The Significant Factors in Refuse Classification Based on Correlation and Multiple Regression Analysis

Liqian Zheng[1, *]

[1] Department of Environment Science, Fudan University, Shanghai, China

*Corresponding author: 18307110163@fudan.edu.cn

Abstract. With the rapid growth of China's economy, great achievements have been made in various constructions. Meanwhile, huge resources and environmental damage have been paid. To reduce pollution and protect environment, the continuous implementation of mandatory waste classification policy in various cities in China, refuse classification has been a practical problem that tightly related to mankind's live. In this article, it finds the significant factors that affect the effectiveness of the refuse classification as well as the people's attitude toward it. Based on the questionnaire, people's will turn positive as the classification commences and guidance takes quite an important role in the process. Refuse classification is a long-term project beneficial to the country and the people. Contemporarily, the government has been advocating waste classification, but it has not been effectively implemented in the actual operation process. On the basis of in-depth analysis of the problems existing in the implementation of garbage classification, key objects are found to form a green and environment-friendly resource conserving society. These results provide some reference for further garbage classification.

Keywords: Refuse classification, linear regression, t-test.

1. Introduction

Refuse classification refers to a series of activities that transform garbage into public resources by classified storage, classified placing and classified handling according to certain regulations or standards. The purpose of classification is to improve the resource value and economic value of garbage and strive to make the best use of everything [1]. In the process of classified storage of garbage, it belongs to people's private goods. After classified placement, the garbage will become a quasi public resource in people's living areas. In addition, after the garbage is classified and transported to the centralized treatment point, it will become a public resource. The advantages of refuse classification are manifold. Firstly, it helps reduce the floor space and improve the land utilization rate. Some domestic wastes contain harmful substances, which are difficult to degrade and damage the land seriously. If the recyclable materials can be removed by classification, the landfill quantity can be reduced by more than 60% [2]. Secondly, it can reduce waste pollution and protect the ecological environment. At present, the garbage disposal in China is mostly in the form of landfill, which occupies a large area of land, and will cause mosquitoes breeding, sewage turbulence, unpleasant smell and other uncomfortable phenomena [3]. Moreover, it can turn waste into treasure and effectively use resources. According to relevant data, the number of plastic fast food boxes, disposable chopsticks and instant noodles bowls used in China each year is up to 4 billion, 1 billion pairs and 500 million to 700 million, accounting for about 15% of domestic waste [4]. In this case, one ton of waste plastic can be used to extract 600kg of diesel oil. Recycling 1500 tons of waste paper can avoid cutting 1200 tons of paper trees [5].

At present, many people use linear regression analysis in the field of environment, Qu used multiple linear regression to analyze the monitoring data of the atmospheric environment [6]. After several times of optimization, the short-term atmospheric environmental quality is accurately predicted, and the mass concentration of future atmospheric pollutants is relatively accurately predicted. Chen used the linear regression model to find the main pollutants of air pollution [7]. Through mathematical analysis, the correlation conclusion between air quality index (AQI) and two dusts and four gases in Dazu District of Chongqing was obtained. The study found that PM2.5, PM10

and O3 in Dazu District of Chongqing were the main pollutants, of which O_3 was the key treatment object. In the aspect of water quality evaluation, Li et al. have established a dynamic prediction model of groundwater quality based on multiple linear regression, for groundwater quality assessment[8], Zhuang et al. used the multiple regression model of water environmental carrying capacity and economic efficiency is to assist environmental and economic decision-making [9]. Zhao et al. used the multiple regression analysis to provide reference for red tide forecast [10]. China is a big garbage producer. As early as 2004, China has surpassed the United States to become the world's first garbage producer. According to the data provided by China Urban Environmental Sanitation Association, in 2015, the average resident in China will produce 440 kg of domestic waste per year, and the growth rate will be up to 10% [11].

