

Influencing Factors of Different Brands with Monofactor Analysis

Chunyan Zhang^{1,2}, Benshuang Liang^{1,2}, Tiantian Wan^{1,2}

¹China Automotive Technology & Research Center Co., Ltd., Tianjin, China;

²China Auto Information Technology (Tianjin) Co., Ltd., Tianjin, China.

Abstract

Nowadays environmental protection is getting stronger and stronger, the appearance of electric vehicles has given people a new driving experience, it also makes a new market prove its technologies by using resources, but it is not fully mature, high cost and other factors have made many users doubtful. The satisfaction of the analysis results and individuality are mainly based on actual expenditures and mathematical model built by other analysis methods to predict the possibility of target customers to purchase vehicles, and provide feasible marketing strategies for the sales department, moreover, explore potential customers, expand the electric vehicle market, and reduce energy use problems.

Keywords

Mathematical Model; Predict; Target Customers.

1. Problem Background

Lucid waters and lush mountains are invaluable assets, resources are increasingly scarce at present, the voice of environmental protection is becoming stronger and stronger, under this background, the advent of new energy vehicles will undoubtedly make the public amazed and have high hopes for sustainable development. However, it is true that the development prospect of electric vehicles is broad, the public have higher acceptance for them [1-4], but when actually considering the purchase intention, many people are still full of doubt. On this basis, scientific decision-making of market sales, full grasp of the user experience, and finding potential customers can not only expand the new energy vehicle market represented by electric vehicles, but also reduce the waste of resources, so as to solve energy and environmental problems, keep the good blue sky, and build a beautiful China.

2. Problem Analysis

2.1. The data is preprocessed

First of all, the full mark of each index satisfaction score is 100, and no reasonable range of other data is given, therefore, it is necessary to study the mutual influence of various indexes by analyzing the meaning of each index, and make reasonable judgment on the outliers of the indexes. After the outliers are obtained, the actual value must be further estimated and replaced. Then, the missing data in the table are filled, ensure that the sample size is large enough, in order to facilitate subsequent statistics and analysis.

2.2. The influencing factors which determine whether the target customer buys an electric vehicle are determined

First, we must analyze the characteristics of existing factors, for example, the value range and category, etc., and then use one-way analysis of variance method to determine the influencing factors of the three brands based on the difference of factor types. In this process, by analyzing the p value obtained, determine whether the independent variable has a significant impact on

the dependent variable, namely purchase intention, so as to determine the main influencing factor of whether the target customers of each brand purchase the electric vehicle.

3. Research Methods

3.1. Data preprocess

3.1.1. Search and process of missing value

According to observation of the sample size, the overwhelming majority of customers with null values are "married/cohabiting with no children (living with parents)", "unmarried, living alone", "married/cohabiting with no children (lovers' world)", "unmarried, living with parents" four categories, so we can determine that the implication is no children.

Processing method: in order to ensure the consistency of the data, these missing values are all assigned to 0.

3.1.2. Search and process of outliers

First of all, according to the data, the full mark of satisfaction score of the experience data is 100, namely, the reasonable range of the data is 0-100. When the data in the table exceeds this range or differs greatly from other data, it is determined as outliers. According to statistics, the satisfaction score of target customer numbers: 1, 100, 48, and 1964 are abnormal. The satisfaction score of customer number 1's battery technical performance (battery durability and convenient charging) is 753.04; the overall satisfaction score of customer number 100's appearance and interior performance is 7.88; the overall satisfaction score of customer number 480's power performance (climbing and acceleration) is 605.63; the overall satisfaction score of customer number 1964's economy (energy consumption and value retention) is 703.

