

Predict the sales of New-energy Vehicle using linear regression analysis

Lingxiao Tang^{1*,a}, and Jia Sun^{2,b}

¹Chongqing Nankai Secondary School; No.1, Sha Nan St., Chongqing Prov., China;

²Johns Hopkins University; 3528 Agate Dr, Apt 4, Santa Clara, CA 95051;

^a +86 13983205605; 1161928003@qq.com

^b +1-571-274-7022; jsun51@jhu.edu

Abstract: New-energy Vehicle is a new form of vehicle that uses environmental friendly energy and becomes a popularity in recent years. It is important for both the consumers and the producers to search for factors that might affect the sales of New-energy Vehicle. In this study, three linear regression models are examined to determine factors that have significant effect on sales of New-energy Vehicle. ANOVA analysis is conducted to test the model validity and to compare the effect of the three models obtained. Result shows that disposable income, number of charging station, subsidies and literacy level positively affect the sales of New-energy Vehicle.

1 Introduction

New-energy Vehicle is a new form of vehicle which utilizes the new energy. New energy refers to energy source that is an alternative to fossil fuel (Dresselhaus and Thomas, 2001). The first New-energy Vehicle was created by Robert Alderson in Scotland. Now New-energy Vehicles are highly rewarded by most of the countries. In this “electrify age”, such vehicle becomes a popular alternation which is advocated by governments and environmentalists. Though the development path varies in different countries and regions, the demand for this product is on the rise.

By employing new energy, the New-energy Vehicles are capable of replacing the conventional vehicles which are condemned by all sectors of community for high carbon dioxide emissions (an important factor in global warming). New-energy vehicles include hybrid electric vehicle(HEV), battery electric vehicle (BEV), plug-in hybrid electric vehicle(PHEV), hydrogen engine vehicle, gas vehicle, alcohol ether vehicle and so on (“Types of Electric Vehicles : BEV, PHEV and HEV”, EVgo.com). For example, in recent years, increasing attention has been paid to the development and utilization of hydrogen energy due to the lack of fossil fuels in the large-scale development. As an important trend of hydrogen energy applications, Nickel-metal hydride batteries are receiving more focus. Meanwhile, Lithium ion battery, another application of new energy, has also been widely used in portable electric appliances such as laptop, camera and mobile communication. The large-capacity lithium ion battery has already been applied to electric vehicles and is expected to become one of the main power sources for electric vehicles in the 21st century.

With the energy shortage and the world's environmental pressure, the evolvement of lithium iron phosphate material battery facilitates the development of lithium battery industry. However, due to the limited accesses to technology and immature market, some factors (such as the competitive price of substitutes and the relatively high cost of exploiting the new energy) are still adversely affecting the sales of the New-energy Vehicles. Thus, this study aims to examine factors affecting the sales of New-energy Vehicle.

2 Literature Review

New-energy Vehicle is a recent popularity and a lot of researches had been done in developing models to predict the sales of it and discussing possible factors which may affect potential users' opinions towards the New-Energy Vehicle.

In 2010, Christensen, L., Kveiborg, O., and Mabit, S. L (2010) focused on the market for electric vehicles about what potential users want. The research indicated that the accessibility of recharging might be one of the most important factors for purchase decision. As increasingly charging stations built up, it could greatly promote consumer's demand for the New-energy Vehicles. Thus, the number of recharging stations could be an important factor on sales of New-energy Vehicles.

Patti Domm (2018), the CNBC markets Editor, states that though today's electric vehicle market is tiny, but it created some very big questions about the future of gasoline demand in her article “Electric vehicles: The little industry that could take a bite out of oil demand” in 2018. Electric cars and light trucks, including hybrids, in the last year displaced only about 50,000 barrels a day of

oil in a world that is using 100 million barrels a day for the first time this year, according to IHS Markit. (Domm, 2018) There's no doubt that when consumers see the benefits of the New-energy Vehicle, which include the sustainable electric transport ecosystem, the cheaper price to drive due to shared mobility and automated technologies and the reduction of the environmental pollution, the sales of New-energy Vehicle will be affected.

