For office use only	Team Control Number	For office use only
T1 ________________	**2020012**	F1 ________________
T2 ________________		F2 ________________
T3 ________________	Problem Chosen	F3 ________________
T4 ________________	**A**	F4 ________________

2020
MCM/ICM
Summary Sheet

Analysis of charging cost based on ridge regression

Summary

In our campus and social life, we charge our electronic devices every day, and the use of electric energy is getting wider and wider. Charging ports are chargeable in some places, but many public places around the world offer "free" charging. Therefore, we discussed the division of power consumption responsibilities and cost savings.

For the first question, we used a non-seasonal time series method to predict the power consumption of public places and found that the power consumption in public places showed an increasing trend year by year and will continue this trend in the next three years. The increase in electricity will always be between 5% and 10%. Then we studied the impact and requirements of increased power consumption on public places, and conducted a path analysis of data collected in the past 20 years that affected ten aspects of public places and consumed power to screen out 3 variables. Finally, it is concluded that the correlation coefficient between GDP and power consumption is the largest. GDP will affect other factors and then affect changes in power consumption, and the increase in energy and charging demand requires GDP in public places. The number of people employed in places and the per capita consumption of public places increased accordingly.

For the two question, in order to build the cost model of power consumption, we selected the variables affected by power consumption in the first question. Considering that there may be a high degree of collinearity among the variables, ridge regression was used for function fitting. Finally, the cost equation Y = 0.003 * X1 + 0.1328 * X2 + 0.9699 * X3 is obtained for public places, where X1 is the gross output of public places, X2 is the number of employees in public places, and X3 is the consumption value of public places. The total degree is 99.89%, and the fitting degree is good. According to the equation, the extent of the cost is mainly provided by consumers, followed by public places, and finally subsidized by the state. Payment methods are mainly generated by consumer consumption.

For the three question, the model is applicable to all public places, and its independent variables are variables common to all industries. The model has a wide range of applications. We take the school as an example to improve the cost model, and change the dependent variable to the electricity consumption of education. The independent variable is education funding X1, the number of employees in the education industry X2, and the per capita investment in education X3. After fitting, the fitting equation is
Y = 0.0585 * X1 + 4.6444 * X2 + 1.3616 * X3-6343.0284. Finally, the model is tested for the degree of fitting, and the degree of fitting is 99.8%, which is good.

For the fourth question, in order to reduce the cost of increasing energy use in public places, we searched the relevant literature on energy saving in public places. According to the cost model of the second question, we found three measures to reduce the cost of energy without affecting productivity. Corresponding parameters, thereby reducing power consumption costs. The three specific measures are detailed in Question 4.

For the fifth question, we wrote an article for the school paper describing our findings and making corresponding suggestion.

At last, we evaluate the strengths and weaknesses of the models according to the background of the research problem.

Kew words: Path analysis, Ridge regression, Energy consumption, public places

Content

1. Introduction

1.1. Background

With the development of electronic equipment, the scope of electronic equipment involves small settings (mobile phones) and large settings (electric cars). Our life is inseparable from electronic devices. Electronic devices are charged every day, and the demand for electricity is increasing. Therefore, in public places, I continuously increase the number of power outlets, charging stations, and even electric cars from the parking lot. In some places, the charging interface needs to be charged, but many places around the world provide free charging, such as airports, train stations, libraries, shopping malls, and so on. But how will these places that provide free electricity affect our electronic equipment? Who will pay for the power consumption?

Energy is an important foundation for the survival and development of human beings. It is also the main driving force for the normal operation and development of society and economy. It is also the key to maintaining world peace and consolidating national security [1]. The increasing demand for electricity will affect the public areas, and how to solve the increase in electricity under the condition of reducing costs.

1.2. Literature review

In 2012, Adom et al. Based on demand-side management guidelines for two econometric methods (ARDL and PAM) and derived total power demand forecasts based on these econometric methods [1].

Comprehensive forecasting technology is the comprehensive application of multiple model methods organically. Compared with the characteristics of power grid prediction, there are many factors that affect the accuracy of the results. Therefore, the practical application effect of this method of this method is better, and it has gradually been improved It has become one of the most concerned topics in the field of prediction theory research [2].

Wolde-Rufael used the cointegration test and Granger causality test to study the long-term causality between per capita power consumption and GDP per capita in 17 African countries from 1971 to 2001 [3].

1.3. Overview of our work

In recent years, with the continuous growth of electric energy, the scope of power facilities has become wider and wider. This has provided us with many conveniences and also increased our demand for charging. What follows is a variety of charging facilities in public places, from mobile phones to charging cars, and most of these charging facilities are "free", which has caused a large amount of electricity and energy consumption. Cost allocation is a problem that has yet to be resolved. Therefore, we have studied the issue of power cost allocation and solved the following problems:

Question 1: We collect data on electricity use from the China Yearbook, observe and predict through time series models that the electricity consumption will increase year by year, and the consumption situation in the next three years will also increase. We used path analysis to determine the impact of increased charging demand and increased power consumption on public places. There are three main exemptions.

Question 2: According to the impact and requirements of the problem, we use ridge regression to build a model of the increased demand for public places and the cost of energy use. According to the model built, the degree of fees and payment methods are described.

Question 3: Based on question two, the model is improved. For different types of public places, we choose schools as an example to improve the demand for public places and the cost of energy use.

Question 4: Based on influencing factors and cost models, explore measures to reduce the cost of increasing energy use in public places, and further improve the model.

Question 5: We write a one-page article for your school newspaper describing our findings and recommendations.

2. Symbol description

symbol	symbol description
X1	is the gross output of public places,
X2	the number of employees in public places,
X3	the consumption value of public places
a	Electricity growth rate
k	ridge coefficient
y	Electricity consumption (100 million kilowatts)
$\hat{\rho}_k$	the correlation coefficient

Notes: Other symbolic descriptions are detailed in the text.

