EV Pricing

Logan Crandall

Advanced Statistical Methods

Le Moyne College

I. Introduction
II. Data
III. Assumptions in MLR
    A. Basic Assumptions
    B. Overview of Assumptions with Plots
    C. Testing Assumptions in R
IV. Full model Multiple Linear Regression and Analysis
    A. Linear Regression Estimation Model
    B. Estimated Regression Equation
    C. Significance Testing
    D. Evaluation of Full Model
V. Possible Interactions Within Model
    A. Creation of New Model with Interaction Terms
    B. New Model Regression Equation
    C. Significance Testing and Evaluation of Interaction Terms
    D. Evaluation of Interaction Term Model
VI. Variable Selection
    A. Method
VII. New Model after Variable Selection
    A. New Model Equation
    B. Estimated Equation
    C. Significance Testing
    D. Evaluation of New Model
VIII. Assumption Testing for New Model
    A. MLR Assumptions
    B. Visual Analysis of New Model Assumptions

I. Introduction

The purpose of my analysis is to attempt to develop a model that predicts the purchase price (in Germany, before tax incentives, in Euros) of electric vehicles (EVs) based on a number of related factors. These factors include acceleration time (in seconds, from 0-100Km/H), top speed of the vehicle (in Km/H), the driving range of the vehicle on a full charge (in Km), energy efficiency of the vehicle (in Watt-hours per Km), charge time while fast-charging (in Km/H), and availability of rapid charger use. This model will be constructed and analyzed in the open-source software R using multiple linear regression.

II. Data

The dataset I will be using is available under public domain on Kaggle, was compiled by user Geoff839, and can be found at link https://www.kaggle.com/datasets/geoffnel/evs-one-electric-vehicle-dataset/data?select=ElectricCarData_Clean.csv. The dataset contains information on over 100 different models of EVs. The dataset is contained in a CSV file, where I have converted the rapid charge measurement from a yes/no binary to a one/zero binary.