Incentives for promoting New-Energy vehicles also become a popular subject of theoretical research. Kristin Y. B., Tom E. and Marianne E. N. (2016) investigated the role of incentives in promoting battery electric vehicles (BEV) in Norway. They found that purchase tax exemption is the most critical factor in promoting BEV sales. Norway is a global forerunner in the field of electro-mobility. The government has strong incentives for promoting purchase of the BEVs. Survey results indicated more than 80% of the respondents considered exemptions from purchase tax and as the most critical incentives for buying BEVs (Bjerkan et al., 2016). As a result, the growth rate of New-energy Vehicle in Norway is quite fast according to the survey from Green Car Inc. Thus, the incentives provided by the government such as the tax exemption and the relatively low vehicle license fee could promote the sales of New-energy Vehicle. Also, Sierzchula et al. (2012) examined the correlation between financial incentives and EV (Electric Vehicle) market shares in 30 countries. They found that financial incentives and number of charging stations are strong and significant predictors for EV adoption.

3 Proposed Model

This research aims to examine factors that affect sales of New-energy Vehicles. Many researches indicate demand trend (Dijk, et al, 2010), economic change (Sigurðsson, 2010), government policy (Cooper, et al, 2018), marketing planning (Cheng and Zhang, 2017), sales policy (Harrison et al., 2017) and production status (Laoonual and Yossapong, 2013) can all become deciding factors in the prediction of sales of New-energy vehicle.

Oil price and oil volume are the measurements of oil, which is a substitute for electrics. They both function as fuels to power the car. As a result, oil could be a crucial factor towards the sale of the new vehicle (Kah, 2018). It's important to examine the relationship between oil price, oil volume and New-energy Vehicle.

H1: Oil price has positive effect on sales of New-energy Vehicles

H2: Oil volume has negative effect on sales of New-energy Vehicles

CPI (Consumer Price Index) is an important indicator that presents whether people can afford their expenditure (Burns et al., 2008). It is a benchmark created by government to measure inflation. The higher index of CPI means that the cash they hold is worth less. In that case, people may be less willing to try New-energy Vehicles with high inflation (Zhang, 2017).

H3: CPI has negative effect on sales of New-energy Vehicles.

Disposable income could be another important factor since it decides how much money people actually have after paying the tax. A high disposable income guarantees that people can afford the expense of buying New-energy vehicles. (Ning Wang, Yafei Liu, 2015)

H4: Disposable income has positive effect on sales of New-energy Vehicles.

Exchange rate is the price of a country's money in relation to another country's money (Britannica Academic, 2010). It is also regarded as the value of one country's currency in relation to another currency. Choosing CNY and USD as the research objects is because increasingly automobile parts are made in China and export to America (Nykqvist et al., 2015). Such automobile parts include bonnets from three new models of Volvo, Buick and Cadillac.

H5: Exchange rate has positive effect on sales of New-energy Vehicles.

Charging station quantity in this study is the number of charging stations in America. The growing number of charging station effectively reflects the development of new energy vehicles (Li, 2016). This study proposes number of charging station might be a factor affecting sales of New-energy Vehicles.

H6: Number of Charging station has positive effect on sales of New-energy Vehicles.

Subsidies is a form of financial aid or support extended to an economic sector generally with the aim of promoting economic and social policy (Qian, 2018). Subsidies decide how much more money people can receive, which stimulate their interests in buying New-Energy Vehicles (Holtmark et al., 2014).

H7: Subsidies have positive effect on sales of New-energy Vehicles.

Lastly, literacy level reflects a country's education level (Barro et al., 1993). Literacy level may decide whether people are willing to buy the new energy vehicles. Higher literacy level offers people professional knowledge to fully understand new energy and boosts their interests in buying New-Energy Vehicles (Kortland, 1993).

H8: Literacy level has positive effect on sales of New-energy Vehicles.

4 Methodology

In statistics, linear regression is a linear approach to modeling the relationship between a scalar response (or dependent variable) and one or more explanatory variables (or independent variables) (Blokhuin, 2019). Basically, researcher assumes there is a straight line that approximates the data set, and makes prediction based on the line. The formula that describes this relationship can be defined as:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i, \quad i = 1, \dots, n$$

In our model, Y represents the sales of New-energy Vehicle, whereas X represents all the factors that might influence the sales of New-energy Vehicle.