On July 1, 2019, Shanghai promulgated Regulations of Shanghai Municipality on the management of domestic garbage [12], which means that Shanghai has become the first city in China to enforce waste classification. In the subsequent times, more and more cities in China began to conduct the policy of mandatory waste classification.Therefore, refuse classification has become an practical problem that tightly related to our lives. In order to help the government make decision more wisely and arrange the proportion of each aspect of work more reasonably. The rest part of the paper is organized as follows. The Sec. 2 will give a brief introduction about multiple linear regression. The Sec. 3 will introduce the methods used in this article and the date resource. The Sec. 4 will show the results of this research. The Sec. 5 is a short discussion and the Sec. 6 is the limitations of this research as well as the future outlooks

2. Statistical Principle

Multiple linear regression analysis is the simplest and most commonly used one in multiple regression analysis. Its basic principles and methods are the same as those of single linear regression analysis.

2.1. Principle of Multiple Linear Regression

In multiple regression, it is necessary to determine the relationship between the dependent variable and multiple independent variables. Its mathematical model is [12]:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_iX_i + \varepsilon \quad (1)$$

The above formula indicates that the dependent variable Y in the data can be approximately expressed as a linear function of the independent variable X_1, X_2, X_3…X_i. b_0 is a constant term, b_1, b_2, b_3…b_i are partial regression coefficients, representing the average variable of Y when Xi increases or decreases by one unit when other independent variables remain unchanged, and ε is the random error (residual error) after removal of i independent variable pairs. Without considering ε, the approximate relationship between dependent variable Y and independent variable X can be obtained:

$$\hat{Y} = b_0 + b_1X_1 + +b_2X_2 + b_3X_3 + \cdots + b_iX_i \quad (2)$$

When the variables X_1, X_2, X_3... X_i take different test values, i group of test data X_{1j}, X_{2j}, X_{3j}, …, X_{ij}, Y_j (j=1, 2, ... i) are obtained, and substitute X_{1j}, $X_{2j,}$ X_{3j}, …X_{ij} into the formula to obtain the corresponding function calculation value $\hat{Y}$, so the sum of residual squares is:

$$SS = \sum_{j=1}^{i}\left(Y_i - \hat{Y}_i\right)^2 = \sum_{j=1}^{i}(Y_i - b_0 - b_1X_1 - b_2X_2 - b_3X_3 - \cdots - b_iX_i)^2 \quad (3)$$

According to the principle of least square method, to minimize SS, conditions shall be met $\frac{\partial(SS)}{\partial b_j} = 0$ (j=1,2,3…i). Thus, the following equations can be obtained [12]:

$$\begin{cases} ib_0 + b_1\sum_{j=1}^{i} X_{1j} + b_2\sum_{j=1}^{i} X_{2j} + \cdots + b_i\sum_{j=1}^{i} X_{ij} = \sum_{j=1}^{i} Y_j \\ b_0\sum_{j=1}^{i} X_{1j} + b_1\sum_{j=1}^{i} X_{1j}^2 + b_2\sum_{j=1}^{i} X_{1j}X_{2j} + \cdots + b_i\sum_{j=1}^{i} X_{1j}X_{ij} = \sum_{j=1}^{i} X_{1j}Y_j \\ b_0\sum_{j=1}^{i} X_{2j} + b_1\sum_{j=1}^{i} X_{1j}X_{2j} + b_2\sum_{j=1}^{i} X_{2j}^2 + \cdots + b_i\sum_{j=1}^{i} X_{2j}X_{ij} = \sum_{j=1}^{i} X_{2j}Y_j \\ \cdots\cdots \\ b_0\sum_{j=1}^{i} X_{ij} + b_1\sum_{j=1}^{i} X_{1j}X_{ij} + b_2\sum_{j=1}^{i} X_{2j}X_{ij} + \cdots + b_i\sum_{j=1}^{i} X_{ij}^2 = \sum_{j=1}^{i} X_{ij}Y_j \end{cases} \quad (4)$$

Hence, the coefficient b_0, b_1, b_2, b_3…b_i of the regression equation can be obtained by solving this equation group.