Processing method: considering that the customers' experience data will be incorrectly when filling manually, and the other satisfaction scores of the four customers are all reasonable values, so the three abnormal data which greater than 100 points are reduced by ten times as true data. An abnormal data whose score is much smaller than other data is expanded ten times as its real data. Then, by analyzing the personal characteristics questionnaire of target customers, according to the analysis of the personal characteristics of target customers, there are attribute classes and attribute values among the 17 types of personal characteristics of customers. For attribute class information, it is necessary to choose among the given classification items. For attribute value information, a reasonable range should be determined based on the actual situation. According to statistics, there are no outliers in the attribute information of the target customer's personal characteristics. The attribute value information has the following outliers: item B17 of customer number 223, namely, the annual car loan expenditure accounts for 300% of the family's total annual income, which is not logical, the processing method is: reducing the ratio by ten times as its true data. Secondly, the customer's marriage and family situation can actually infer whether the customer has children or not. When there is inconsistent, it will be processed as an outlier. According to statistics, in the customers who choose to have no children in the marriage and family situation, 26 customers fill that they have 1 child, and 4 customers fill that they have two children; the processing method is:

the data that have children is modified to have no children. In the customers who choose to have children in the marriage and family situation, two customers do not fill in the results. The processing method is as follows: on the basis of data statistics, customers with one child account for as high as 85% of customers with children, therefore, it is conservatively assumed that the two customers who do not fill in the results have one child as their true data.

Finally, it can be known that the family's annual income must be greater than the household disposable annual income by consulting the data. According to statistics, the annual household income of nearly 70 customers among all customers is less than the household disposable

annual income. These data are outlier, and the processing method is as follows: considering that the concept of disposable annual income is not clear when the customer fills in the questionnaire, as a result, the phenomenon of conceptual confusion occurs. In order to preserve these sample data for subsequent analysis, we exchange the data on the annual household income of these customers with the household annual disposable income as the real data.

3.1.3. Statistical analysis

For attribute value information B2, B4, B5, B7, B8, B10, B13, B14, B15, B16, B17, count the minimum value, maximum value, mean value, standard deviation, and median are counted, and the results are shown in Table.1:

Table 1. Statistics of target customer attribute value information

personal characteristic name	minimum value	maximum value	mean value	standard deviation	median
B2	1	60	21.38	11.44	20
B4	1	30	7.64	4.13	7
B5	1	6	3.46	1.08	3
B7	0	3	0.85	0.61	1
B8	21	61	34.30	5.01	34
B10	1	38	10.08	4.91	10
B13	7	100	27.23	12.46	25
B14	3	95	16.59	10.91	14
B15	2	80	15.97	9.79	13
B16	0	60	15.28	13.14	15
B17	0	60	9.45	10.59	6

It can be seen from the table that the maximum values of the 9 items of B2, B4, B8, B10, B13, B14, B15, B16, B17 exceed the average value by 3 standard deviations, it suggests that the data fluctuates greatly, it is more appropriate to use the median to describe the overall level. As can be seen that:

In the target customers, 69.30% of customers' accounts are in this city, it suggests that the target customers are mainly locals. In addition, only 0.20% of customer accounts chose others, the sample size is small, and it has no reference value;

In the target customers, 51.43% of the customers live in the city center, 39.21% of the customers live in the non-city center, and the total proportion of customers living in the rural-urban fringe zone, county and rural areas is less than 10%, it suggests that the target customers mainly live in the urban area.

In the target customers, 49.39% of the customers are married, have children and do not live with their parents, and 22.71% of the customers are married, have children, and live with their parents, the total proportion is as high as 72.1%, the proportion of other marital status is relatively small, it suggests that the target customers are mainly married customers with children.

In the target customers, the proportion of college and undergraduate customers is as high as 84.93%, the proportion of low-educated customers is very small, there is not even customers who have a primary school educational background or below, it shows that the target customers are mainly people with higher knowledge levels.

In the target customers, private enterprises account for the largest proportion of customers' units, accounting for 35.13%, the next state-owned enterprises, foreign-funded enterprises, and joint ventures, accounting for 10-20%, however, the total proportion of government agencies, government departments, grass-roots organizations, scientific research, education, health and medical institutions, freelancers and non-working institutions is less than 10%, it suggests that the target customers are mainly customers whose units are private enterprises.