- Y: sales of New-energy Vehicle (per unit)
- X₁: Oil Price (USD/barrel)
- X₂: Oil Volume (per barrel)
- X₃: CPI
- X₄: Disposable Income (per person)
- X₅: Exchange Rate (USD/CNY)
- X₆: Number of Charging Station (per unit)
- X₇: Subsidies (billion)
- X₈: Literacy level (%)

4.1. Data collection

The data of Monthly car sales is collected from InsideEVs, a company which devotes to examine the how, what, and why of electric vehicles, ranging from 2010-2019. Both oil price and oil volume are gathered from Investing, a website mainly makes stock markets quotes and publishes financial news, ranging from 1970-2019. Statistics of CPI comes from Inflation Data, ranging from 1913-2019. The Data of Disposable income is collected from FRED, a database maintained by the Research division of the Federal Reserve Bank of St. Louis, ranging from 1992 to 2017. The data of exchange rate is gathered from Investing, ranging from 1970 to 2019. The quantity of Charging station is come from Statista, ranging from 2010-2019. The historical data of subsidies is also collected from FRED, ranging from 1947 to 2019. Finally, the statistics of literacy level is gathered from an international financial institution that provides interest-free loans and grants to the governments of poorer countries for the purpose of pursuing capital projects, ranging from 1970-2018 (The source of data collection is attached in **Attachment**).

4.2 Procedure

We compiled all the data into one CSV file including independent variables and dependent variable. Null data or incomplete data was excluded before importing to R. Then we imported data into R and summarized the descriptive analysis. Next step, we ran a full model to examine the effect of all independent variables on sales of New-energy Vehicles. According to full model, we removed insignificant factors and run a reduced model again. Finally, we examined the model results with ANOVA analysis.

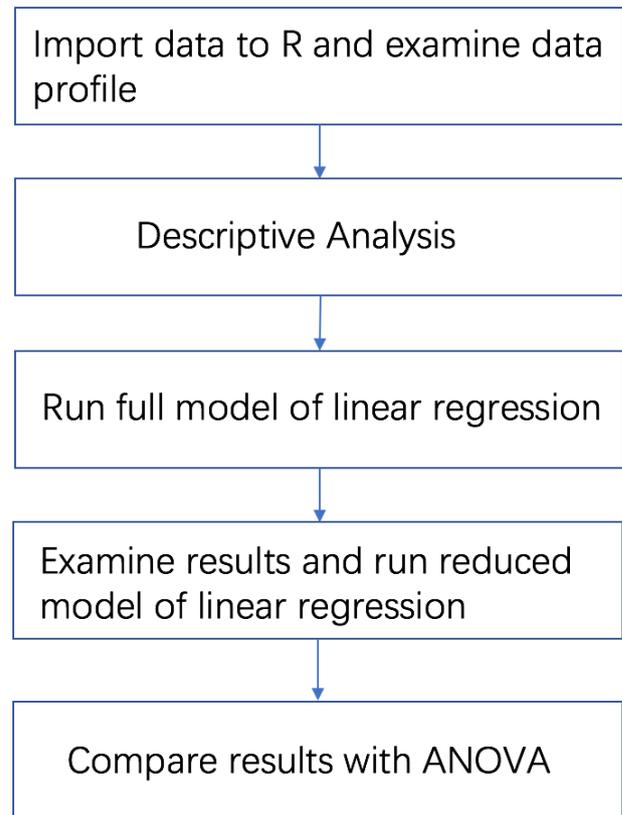


Fig. 1. Procedure

5 Results

5.1. Descriptive analysis

Distribution	Car sale	Oil Price	Oil Volume	CPI	Disposable Income	Charging Station	subsidies	Literacy	Currency
Maximum	286,367	113.93	5,300,000	252.9	14,640	61,067	85.3	86.4	6.976
Minimum	345	33.62	3,590,000	219.2	12,027	3,394	56.01	84.56	6.054
Mean	50,141	73.94	8,129,072	237	13,119	27,251	60.29	85.78	6.424
Median	13,038	73.25	7,170,000	237.1	13,139	25,602	59.39	85.81	6.357
Interquartile	73,676	46.71	5,580,000	10.4	1,327	28,637	2.4	0.89	0.388
Range	286,022	80.31	1,710,000	33.7	2,613	57,673	29.29	1.84	0.922

