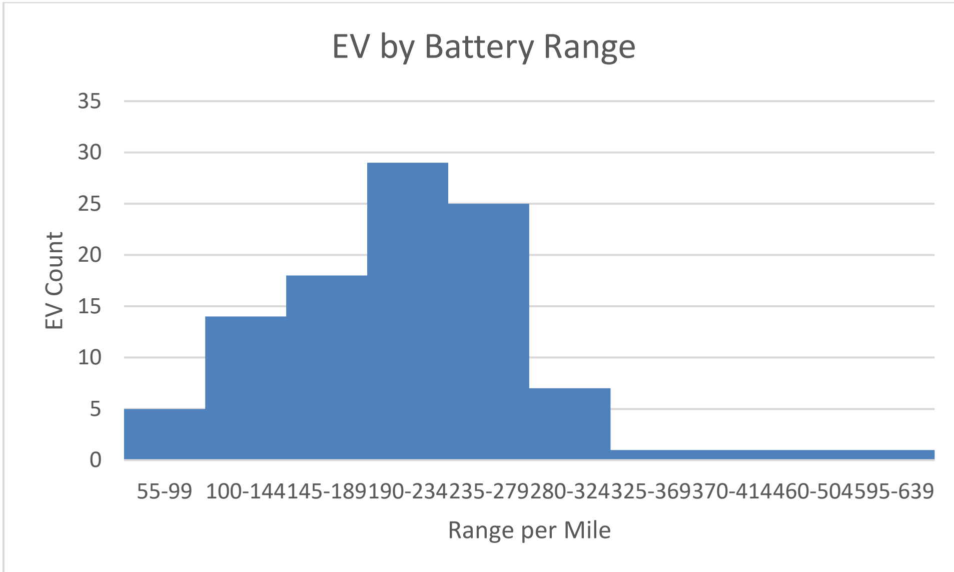


Table 4 shows that the median battery range is 211 miles. There is a variation shown by the standard deviation of 79 miles. 59 miles is the lowest mile range that an EV has. Not many people would purchase this EV because they will be constantly charging the battery and they won't be able to travel long distance away from a charging station or home. It would be less likely that customers would purchase EV that fall in the 25 percentile which is 155 miles because people are wanting the highest possible battery range. The highest battery range is 603 miles, that would be the ideal EV. 75% of EVs have a battery range of 251 miles. The interquartile is 96 miles. The upper limit is 395 miles, and the lower limit is 11. Figure 5 shows that there are 2 outliers.

The average battery range is 210 miles that is a somewhat disappointing number. I would think this number would been higher because companies have been working on battery technology for quite a bit.

Table 4: EV by Battery Range Per Mile

Summary Statistics - Vehicle by Battery Range	
Mean	210
First Quartile	155
Median	211
Third Quartile	251
Smallest Value	59
Largest Value	603
Standard Deviation	79

Battery Capacity

For an EV the size of the battery pack is important because the bigger the battery pack more energy the battery will hold. The battery pack to an EV is important as a gas tank to an ordinary vehicle. The battery pack is capacity is measured by kilowatt-hour.

Figure 6 shows that higher battery capacity is associated with the higher battery range. This makes sense the bigger the battery, the more battery capacity it can hold, and it can go a longer range. This can hold true to an ordinary vehicle; larger the gas tank the higher range you can go. There are 2 outliers that Figure 6 shows at the largest values of battery capacity.