3. Assumptions and Justifications

Assumption 1: Assume that there is no major disaster in the year to which the data belongs, and ensure that economic data will not be abnormal.

Assumption 2: Assume that the power consumption data conforms to a normal distribution to ensure that time series predictions can be made.

4. Forecast of power consumption

4.1. Problem analysis

Regarding question one, we first collected data on power consumption at the National Bureau of Statistics and adopted a time series forecast method to predict the power consumption in the next three years. Ten data were collected, path analysis was performed using these data and electricity, and all data that did not meet the requirements were screened out. Finally, the impact on public places and three requirements were obtained.

4.2. Data collection

From the China Yearbook, we found the power consumption from 2000 to 2017, and collected the following data from the National Bureau of Statistics: total power consumption of the tertiary industry, production value of the tertiary industry, contribution rate of the tertiary industry, GDP , Education funding, employment in public places, electric power fixed asset investment, electric power investment, per capita public place consumption, 10 library-related building areas per 10,000 people have indicators related to public places.

4.3. Time series models

4.3.1. Theoretical basis for time series models

The time series model is a quantitative analysis method. The main content of the analysis is the statistical analysis of stationary random sequences. The ARIMA model is applied to d-order differential stationary sequence fitting.

4.3.2. Time series model

(1) Stationarity test

① Timing chart inspection

Mean of Stationary Time Series:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_t \tag{4-1}$$

Variance：

$$\hat{\gamma}_k = \frac{1}{n-k}\sum_{t=1}^{n-k}\left(X_t - \overline{X}\right)\left(X_{t+k} - \overline{X}\right) \tag{4-2}$$

As a constant property, the time series of a stationary sequence should show that the sequence always fluctuates randomly around a constant value, and the range of the fluctuation is bounded, without obvious trends and periodic characteristics.

② Autocorrelation diagram test

The stationary sequence has short-term correlation. It describes that as the number of delay periods increases, the autocorrelation coefficient of the stationary sequence will quickly decay to zero, that is, the correlation coefficient:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\sqrt{DX_t}\sqrt{DY_{t-k}}} \tag{4-3}$$

Short-term correlation.

Based on the sample data found, the power consumption is tested for stability. The timing diagram is as follows:

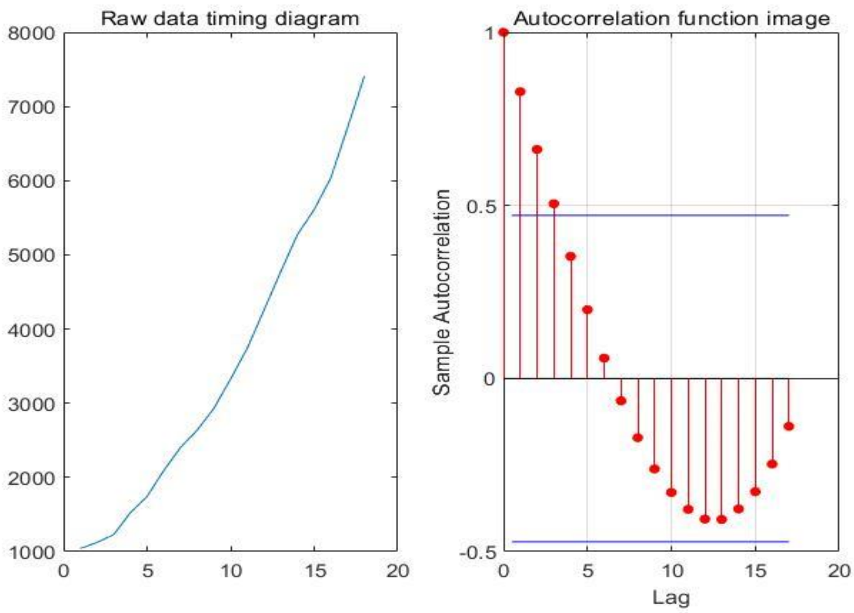

Figure 1: Timing diagram of raw data and autocorrelation function diagram

It can be seen from the timing diagram that the sequence has a significant trend, and it is initially judged as a non-stationary sequence. According to the autocorrelation function diagram, the autocorrelation coefficient has certain periodic characteristics, so the sequence is a non-stationary time series.

(2) Variance operation

The essence of the variance operation is to extract deterministic information by means of autoregression, and obtain a new sequence after performing first-order difference.