The sales of New-Energy Vehicle (per unit) ranges from 345 to 286,367 with an average of 50,141. The number is increasing over the year. The oil price (USD/BBL) ranges from 33.62 to 113.93 with an average of 73.94, following a steady rise of sales each year. The oil volume (per BBL) ranges from 3,590,000 to 8,129,072 with an average of 530,000. There has been a slight surge in sales over the year. The CPI ranges from 219.2

to 252.9 with an average of 237. The number has been growing steadily over the last years. The disposable income (per person) ranges from 12,027 to 14,640 with an average of 13,119. There was an improvement in numbers each year. The number of charging station ranges from 3394 to 61,067 with an average of 27,251. The number is increasing over the year. All subsidies (per billion) ranges from 56.01 to 85.3 with an average

of 60.29, with a steadily climbed number each year. Literacy level (%) ranges from 84.56 to 86.4 with an average of 85.78. The number is increasing over the year. The exchange rate (CNY/USD) ranges from 6.054 to 6.976 with an average of 6.424. The number has been growing steadily over the year.

5.2 Linear Regression Model

We ran a full model as Model₀ with all the independent variables. Per results, we cannot reject the null hypothesis of oil price (p=0.21), oil volume (p=0.28), CPI (p=0.76) and currency (p=0.21). The R² of Model₀ is 0.9206, which measures how close the data are to the fitted regression line (Frost, 2013).

	Estimate	Std. Error	t-value	Pr(> t)	
(intercept)	3.274e+06	9.072e+05	3.608	0.000511	***
Oil Price	2.432e+02	1.928e+02	1.261	0.21049	
Oil Volume	-1.350e-03	1.235e-03	-1.093	0.277345	
CPI	4.536e+02	1.497e+03	0.303	0.762537	
Disposable Income	5.073e+01	1.342e+01	3.781	0.000284	***
Currency	1.467e+04	1.168e+04	1.256	0.212473	
Charging Station	2.239e+00	8.376e-01	2.672	0.008971	**
All subsidies	3.045e+03	4.601e+02	6.617	2.77e-09	***
Literacy	-5.062e+04	1.086e+04	-4.663	1.10e-05	***

Fig. 2. Full Model

According to model results, p-value of oil price, oil volume, CPI and Currency is larger than 0.05. While p value of oil price, oil volume and currency is close to 0.1. Thus, we only exclude CPI and run a reduced model as Model₁. The results show that we cannot reject null hypothesis of oil volume (p=0.28), currency (p=0.22) and oil price (p=0.08). The R² of Model₁ is 0.9214.

	Estimate	Std. Error	t-value	Pr(> t)	
(intercept)	3.264e+06	9.021e+05	3.618	0.000492	***
Oil Price	2.773e+02	1.558e+02	1.780	0.078466	
Oil Volume	-1.338e-03	1.228e-03	-1.089	0.278954	
Disposable Income	5.162e+01	1.302e+01	3.964	0.000149	***
Currency	1.358e+04	1.106e+04	1.228	0.222649	
Charging Station	2.407e+00	6.248e-01	3.852	0.000221	***
All subsidies	3.064e+03	4.532e+02	6.762	1.39e-09	***
Literacy	-4.941e+04	1.004e+04	-4.920	3.95e-06	***

Fig. 3. Model₁

According to the results of Model₁, we exclude the three factors oil price (p=0.08), oil volume (p=0.28) and currency (p=0.22). Then we run a further reduced model excluding these three factors as Model₂. The result of the reduced model₂ indicates that all the factors have significant effect on the sales of New-energy Vehicle. And the R² of model₂ is 0.92.