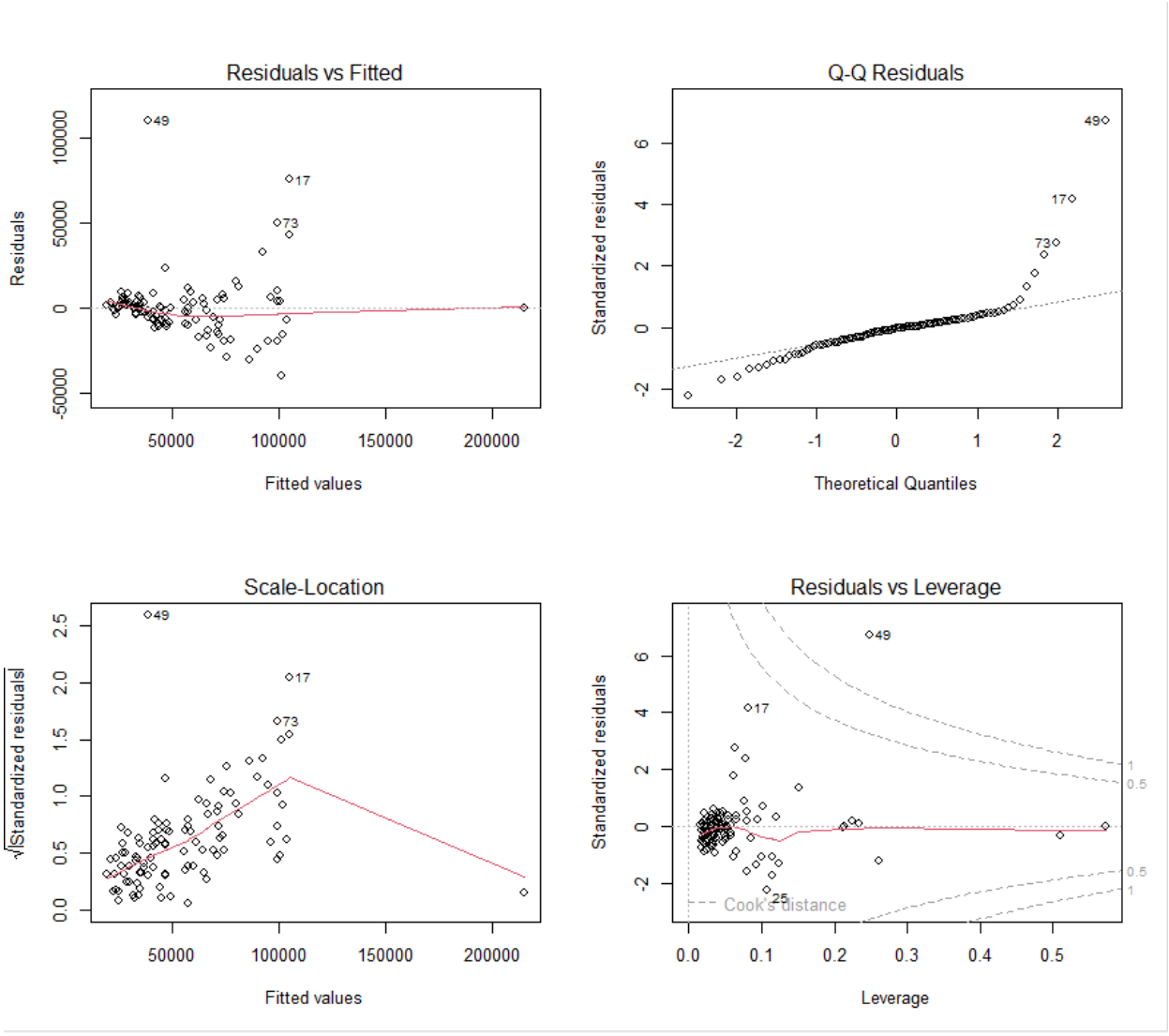
III. Assumptions in MLR

A. Basic Assumptions

In multiple linear regression, there are four primary assumptions a researcher must consider in order to analyze the quality of the regression. These assumptions can be categorized under the acronym LINE, and are as follows: a (L) linear relationship exists between the response variable and the explanatory variable, (I) errors are independent of one another, (N) there are normally distributed residuals, and (E) there is equal variance and standard deviation among the residuals at all levels of x. It is also worth checking to see if there are any outliers in the dataset, or data points with a Cook's distance (a measurement of the influence of a data point) of greater than one. I will also be examining multicollinearity, or the correlation of two explanatory variables such that one can reliably be predicted from the other.

B. Overview of Assumptions with Plots

I used R to create plots in order to get a general overview of some of the assumptions. The graphs are as follows:



Firstly, I examined the "Residuals vs. Fitted" graph to analyze the linearity assumption. The plots on this graph seem to be clustered on the left side, but that seems to be mainly caused by a few outliers expanding the graph on the right and top to give it that appearance. What clustering exists is likely to be expected from a small sample size, and I do not believe it to be a cause for concern.

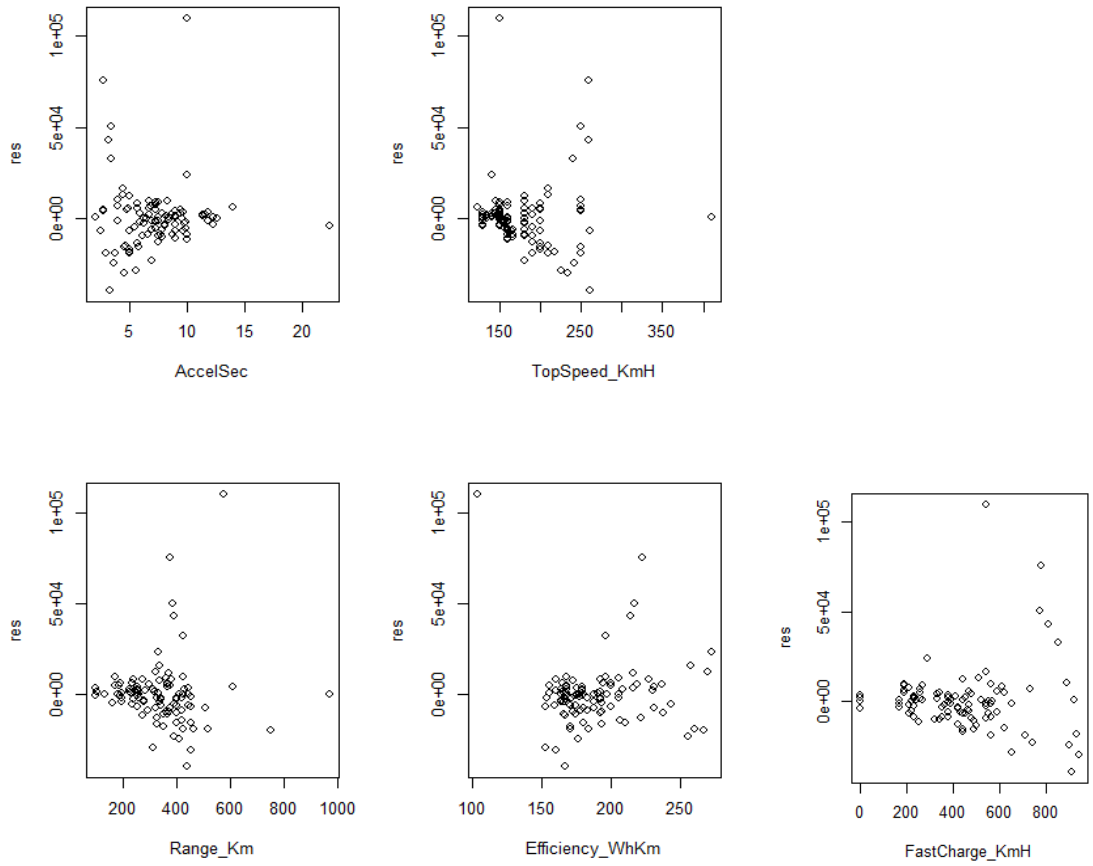
Secondly, I looked at the "Q-Q Residuals" plot to test the assumption that the residuals are normally distributed. The points on the graph deviate from the line on both the upper and lower section. These points that deviate also seem to be the same data points that appear to be outliers in the "Residuals vs. Fitted" plot, and happen to be some of the most expensive premium brands of EVs. This is possibly a result of a brand premium that people are willing to pay to own the more expensive EVs. Either way, I cannot conclude that there is a normal distribution among the residuals.