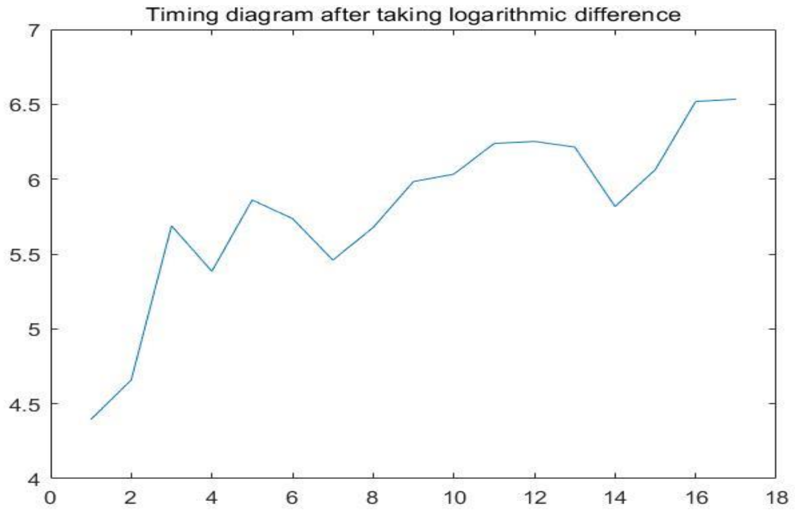

Figure 2: First-order variance square diagram

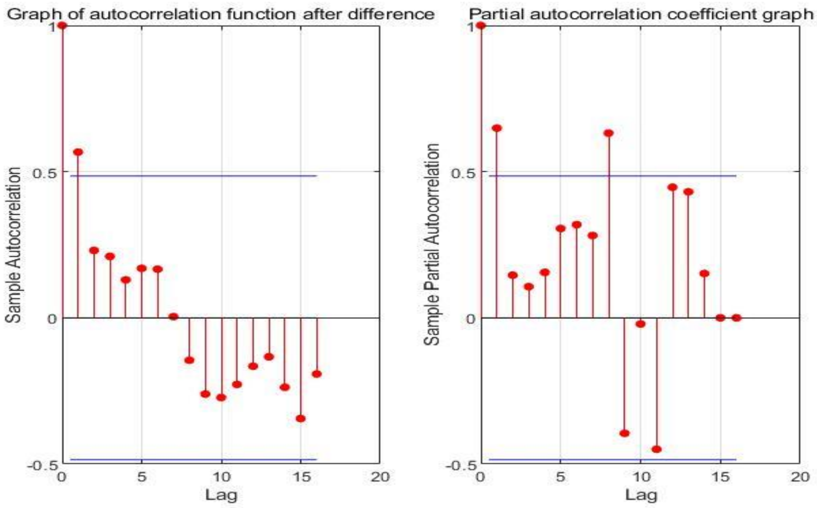

Figure 3: Autocorrelation and partial correlation coefficients after taking the log

It can be known from Figures 2 and 3 that the data fluctuates between certain values without obvious trends and periodicity. The sequence is stable after judging the first order variance.

(3) White noise test

In delay orders of 6 and 12 χ^2, the p-values of the statistics are 0.0994 and 0.0360, which are less than 0.05. Therefore, the post-variance sequence is considered to be a non-white noise sequence, and the post-variance sequence also contains relevant information that cannot be ignored.

(4) The order of the model is determined by the AIC criterion

According to the ARIMA (p, d, q) model, when p = 2 and q = 2, AIC = 5.9045 and BIC = 10.9038 are the minimum values, so the ARIMA model is used to fit the original sequence.

(5) Parameter estimation and testing

Under the 6th and 12th order delays, the p-values of the fitting statistics are 0.1349 and 0.2008, respectively, and both are significantly greater than 0.05. It can be considered that the residual sequence is a white noise sequence, and the coefficient significance test shows that both parameters are significant. This shows that the ARIMA model has a good fit to the sequence.

4.3.3. Forecast and result analysis

According to the data after the variance, the data of 2018-2020 is predicted from the data of power consumption in 2000-2017, and the specific results are as follows:

Table 2: Forecast value of power consumption in 2018-2020

years	2018	2019	2020
Forecast result (Billion kWh)	7850.1	8339.4	8850.5

A clearer and more intuitive expression of its change trend is shown in the following histogram:

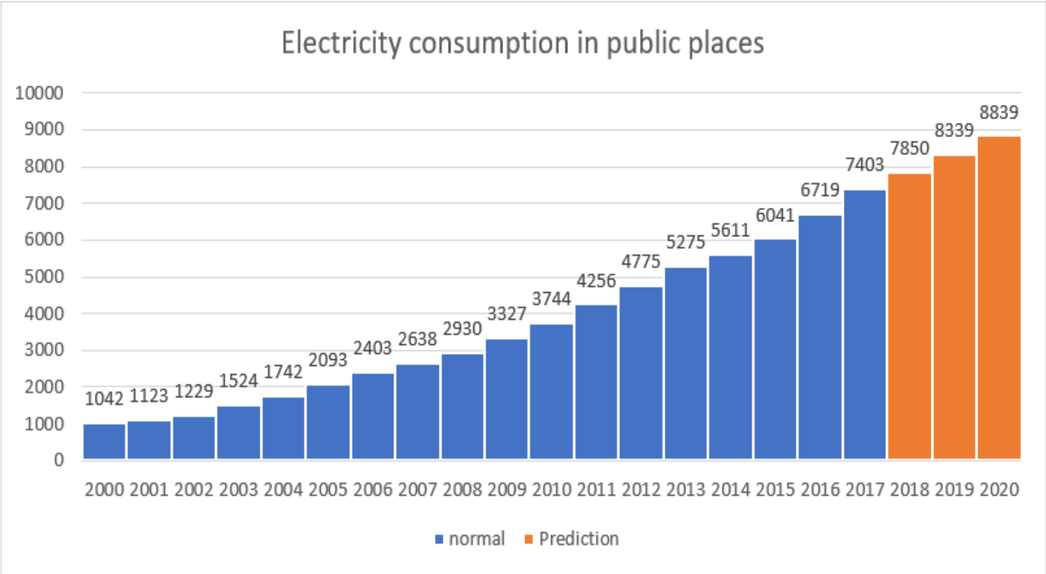

Figure 4: Consumption change from 2000 to 2020

As can be seen from the histogram, power consumption has increased year by year, and it is predicted that the power consumption in the next three years will also increase. People's demand for electricity is increasing, and the scope of use is wider. The change in power growth rate is shown in the following figure:

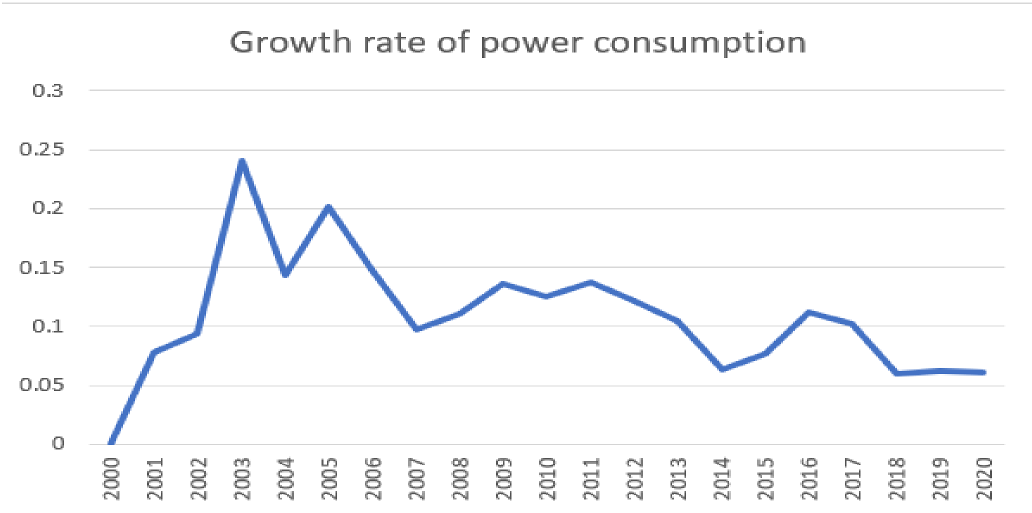

Figure 5: Changes in the growth rate of power consumption

In the next three years, from 2018 to 2020, the power growth rate will be between 5% and 10%. In 2003, the peak electricity consumption reached a peak of 24%.

4.4. Impact of increased power consumption on public places

According to the 10 related indicators that have an impact on public places, the data is gradually regressioned and path analysis screened to 3 impacts and demand, which are GDP, employment in the tertiary industry, and per capita cultural and educational entertainment for urban residents The correlations between the three consumers and the electricity consumption in public places were 0.9871, 0.9920, and 0.9981.

4.4.1. Path analysis

The path model is composed of a set of linear equations, reflecting the direct and indirect effects of each factor. It is a model based on multiple linear regression equations.

To analyze the impact of increasing electrical and charging needs on public places, conduct a path analysis based on the 10 indicators we selected. The specific steps are as follows:

$Step1$: Correlation coefficient:

$$r_{(m-1)m} = \frac{COV_{(m-1)m}}{S_{x_{m-1}} S_{x_m}} \tag{4-4}$$

Where r is the correlation coefficient between m-1 and m, and S is the variance of the independent variable.

$Step2$: If there is a linear relationship between related variables, the complex regression equation is:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_m x_m \tag{4-5}$$

and $x_1, x_2, x_3, \cdots, x_m$ is irrelevant, that is, $r_{12} \neq 0, r_{13} \neq 0, \cdots, r_{(m-1)m} \neq 0 (m \neq 0)$,

The sum of the determination coefficients of the result y plus the pairwise correlation causes the determination coefficient of the result y is equal to 1, that is,

$$\sum_{i=1}^{m} d_{y.i} + \sum_{i<j}^{m} d_{y.ij} = 1 \tag{4-6}$$

among them,

$$d_{y.i} = P_{y.i}^2, d_{y.ij} = 2 p_{y.i} p_{y.j} r_{ij} (i, j = 1,2,\cdots,m; i < j) \tag{4-7}$$

$Step3$: By solving the coefficient b through the system of equations, we can get a case of dependent variables y and m independent variables, then,

$$\begin{cases} r_{11}P_{y.1} + r_{12}P_{y.2} + r_{13}P_{y.3} + \cdots + r_{1m}P_{y.m} = r_{1y} \\ r_{12}P_{y.1} + r_{22}P_{y.2} + r_{23}P_{y.3} + \cdots + r_{2m}P_{y.m} = r_{2y} \\ \cdots\cdots \\ r_{m1}P_{y.1} + r_{m2}P_{y.2} + r_{m3}P_{y.3} + \cdots + r_{mm}P_{y.m} = r_{my} \end{cases} \tag{4-8}$$

4.4.2. Impact of increased electricity on public places

(1) Parameter estimation and testing

Table 3: Parameter estimation

Variable	estimate	standard error	t-value	p-value
Constant term	0.0000	0.0093	0.0000	1.0000
X1	0.5976	0.0792	7.5434	0.0000
X2	0.2100	0.0692	3.0322	0.0090
X3	0.1948	0.0878	2.2177	0.0436

As can be seen from the above table, the test probability of X1, X2 is less than 0.05, indicating that X1, X2 and y have a significant linear relationship, and the test probability of X3 is close to 0.05, indicating that X3 and y have a linear relationship.

(2) Analysis of variance

The purpose of the analysis of variance is to test whether the linear regression equation is significant. The MATLAB operation results are F = 3665.3455, p <0.0001, the multivariate determination coefficient is 0.9987, and the corrected multivariate coefficient is 0.9985. This shows that the linear regression equation is sought. Is very stable.

(3) Specific results and analysis

1) Preliminary results

Path coefficient is a statistic used to express the causality of related variables. It is a standardized partial regression coefficient, also called path weight. The path coefficient is generally estimated by the least square method (OLS) or the maximum likelihood estimation method (MLE). According to the partial regression coefficient with the path coefficient as the standard, the path coefficients of the independent and dependent variables can be obtained.

According to:

Standard partial regression coefficient = independent variable partial regression coefficient * (standard deviation of independent variable / standard deviation of dependent variable)

So the correlation coefficient between variables can be obtained, as shown in Table 4:

Table 4: Correlation coefficients between variables

	X1	X2	X3	Y
X1	1.0000	0.9871	0.9920	0.9981
X2	0.9871	1.0000	0.9895	0.9926
X3	0.9920	0.9895	1.0000	0.9954
Y	0.9981	0.9926	0.9954	1.0000

As can be seen from the table above, the effect of the respective variables on power consumption in order from large to small is:

GDP> per capita consumption in public place>Employment in public places

① The correlation coefficient between GDP and power consumption is the largest (R = 0.9981), and the direct path coefficient is also the largest ($P_{y.1}$ = 0.5976), reaching a significant level. Although the effect of GDP through other factors can not be ignored, the effect of GDP on power consumption mainly comes from itself. If GDP is increased by one standard deviation unit (S_{x_1} = 239150 billion yuan), power consumption can be increased by 0.9981 standard deviation units (S_y = 202.95 (billion kWh)).