Figure 6: Battery Range Per Miles by Capacity KWH

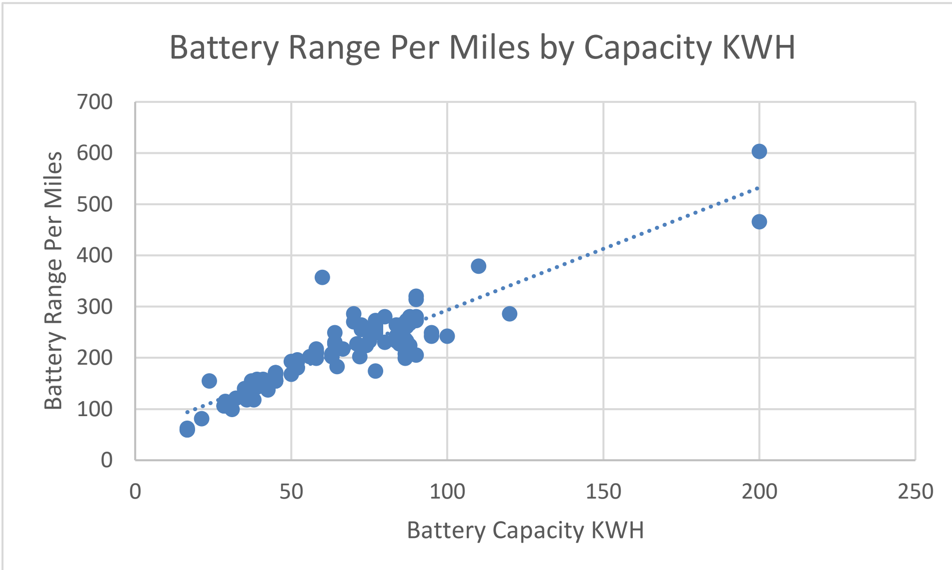


Table 5 shows the covariance is 2148 and correlation coefficient is 0.91, the relationship between battery capacity and battery range has a strong positive relationship. In Figure 6 you can see the relationship of both variables by noticing that the dots are increasing with the trendline. Also, you can see that the dots are highly clustered together and to the trendline.

Table 5: Battery Range Per Miles by Battery Capacity KW/H

Summary Statistics - Battery Range By Capacity	
Covariance	2147.869715
Correlation Coefficient	0.91038647

Battery Efficiency Watt-Hours/Mile

The efficiency of a battery is as important as miles per gallon is to an ordinary vehicle. Battery efficiency is known as miles per gallon equivalent (MPGe). MPGe is a metric to compare the energy efficiency of an EV to an ordinary vehicle by showing how many miles the vehicle can travel on the energy contained in one gallon of gas.

When we reference efficiency, we are referring to how much energy a battery uses. Battery uses energy when is moving the EV. Other things to consider are that the battery can lose energy by being charged, weather or when the EV is idle. In these cases, the EV is not moving and thus the battery efficiency will not always determine the battery range per miles.

Figure 7 shows that battery efficiency increases, battery range slightly decreases. Figure 7 shows there is some clustering of data but there is a clear distance of data way from the trendline. Table 6 shows the covariance of -352 and the correlation coefficient of -0.26 indicates the negative relationship between battery range and battery efficiency. The negative relationship can be because the battery efficiency is more complex than just calculating the range. More variables come into play as the battery uses the energy.

Figure 7: Battery Range Per Miles by Efficiency MPGe

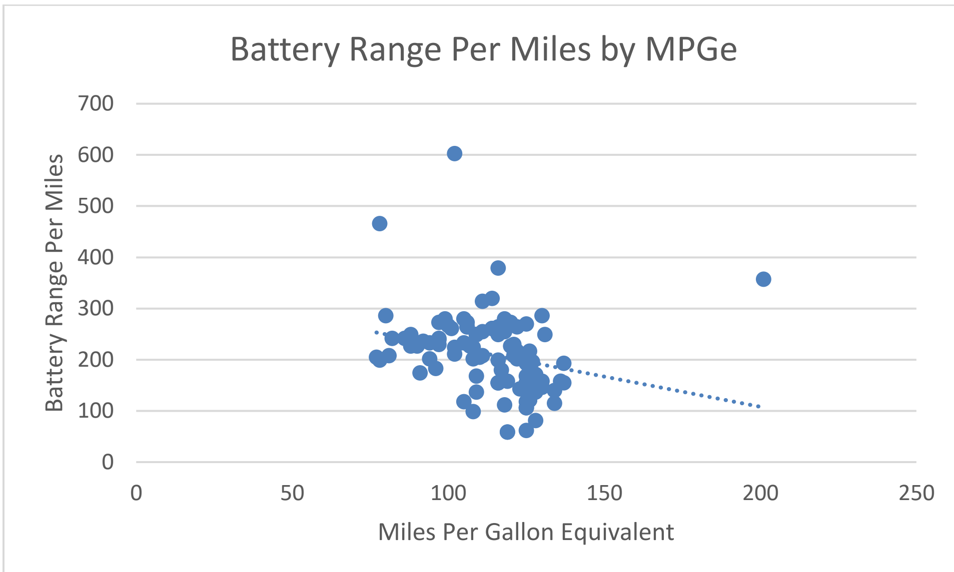


Table 6: Battery Power Per Miles by Efficiency MPGe

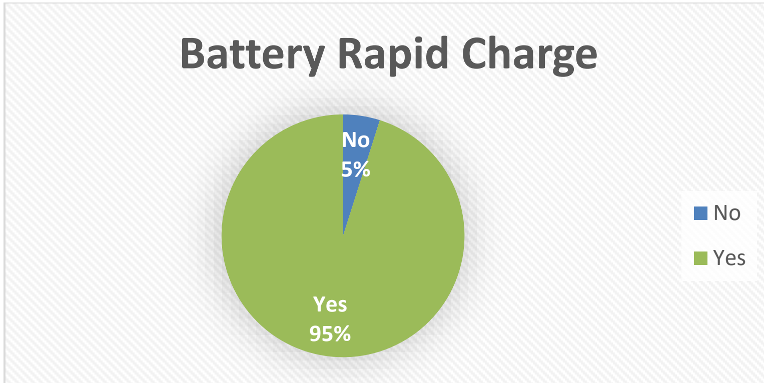
Summary Statistics – Battery Range by MPGe	
Covariance	-352.0302854
Correlation Coefficient	-0.258754188

Battery Rapid Charge

Customers want EV batteries to charge as fast as possible and keep charge as long as possible. It is important for an EV battery to have the ability to rapidly charge because it reduces the time of charge for the battery. There is not a definite time charge decrease because the environment in which the battery is in, battery condition, battery type and the size of the battery are some of the factors that affect the charge time.