5.3 ANOVA

At the end of the examination, we run the ANOVA analysis to test the model validity and compare the effect of the three models obtained. According to definition, ANOVA test is a way to find out if survey or experiment

results are significantly different. In other words, they help you to figure out if there is significant difference between models. ("ANOVA Test: Definition, Types, Examples" 2019)

	Res.Df	RSS	DF	Sum of Sq	F	Pr(>F)
1	88	2.6405e+10				
2	89	2.6432e+10	-1	-27563732	0.0919	0.7625
3	92	2.8500e+10	-3	-2067590171	2.2969	0.08314

Fig. 4. Model₂

Both p-values (p=0.76 and p=0.08) are larger than 0.05, which means we cannot reject the null hypothesis of significant differences between models. Assuming that this is a random and independent sample which is selected from a normal population and the sample variance is the same, we can conclude that all the three models have no significant difference. We are confident to accept model₂ as our final model in deciding the factors that will affect the sales of the New-energy Vehicle.

6 Conclusion and Limitation

This research utilizes linear regression analysis to examine factors that significantly affect sales of NEV, and uses ANOVA test to compare model results. According to our final model, oil price, disposable income, number of charging station, subsidies and literacy level have significant effect on sales of New-energy Vehicles. Since oil is the complement of the fuel used by the New-energy Vehicle, a rising price of oil leads to the decreasing of quantity demanded of oil. Consumers make decisions based on the opportunity cost of the products. It's an economical way for them to buy the New-energy Vehicle (Munson 2014).

More charging stations also indicate the promotion of infrastructure. When consumers realize that their governments are investing the industry of New-energy Vehicle, specifically the charging station, they are more assured with the purchase of NEV because it is more convenient for them to charge their vehicles. With more disposable incomes, families are more aware of environmental protection, thus are more inclined to purchase environmental friendly vehicles, such as NEV.

Subsidy is also an important factor that directly affect people's purchase because people will consider it as a promotion and a promising trend in the future. Finally, literacy will significantly influence whether people are willing to buy New-energy Vehicles. As stated by Keith Barry on Consumer Reports in 2018, "Reliability is a driving factor, and range anxiety has become less of a concern." (Barry, 2018) Literacy leads more consumers to believe the new technology is safety and eco-friendly, which are positively affect the sale of New-energy Vehicles.

Since new energy car is a relatively new technology, most collected data were dated from 2009 to recently—a relatively short data range. In the future, with the testing of more factors and stronger database, the research will gain a stronger prediction power.

