Team Number:	apmcm2309958
Problem Chosen:	C

2023 APMCM summary sheet

This study comprehensively investigates the new energy electric vehicle market in China and delves into six key questions, namely: question one explores the development factors, question two forecasts the development trend, question three analyzes the global impacts, question four assesses the policy impacts, question five simulates and regresses, and question six communicates the research results to the public.

For question one, we analyze the relationship between five main factors and the sales volume of new energy electric vehicles through **multiple linear regression**. The results show that the number of charging piles and energy efficiency have a critical impact on sales volume, providing key insights for marketing.

Problem two utilizes a **time series model** to predict sales volume over the next ten years, and the results show that the market for new energy electric vehicles will experience explosive growth.

Problem three explores the impact of new energy electric vehicles on the global conventional energy automotive industry through **gray correlation analysis**, and the results show a strong positive correlation between the two.

Question four evaluates the impact of the EU policy on China's new energy electric vehicle exports using a **difference-in-difference model**, and the results show that the implementation of the policy has imposed certain constraints on exports.

Question five analyzes the comprehensive impact of new energy electric vehicles in urban ecosystems through **simulation and regression model**, revealing their great potential for reducing carbon emissions.

In question six, we communicate the research results to the public, emphasizing the positive impacts of new energy electric vehicles, including improved air quality, reduced carbon emissions, and increased energy sustainability, and calling for active public support and participation.

By comprehensively answering these six questions, this study provides insights for understanding and promoting the sustainable development of China's new energy electric vehicle market.

Keywords: multiple linear regression, time series models, gray correlation analysis, difference-difference models, R language, Python and MATLAB.

Contents

I. Introduction

1.1 Background

New Energy Vehicles (NEVs) utilize advanced technological principles and structural features, using non-conventional automotive fuels as their power source, which has the advantages of high energy density and long cycle life. In recent years, driven by the rapid development of new energy vehicles, the global automotive industry is transforming towards sustainable transportation. China, as a key participant in the international automotive sector, has been at the forefront of these innovations.

1.2 question restatement

Question 1: Analyze the main factors affecting the development of new energy electric vehicles in China and explain the impact of these factors.

Question 2: Describe and predict the development of new energy electric vehicles in China in the next 10 years.

Question 3: Analyze the impact of new energy electric vehicles on the global traditional energy automobile industry.

Question 4: Analyze the impact of some countries' policies to resist the development of new energy electric vehicles in China on the development of new energy electric vehicles in China.

Question 5: With a city population of 1 million, analyze the ecological impact of electrifying the city with new energy electric vehicles.

Question 6: Based on the above analysis, write an open letter to the public.

II. The Description of the Problem

2.1 Question 1

Question 1 hopes that after understanding the main factors affecting the development of China's new energy electric vehicles, we can get the influence on the development of China's new energy electric vehicles through the establishment and analysis of mathematical models.

Based on the fact that the total annual sales of new energy electric vehicles in China is an important expression of the development of new energy electric vehicles in China, we choose it as an explanatory variable. Meanwhile, based on the great influence of per capita disposable income, the number of charging piles, the average oil price, the new Energy Vehicle Energy Efficiency, and the government subsidy budget on people's car purchases, we consider them as indicators and use them as quantitative influencing factors.

Using R language programming, we use the above data to establish a multiple linear regression model for the total annual sales of new energy electric vehicles in China, and use the T-test to test the significance of the data to derive the effect of the above factors on the development of the electric vehicle industry.

2.2 Question 2

Question 2 wants to analyze the data collected in the first question and forecast the sales of new energy electric vehicles in China in the next 10 years based on the first question.

We choose the two most influential factors derived from Question 1, charging piles number and New Energy Vehicle Energy Efficiency, to build and analyze a time series model to achieve the purpose of prediction.

Using the python programming, we use the above data to build a TSLM time series model of the total annual sales of new energy electric vehicles in China, and derive the forecast value of the sales of new energy vehicles in the next ten years.

2.3 Question 3

Question 3 wants to analyze the impact of new energy electric vehicles on the global traditional energy automobile industry with the help of a mathematical model.

