WhatToCharger Model

Ethan Machado

January 15th, 2022

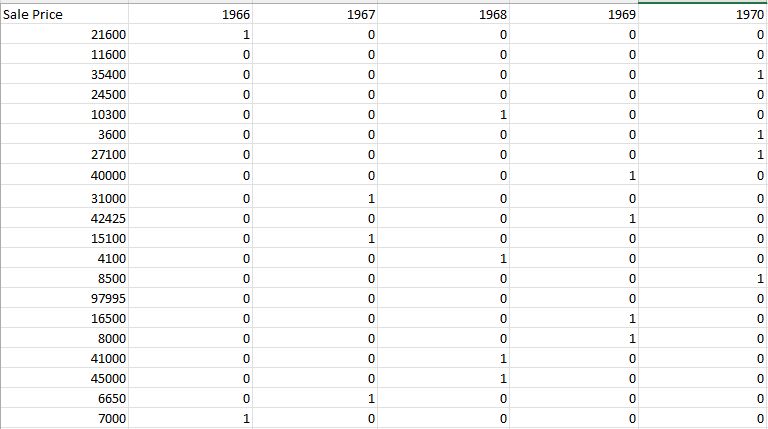
An OLS linear regression model that forecasts the values of Dodge Chargers from 1966-1971.

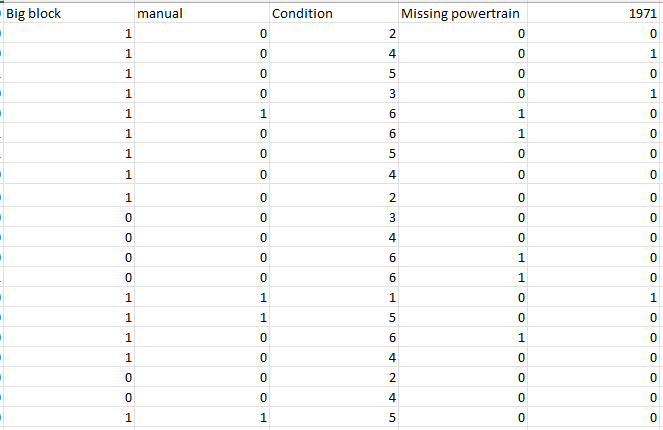
Method: I went on eBay to collect data for this model. I collected the sale price of each car, engine type: small (0) or big block (1), transmission type: automatic (0) or manual (1), model year, missing powertrain: yes (1) or no (0), condition: based on [Hagerty’s 1-6 scale](https://www.hagerty.com/media/buying-and-selling/car-conditions-what-the-numbers-mean/).

Note: Hagerty condition scale link: <https://www.hagerty.com/media/buying-and-selling/car-conditions-what-the-numbers-mean/>

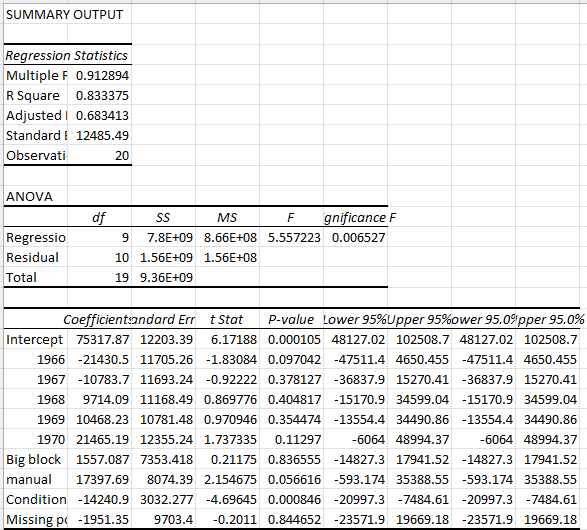
Note: I collected data for at least two cars of each model year, and they all had to be sold between 11-1-2021 and 1-16-2022 to partially control for time of year being a factor in sale price. All data was added onto an Excel spreadsheet and all calculations were executed on Excel.

Data: The data is split into two screenshots to make the text readable.





Results of the OLS linear regression 1:



The full equation of the linear regression is:

Sale price = 75317.87 - 21430.50(1966) - 10783.70(1967) + 9714.09(1968) + 10468.23(1969) + 21465.19(1970) + 1557.09(Big block) + 17397.69(Manual transmission) - 14240.90(Condition) - 1951.35(Missing powertrain)

As can be seen from the results, this model’s adjusted R squared is .683. This means that the regressors make up 68.3% of the variation in sale price. While the 68.3% is high, R squared increases whenever more regressors are added so this metric must be scrutinized.

The standard error of regression is also high at $12,485.49. This means that there was an average mistake of $12,485.49 made by the linear regression for every observation in the model.