2.2. Significance Test of Multivariate Linear Regression

In the study of practical problems, one cannot conclude that there is a linear relationship between random variable Y and variable X_1, X_2, X_3…X_i in advance. Before estimating the regression parameters, multiple linear regression equation was used to fit the relationship between random variable Y and variable X_1, X_2, X_3…X_i is only a hypothesis based on some qualitative analysis. Therefore, it is of great importance to test the significance of the obtained linear regression equation. This paper uses the correlation coefficient test method to test the significance of the regression equation. The complex correlation coefficient R in the significance test of multiple regression equation is similar to the correlation coefficient r of linear regression, which can be calculated by the following method.

It is assumed the total sum of squares be expressed in SS_T, which reflects the total fluctuation degree of the observed data and measures the fluctuation degree or uncertainty degree of the explained variable Y. The sum of squares of regression is expressed in SS_R, which reflects the sum of squares of deviations in the degree of correlation between the independent variable and the dependent variable. It measures the part of the uncertainty degree that can be explained by the explained variable X, which is the explained variable Y. The partial regression coefficient R can be expressed by SS_T and SS_R, and the formula is as follows:

$$R = \sqrt{\frac{SS_R}{SS_T}} \quad (5)$$

For a given significance level, the significance test requires that when $R>R_i$, it can show that there is a close linear relationship between Y and variable X_1, X_2, X_3…X_i. R_{min} is called the critical value of complex correlation coefficient l, and its value is obtained from the critical value table in the statistical book.

3. Data & Method

The data came from a questionnaire and a total of 56 effective observations were collected in this questionnaire survey. The questionnaire includes 25 questions, including the subjects' attitude towards garbage classification, their understanding level of garbage classification, the situation of garbage classification points and their own and others' behaviors when taking out garbage. In addition to basic information, the subjects will score according to the content of the question, which is divided into seven scores from 1 to 7. Besides, they will also be asked to do a small test about refuse classification. The average age is 22.43, the maximum value is 50, and the minimum value is 18, indicating that most of the subjects are young. In combination with the curriculum survey background, most of the subjects are college students. The mean value of SH variable is 1.36, which indicates that the sample includes both the subjects who have carried out garbage classification in their hometown and some subjects who have not carried out garbage classification in their hometown. The descriptive statistics are shown in the Table. 1.

The subjects had a higher understanding of the three basic information of "garbage placement place, garbage placement time and garbage classification method", while they had a lower understanding of the information that needed to be further explored, such as "garbage further treatment and illegal treatment method". There are 6 attitude questions in total, and the descriptive

statistical results have little difference. The score is 4-6 on a 7-point scale, indicating that the subjects are generally recognized for garbage classification. Relatively speaking, “nece” (the level of the necessity of classification) is the largest, indicating that the subjects actively recognize the necessity of waste classification overall, but the scores of rationalities and effectiveness in the attitude variables are relatively low, indicating that the subjects generally believe that the current implementation of waste classification does not match the ideal situation. The above results are consistent with this research results. The results show that the minimum values of "whether they leave garbage for the next day", "whether they classify garbage incorrectly" and "whether there are people around who do not classify garbage" are greater than 1, indicating that these three situations are more common than other behaviors. The mean values of other behaviors are also greater than 3, indicating that the frequency of occurrence is not low. It should be noted that although the average value of garbage accumulation in the corridor in the survey interview is 3.64 (the lowest value in the behavior test), its standard deviation is also the largest, indicating that this situation varies from building to building, indicating that it may not be caused by the implementation of garbage classification. This paper divides three demands from the demand side: convenience demand, environmental protection demand and economic demand. The survey results show that convenience is the most demanding, followed by environmental protection, and finally economic. This first shows that the subjects do not mind increasing the cost of garbage classification. It also shows that the subjects have a certain awareness of environmental protection, and the most important thing is the convenience of the procedure.