Firstly, we collected the total annual sales of new energy electric vehicles and traditional energy vehicles in the world, and used them as the two variables of the inquiry. Matlab programming was used to build a gray correlation model to explore the degree of closeness between the two with the help of the calculation of correlation coefficient. After obtaining the degree of relationship, through SWOT analysis, we can comprehensively analyze the Strengths, Weaknesses, Opportunities and Threats of the traditional energy automobile industry after the development of new energy industry, thus, we obtain a comprehensive understanding of the problem.

2.4 Question 4

Question 4 seeks to analyze the impact of some countries' policies to resist the development of Chinese new energy electric vehicles on the development of China's new energy electric vehicle industry.

We take the European Union's anti-dumping policy on Chinese new energy vehicles as an example, and collect the number of Chinese new energy vehicles imported into Belgium, an EU member state, and use a double-difference model to evaluate the change in sales volume in the case of the policy occurring and not occurring, so as to explore the impact of other countries' policies on the development of China's new energy vehicle industry.

2.5 Question 5

Problem 5 wants to analyze the ecological impact of electrification of new energy electric vehicles (including electric buses) in a city with a population of 1 million.

We collect population data and data on the ownership of new energy vehicles and traditional fuel vehicles from more than 20 cities across China, fit them using a function, derive the most suitable fitting scheme with the help of correlation coefficients, and calculate the ownership of fuel vehicles and new energy vehicles in a city with a population of 1,000,000 people. Finally, it is assumed that all fuel vehicles are replaced by new energy vehicles to calculate the final carbon emission reduction.

III. Model Assumptions

1. It is assumed that per capita disposable income, the number of charging piles, the average oil price, the government's subsidy budget, and the energy efficiency of new energy vehicles jointly affect the sales of new energy EVs in China.

2. It is hypothesized that the TSLM time series model can better predict the future development of new energy EVs in China based on historical data.

3. It is assumed that the gray correlation analysis model will quantify the impact of new energy EVs on the traditional energy automobile industry.

4. Assume that Belgium's import and export data can describe the impact of the policy after its introduction, while Thailand's new energy import and export are not affected by it.

5. it is assumed that the cities analyzed for ecological impact are statistically similar to cities with a population of 1 million.

IV. Description of symbols

Symbols	Meaning
β_i	The regression coefficient is a parameter in a regression equation that indicates the magnitude of the effect of the independent variable (x) on the dependent variable (y).
σ	Represents the standard deviation of X_i.
Q	The minimum residual between the original data and the multiple linear regression model.
ϵ_t^2	The minimum residual between the original data and the TSLM model.
φ_i	The regression coefficient is a Coefficient of autoregressive model.
CV	Cross-validation statistics for TSLM models.
$\widehat{\sigma}_e^2$	Residual variance between predicted and actual values.
ρ	Resolution coefficient used to obtain grey correlation coefficient (generally 0.5).
$Y(x\_0(k), x\_i(k))$	The grey correlation coefficient is the value of the degree of correlation between the comparison series and the reference series at each moment (that is, at each point in the curve).
Y_{it}	China's electric vehicle export volume variable at different times.
Di	Policy grouping dummy variable (0 or1).
T_i	A dummy variable that marks whether it is subject to policy intervention.
δ, τ, β	DID model coefficients belonging to T_i Di and D_iT_i.
C, C_e, C_f	The carbon emissions from automobiles, carbon emissions from fuel vehicles, and carbon emissions from electric vehicles in the whole city.
O_f, O_e	The number of fuel vehicles in the whole city and the number of electric vehicles in the whole city.
ΔC	The amount of carbon emissions reduced by the hypothetical scenario.