In the target customers, 21.59% of the customers are middle managers, 19.40% of the customers are middle technician, and the proportion of customers of the rest positions is basically less than 10%, it suggests that the target customers are mainly middle managers and technician.

3.2. Based on variance analysis

3.2.1. Model building

In a complex thing, there are generally multiple factors that restrict and depend on each other. The purpose of the variance analysis is to find the factors that have a significant impact on the matter through data analysis, and the optimal level of the significant impact factors. Specifically, in production and scientific research, it is often necessary to analyze whether changes in production conditions or experimental conditions have an impact on product quality and experimental results. The basic idea is to determine the influence of controllable factors on the research results by analyzing and studying the contribution of different sources of variation to the total variation.

The factors influencing customers to purchase electric vehicles are complex and diverse, there are mainly 8 factors in satisfaction score and 17 factors in customer personal characteristic information, if the multi-factor analysis of variance model is built, its modeling is very difficult and it is very likely to be difficult to solve, however, we found that there is no interaction between the above factors, namely, the variables will not affect each other, for example, the battery technical performance score a1 does not affect the comfort score a2, etc., based on this, we can determine whether each factor has a significant impact on the intention to buy vehicle and the size of the impact by building single-factor variance analysis model many times.

Suppose that single factor a1 has n levels, namely, purchasing vehicle and without purchasing vehicle two categories, they are recorded as $A_i (i = 1, 2, \dots, n)$ below each level $A_i (i = 1, 2, \dots, n)$, respectively, the index to be investigated can be regarded as a whole, so there are n totals, and suppose:

- (1) Each population obeys the normal distribution, namely $x_i \sim N(\mu_i, \sigma^2), i = 1, 2, \dots, n$
- (2) The variance of each population σ^2 is the same;
- (3) The samples $x_{i,1}, x_{i,2}, \dots, x_{i,m}$ drawn from each population are independent of each other.

x_{ij} is the j-th group of data at the i-th level, our purpose is to judge whether there is a significant difference between the two levels of a1, it is equivalent to carry out the following hypothesis test:

3.2.2. Results and analysis

(1) Joint venture brand

The satisfaction score without purchase is about 10 points lower than the satisfaction score with purchase, it shows that different purchase intention samples show significant differences for a1, a2, a3, a4, a5, a6, a7, and a8. Namely, the eight factors of electric vehicles themselves are all influencing factors that determine whether the target customer will purchase electric vehicle of joint venture brand.

The variance analysis of customer characteristic factors is shown in Fig.1 below:

As shown in Fig.1, samples of different purchase intentions will not show significance for 13 items of B1, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13 and B14, it means that samples of different purchase intentions show consistency for B1, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B14, and there is no difference. However, the purchase intention sample is significant for B2, B15, B16 and B17, it means that different purchase intention samples are different for B2, B15, B16, and B17. Specific analysis shows that: the number of living years in this city B2, household disposable annual income B15, annual mortgage ratio B16, annual car loan ratio B17

are the influencing factors in determining whether target customers will purchase joint venture brand electric vehicles.