Table 1. Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
time	56	134.21	89.17	43	516
age	56	22.43	5.56	18	50
sh	56	1.36	0.48	1	2
gender	56	0.32	0.47	0	1
location	56	5.43	1.63	1	7
ruletime	56	4.96	1.80	1	7
standard	56	4.57	1.54	1	7
disposal	56	3.39	1.56	1	7
rule	56	3.36	1.58	1	7
conv	56	4.71	1.59	1	7
reasonable	56	4.61	1.60	1	7
effect	56	4.00	1.35	1	7
environ	56	5.21	1.53	1	7
econ	56	4.54	1.46	1	7
nece	56	5.57	1.19	3	7
tomorrow	56	5.36	1.38	2	7
mixerror	56	5.32	1.35	2	7
divid_near	56	5.50	1.33	3	7
divid_guide	56	4.29	1.70	1	7
corridor	56	3.64	2.00	1	7
dirty	56	4.43	1.86	1	7
type	56	3.93	1.66	1	7
conv_demand	56	78.11	22.51	34	100
environ_de~d	56	69.54	27.67	0	100
econ_demand	56	54.64	28.47	0	100

4. Results

4.1. T-tests

In this paper, whether garbage classification has been started in hometown is grouped, and the difference t-test results are shown in Table 2. The results in the table show that there are significant differences. The subjects whose their hometown have been carried out garbage classification know more about the garbage classification standards. For the environmental benefits of waste classification, the subjects who have carried out waste classification in their hometown are more active. Regarding the necessity of garbage classification, the subjects who have carried out garbage classification in

their hometown are more active. Whether the garbage sorter is not in compliance, the subjects who have carried out the test report more, which is in line with the actual situation. Whether there is garbage accumulation in the corridor, and there are fewer places for garbage classification. For the economic demand of waste classification, the subjects who have been carried out in their hometown think it is more important. Then, this paper classifies the test results of the garbage classification problem. The classification basis is the group with high scores and the group with low scores. Seen from the results, there were significant differences in the t-test of groups. The understanding degree of garbage placement location, garbage placement time and garbage classification standard, and the understanding degree of high groups was significantly higher. The sanitation of the garbage dumping site and the unreasonable setting of garbage cans are more reflected by the low group. In combination with the above analysis, it should be noted that no matter what kind of grouping is carried out, there is no significant difference in the garbage classification effect (effectiveness and rationality), which indicates that the current garbage classification has not had an important impact on the attitude.

Table 2. T-test according to two classification methods

Variables	Mean of started	Mean of not started	MeanDiff 1	Mean of low score	Mean of high score	MeanDiff 2
location	5.556	5.2	0.356	5.067	5.846	-0.779*
ruletime	5.056	4.8	0.256	4.467	5.538	-1.072**
standard	4.889	4	0.889**	4	5.231	-1.231***
disposal	3.167	3.8	-0.633	3.533	3.231	0.303
rule	3.278	3.5	-0.222	3.6	3.077	0.523
difficulty	5.111	4.9	0.211	5.267	4.769	0.497
conv	4.889	4.4	0.489	4.933	4.462	0.472
reasonable	4.778	4.3	0.478	4.333	4.923	-0.59
effect	3.944	4.1	-0.156	4	4	0
environ	5.556	4.6	0.956**	5.4	5	0.4
econ	4.667	4.3	0.367	4.533	4.538	-0.005
nece	5.833	5.1	0.733**	5.467	5.692	-0.226
tomorrow	5.333	5.4	-0.067	5.267	5.462	-0.195
mixerror	5.278	5.4	-0.122	5.067	5.615	-0.549
divid_near	5.667	5.2	0.467	5.333	5.692	-0.359
divid_gu~e	4.889	3.2	1.689***	4.4	4.154	0.246
corridor	3.167	4.5	-1.333**	3.8	3.462	0.338
dirty	4.556	4.2	0.356	5.133	3.615	1.518***
type	3.944	3.9	0.044	4.6	3.154	1.446***
conv_dem~d	81.222	72.5	8.722	75.4	81.231	-5.831
environ_~d	73.556	62.3	11.256	70.267	68.692	1.574
econ_dem~d	60.444	44.2	16.244**	50.533	59.385	-8.851

*, **, *** represets for significant value at 90%, 95% and 99%.