V. Models and Solves

5.1 Question 1: Analysis of Development Factors

5.1.1 Modeling: multiple linear regression models

In order to initially understand the impact of each major factor on the development of new energy electric vehicles in China, we collect information from the following five aspects:

- **Per Capita Disposable Income:** Per capita disposable income is related to individual purchasing power. Higher per capita disposable income may encourage more people to consider purchasing new energy vehicles, thereby boosting market demand.
- **Charging Piles Number:** The number of charging piles is crucial for the convenience of electric vehicle usage. A higher number of charging piles helps address range anxiety issues, enhancing user acceptance of electric vehicles.
- **Average Oil Price:** The average oil price can impact the choice between traditional fuel vehicles and new energy electric vehicles. Higher oil prices may serve as a catalyst for the shift towards new energy vehicles, lowering overall vehicle operating costs.
- **Government Subsidy Budget:** Government subsidies serve as a policy tool to drive the development of the new energy vehicle market. A larger government subsidy budget is expected to stimulate market demand, reduce the cost of vehicle acquisition, and facilitate the widespread adoption of new energy vehicles.
- **New Energy Vehicle Energy Efficiency:** Energy efficiency in new energy vehicles directly influences their performance in real-world usage. Higher energy efficiency is anticipated to enhance user experience and contribute to an increased market acceptance of new energy vehicles.

We collected data from 2011 through 2022 in the table below:

Table 5-1-1 2011-2022 New Energy Vehicle Related Data

Year	sale volume	per capita disposable income	charging piles number	average oil price	government subsidy budget	New Energy Vehicle Energy Efficiency
2011	0.9	14551	0.68	7.63	0.7	4
2012	1.3	16510	1.17	7.83	1.3	4.1
2013	1.8	18311	2.3	7.73	2.1	4.2
2014	7.5	20167	3.96	7.51	4.2	4.5
2015	32.9	21966	5.78	6.07	28	4.8
2016	50.2	23821	20.4	5.95	213	5.1
2017	76.8	25974	44.6	6.42	376	5.4
2018	124.7	28228	77.7	7.22	471	5.7
2019	120.6	30733	121.9	6.81	274	6
2020	132.3	32189	168.1	5.71	153	6.3
2021	350.7	35128	261.7	6.83	195	6.6
2022	385.1	36883	521	7.61	199	6.9

Python programming is used to draw heat maps for preliminary data analysis, as in Fig. 5-1-1.