Figure 8 shows that only 5% of the EVs do not have rapid charge. The technology for charging EVs has improved over time as well as customers demanding to have faster charge times. This demand has meant that 95% of EVs have rapid charge. EV companies are aware of customers’ demand and will continue developing faster charge technology than rapid charge.

Figure 9: Battery Rapid Charge



Inferential Statistics

Inferential statistics help to explain situations and allows to draw conclusions of the variables that are being compared and the dataset. Measurements of the characteristics of a population are being made on samples that are taken from the population. The dataset that the project is working with is a sample of a population. The population parameters are unknown and will be estimating the population parameters. The assumptions for the t test for the next sections are that the scale of measurement applied to the data collected follows a continuous scale. The observations are independent of one another. The data is approximately normally distributed and homogeneity of variance.

In the following sections, conclusions about the variable's acceleration, top speed, battery range, price and relationship between battery range and battery capacity, battery range and battery efficiency will be drawn using inferential statistics techniques.

EV Price

The price of the EV is the most looked variable and it should be. Vehicles are the second most expensive thing a person buys. Customers want to invest in a vehicle that they will enjoy driving. EV prices are known to be high. The price of the EV should be reasonable. The hypothesis that will be diving into is, can an individual with an average salary of $35,977 (Alex Kopestinsky, 2021) afford an EV? The null hypothesis is the mean equal or less than $35,977. The alternative hypothesis is the mean greater than $35,977.

The significate level that will be using is 0.05 or $\alpha = 0.05$. The population parameters are unknow so we will be using the t test statistic. This is a right tail test for one mean. The critical t value is 1.66 and the t test value is 4.416. The t test value is higher than the critical t value, with 5% significance level we reject $H_0: \mu \leq \$35,977$.

The results suggest that an average EV's price is higher than $35,977. The average EV's price is higher than an ordinary mid-size car at $27,545 (Ben Luthi, 2020). EV price should keep coming down when the technology improves. The more affordable the cost of the EV, the more people will consider buying an EV. There are 48 or 47% of EVs that are at $37,500 or below.

EV Acceleration

Acceleration is one of the import aspects of a vehicle. The faster a vehicle can accelerate 0-60 the more favorable the vehicle is to customers. Customers compare EV acceleration data with ordinary vehicle acceleration data because they are familiar with ordinary vehicles. Customers want similar or better performance when they are considering investing their hard-earned money in an investment such as an EV.

Average acceleration of an ordinary mid-size car is around 6 seconds. For customer to consider an EV, EV needs to accelerate 0-60 in 6 seconds or better. Hypotheses testing will get a closer look at the hypothesis of EV having better accelerate of 0-60 than ordinary mid-size car. The null hypothesis is the mean equal or less than 6 seconds. The alternative hypothesis is the mean greater than 6 seconds.

A right-tailed test will be performed with a t-test for one means. The population standard deviation is unknown and will be used as sample standard deviation. Will continue working with significate level of 0.05 or $\alpha = 0.05$. The critical t value is 1.66 and the t test value is 4.634. The t test value is higher than the critical t value. There is significant evidence to reject the null hypothesis of $H_0: \mu \leq 6$.

The results suggest that an average of EVs is slower than an average ordinary mid-size car of 6 seconds. There are 35 or 34% EVs that accelerate 6 seconds or faster. The EV percentage of acceleration 0-60 in 6 seconds should be closer to 50%. EV companies are wanting to compete with ordinary vehicle acceleration numbers, they need to improve on the acceleration of EVs.

EV Top Speed

EVs companies are aware of speed limits around the United States. The highest speed limit in the United States is 85 mph. For an EV to have all the potential customers, EV top speed needs to be above 85 mph. The hypothesis that we will be looking into more is EV are able to be driven anywhere in the United States? The null hypothesis is the mean equal or less than 85 mph. The alternative hypothesis is the mean greater than 85 mph.

We will continue using significate level of 0.05 or $\alpha = 0.05$. The population parameter is unknown and will use a sample standard deviation. Will use a right tail with the t test for one sample mean. The critical t value is 1.66 and the calculated test value is 9.771. It is observed that the t test value is higher than the critical t value, at 5% significance level we reject $H_0: \mu \leq 85$. The results suggest that the average EV has a higher top speed of 85 mph. There are 11 or 11% of EVs that their top speed is 85 mph or less. 11 EVs have limitations of where they could be driven because they do not meet the requirement of top speed of 85mph.