Thirdly, I reviewed the "Scale-Location" plot to examine the homoscedasticity, or equal variance among residuals at all levels of x, assumption. The points in the graph do seem to be

randomly distributed along the line. However, the line is not horizontal, so I cannot conclude that the homoscedasticity assumption holds.

The fourth graph I examined, “Residuals vs. Leverage”, I used to check if there are any obvious outliers. While most of the data seems to fall under a Cook’s distance of one, point 49 seems to be an obviously troublesome outlier, as I have seen in the other graphs. There are also a few points approaching a Cook’s distance of 0.5 to be wary of. Point 49 will likely be an influential point in the dataset.

In order to test the assumption of no autocorrelation, or the independence assumption, I used R to plot all of my continuous variables against the residuals as follows:



Note that I did not plot the variable RapidCharge against the residuals because it is a binary variable. These graphs do appear to follow a pattern, but that seems to likely be because of the few outlier data points changing the domain and range values of the graphs. It does not seem to be a cause for concern, as will be shown later in the test for autocorrelation.

C. Testing Assumptions in R

Looking at the graphs, while helpful, is usually not enough to conclude the validity of the assumptions. So, I decided to run some tests in R to test the assumptions mathematically, as well.

I used the Shapiro-Wilk test to determine if the assumption of normal distribution in the residuals holds. The output in R is as follows:

```
        Shapiro-Wilk normality test

data:  res
W = 0.73215, p-value = 2.037e-12
```

The p-value is well below 0.05 in the Shapiro-Wilk test, so I rejected the null hypothesis that the residuals are normally distributed and concluded that the assumption has been violated. This is a major problem in analysis, and is likely to be the result of a specification that has been left out of the model. I believe it is likely due to leaving brand and model out, as I would expect brand and model preference (especially in luxury and high-end EVs) would have a significant impact on price. I have decided to continue with regression analysis, anyway.

Next, I used the Breusch-Pagan test to examine the homoscedasticity assumption. The R output is as follows:

```
        studentized Breusch-Pagan test

data:  fit
BP = 16.4, df = 6, p-value =
0.01176
```

The Breusch-Pagan test I conducted supports the idea that the homoscedasticity assumption does not hold. A p-value of less than 0.05 would suggest that I should reject the null hypothesis that all error variances are equal. I concluded that heteroscedasticity, or non-constant variance, is likely present. This means that the linear regression is no longer BLUE (Best Linear Unbiased Estimator), but I will continue with linear regression nonetheless.

In order to test for autocorrelation, I used a Durbin-Watson test. The R output is as follows:

```
        Durbin-Watson test

data:  fit
DW = 2.0698, p-value = 0.6529
alternative hypothesis: true autocorrela
tion is greater than 0
```

The Durbin-Watson test I conducted yielded a p-value of greater than 0.05, so I did not reject the null hypothesis that autocorrelation is not present. I concluded that the independence assumption holds. There is no autocorrelation present.

In order to test for multicollinearity in the data, I used R to calculate the Variance Inflation Factors (VIF) for each of my variables. A VIF is a measure of multicollinearity in multiple regression variables, and if a variable's VIF is greater than 10, it is a cause for concern. The VIF values are as follows:

```
       AccelSec     TopSpeed_KmH
       4.115324         4.297894
      Range_Km Efficiency_WhKm
       2.829879         1.197887
FastCharge_KmH      RapidCharge
       3.695125         1.657280
```

Since none of the variance inflation factors are greater than 10, I concluded that there is no multicollinearity present between the independent variables.

IV. Full Model Multiple Linear Regression and Analysis

A. Linear Regression Estimation Model

The model for multiple linear regression can be represented by the equation as follows:

$$y_i=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3+...\beta_ix_i+\varepsilon_i$$

Where y_i = the dependent variable, β_0 = the intercept, β_i = the coefficient of the ith variable, x_i = the ith independent variable, and ε_i = the error term. I will be substituting my dependent variable (PriceEuro) for y, and my independent variables (AccelSec, TopSpeed_KmH, Range_Km, Efficiency_WhKm, FastCharge_KmH, and RapidCharge) for the x variables in the equation in order to estimate the regression line. I can use this regression line in order to estimate the dependent variable based upon measurements of my independent variables, as described in the "Introduction" and "Data" sections of this paper.