② The correlation coefficient between employment in public places and power consumption is significant, although the direct path coefficient is close to a significant level, but The indirect benefit of GDP exceeds its direct path coefficient, which indicates that the effect of employment in public places on power consumption is due to the combined effect of itself and GDP.

③ The correlation coefficient between urban per capita cultural and educational expenses and power consumption is significant, and the direct path coefficient is close to a significant level, but the indirect impact of employment in public places and GDP exceeds its own impact on power consumption.

2) The final result

According to the relationship between charging demand in public places and GDP, employment in public places, and per capita consumption in public places, charge demand is used as the dependent variable, and the other three are independent variables, that is, changes in the other three drive charging. The change in demand also means that the requirement is a necessary condition. It is known that the demand for charging is increasing. Then the requirements for public places are that these three independent variables have changed. This is the demand for charging places for public places. They change, so charging needs can change. Then add the impact of energy and charging demand growth on public places on the basis of the first question, and also use the first ten data obtained, but this time as the dependent variable, energy and charging demand growth as independent variables, Path analysis is performed in turn. Only the public sector's gross domestic product has passed the test. The increase in energy and charging demand will also increase the public sector's gross domestic product.

Table 5: Electricity increase and total value

	Charging demand	Energy growth	Gross domestic product
Charging demand	1.4322	-0.4359	0.9963
Energy growth	1.4178	-0.4403	0.9775

4.4.3. Summary

Increased demand for energy and charging affects the gross domestic product; increased demand for energy and charging requires changes in the gross domestic product of public places, employment in public places, and per capita public place consumption.

5. Problem two Energy use and cost model

5.1. Problem Analysis

According to the results of question one, we can see that the three factors have a great correlation with power consumption. It can also be seen that only some of X2 and X3 are directly acting on the dependent variable, and there are also indirect factors acting on the dependent variable. Therefore, considering the co-linearity between independent variables, ridge regression was used for fitting. First, perform a collinearity test on the independent variable and the dependent variable. If collinearity exists, consider whether there is multicollinearity. If there is no multicollinearity, perform ridge regression fitting. Select the ridge parameters according to the image to obtain the ridge regression equation. Test the fitness of the ridge regression equation.

5.2. Ridge regression

Ridge regression is a biased estimation regression method used for collinear data analysis. It sacrifices unbiasedness in exchange for a substantial reduction in the variance part, and ultimately reduces the mean square error.

Suppose the dependent variable y is linearly related to the independent variable x. The n sets of data collected meet the following multiple linear regression matrix model:

$$\begin{cases} y = X\beta + \xi \\ \xi \sim N(0, \sigma^2 I_n) \end{cases} \tag{5-1}$$

among them, β_0, $\beta_1, \cdots$, β_p, σ^2 is an unknown parameter; β_0, $\beta_1, \cdots$, β_p is a regression coefficient; I_n is an identity matrix of order n.

When there is multicollinearity in the independent variable, a diagonal matrix is added $kI(k \geq 0)$, so some characteristic roots close to 0 will be improved and the original multicollinearity will be broken.

So defined:

$$\hat{\beta} = (x^T x + kI)^{-1} x^T Y \tag{5-2}$$

It is called ridge regression parameter estimation, where k is the ridge parameter. When k = 0, the ridge regression estimate is the ordinary least squares estimate. Because the ridge parameter k is not uniquely determined, the ridge regression estimate obtained is actually an estimate of the regression parameters.

5.3. Results analysis

Take the GDP, employment in public places, per capita consumption in public places as independent variables, and charging demand in public places as dependent variables.

The analysis of variance shows that F =3665.3455, p = 0<0.05, and the adjusted judgment coefficient is 0.9985, which indicates that the overall model fitting effect is better.

According to the parameter estimation results in Table 6, it can be seen that the regression coefficients have passed the t-value test, but the variance expansion factors of the variables are all greater than 10, indicating that there is colinearity between the independent variables.

Table 6: Parameter estimation

Variable	estimate	standard error	t-value	p-value	variance expansion factor
Constant term	-1718.5100	509.5016	-3.3729	0.0046	0.0000
X1	0.0051	0.0007	7.5434	0.0000	69.0921
X2	0.0910	0.0300	3.0322	0.0090	52.7864
X3	0.5983	0.2698	2.2177	0.0436	84.9624

The condition number is 19.7713, which is less than 30, indicating that there is no serious multicollinearity among the independent variables.

The ridge trace is drawn according to the ridge trace method. When the ridge parameter k changes within (0, + ∞), the ridge regression estimate $\hat{\beta}_i(k)$ of each independent variable x is a function of k, and $\hat{\beta}_i(k)$ is plotted on the plane coordinate system. The curve drawn is called a ridge trace. Draw the following ridge trace according to the relevant data:

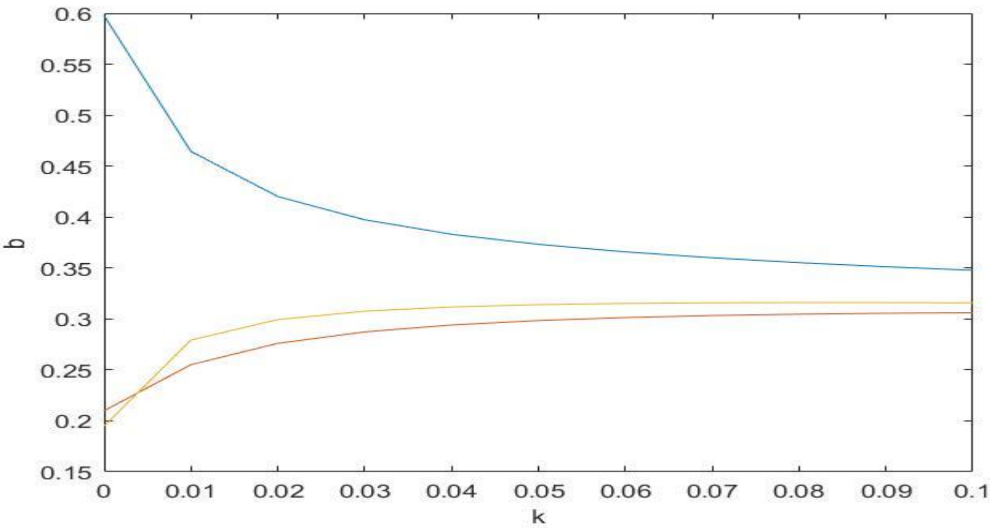

Figure 6: Ridge trace

It can be seen from the ridge trace graph that after k is greater than 0.02, the ridge trace curve of each regression coefficient tends to be stable, so the ridge parameter 0.1 is selected to establish the ridge regression equation.