EV Battery Range

Battery range is an import part of an EV. People have been hesitant to invest in an EV because the range that a person could drive away from a charging station or home was short. An average person drives 30 miles per day. 7 days is in week, doing the calculation of 30 * 7 = 210 miles per week. We want to know what is the probability that you will need to charge the electric car only once a week. The $p(x \leq 210) = 0.5$. Based on an average driving week there are 50% EVs that will need to charge once a week. It is interesting that the average miles driving per week is the same as the mean and one mile less than the median. It seems that companies are striving for those average miles driving per week number.

It will be ideal to charge the EV every two working weeks. If the person does not drive on the weekends, then that would be 30 * 10 = 300, 300 miles every two working weeks. The null hypothesis is the mean equal or greater than 300. The alternative hypothesis is the mean less than 300.

A left-tailed test will be performed. The t test will be with one sample mean. Population standard deviation is unknown and will use the sample standard deviation. We are continuing to use a significate level of 0.05. The critical t value is -1.66 and the statistic t value is -11.488. Statistic t value is less than the critical t value thus there is significant evidence to reject the null hypothesis of H_0: $\mu \geq 300$.

Results suggest that the average EV battery range per mile is less than 300 miles. Based on an average of two workings weeks people would need to charge the EV more often than two working weeks. There are 6 of 6% of EVs that will last at least two working weeks without needing to charge. I believe that around this battery range, people would consider buying an EV.

Confidence Interval

As mentioned before the population parameters are unknown and will need to get a better sense of the population parameters. The significate level that will be using is 0.05 or α = 0.05. Table 7 shows the confidence interval for all the variables.

Table 7: Confidence Interval and Sample Mean

	Acceleration	Top Speed	Battery Range	Battery Capacity	MPGe	Price
Upper	7.99	116.68	225.95	71.33	116.62	55038.81
Lower	6.79	105.94	194.87	59.51	109.81	43180.96
Sample Mean	7.39	111.31	210.41	65.42	113.22	49109.88

Table 7 shows that we are 95% confident that the population mean of each variable of EV is between the confidence interval that Table 7 shows. The sample mean in each of the variables for the EV are between their perspective confidence intervals.

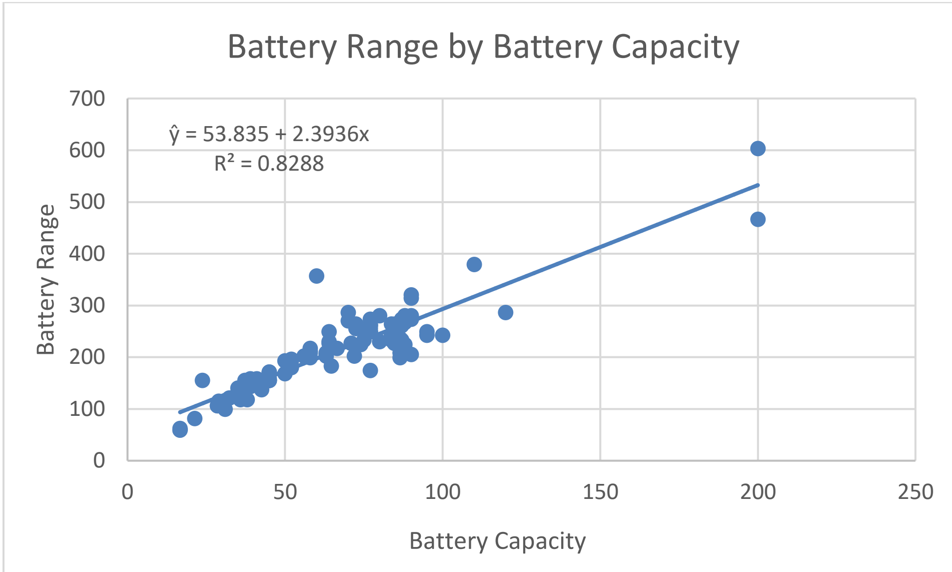
Linear Regression

In the descriptive statistics we looked at the correlation of battery range and battery capacity. From Figure 6 and Figure 7 we got a sense of the relationship between battery range and battery capacity and battery efficiency. Our interest is to know what impact battery range has from the predicators, such as battery capacity and battery efficiency. Performing linear regression allows one to confidently determine which factors matter and how these factors influence each variable.

Relationship Between Battery Range and Battery Capacity

Figure 9 shows that battery range and battery capacity have a positive relationship. Is expected a large battery allows the EV to have a large driving range. To provide a better understanding of the relationship between these two variables will be performing a single linear regression.