After running a multiple regression analysis in R for that purpose. The output is as follows:

```
Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)     -1.047e+05  3.057e+04  -3.427   0.0009 ***
AccelSec         1.550e+03  1.259e+03   1.232   0.2211
TopSpeed_KmH     6.066e+02  8.908e+01   6.810  8.5e-10 ***
Range_Km         3.605e+01  2.499e+01   1.442   0.1525
Efficiency_WhKm  1.436e+02  6.930e+01   2.072   0.0410 *
FastCharge_KmH   4.014e+00  1.638e+01   0.245   0.8070
RapidCharge     -7.544e+02  1.116e+04  -0.068   0.9463
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 18910 on 96 degrees of freedom
Multiple R-squared:  0.7112,    Adjusted R-squared:  0.6932
F-statistic:  39.4 on 6 and 96 DF,  p-value: < 2.2e-16
```

B. Estimated Regression Equation

Based on the above output in R, the estimated regression line to predict the price of EVs can be stated as follows:

$PriceEuro_i$ = -104,700 + 1,550AccelSec + 606.2TopSpeed_KmH + 36.05Range_Km + 143.6Efficiency_WhKm + 4.014FastCharge_KmH - 754.4RapidCharge

Where PriceEuro = Price (in Germany, before tax incentives, in Euros), AccelSec = seconds it takes the vehicle to reach 100 Km/H from 0Km/H, TopSpeed_KmH = top speed of the vehicle (in Km/H), Range_Km = distance the vehicle can drive on a full charge (in Km), Efficiency_WhKm = efficiency of the vehicle (in Watt-Hours per Km), FastCharge_KmH = the speed at which the battery is charged while fast-charging (in Km/H), and RapidCharge = a binary categorical variable which measures the vehicles rapid charge compatibility (equals one if the car has rapid charge capability, zero otherwise).

C. Significance Testing

In order to test the null hypothesis that none of the variables are significant, we use an F test to test the null hypothesis that none of the variable coefficients are significantly different from zero. In other words, it is testing whether the full model is better than a model with only the intercept, β_0.

In the R output above, the F statistic is 39.4, with a p-value of close to zero, so I reject the null hypothesis that none of the variables are significant. I can conclude that there is at least one significant variable in the model.

Similarly to the F test, we can use a T test to test the significance of specific variables within MLR. In support of the F test conclusion above, the R output would lead me to believe that there are two variables that are significant in the full model. For top speed and efficiency the t values were 6.810 and 2.072, respectively. The coinciding p-values are close to zero and 0.4, respectively. Since both the p-values were below 0.5, I rejected the null hypothesis that the variables were insignificant, and concluded that both top speed and efficiency are significant predictor variables in respect to EV price.

D. Evaluation of Full Model

Firstly, the R^2 value for this full model is shown in the R output above to be 0.7112. This means that the model can predict 71.12% of the variability in price. This is a fairly good starting point. I also used R to calculate the AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion). The R output is given below. AIC and BIC are measurements of the goodness of fit of a statistical model which penalize for over-complexity, the lower the score being the better model. The AIC and BIC scores, as shown below, are 2329.609 and 2350.687, respectively.

```
> AIC(fit)
[1] 2329.609
> BIC(fit)
[1] 2350.687
```

V. Possible Interactions Within Model

A. Creation of New Model with Interaction Terms

There are a few terms within the complete model that I believe have significant effects on each other. I wanted to test to see if there is an interaction between AccelSec and TopSpeed_KmH, and also if there is an interaction between Rank_Km and Effeciency_WhKm. I created interaction terms for these, and the resultant model would be as follows:

$PriceEuro_i = \beta_0 + \beta_1 AccelSec_i + \beta_2 TopSpeed\_KmH_i + \beta_3 Range\_Km_i + \beta_4 Efficiency\_WhKm_i + \beta_5 FastCharge\_KmH_i + \beta_6 RapidCharge_i + \beta_7 AccelSec_i TopSpeed\_KmH_i + \beta_8 Range\_Km_i Efficiency\_WhKm_i + \varepsilon_i$

I ran an MLR analysis on this new model in R. The resultant output is as follows:

```
Coefficients:
                             Estimate Std. Error t value Pr(>|t|)
(Intercept)                -2.728e+05  3.944e+04  -6.916 5.56e-10 ***
AccelSec                    1.415e+04  3.099e+03   4.566 1.50e-05 ***
TopSpeed_KmH                8.160e+02  8.872e+01   9.197 9.38e-15 ***
Range_Km                    4.611e+02  8.608e+01   5.357 5.99e-07 ***
Efficiency_WhKm             1.032e+03  1.763e+02   5.854 6.98e-08 ***
FastCharge_KmH             -4.343e-01  1.459e+01  -0.030    0.976
RapidCharge                 2.540e+03  9.775e+03   0.260    0.796
AccelSec:TopSpeed_KmH      -1.063e+02  2.352e+01  -4.519 1.80e-05 ***
Range_Km:Efficiency_WhKm   -2.237e+00  4.276e-01  -5.233 1.01e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 16120 on 94 degrees of freedom
Multiple R-squared:  0.7946,    Adjusted R-squared:  0.7771
F-statistic: 45.44 on 8 and 94 DF,  p-value: < 2.2e-16
```