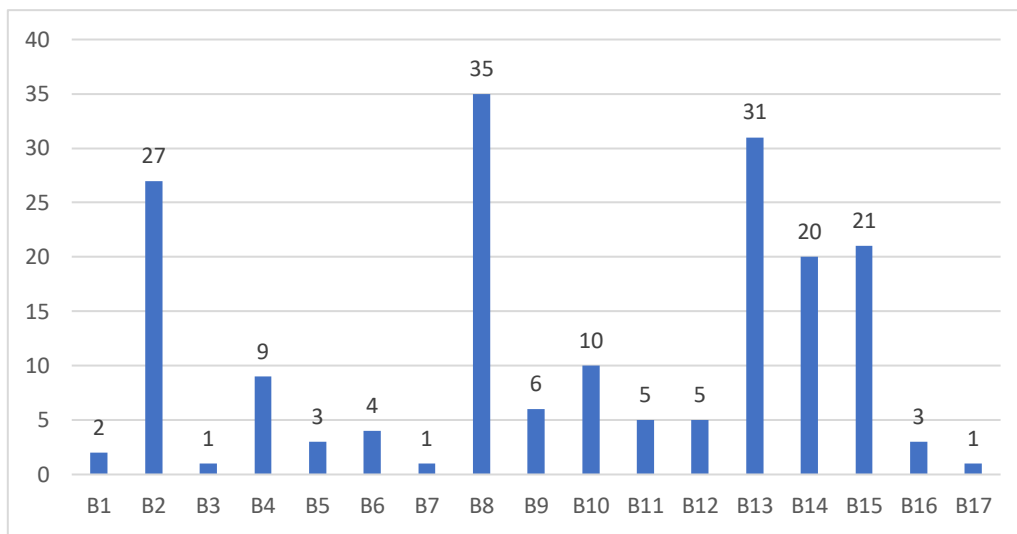


Fig. 1 Comparison diagram of variance analysis of personal characteristics of joint venture brand target customers

(2) Self-owned brand

The satisfaction score without purchase is 8 to 11 points lower than the satisfaction score with purchase. It shows that different purchase intention samples show significant differences for a1, a2, a3, a4, a5, a6, a7, and a8. Namely, the eight factors of electric vehicles themselves are all influencing factors that determine whether the target customer will purchase the self-owned brand electric vehicle.

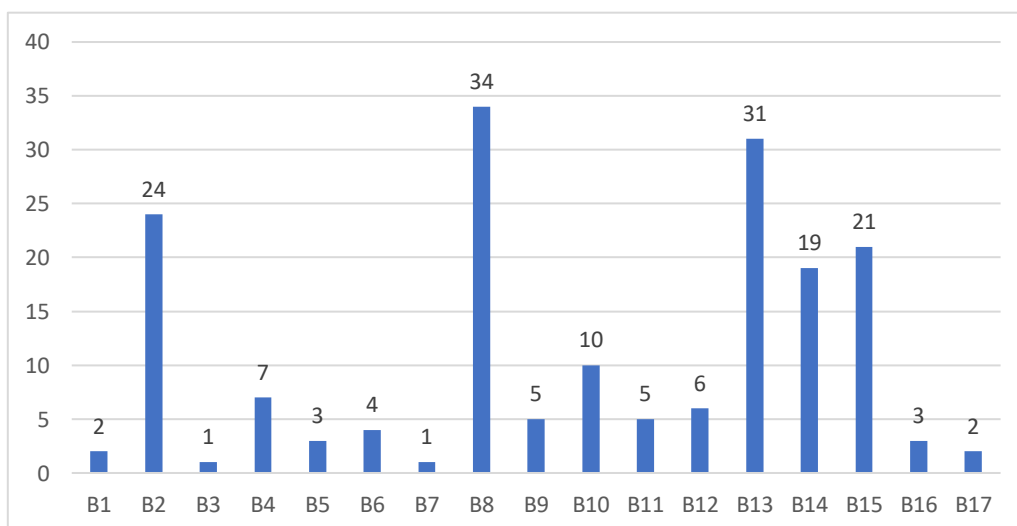


Fig. 2 Comparison diagram of variance analysis of individual characteristic factors of self-owned brand target customers

As shown in Fig.2, samples of different purchase intentions will not show significance for B1, B3, B4, B5, B6, B7, B8, B9, B10 and B12, it means that different purchase intention samples do not show consistency for B1, B3, B4, B5, B6, B7, B8, B9, B10 and B12, and there was no difference. In addition, the purchase intention samples show significance for B2, B11, B13, B14, B15, B16, and B17, it means that different purchase intention samples are different for B2, B11,

B13, B14, B15, B16, and B17. Specific analysis shows that the number of living years in this city B2, the nature of the unit B11, the annual household income B13, the personal annual income B14, the disposable annual household income B15, the annual housing loan proportion B16, and the annual car loan proportion B17 are the influencing factors that determines whether the target customer will buy electric vehicles of self-owned brand.