Figure 9: Relationship Between Battery Range and Battery Capacity



The assumption of ε is a random variable with mean of zero, variance of ε is the same for all values of the independent variables, the values of ε is independent, ε is a normally distributed random variable for all independent variables.

The null hypothesis states that battery capacity does not have a relationship with battery range. The alternative hypothesis states that battery capacity does have a relationship with battery range.

H_0: $ß_{\text{battery capacity}} = 0$
Ha: $ß_{\text{battery capacity}} \neq 0$

We will be able to reject the null hypothesis if *p*-value $\leq$ .01 or |t| > 2.626 with 100 degrees of freedom. With a significate level of 0.01 and calculating the t test of 22.003. With 99% confidence we reject the null hypothesis because the t test of 22.003 is greater than the critical t value of 2.626. The p value is 0.000. P value is less than 0.01 and this also suggests rejecting the null hypothesis.

The confidence interval with a significate of 0.01 is between 2.1079 and 2.679. Hypothesized $ß_{\text{battery capacity}}$ of 0 is not in the confidence interval, there for with 99% confidence level we can conclude that $ß_{\text{battery capacity}}$ is not equal to 0. With that evidence we can reject the null hypothesis and say that battery capacity and battery range have a relationship.

The single linear regression equation for battery capacity is.

$\hat{y} = 53.835 + 2.394x_{\text{battery capacity}}$

When battery capacity unit increases by 1 the battery range will increase by 2.394 miles. The coefficient of battery capacity is positive. This shows that the relationship between the battery capacity and battery range is positive.

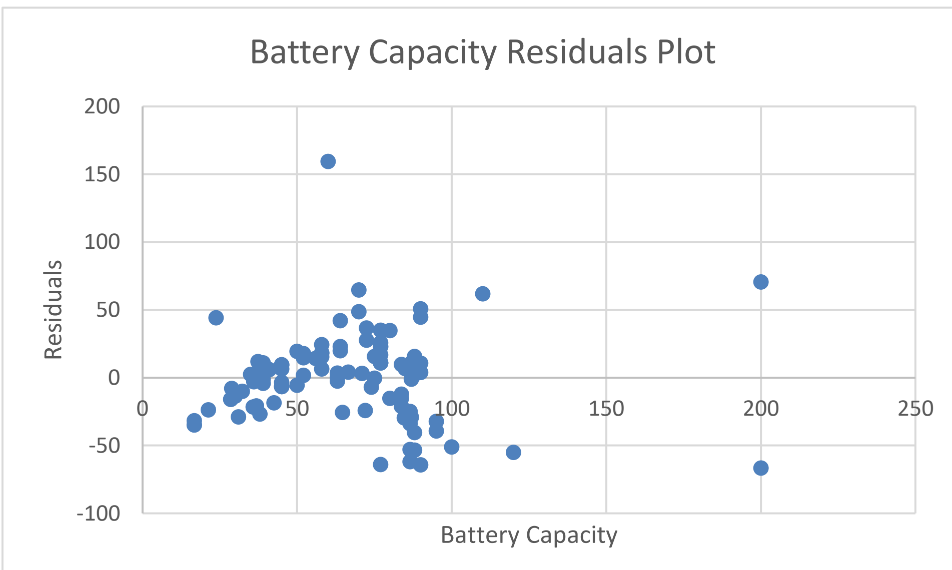
The correlation coefficient (r) between battery range and battery capacity of 0.910 has a strong positive relationship. As battery capacity goes up as well does battery range.

Looking at the coefficient of determination (R^2) of battery range and battery capacity of 0.829. 82.9% of variation in the battery range is explained by the variation in battery capacity. 17.1% of the variance is caused by other factors such as measurement errors. This tells us that the data is highly clustered around the regression line and the goodness of fit of the model is good.

A residual plot of battery capacity is displayed in Figure 10. This residual plot shows a pattern that there is a non-constant variance about the regression line because larger values of battery capacity equal, larger variability. The residual plot provides evidence that the assumption of the regression model is not an adequate representation of the relationship between the battery range and battery capacity. A multiple regression model should be considered.