B. New Model Regression Equation

The R output above allows me to fit an estimated equation to attempt to predict price with the new interaction terms present. The equation is as follows:

PriceEuro = -272,800 + 14,150AccelSec + 816TopSpeed_KmH + 461.1Range_Km + 1,032Efficiency_WhKm - 43.43FastCharge_KmH + 2,540RapidCharge - 106.3AccelSec*TopSpeed_KmH - 2.237Range_Km*Efficiency_WhKm

Where PriceEuro = Price (in Germany, before tax incentives, in Euros), AccelSec = seconds it takes the vehicle to reach 100 Km/H from 0Km/H, TopSpeed_KmH = top speed of the vehicle (in Km/H), Range_Km = distance the vehicle can drive on a full charge (in Km), Efficiency_WhKm = efficiency of the vehicle (in Watt-Hours per Km), FastCharge_KmH = the speed at which the battery is charged while fast-charging (in Km/H), RapidCharge = a binary categorical variable which measures the vehicles rapid charge compatibility (equals one if the car has rapid charge capability, zero otherwise), AccelSec*TopSpeed_KmH = the interaction between acceleration and top speed, and Range_Km*Efficiency_WhKm = the interaction between range and efficiency.

C. Significance Testing and Evaluation of Interaction Terms

Again, an F test can be used to determine if there are significant variables present in the model. As seen above, the F statistic is 45.44, and the resultant p-value is close to zero. So, we reject the null hypothesis that there are no significant variables and conclude there is at least one significant variable in the analysis.

To test specific variables, we again use a T test. The p-value obtained from a t-test on a specific variable tells us if it is statistically significant. If a variable's p-value is less than the alpha value set for this test (0.05), we reject the null hypothesis that it is not significant, and conclude the variable is statistically significant. Based on this rule, I decided that all of the variables present, except FastCharge_KmH and RapidCharge, were statistically significant, as the remainder had p-values of less than 0.05.

D. Evaluation of Interaction Term Model

Interaction terms for this model are significant. Accelsec and Range also become significant in this model, when they were not before. P-values have also decreased across the board for the singular variables, except for FastCharge_KmH, which has increased in p-value. This indicates that it might be advisable to switch to this new interaction term model in analysis.

R-Squared has also increased from the full model to the interaction term model, from 0.7112 to 0.7946. This indicates this new model can predict 8.34% of the variation in the price that the previous model could not.

I also used R to calculate AIC and BIC, as in the previous model. The results are as follows:

```
> AIC(fit2)
[1] 2298.533
> BIC(fit2)
[1] 2324.881
```

As you can see, the AIC decreased from 2329.609 to 2298.533, and the BIC decreased from 2350.687 to 2324.881. That means that even given the increased complexity of this new model, it is still a better fit to the data than the old model.

VI. Variable Selection

A. Method

Based on the previous evaluations of the models, I determined that I should continue my analysis with the model including the interaction terms. In order to select which variables to keep in my analysis, I used stepwise backward selection. I used this because, based on p-values, I believed that I would be removing less variables than I was keeping, so it would have taken less time had I done it outside of R. I believed I would be removing the FastCharge_KmH and RapidCharge variables. I performed stepwise backward regression in R to make sure I was correct. The output is as follows:

Elimination Summary

Step	Variable Removed	R-Square	Adj. R-Square	C(p)	AIC	RMSE
1	FastCharge_KmH	0.7946	0.7794	7.0009	2296.5342	16031.8587
2	RapidCharge	0.7944	0.7816	5.0675	2294.6072	15953.7947

R also eliminated the aforementioned variables based on p-value. These variables are insignificant in the model, so this is as to be expected.

VII. New Model after Variable Selection

A. New Model

After removing variables in the last section, the equation for the model can be shown as follows:

$$PriceEuro_i = \beta_0 + \beta_1 AccelSec_i + \beta_2 TopSpeed\_KmH_i + \beta_3 Range\_Km_i + \beta_4 Efficiency\_WhKm_i + \beta_5 AccelSec_i TopSpeed\_KmH_i + \beta_6 Range\_Km_i Efficiency\_WhKm_i + \varepsilon_i$$