Table 3. Factor analysis rotation matrix results

Variable	Factor1	Factor2	Factor3	Factor4	Factors	Factor6
location	0.03172	0.07801	-0.05246	0.38362	-0.12679	0.03798
ruletime	-0.01928	-0.03776	0.09999	0.32220	0.00489	-0.02935
standard	0.00809	0.01753	-0.08636	0.34463	0.02243	0.11318
disposal	-0.02154	-0.02039	0.02034	-0.04651	0.48667	0.11055
rule	0.02059	0.07295	-0.05176	-0.03125	0.45138	-0.11816
conv	0.01016	0.01672	-0.01383	-0.07943	0.09292	-0.11629
reasonable	0.18070	0.02326	0.09070	0.04411	-0.06278	-0.04314
effect	0.02042	-0.22921	0.14764	-0.09735	0.15691	0.06121
environ	0.29940	0.14898	-0.05834	-0.00140	-0.02067	-0.13133
econ	0.18835	-0.01495	0.03060	-0.06847	0.11587	-0.04283
nece	0.30745	0.11770	0.03755	0.05507	-0.13348	-0.16365
tomorrow	0.05171	-0.07787	0.39801	0.04164	-0.05566	-0.18832
mixerror	-0.00767	-0.09229	0.36282	-0.04715	0.04079	-0.00053
divid near	0.06511	0.04025	0.25020	0.01236	-0.05352	0.08560
divid guide	-0.09792	0.07000	-0.10642	0.02131	0.00470	0.61771
corridor	-0.15158	-0.08227	0.28051	-0.04602	0.06858	0.00954
dirty	0.18782	0.65014	-0.11681	0.06514	0.03713	-0.01125
type	-0.00493	0.21797	-0.02045	-0.11606	0.14131	0.21555

4.2. Factor analysis

According to the questions in the questionnaire, this paper conducts factor analysis to form the variables that will be used for analysis. The fitting results are summarized in Table 3. After matrix rotation, six factors were screened out, namely factor1 to factor6. The meanings of each factor are summarized as follows:

- Factor1 The overall attitude of the subjects to garbage classification;
- Factor2 Improper phenomenon of garbage delivery point;
- Factor3 Relevant behaviors of throwing garbage by oneself and others around;
- Factor4 Understanding level of basic knowledge of waste classification;
- Factor5 Understanding level of advanced knowledge of waste classification;
- Factor6 Behavior of garbage sorting instructors.

4.3. Regression

The results of multiple linear regression including the correlation coefficients and significant level are shown in Table 4. This paper used the scores of garbage classification test, overall garbage classification demand, convenience demand, environment demand, economic demand, and people's attitude to garbage classification effect as the outcome variables. The results of regression analysis are presented in the Table 5. The plus sign here means this factor has positive correlation to the result variable, while the minus sign means they negative correlation.

Table 4. The results of multiple linear regression

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Test	Demand	Conv_D	Envrion_D	Econ_D	Effect
f1	0.192	32.293***	3.047	17.517***	11.729***	0.678***
	(1.20)	(6.21)	(0.86)	(6.37)	(5.27)	(7.05)
f2	-0.782***	-25.384***	-3.315	-8.534***	-13.535***	-0.894***
	(-5.19)	(-5.17)	(-0.98)	(-3.28)	(-6.43)	(-9.84)
f3	-0.344***	2.563	2.723**	-2.662	2.503	0.015
	(-2.37)	(0.54)	(0.84)	(-1.06)	(1.23)	(0.17)
f4	0.243	5.862	1.528	-0.255	4.589**	-0.198**
	(1.67)	(1.24)	(0.47)	(-0.10)	(2.26)	(-2.26)
f5	0.421***	1.140	-2.872*	5.963**	-1.951	0.334***
	(2.94)	(0.24)	(-1.90)	(2.41)	(-0.97)	(3.86)
f6	0.113	30.343***	5.202	7.120**	18.020***	0.179*
	(0.69)	(5.67)	(1.42)	(2.51)	(7.86)	(1.80)
gender	0.996***	4.007	-4.083	3.754	4.335	-0.098
	(2.82)	(0.35)	(-0.52)	(0.62)	(0.88)	(-0.46)
sh	-0.947***	-0.697	-3.185	2.171	0.317	0.183
	(-2.70)	(-0.06)	(-0.41)	(0.36)	(0.06)	(0.87)
age	-0.075**	0.966	0.268	0.507	0.191	0.020
	(-2.51)	(0.99)	(0.40)	(0.98)	(0.46)	(1.10)
time	0.000	-0.042	0.011	0.003	-0.056**	0.000
	(0.03)	(-0.72)	(0.28)	(0.09)	(-2.24)	(0.18)
Constant	9.789***	185.966***	76.226***	53.651***	56.089***	3.310***
	(11.46)	(6.67)	(3.99)	(3.63)	(4.70)	(6.42)
Observations	56	56	56	56	56	56
R-squared	0.617	0.761	0.183	0.677	0.800	0.834