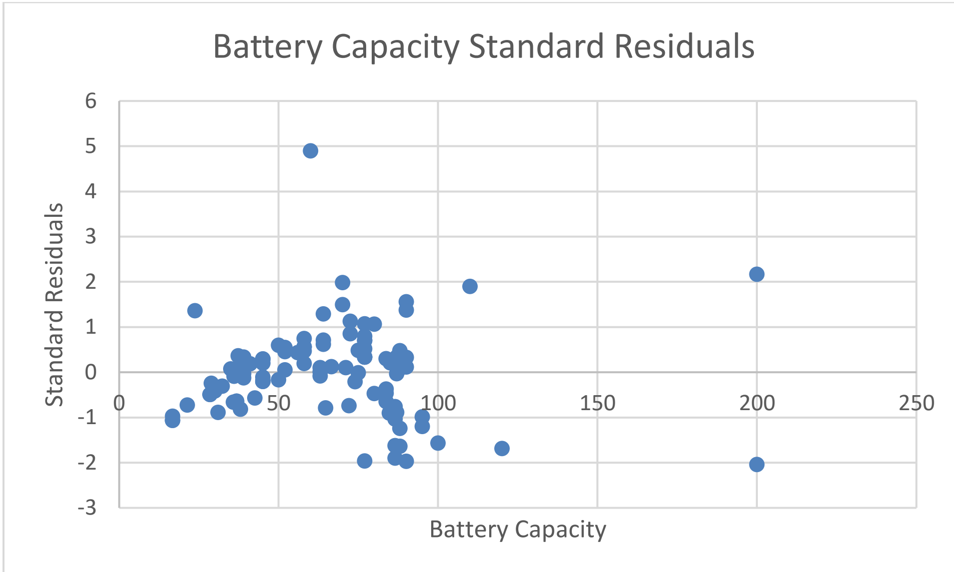
Residual plot of $\hat{y}$ will look the same as Figure 10 because this is a single regression model. When we investigate multiple regression will evaluate residual plot of $\hat{y}$.

Figure 10: Battery Capacity Residual Plot



The standard residual provides insight about the battery capacity values. The sample mean is 65.416 and sample standard deviation is 29.956. Figure 11 shows that 98% of battery capacity kwh values land between -2 and 2 standard deviations. This tells us that residuals appear to come from a standard normal probability distribution. Therefore, Figure 11 gives us no reason to question the assumption that ε has a normal distribution.

Figure 11: Battery Capacity Standard Residual Plot

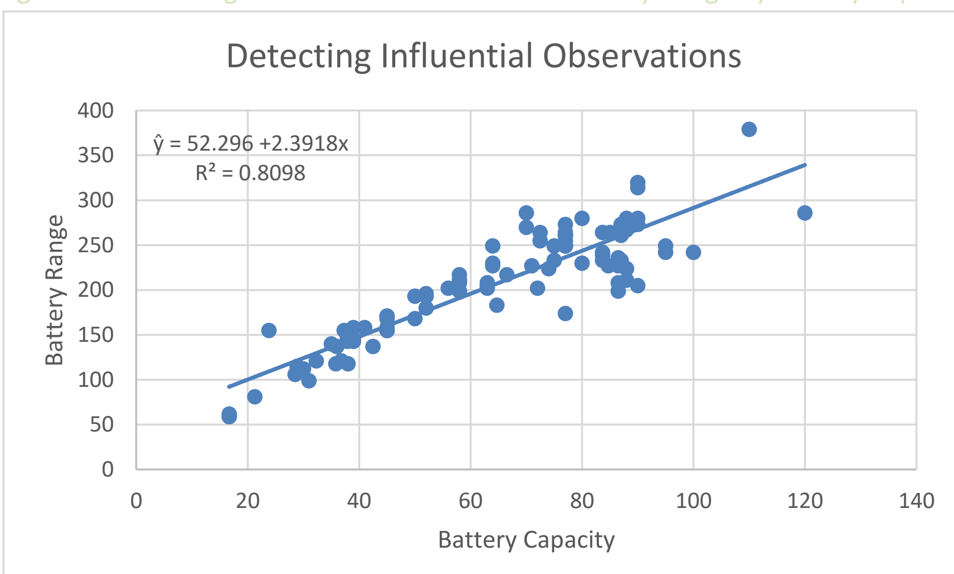


There are 3 outliers at (200, 2.16), (200, -2.04) and (60, 4.90) that Figure 11 shows. After removing this observations Figure 12 shows the outcome. With these observations deleted the equation is

$\hat{y} = 52.296 + 2.3918x_{\text{battery capacity}}$

The y intercept decreased from 53.835 to 52.296. The slope of the line decreases slightly from 2.394 to 2.3918. The effect of deleting these observations results is not influential. You can see that in comparing Figure 9 with Figure 12.

Figure 12: Detecting Influential Observations in Battery Range by Battery Capacity



Relationship Between Battery Range to Battery Capacity and Battery Efficiency

Performing a multiple regression by adding battery efficiency into the previous single linear regression model can provide a better understanding of the strength and importance of the relationship between battery range and predictor variables as battery capacity and battery efficiency.

When considering EV customers are presented with Monroney stickers on the windows of each EV. This sticker provides the customer with details about the EV such as how far you can drive, miles per gallon equivalent and how many kilowatts a battery uses per 100 miles. These variables are so important that they must be on a new EV window.