References and Attachments

1. Dresselhaus, M. S., and I. L. Thomas. "Alternative energy technologies." *Nature* 414.6861 (2001): 332.
2. "Types of Electric Vehicles: BEV, PHEV and HEV", EVgo.com
3. Christensen, Linda, Ole Kveiborg, and Stefan Lindhard Mabit. "The Market for electric vehicles—what do potential users want." *12th World Conference on Transportation Research*. 2012.
4. Domm, P. (2019). *The little industry that could take a bite out of oil demand*. [online] Nightly Business Report. Available at: <http://nbr.com/2018/03/02/the-little-industry-that-could-take-a-bite-out-of-oil-demand/> [Accessed 19 Jul. 2019].
5. Bjerkan, Kristin Ystmark, Tom E. Nørbech, and Marianne Elvsaaas Nordtømme. "Incentives for promoting battery electric vehicle (BEV) adoption in Norway." *Transportation Research Part D: Transport and Environment* 43 (2016): 169-180.
6. Sierzchula, William, et al. "The influence of financial incentives and other socio-economic factors on electric vehicle adoption." *Energy Policy* 68 (2014): 183-194.
7. Dijk, Marc, and Masaru Yarime. "The emergence of hybrid-electric cars: Innovation path creation through co-evolution of supply and demand." *Technological Forecasting and Social Change* 77.8 (2010): 1371-1390.
8. Sigurðsson, Jóhann. Economic Effect of Implementing Electric Cars. Diss. 2010.
9. Cooper, Adam, and Kellen Schefter. "Electric Vehicle Sales Forecast and the Charging Infrastructure Required Through 2030." *Edison Electric Institute and Institute for Electric Innovation* (2018).
10. CHEN, RU, and WEI ZHANG. "Research on the Marketing Strategy of Electric Vehicle based on Consumer Behavior." *DEStech Transactions on Social Science, Education and Human Science* msic (2017).
11. Harrison, Gillian, and Christian Thiel. "An exploratory policy analysis of electric vehicle sales competition and sensitivity to infrastructure in Europe." *Technological Forecasting and Social Change* 114 (2017): 165-178.
12. Laonual, Yossapong. "EV current status in Thailand." (2013).
13. Kah, M. (2018). *Electric Vehicles and Their Impact on Oil demand: Why Forecasts Differ*. [online] energypolicy.columbia.edu.
14. Burns, Cate, Gary Sacks, and Lisa Gold. "Longitudinal study of Consumer Price Index (CPI) trends in core and non-core foods in Australia." *Australian and New Zealand Journal of Public Health* 32.5 (2008): 450-453
15. Zhang, Yong, et al. "Forecasting electric vehicles sales with univariate and multivariate time series models: The case of China." *PloS one* 12.5 (2017): e0176729.
16. Wang, Ning, and Yafei Liu. "Key factors influencing consumers' willingness to purchase electric vehicles in China." *The 28th International Electric Vehicle Symposium and Exhibition*. 2015.
17. "Exchange rate." *Britannica Academic*, Encyclopædia Britannica, 1 Nov. 2010. academic-com.proxy.library.upenn.edu/levels/collegiate/article/exchange-rate/33406. Accessed 26 Jul. 2019.
18. Nykvist, Björn, and Måns Nilsson. "Rapidly falling costs of battery packs for electric vehicles." *Nature climate change* 5.4 (2015): 329.
19. Li, Jing. "Compatibility and investment in the us electric vehicle market." *Job Market Paper* (2016).
20. Qian, Jing. Evaluating the Electric Vehicle Subsidy Program in China. Working paper, Cornell University, 2018.
21. Holtsmark, Bjart, and Anders Skonhøft. "The Norwegian support and subsidy policy of electric cars. Should it be adopted by other countries?" *Environmental science & policy* 42 (2014): 160-168.
22. Barro, Robert J., and Jong-Wha Lee. "International comparisons of educational attainment." *Journal of monetary economics* 32.3 (1993): 363-394.
23. Kortland, J. "Scientific Literacy and Context-Based Curricula: Exploring the Didactical Friction between Context and Science Knowledge." *D. Höttecke, Naturwissenschaftliche Bildung als Beitrag zur Gestaltung partizipativer Demokratie* (2011): 17-31.
24. Andriy Blokhin. "What Is the Difference Between Linear and Multiple Regression?" *Investopedia*, Investopedia, 21 Apr. 2019, www.investopedia.com/ask/answers/060315/what-difference-between-linear-regression-and-multiple-regression.asp.
25. Frost, Jim. "Regression analysis: How do I interpret R-squared and assess the goodness-of-fit." *The Minitab Blog* 30 (2013).
26. "ANOVA Test: Definition, Types, Examples." *Statistics How To*, www.statisticshowto.datasciencecentral.com/probability-and-statistics/hypothesis-testing/anova/.
27. Munson, Isabel. "Economics Explained: Complements, Substitutes, and Elasticity of Demand." *EconoGIST*, EconoGIST, 19 Sept. 2014, www.econogist.com/home/complements-and-substitutes
28. Barry, Keith. "More American Drivers Want Electric Cars, AAA Survey Says." *Consumer Reports*, www.consumerreports.org/hybrids-evs/more-american-drivers-want-electric-cars-aaa-survey-says/.

29. Attachments

No.	Factors	Duration(year)	Websites
1	Oil Price	January, 2010- March,2019	https://www.investing.com/commodities/crude-oil-historical-data
2	Oil Volume	January, 2010- March,2019	https://www.investing.com/commodities/crude-oil-historical-data
3	CPI	January, 1913- March,2019	https://inflationdata.com/Inflation/Consumer Price Index/HistoricalCPI.aspx?reloaded=true
4	Disposable Income	January, 1992- March,2017	https://fred.stlouisfed.org/series/W388RC1A027NBEA
5	Exchange Rate	January, 1970- March,2019	https://www.investing.com/currencies/cny-usd
6	Number of Charging Station	January, 2010- March,2019	https://www.statista.com/statistics/416750/number-of-electric-vehicle-charging-stations-outlets-united-states/
7	Subsidies	January, 1947- March,2019	https://fred.stlouisfed.org/series/B096RC1Q027SBEA
8	Literacy level	January, 1970- March,2018	https://data.worldbank.org/indicator/se.adt.litr.zs?end=2018&start=1970&view=chart