*, **, *** represets for significant value at 90%, 95% and 99%.

The first column shows the regression results using the total demand for garbage classification as the result variable. In addition to the control variables, the significant items include factor1, factor2 and factor6. Combining the significance level and coefficient estimation symbols, it can be found that the more agreement with the rationality and necessity of waste classification, the better the

environment of the garbage classification sites in the community, and the more irresponsible the garbage classification instructors in the community, the higher the demand for garbage classification.

The second column shows the regression results using the environmental demand of waste classification as the result variable. In addition to the control variables, the significant items are factor1, factor2, factor5 and factor6. The results show that the more agreement with the reasonability and necessity of waste classification, the better the environment of waste classification sites in the community, the more likely know about advanced knowledge of waste classification, and the more irresponsible the waste classification instructors in the community are, the higher the demand for environmental protection of waste classification is.

The third column shows the regression results using the economic demand of waste classification as the outcome variable. In addition to the control variables, the significant items are factor1, factor2, factor4 and factor6. The results show that the more agreement with the rationality and necessity of waste classification, the better the environment of the garbage classification sites in the community, the more command about the basic knowledge of garbage classification, and the less responsible the garbage classification instructors in the community are, the more economic needs for garbage classification.

The last column shows the regression results using the overall garbage classification effect as the result variable. Except for the control variables, the significant items include all factors except factor3. The results show that the more agreement that garbage classification is reasonable and necessary, the better the environment of the garbage classification sites in the community, the less acquaintance about the basic knowledge of garbage classification, the more command about the advanced knowledge of garbage classification, and the more conscientious the garbage classification instructors in the community, the more that garbage classification is effective.

Table 5. The summary of the results of regression analysis

Result variable	Overall demand for garbage classification	Economy demand for garbage classification	Environment demand for garbage classification	garbage classification effect
significant factors	factor1(+) factor2(+) factor6(-)	factor1(+) factor2(+) factor4(+) factor6(-)	factor1(+) factor2(+) factor5(+) factor6(-)	factor1(+) factor2(+) factor4(-) factor5(+) factor6(+)

5. Discussion

According to the results of grouping t-test, no matter which grouping is carried out, there is no significant difference in the garbage classification effect (effectiveness and rationality), which indicates that the current garbage classification has not had an important impact on the attitude of the subjects. Guidance is quite effective. It is found if there is a responsible instructor at the point of taking out garbage, the effectiveness of the refuse classification will be much better and the people's attitude to it will also be more positive. Moreover, people's attitude will turn positive as the classification commences, and the effectiveness of refuse classification has little relationship with the subjects' understanding of garbage classification. Therefore, it can be concluded that when government decide to conduct the policy of refuse classification,there is no need to invest a lot of energy in the preparatory work to educate the people. Hence, they can just carry it out earlier and put more attention in the process. In addition, people have high demand for convenience and demand for low economy in refuse classification, which means they do not mind increasing the cost for waste classification and also shows that the subjects have a certain environmental awareness on the whole, and the most important thing is the convenience of the procedure.

6. Limitations & Future outlooks

This article only uses multiple linear regression for regression analysis. Thus, it cannot avoid some limitations.First of all, regression analysis is limited to revealing the linear relationship between variables. If the variables do not change significantly in a straight line, it is not suitable for regression

analysis unless it is transformed into a linear relationship such as the relationship between convenience demand and the six factors, the R^2 is only 0.183. Secondly, the regression equation may not be able to determine the causal relationship between the dependent variable and the independent variable. The causal relationship must be derived from previous experience and relevant professional knowledge. The mathematical equation is just a verification. Therefore, the dependent variable and independent variable should be determined by the researchers themselves, rather than relying solely on statistical tools. Third, the reliability of the regression equation for the prediction of dependent variables corresponding to independent variables beyond the given range is not high, that is, "extrapolation prediction". Fourth, this article's analysis based on 56 samples, sothe number of samples should be increased in order to find the relationship more accurately.