A large capacity battery should last longer than one that has less capacity. A battery that is more efficient should last longer. A larger battery and an efficient battery should provide a longer driving range. The hypothesis says that there is no linear relationship between battery range and $b_{\text{battery capacity}}$ or $b_{\text{battery efficiency}}$. If at least one of these independent variables is not equal to 0, there would be significate evidence to reject the null hypothesis and conclude that there is a relationship.

H_0: $b_{\text{battery capacity}} = b_{\text{battery efficiency}} = 0$
Ha: $b_i \neq 0$

To show a potential good model the project will use in this instance a significate level of 0.01 or $\alpha = 0.01$. The F test statistic is 908.297 and significate F is 0.0000. The significate F is smaller than significate level of 0.01. There is significate evidence at 99% confidence that at least one of coefficients in the regression model is not equal to zero. This tells us that there are at least two variables that have a relationship, there for at 99% confidence to reject the null hypothesis. This represents that this model is a good linear regression.

Looking at each individual independent variable we have the following hypothesis.

H_0: $b_{\text{battery capacity}} = 0$
Ha: $b_{\text{battery capacity}} \neq 0$

H_0: $b_{\text{battery efficiency}} = 0$
Ha: $b_{\text{battery efficiency}} \neq 0$

The P value of battery capacity and battery efficiency are 0.0000. The project will use in this instance a significate level of 0.01 or $\alpha = 0.01$. P value of 0.0000 is less than the significate level of 0.01. With 99% confidence there is evidence to reject the null hypotheses. This shows that the coefficients are good.

The regression equation for $b_{\text{battery capacity}}$ and $b_{\text{battey efficiency}}$ is.

$\hat{y} = -210.841 + 3.059x_{\text{battery capacity}} + 1.953x_{\text{battery efficiency}}$

For every unit increase in the battery capacity and battery efficiency is stays constant, the battery range will go up 3.059 miles and if the battery capacity was stayed constant and battery efficiency unit increased by 1, battery range would increase be 1.953 miles. Coefficient of battery capacity and battery efficiency are both positive. This shows that the relationship between the coefficient and battery range is positive.

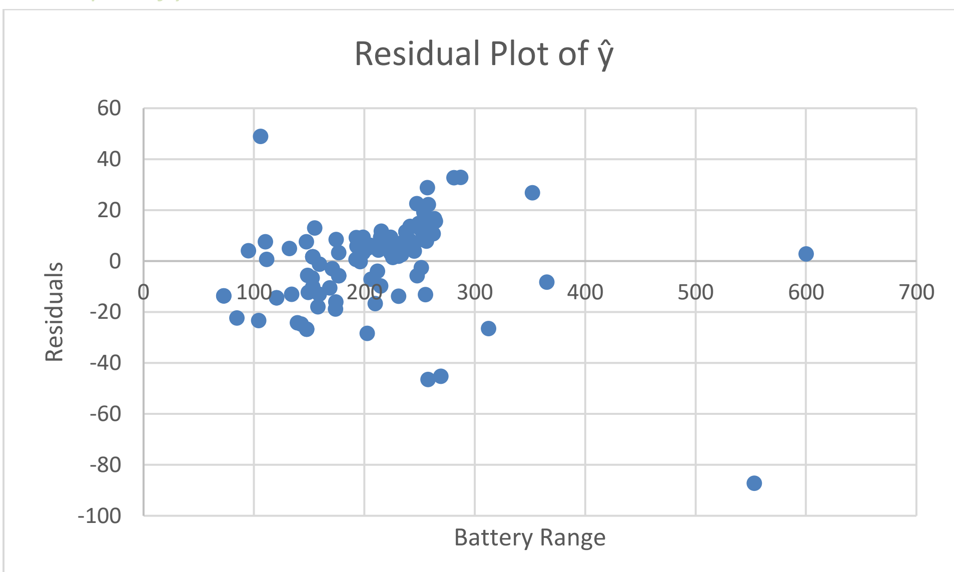
Looking at correlation coefficient (r) between battery range, battery capacity and battery efficiency at 0.974 has a strong positive relationship. As battery capacity and battery efficiency goes up as does battery range.

R^2 of battery range, battery capacity and battery efficiency are 0.948, there for 94.8% of variation in the battery range explains by the variation in battery capacity and battery efficiency. 5.2% of the variance is caused by other factors such as measurement errors. This tells us that the data is highly clustered around the regression line and the goodness of fit of the model is good.

The regression model has 2 independent variables which can cause unexpected outcomes of R^2 such as bias estimates. Considering the impact of the R^2 results will look into adjusted R^2 to have better reliability. 0.947 is the adjusted R^2 of battery range, battery capacity and battery efficiency. Adjusted R^2 of 94.7% accounts for the total variability. R^2 is 0.001 higher than adjusted R^2. Adding an additional variable to this regression model slightly added value.