I ran a regression analysis on this newest model in R. The output is as follows:

```
Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)               -2.689e+05  3.598e+04  -7.473 3.65e-11 ***
AccelSec                   1.389e+04  2.763e+03   5.028 2.30e-06 ***
TopSpeed_KmH               8.060e+02  7.854e+01  10.262  < 2e-16 ***
Range_Km                   4.638e+02  8.104e+01   5.723 1.19e-07 ***
Efficiency_WhKm            1.033e+03  1.717e+02   6.017 3.23e-08 ***
AccelSec:TopSpeed_KmH     -1.051e+02  2.247e+01  -4.679 9.46e-06 ***
Range_Km:Efficiency_WhKm -2.245e+00  4.135e-01  -5.429 4.25e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15950 on 96 degrees of freedom
Multiple R-squared:  0.7944,    Adjusted R-squared:  0.7816
F-statistic: 61.82 on 6 and 96 DF,  p-value: < 2.2e-16
```

B. Estimated Equation

The R output above allows me to fit an estimated equation to my newest model. The estimated regression line can be stated as follows:

PriceEuro = -268,900 + 13,890AccelSec + 806TopSpeed_KmH + 463.8Range_Km + 1,033Efficiency_WhKm - 105.1AccelSec*TopSpeed_KmH - 2.245Range_Km*Efficiency_WhKm

Where PriceEuro = Price (in Germany, before tax incentives, in Euros), AccelSec = seconds it takes the vehicle to reach 100 Km/H from 0Km/H, TopSpeed_KmH = top speed of the vehicle (in Km/H), Range_Km = distance the vehicle can drive on a full charge (in Km), Efficiency_WhKm = efficiency of the vehicle (in Watt-Hours per Km), AccelSec*TopSpeed_KmH = the interaction between acceleration and top speed, and Range_Km*Efficiency_WhKm = the interaction between range and efficiency.

C. Significance Testing

As with before, I used the F statistic above, 61.82 with a resulting p-value of almost zero, to determine that there is at least one statistically significant variable present. I also used T tests to determine if each variable was significantly different from zero. All of the present variables are significant, as they all have a corresponding p-value of less than 0.05, so we reject the null hypothesis that they are not statistically significant.

D. Evaluation of New Model

Compared to the previous model in which the two discluded variables were still present, p-values have decreased for every variable still present in this new model.

R^2 has stayed about the same between the two models, decreasing very slightly from 0.7946 to 0.7944. Adjusted R^2, however, which penalizes a model for excess variables, has increased from 0.7771 to 0.7816, indicating this new model is superior.

I have also used R to calculate the AIC and BIC for the new model. The output is shown below:

```
> AIC(fit5)
[1] 2294.607
> BIC(fit5)
[1] 2315.685
```

AIC has decreased from 2298.533 to 2294.607, and BIC has decreased from 2324.881 to 2315.685, indicating that this model is a better fit for the data after variable selection. I have concluded that the model post-variable selection is a better model overall, and will be continuing with it.