With the increasing complexity of environmental problems, the environmental data is also becoming more numerous and diverse. This requires more sophisticated, accurate and efficient analysis methods to assist the analysis of pollution sources: regression analysis methods still need to be improved and innovated in application to continuously meet the needs of environmental analysis. In the future, an analysis will be conducted based on more samples. Besides, more method including random forest, Lasso regression, Logical regression will be used in order to find a model with a higher degree of fit, which provides great convenience for the management and analysis of environmental pollution data. Seek the best way to effectively solve various environmental problems.

7. Conclusion

In summary, regression analysis mainly studies the relationship between objective things. It is a traditional scientific method with strong applicability and an important branch of modern statistics. This article analyzes significant factors in refuse classification based on correlation and multiple regression analysis, find significant factors and their partial regression coefficients. The results indicates: a guidance is a significant factor in all dependent variables, which means the classification effect of delivery points with staff guidance is significantly improved. The degree of understanding of garbage classification standards, and the subjects who have carried out garbage classification in their hometown know more; people focus on convenience demand more than economy demand in refuse classification. Though this article has some limitations, such as the number of subjects in this paper is small, only a multiple linear regression model is used, and the linear relationship between some dependent variables and independent variables is not significant, it can still reveal some the answers to some questions. This article finds out the factors that have a significant impact on the garbage classification effect and the public's attitude towards garbage classification, so as to help the government make better decisions in the process of implementing policies, rationally allocate human resources and economic resources, and obtain better classification results. Refuse classification is seemingly a small matter, but it is really a big thing that benefits the country and the people. Everyone can do it, starting from himself and every family, with a new look and good civilized habits, so as to protect the living environment and build a better home together. In the future, more methods and models will be used to offer an more accurate and reasonable results. Overall, these results offer a guideline for refuse classification.

References

[1] Cao C. Economic Analysis of Waste Classification. Green technology, 2022, 24 (4): 237-240245.

[2] Tencent New. I am the first to classify garbage, 2022.1.26, Tencent, Retrieved from: https://new.qq.com/rain/a/20220126a0bc6300.

[3] Haikou Daily: Tips on Domestic Waste Classification, 2021.11.22, Retrieved from: http://szb.hkwb.net/szb/html/2021-11/22/content_554319.html.

[4] Wang Zheng. Study on Countermeasures for Urban and Rural Waste Classification. Rural Economy and Technology, 2021, 32 (14): 30-32.

[5] Wei Zhai. Analysis and research on environmental monitoring data based on multiple linear regression. Microcomputer Application, 2022, 38 (3): 186-188, 192.

[6] Chen X. Analysis on influencing factors of air quality index based on linear regression model: taking Dazu District of Chongqing as an example. Environmental Impact Assessment, 2021, 43 (5): 79-82.

[7] Li C, Yu F, Liu J, et al. Comprehensive evaluation of water quality based on multivariate statistical analysis. Journal of Water Resources and Water Engineering, 2006, 17 (4): 36-40.

[8] Zhuang Z, Hu X, Ma X. Multivariate regression model and application of water environment carrying capacity and economic efficiency. Resources and Environment in Arid Areas, 2007, 2l (9): 4l-45.

[9] Zhao M, Li K, Zhang Z. Study on the relationship between red tide characteristic organics and red tide by multiple regression method. Journal of Sun Yat sen University, 2003, 42 (1): 35-38.

[10] Kong B. Discussion on waste classification. China Science and Technology, 2019 (19): 6-7.

[11] Garbage classification is imperative! Sohu, August, 2019, Retrieved from: https://www.sohu.com/a/332476564_120055797.

[12] James G, Witten D, Hastie T, et al. An introduction to statistical learning. New York: springer, 2013.