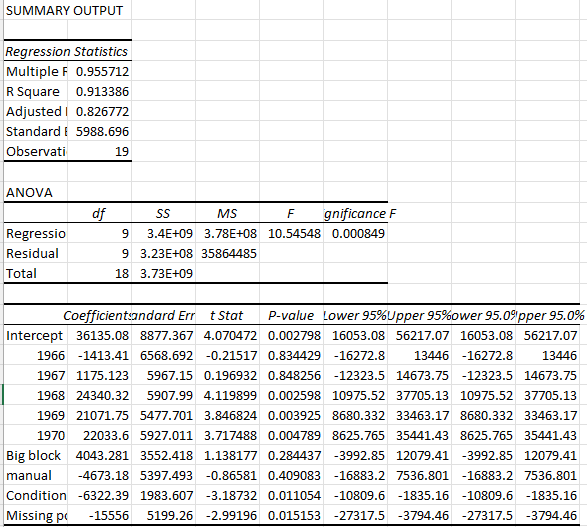
The regression as a whole is significant, as the significance F value is less than 0.05. A joint hypothesis would prove that this model as a whole is statistically significant.

This model is not perfect by any means, as it suffers from a small sample size. I could only find 20 observations, which means that this model does not approach normal distribution. If the model had more observations, then the central limit theorem and the law of large numbers would state the model would approach normal distribution. This is optimal.

This model should be tightened up to have fewer regressors and more observations to improve its results. This WhatToCharger model is a new idea for a quantitative project that I have recently come up with, so that is why the model lacks a larger sample size as well as better regressors.

Also, the 1971 Dodge Charger that sold for $97,995 is an outlier in the data. The car is a rare Dodge Super Bee that is one of about 300 produced. That is why its value is much higher than all the others on the data sheet. This rarity is unaccounted for in the model and it shifts the linear regression plot.

When the outlier 1971 Dodge Charger Super Bee is removed from the model, the results are dramatically improved:



The Adjusted R squared value increases to 82.6% from 68.3%, and the standard error of regression decreases from $12,485.49 to $5,988.70.

The overall model is still statistically significant as the significance F value is still less than 0.05.

The model still needs to be adjusted to include a larger sample size and less regressors, but this is an improvement, nonetheless. Additionally, regressors like Manual transmission need to be removed because they are not statistically significant and one would think a manual transmission increases the value of a car, not decreases it like the model predicts.

I plan on updating my data sheet every month with new Chargers that sell on eBay to increase my sample size. I am thinking of keeping an overall model and a monthly model to see the changes in sale price based on the time of year.

Forecasting using the WhatToCharger model:

I have a 1971 Charger base model. I would rate its condition a 5 (needs complete restoration, but not a parts car). It is a small block, automatic car with the drivetrain still in the car.

Here is a picture of the car:



Using the updated equation to forecast a value for my 1971 Charger:

Sale Price = 36135.08 - 1413.41(1966) + 1175.12(1967) + 24340.32(1968) + 21071.75(1969) + 22033.60(1970) + 4043.28(Big block) - 4673.18(Manual transmission) - 6322.39(Condition) - 15556(Missing powertrain)

Updating the equation to fit my car’s specifications:

Sale Price = 36135.08 - 1413.41(0) + 1175.12(0) + 24340.32(0) + 21071.75(0) + 22033.60(0) + 4043.28(0) - 4673.18(0) - 6322.39(5) - 15556(0)

Sale price = 36135.08 - 6322.39(5)

Sale price = $4,523.10

The WhatToCharger model predicts that my 1971 Charger would sell for $4,523.10 if I posted it on eBay as an auction item. I believe this is a fairly accurate sale price because I have been told by some experts that my car is worth $4,000-$5,000.

Update 02/04/2022:

I have added 5 more observations (cars sold via eBay auctions) and created another linear regression model that has the updated results. At this point, I am still trying to input more observations into my data set to improve the statistical integrity of my WhatToCharger model. With the sample size only being 23 cars, the data is susceptible to being influenced by chance or statistical anomalies that a larger data set would not be. The model will be constantly improving as more cars sell and my data set becomes more robust. I am also considering updating the model to remove the 1966 and 1967 model years because these years do not see many cars sold. Additionally, adding a rarity regressor would help to explain more of the variation in the sale price variable, so I am currently trying to figure out a methodology to include it.

FUTURE IMPLEMENTATION UPDATE: I would like to leverage my web development skills to construct a webpage that can hold my WhatToCharger regression model. I plan on creating input fields that users will use to input data about their 1966-1971 Charger(s). From there, the user will click a button to execute the forecasting calculation to get a sale price (approximate value) for their Dodge Charger. The sale price will be printed on the screen with an option to enter in another 1966-1971 Charger for evaluation.

Here is the regression output (02/04/2022):

Table

Description automatically generated