The ridge regression equation with standardized data based on the obtained data is:

$$y^* = 0.4201x_1^* + 0.2761x_2^* + 0.2994x_3^*$$

Reducing the normalized variable to the original variable, the equation is:

$$y = -2396.652 + 0.0036x_1 + 0.1197x_2 + 0.9195x_3$$

It can be seen that the degree of cost is mainly provided by consumers, followed by stores, and finally the country. Payment methods are mainly generated by consumer consumption.

It can be seen that the degree of cost is mainly provided by consumers, followed by stores, and finally the country. Payment methods are mainly generated by consumer consumption.

6. Problem three ridge regression model improvement

6.1. Problem Analysis

For the third question, it is required to test the applicability of the model to different occasions. We adopt the method of modifying the variable data to reduce the independent variable range in the cost model from large to small, and transform the entire public place into data of specific locations. Therefore, we collected the school-related data again and fitted it with the second question method to obtain the main payer.

6.2. Model improvement

Because the use value of different places is different, for different places, the impact on power consumption value is different. Let's take the school as an example, and change the influencing factors into education funding, education employment, and per capita education investment.

6.3. Results analysis

We collected these data from the National Bureau of Statistics and put them into the model for testing, and performed the analysis of variance F = 1193.3740, p = 0, p value less than 0.05, and the adjusted judgment coefficient was 0.9953, which indicates the overall fitting effect of the model.

Table 7: Parameter estimation

Variable	estimate	standard error	t-value	p-value	variance expansion factor
Constant term	-8166.6822	2449.0723	-3.3346	0.0049	0.0000
X1	0.0498	0.0183	2.7238	0.0165	33.6363
X2	5.7361	1.6242	3.5316	0.0033	49.5682
X3	1.3610	0.5688	2.3926	0.0313	66.2566

It can be known from the parameter estimates in Table 7 of the Appendix that the variance expansion factors of the three independent variables are all greater than 10, indicating that there is colinearity between the independent variables. Therefore, the results of collinearity diagnosis are as follows:

Table 8: Colinearity diagnosis

Colinearity diagnosis	Serial number	Eigenvalues	X1	X2	X3
1	2.9686	1.0000	0.0033	0.0023	0.0017
2	0.0217	11.7056	0.8512	0.3146	0.0292
3	0.0097	17.4835	0.1455	0.6832	0.9691

As can be seen from Table 8, the condition index is 17.4835, which is less than 30, and there is no serious multicollinearity. Draw a ridge trace, as shown in Figure7:

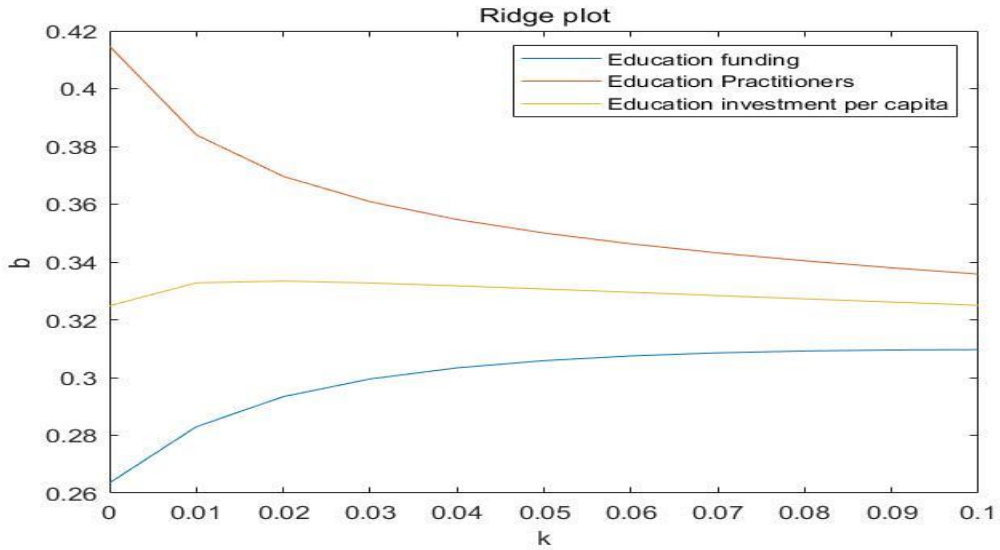

Figure 7: Ridge map of model improvement

As can be seen from the figure above, when k> 0.03, the ridge trace of each regression coefficient tends to be stable. Selecting k = 0.03 as the ridge parameter, the final fitting equation is:

$$y = -6973.4353 + 0.0566x_1 + 4.9908x_2 + 1.3940x_3$$

6.4. Model checking

The model is tested for the degree of fit, and the value of the degree of fit is 99.8%, so the degree of fit is good.

7. Question four Ways to Reduce Energy Costs

7.1. Problem Analysis

For the fourth question, through some measures to reduce the cost of energy use, we need to reduce the cost based on the model of question two. From the original cost model, we can know that the total value of public places 'production, the number of people employed in public places, There is a positive correlation between per capita consumption and cost. Therefore, consider taking measures in these three variables to reduce the increased cost of energy use in public places.