(3) New power brand

The satisfaction score without purchase is 8 to 14 points lower than the satisfaction score with purchase. It shows that different purchase intention samples show significant differences for a1, a2, a3, a4, a5, a6, a7, and a8. Namely, the eight factors of electric vehicles themselves are all influencing factors that determine whether the target customer will buy electric vehicles of new power brands. The variance analysis of the target customer's personal characteristic factors is shown in Fig.3 below:

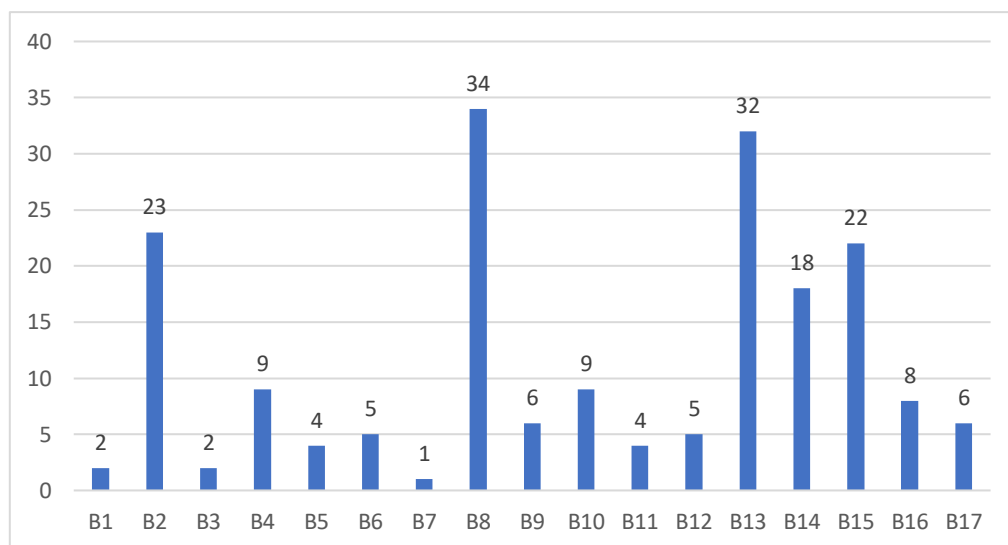


Fig. 3 Comparison diagram of variance analysis of the personal characteristics of new power brand target customers

As shown in Fig.3, samples of different purchase intentions will not show significance for B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B14, B15, and B17 16 items, it means that samples of different purchase intentions show consistency for B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B14, B15, B17, and there is no difference. In addition, the purchase intention sample show significance for B16, it means that different purchase intention samples have difference for B16. Specific analysis shows: the proportion of mortgage in the whole year, B16, is the influencing factor in determining whether target customers will purchase electric vehicles from new power brand.

References

- [1] Hongchao Shen. Analysis of Influencing Factors of China's Electric Vehicle Demand based on Analytic Hierarchy Process[J]. Journal of Jilin Business and Technology College, 2013.
- [2] Mengkai Yan. 6-Meter Light Truck Development Strategy Based on User Requirement Model[D]. Hebei University of Technology, 2018.
- [3] Leshan Li. Design Survey[M]. Beijing: China Machine Press, 2007.
- [4] Wenping Zhang. Setting Theory of Qualitative Research Sample Amount and Practice in Vehicle Market Research[J]. Auto Industry Research, 2015(04): 22-26.