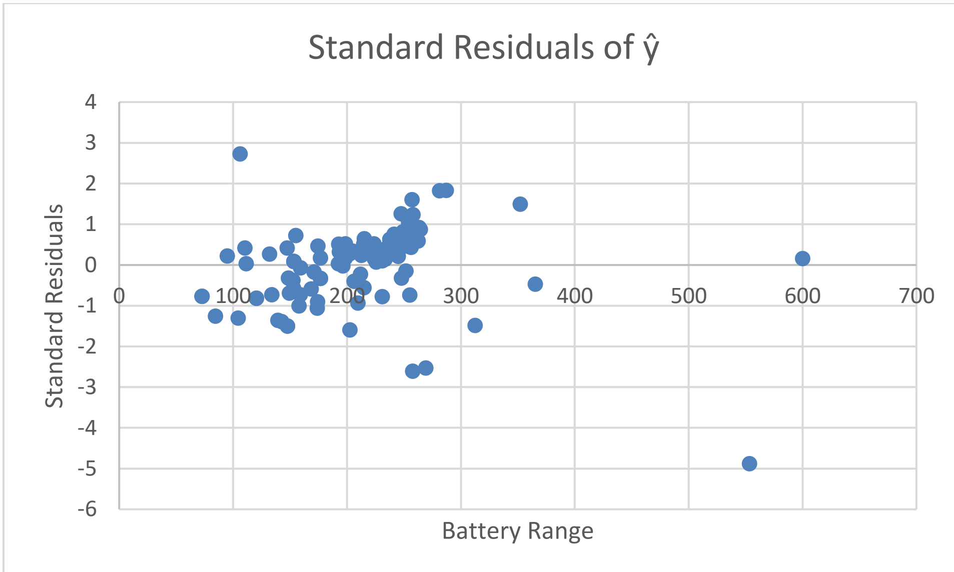
On the single regression we looked at battery capacity residual plot. For the multiple regression will dive into the residual plot of ŷ. Figure 13 provides evidence that the regression model has a good pattern. At this point, we are confident in the conclusion that multiple linear regression model is valid.

Figure 13: Residual plot of ŷ



The standard residual for the multiple linear regression is shown in Figure 14. The sample mean of battery range is 210.4117 and sample standard deviation is 78.759. Figure 14 shows that 96% of battery range miles values land between -2 and 2 standard deviations. This tells us that residuals appear to come from a standard normal probability distribution. Therefore, Figure 14 provides evidence that ε is a normal distribution.

Figure 14: Standard Residual plot of ŷ



There are 4 outliers at (257.586, -2.602), (269.306, -2.530), (553.300, -4.876) and (106.133, 2.729) that is displayed on Figure 14. After removing these observations, the equation is

$\hat{y}$ = - 233.733 + 3.315$x_{\text{battery capacity}}$ + 2.023$x_{\text{battery efficiency}}$

The y intercept decreased from -210.841 to -233.773. The slope of battery capacity stayed at 3.315 and the slope of battery efficiency increased from 1.953 to 2.023. The effect of deleting these observations has an influential effect.

Conclusion

Successful companies stay competitive by being ahead of other companies by developing new products, new methods, new systems, and new technology. Before considering something new, it is desirable to conduct research to determine whether there is statistical support for the conclusion that the new approach or product is indeed better. Comparing ordinary vehicle with an EV is comparing two different groups thus there are many other variables and techniques to consider that this project did not consider. However, people try to associate ordinary vehicle information with the EV information, and they have their own perspective and importance of each EV variable.

Most EVs acceleration are slower than people are used to around 6 seconds, but people can reach traffic speeds safely. Top speed is not an issue with EVs because most all EVs can maintain a highway top speed of 85 mph. People won't be happy with the battery range until it reaches ordinary vehicle gas tank ranges. However, the results suggest that 50% of EV will last longer than a week without a charge. MPGe is the variable that the EV excels in. Electricity is cheaper than gas and people are saving money by driving EV. The savings can be 3-4 times. The results suggest that a larger battery and high efficiency will have a larger battery range. EV must have rapid charge to stay competitive in the EV market and that is why so many EV have them. Price is a sensitive variable in terms of people's budgets. A higher budget a person has more EV options that person will have. Results suggest with our assumptions that a person would need a budget around $45,000 - $60,000 for an EV.

References

Divyanshu Gupta, (2021, June). *"Cars Dataset With Battery Pack Capacity*": keggle.com. https://www.kaggle.com/divyanshugupta95/cars-dataset-with-battery-pack-capacity/

Ben Luthi, (2020, December). "*What Is the Average Price for a New Car?*": experian.com. https://www.experian.com/blogs/ask-experian/what-is-the-average-price-for-a-new-car/

Alex Kopestinsky, (2021, August). "*What Is the Average American Income in 2021?*": policyadvice.net. https://policyadvice.net/insurance/insights/average-american-income/