7.2. Analysis of factors affecting cost

(1) The gross output of public places in this model represents the basic electricity consumed by the operation of the facility

(2) The number of people employed in public places represents the power consumption caused by the work of employees in this model

(3) Per capita consumption in public places in this model represents the power consumption caused by consumers in the process of using the device to charge

7.3. Specific measures

Without affecting production, reduce the cost of increased charging of mobile devices, that is, reduce the parameters of per capita consumption in public places in the model, that is:

(1) Control the charging efficiency of the charging interface of mobile devices. Consider that in most coffee shops, airports, railway stations, shopping malls, consumers' main purpose is to receive services at the place, and charging is only a derivative of the service. Reduce the charging efficiency of the charging interface of the mobile device, and the time when the consumer arrives at the store does not change, so the cost of energy use is reduced. Therefore, the efficiency of the charging interface can be controlled to ensure that the consumer uses it

during the charging process As long as the mobile device does not reduce the power of the mobile device.

(2) Vigorously promote the use of new energy power generation equipment, such as solar power generation equipment, so that the cost of electricity consumption is reduced; after the implementation of this measure, the cost model will include the coverage factor of new energy equipment.

(3) Energy conservation education for consumers or charging measures for this equipment effectively reduces the frequency of consumers' use of free charging equipment.

8. Write an article for school

Who pays for "free" charging?

In recent years, with the continuous growth of electric energy, the scope of power facilities has become wider and wider. This has provided us with many conveniences and also caused some problems. Charging facilities in a variety of public places, from mobile phones to charging cars, continue to increase the consumption of electrical energy, and most of these charging facilities are "free," so who pays? Therefore, the cost allocation problem is a problem that has yet to be resolved. Therefore, we have studied the problem of power cost allocation and obtained the following conclusions and put forward the following suggestions.

First, we collected a large amount of data from the National Bureau of Statistics, and made predictions on power consumption based on past power consumption. It can be seen from the forecast chart that power consumption is increasing year by year, and the growth rate of power consumption has gradually stabilized between 5% and 10% in recent years. Then we explored the impact and requirements of electricity consumption on public places. We analyzed the electricity consumption after filtering out a large amount of data and irrelevant data, and obtained the factors that affect the electricity consumption. Employment in public places (X2) and per capita consumption in public places (X3). Who accounted for more of these factors?

We further processed these data and got the fitting equation:

$$y = -2396.652 + 0.0036x_1 + 0.1197x_2 + 0.9195x_3$$

It can be seen that consumers will pay most of the cost, so how do consumers pay for "free" facilities? We did the following discussion:

While charging the device, consumers are more likely to consume from nearby businesses, and the emergence of "free" charging facilities is more likely to cause consumers to go out longer and increase the possibility of consumption. This is also the cause of public place production One aspect of the increase in total value. So if it is impossible to consume around the charging device, then who pays? Therefore, we used the school as a specific object for further research. Using the cost model we constructed, we obtained the modified cost equation:

$$y = -6973.4353 + 0.0566x_1 + 4.9908x_2 + 1.3940x_3$$

Among them, X1 represents national education funding, X2 represents school employees, and X3 represents per capita education expenditure. It can be seen that schools will be the main payers, which is also consistent with our daily observations.

Studying this issue, we found another phenomenon, these "free" facilities are easy to cause waste of resources. For example, many people do n't shut down the charging facilities when they run out, or keep charging their equipment. This will undoubtedly waste a lot of electricity and increase the cost of electricity. So we made a few suggestions:

1. Control the charging efficiency of the mobile device's charging interface. Consider that in most coffee shops, airports, train stations, shopping malls, the main purpose of consumers is to receive services at the place. Charging is just a derivative of the service. Reduce the charging efficiency of the charging interface of the mobile device, and the time when the consumer arrives at the store does not change, so the cost of energy use is reduced. Therefore, the efficiency of the charging interface can be controlled to ensure that the

consumer uses it during the charging process As long as the mobile device does not reduce the power of the mobile device.

2. Vigorously promote the use of new energy power generation equipment, such as solar power generation equipment, which will reduce the cost of electricity consumption.

3. Energy conservation education for public place energy users or charging measures for this equipment will effectively reduce their use of "free" charging equipment.

9. Model extensions

Ridge regression is used to build a model for the increased demand for public places and the cost of energy use. By choosing a stable ridge parameter k, the problem of co-linearity among independent variables is solved. We can change the independent factors by different factors in different places. The choice of variables keeps the form of the equation unchanged, and obtains different cost equations.

10. The strengths and weaknesses of the mode

10.1. Strengths

1. Ridge regression analysis method to determine the determination of the k value of this subjectivity is helpful to achieve the organic combination of qualitative analysis and quantitative analysis.

2. Using the path analysis to decompose the correlation coefficient, we can see the direct and indirect effects of the independent variable and the dependent variable, and better the relationship between the influencing factors.

10.2. Weaknesses

1. The ridge regression is used to fit the cost equation. The selection of ridge parameters is subjective.
2. There are only 10 types of original data collection. There may be more factors but not collected, and the accuracy of the model can be improved again.

11. References

[1] Tian Yanran. Research on the Energy Consumption Forecast and Influencing Factors of Dingxi City [D]. Lanzhou University, 2019.
[2] Feng Hao. Design and implementation of power consumption prediction system based on .Net platform [D]. Chengdu University of Technology, 2014.
[3] Yang Chenghui, Yang Taihua. Co-integration Model Analysis of Electric Power Investment in the New Era [J]. Energy and Environment, 2019 (01): 2-4 + 9.
[4] Chen Lishun. Research on energy consumption behavior of urban residents [D]. Dalian University of Technology, 2009.
[5] Li Mingyang. Research on Energy Distribution of Public Buildings Based on Contract Energy Management [D]. Beijing Jiaotong University, 2019.
[6] Jose Eduardo, Jin Bei. Management of electric energy in hospital environment [J]. China Hospital Architecture and Equipment, 2004 (04): 42-43.
[7] Fan Siqi, Sun Jianting. Research on Energy Saving in Public Places [J]. China Collective Economy, 2011 (21): 183-184.
[8] Xu Yonghui, Chen Hailong, Kuang Jianchao. "Coupling Research and Forecast of Energy, Economy, Environment and Public Health in Western Regions [J]. Journal of Tianjin Institute of Business and Technology, 2019, 7 (02): 3-11.
[9] He Shu. Study on the layout optimization of public energy charging piles for new energy vehicles [D]. South China University of Technology, 2019.