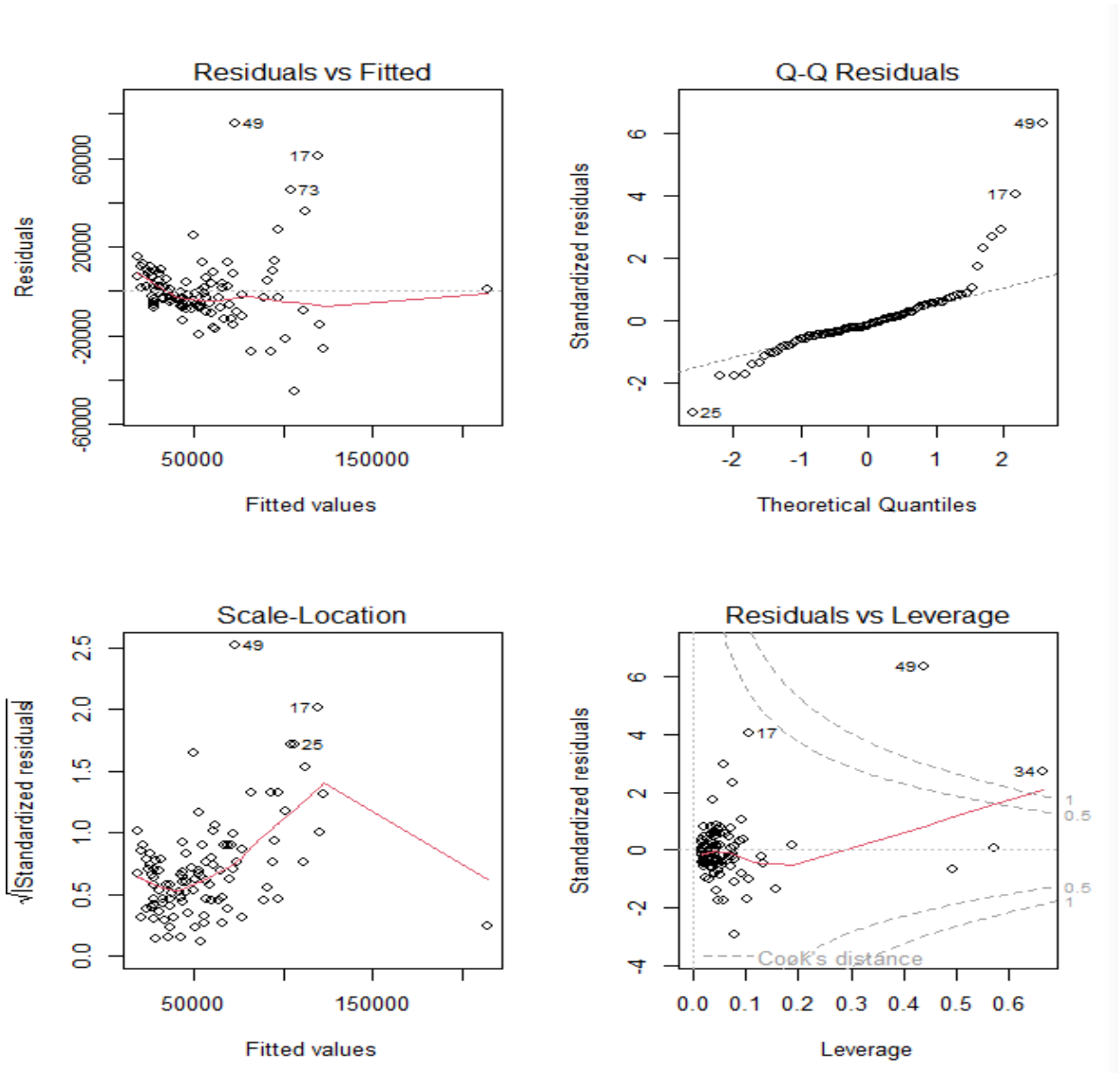
VIII. Assumption Testing for New Model

A. MLR Assumptions

Please refer to section III-A for basic assumptions being tested for MLR.

B. Visual Analysis of New Model Assumptions

For this new model, I again used R to create some plots to get a general overview of some of the assumptions. The graphical output is as follows:



The “Residuals vs. Fitted” graph can be analyzed to get an idea of how the linearity assumption holds. In this case, the graph looks about the same as the original model, with slightly less clusters. Overall, an improvement from the original model. As I said before, this clustering is likely to be expected from this sample size, so it is likely not a cause for concern.

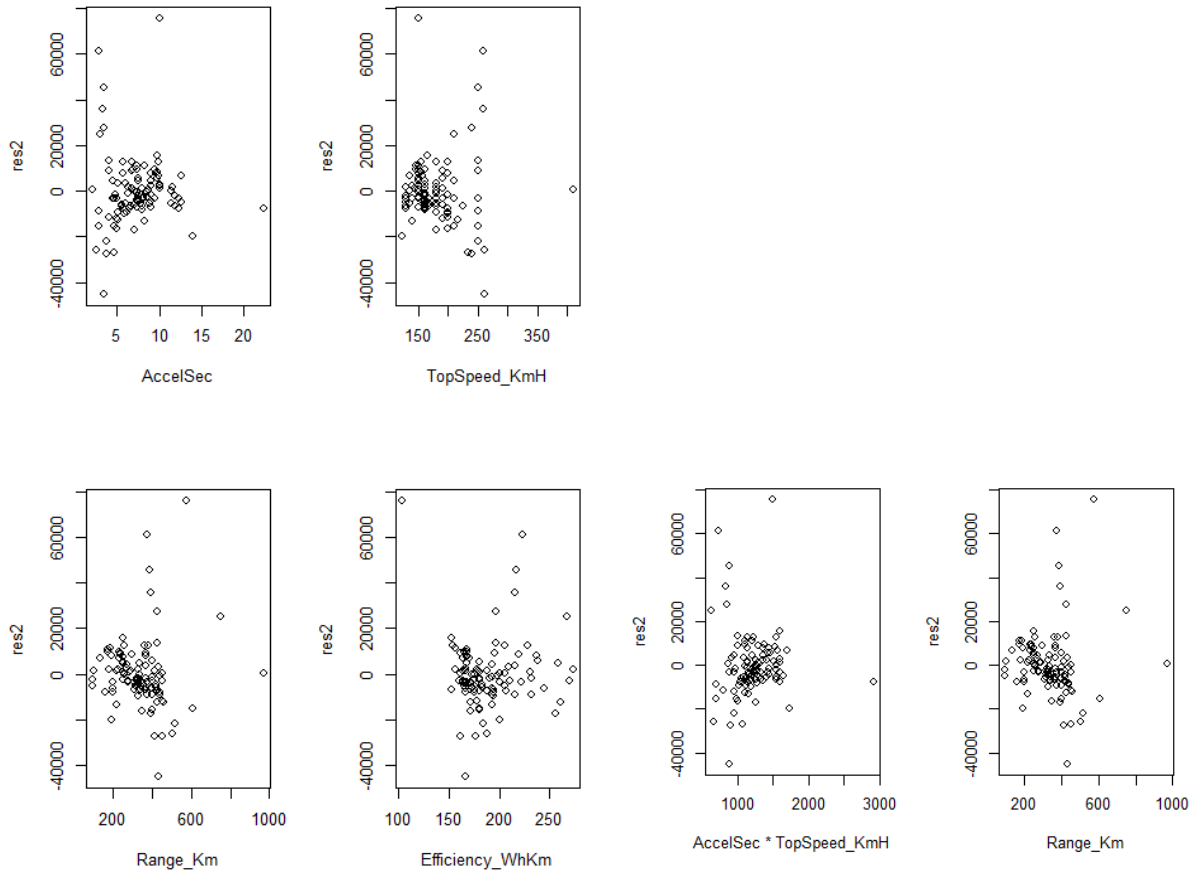
To check the assumption that the residuals are normally distributed, I again looked at the “Q-Q residuals” plot. This plot looks almost exactly the same as the original model. Again, the deviated points are the outliers from the “Residuals vs. Fitted” graph, and are some of the more expensive EVs. I, again, expect this to be a result of brand premium that people are willing to pay, skewing the data. I cannot conclude normal distribution based on this Graph.

I then examined the “Scale-Location” graph to analyze the homoscedasticity assumption, or the assumption that there is equal variance among residuals at all levels of x. Like in the

original model, while the points seem to be randomly distributed along the line in the graph, the line is not straight and thus we cannot assume the Homoscedasticity assumption holds.

In order to test for outliers, I looked at the "Residuals vs. Leverage" graph. This time, points 34 and 49 have a Cook's distance of greater than one, and can be considered outliers. There are likely to be influential points in this model.

I again tested for autocorrelation to check the independence assumption by plotting the continuous variables against the residuals, which are named "res2" in this model. The graphs are as follows:



As with the original model, the graphs visually seem to be clustered. However, that is likely due to the range and domain of the graphs being stretched as a result of outlying points, as will be shown in the following Durbin-Watson test for autocorrelation.

C. Mathematical Analysis of New Model Assumptions

I, again, wanted stronger evidence of the assumptions than a visual examination of related plots. So, I ran some tests in R to find out whether the assumptions held.

I used the Shapiro-Wilk test to determine if the normal distribution assumption is held. The output is as follows:

```
        Shapiro-Wilk normality test

data:  res2
W = 0.83322, p-value =
2.028e-09
```

Again, the p-value is very low and we reject the null hypothesis that the residuals are normally distributed, and conclude that the assumption has been violated. Like described before, this is a major problem, and is likely the result of misspecification of the model, likely in performing analysis that does not account for the brand or model, as they would have significant impact on price in the form of 'brand premiums'. I will continue with analysis of the model, regardless.

I used the Breusch-Pagan test to analyze the homoscedasticity assumption. The R output is as follows:

```
        studentized Breusch-Pagan test

data:  fit5
BP = 29.458, df = 6, p-value = 4.983e-05
```

The associated Breusch-Pagan test supports the idea that the assumption does not hold. A very small P-value leads us to reject the null hypothesis that the error variances are all equal. We can conclude that heteroscedasticity is present, and that this linear regression is no longer BLUE (best linear unbiased estimator), but I will continue with the analysis.

I used the Durbin-Watson test to test for autocorrelation. The R output is as follows:

```
        Durbin-Watson test

data:  fit5
DW = 2.2724, p-value = 0.923
alternative hypothesis: true autocorrelation is greater than 0
```

Based on the high p-value of the Durbin-Watson test, I fail to reject the null hypothesis that autocorrelation is not present. I concluded that autocorrelation is not present, and the independence assumption holds.

Because I am using interaction terms based directly on the base variables, with low VIF scores for the base variables and statistically significant interaction terms, I decided it would be best to ignore high VIF values for multicollinearity levels.

I. Conclusions

A. Final Model

Based upon my multiple linear regression analysis, I have arrived at the following model in order to predict EV price:

PriceEuro = -268,900 + 13,890AccelSec + 806TopSpeed_KmH + 463.8Range_Km + 1,033Efficiency_WhKm - 105.1AccelSec*TopSpeed_KmH - 2.245Range_Km*Efficiency_WhKm

Where:

PriceEuro = Price (in Germany, before tax incentives, in Euros)
AccelSec = seconds it takes the vehicle to reach 100 Km/H from 0Km/H
TopSpeed_KmH = top speed of the vehicle (in Km/H)
Range_Km = distance the vehicle can drive on a full charge (in Km)
Efficiency_WhKm = efficiency of the vehicle (in Watt-Hours per Km)
AccelSec*TopSpeed_KmH = the interaction between acceleration and top speed
Range_Km*Efficiency_WhKm = the interaction between range and efficiency

I. Example Prediction of Price

I can use this model to predict the price of EV models fairly accurately using the aforementioned model. For example, A car with a 0-100 Km/H time of 3 seconds, a top speed of 260 Km/H, a range of 480 Km, and efficiency of 180 Wh/Km should have an approximate price of €114,948.