[10] Cheng Wei, Li Changwu, Jiang Jianlin, Ni Jinchun. Standards lead informatization support to build a new model of energy resource measurement management for public institutions [J] .China Metrology, 2019 (04): 41-43 + 94.

12. Appendix

Appendix 1: Power consumption forecast

```
clc, clear
%%
x = [7407 6719 6041 5611 5275 4775 4256 3744 3327 2930 2638 2403 2093 1742 1524 1228.5 1123 1041.9] ';
x = flip (x);
figure (1)
subplot (1,2,1)
plot (x)
title ('Raw data timing diagram')
subplot (1,2,2)
autocorr (x)% draw autocorrelation coefficient plot
title ('Autocorrelation function image')
%%
x1 = diff (x)
x1 = log (x1);
figure (2)
plot (x1)
title ('Timing diagram after taking logarithmic difference')
figure (3)
subplot (1,2,1)
autocorr (x1)% draw autocorrelation coefficient
title ('Graph of autocorrelation function after difference')
set (gca, 'Fontname', 'Times newman', 'Fontsize', 8);
subplot (1,2,2)
parcorr (x1)% plots partial autocorrelation coefficients
title ('Partial autocorrelation coefficient graph')
set (gca, 'Fontname', 'Times newman', 'Fontsize', 8);
%%
[h1, pValue1, stat1, cValue1] = lbqtest (x1, 'lags', [6,12])% white noise test
%%
k = 0; n = length (x1);
for i = 0: 3
   for j = 0: 3
      k = k + 1;
      Mdl = arima (i, 0, j);% specifies the structure of the model
      [EstMdl1, EstParamCov1, logL (k), info1] = estimate (Mdl, x1, 'Display', 'off');
      % Parameter estimation and parameter testing
```

```
        numParams (k) = sum (any (EstParamCov1));% number of parameters to be evaluated
        [aic (k), bic (k)] = aicbic (logL (k), numParams (k), n);% Calculate AIC and BIC
        fprintf ('p =% 1d, q =% 1d, AIC =% 8.4f, BIC =% 8.4f \ n', i, j, aic (k), bic (k));
        % Show results for AIC and BIC
    end
end
fprintf ('\ n');
%%
p = input ('input order p =');
q = input ('input order q =');
Md2 = arima (p, 0, q);% specifies the structure of the model
[EstMdl2, EstParamCov2, logL2, info2] = estimate (Md2, x1);% parameter estimation and parameter test
%%
% Model test
res = infer (EstMdl2, x1);%
[h2, pValue2, stat2, cValue2] = lbqtest (res, 'lags', [6,12])% residual white noise test
%%
[dx_forecast, YMSE] = forecast (EstMdl2,3, 'Y0', x1);
dx_forecast1 = exp (dx_forecast);
x_forecast = x (end) + cumsum (dx_forecast1)% Calculate the 10-step prediction value of the original data
figure (4)
plot ([x; x_forecast]);
title ('2000-2020 power consumption');
x_z = [x; x_forecast] ';
zzl = zeros (1, length (x_z));
for i = 2: length (x_z)
zzl (i) = (x_z (i) -x_z (i-1)) / x_z (i-1);
end
figure (5)
plot (zzl);
title ('2000-2020 power consumption growth rate');
```

Appendix 2 : Path analysis of its impact

```
clc, clear
%%
[data, name] = xlsread ('Data.xlsx');
[n, m] = size (data); year = data (:, 1) ';
electric = data (:, 2);
others = data (:, 3: end);
stepwise (others, electric)
Others = [others, electric];
[table, Coeffients] = Diam (Others)
s = zeros (1,4);
for i = 1: 4
    s (i) = std (Others (:, i));% standard deviation
```

```
end
```

Appendix 3: Ridge regression model

```
clear all
clc
%%
[data, name] = xlsread ('Data.xlsx');
[n, m] = size (data); year = data (:, 1) ';
electric = data (:, 2);
others = data (:, 3: end);
%% Test collinearity Variance expansion factor greater than 10 is collinear
xishu1 = regression (electric, others);

%% Tests for multicollinearity. Condition index greater than 30 has multicollinearity.
xishu2 = gongxianxing (electric, others);

%% Perform Ridge Regression Ridge parameter selection 0.1 to get fitted equation
xishu3 = linghuigui (electric, others);

%% Tests the fitness of the fitted equation
xishu = xishu3 (2,2: end) ';
new_electric = others * xishu + xishu3 (2,1);
R = corrcoef (electric, new_electric)% is 99.89%
```

Appendix 4 : Improved Ridge Regression Model (Taking School as an Example)

```
clear
clc
%%
[data, name] = xlsread ('school.xlsx');
[n, m] = size (data); year = data (:, 1) ';
electric = data (:, 2);
others = data (:, 3: end);
%% Test collinearity Variance expansion factor greater than 10 is collinear
xishu1 = regression (electric, others);

%% Tests for multicollinearity. Condition index greater than 30 has multicollinearity.
xishu2 = gongxianxing (electric, others);

%% Perform Ridge Regression Ridge parameter selection 0.1 to get fitted equation
xishu3 = linghuigui (electric, others);

%% Tests the fitness of the fitted equation
xishu = xishu3 (2,2: end) ';
new_electric = others * xishu + xishu3 (2,1);
R = corrcoef (electric, new_electric)% is 99.89%
```