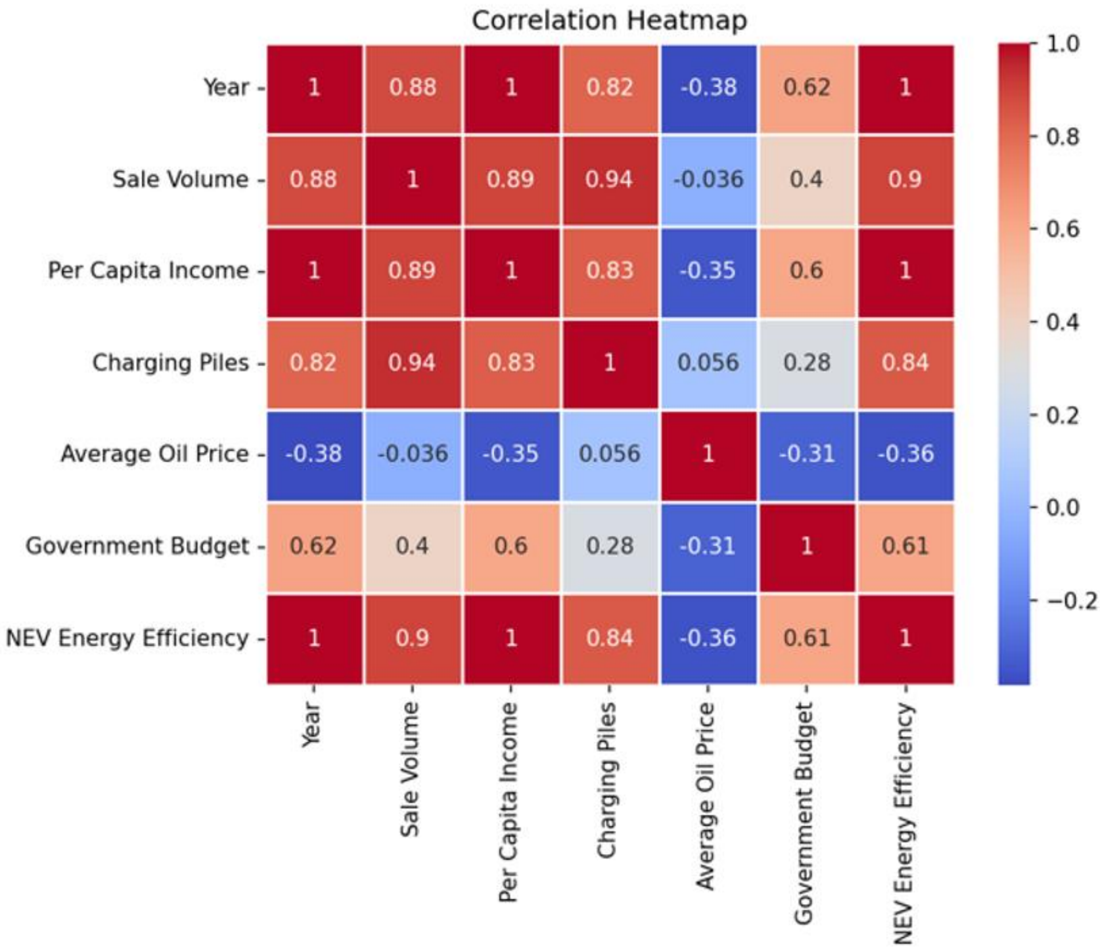

Figure 5-1-1 Development Factors Correlation Heatmap

Cleaning the data and dealing with missing values and outliers is necessary in order to study the main effects affecting the multiplicity performance more precisely. Therefore, we standardized the data to ensure consistent magnitudes:

$$X_i^{'} = \frac{X_i - \overline{X_i}}{\sigma}$$

Meanwhile, visualization is carried out through the R language, and the following Fig. 5-1-2 is obtained:

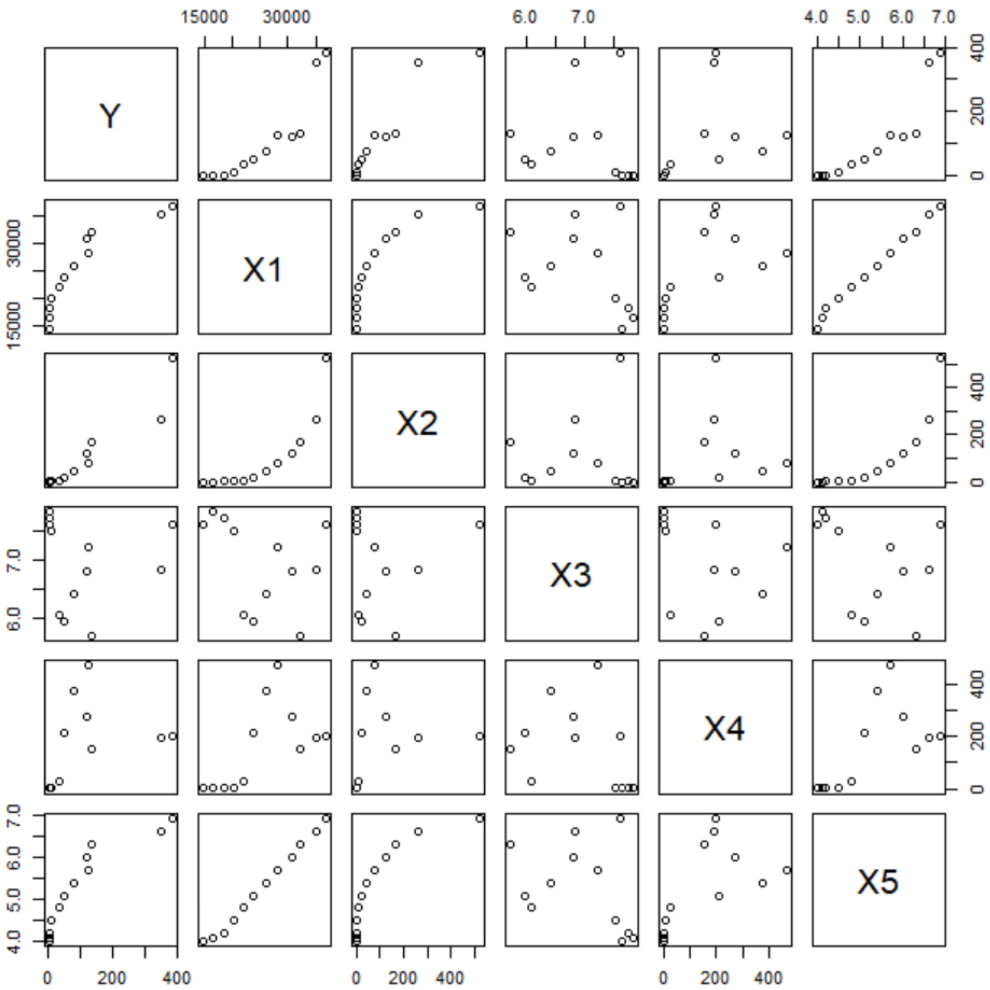

Figure 5-1-2 Multi-factor vs. total annual sales line chart

In order to solve for the influence relationship, we establish a multivariate linear regression model for the total annual sales of new energy electric vehicles in China. First, we construct the regression equation and list:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \epsilon$$

Based on the OLS equation:

$$Q = \min \sum e_i^2 = \min \sum (Y_i - \overline{Y_i})^2$$

Solve:

$$\frac{\partial Q}{\partial \beta_k} = \sum 2X_{ik}(Y_i - \beta_1 X_1 - \beta_2 X_2 - \cdots - \beta_k X_k) = 0$$

Solved in general matrix form:

$$\beta = (X'X)^{-1}X'Y$$

5.1.2 Model solving

Step1. Building Multiple Linear Regression Models: The R language provides a wide range of multiple linear regression analysis functions that allow for easy fitting and evaluation of multiple linear regression models. By using R for analysis, we can build appropriate multiple linear regression equations based on regression principles and OLS methods.

Step2. Data Processing and Integration: We create a table by collecting data to calculate the correlation coefficients between the independent and dependent variables in the form of sample data. The purpose of this step is to determine whether other indicators need to be included in the regression equation for the development of new energy electric vehicles in China. The data were cleaned and integrated in order to make the data preparation process simple and efficient.

Step3. Calculation and Interpretation of Statistical Results: Utilizing the calculation functions provided by the R language, we can obtain various statistical indicators of the regression results, including regression coefficients, goodness-of-fit, and standard errors. These results help to interpret and evaluate the multiple linear regression model.

Step4. Stepwise regression analysis: Stepwise regression is used to find the main influencing factors by constantly introducing more important variables into the regression or eliminating less important ones.

Step5. Storage of results and subsequent analysis: analyze the conclusions obtained by R language, derive the key factors and deduce their intrinsic connection.

5.1.3 Drawing conclusions

Analyzing the data through the R language, we derived the standard regression coefficients, integrated in the Figure 5-1-3:

	Non-standardized coefficient		standardized coefficient	t	P	VIF	R^2	Adjusted R^2	F
	B	standard error	Beta						
	0	0.106	-	0	1.000	-			
X1	0.102	1.51	0.102	0.068	0.948	202.821			
X3	0.143	0.191	0.143	0.748	0.483	3.256			
X2	0.431	0.398	0.431	1.084	0.320	14.066	0.933	0.876	F=16.587 P=0.002***
X4	-0.05	0.189	-0.05	-0.266	0.799	3.174			
X5	0.514	1.765	0.514	0.291	0.781	277.049			

Note: ***, **, * represent 1%, 5%, and 10% significance levels, respectively.

Figure 5-1-3: Plot of results of multiple linear regression analysis

The multiple regression model obtained is:

$$Y = -9.075e^{-17} + \ 1.023e^{-1}x_1 + 4.312e^{-1}x_2 + \ 1.432e^{-1}x_3 + \ -5.026e^{-2}x_4 + \ 5.135e^{-1}x_5$$

Meanwhile the analysis of the results of the F-test can be obtained, the significance p-value of 0.001864, which presents significance at the level and rejects the original hypothesis that the regression coefficient is 0. However, it is not possible to derive a significant situation for the independent variables because the p-value of each factor is too large (all greater than 0.05).

Impact Fac	p value
X1	0.068
X2	1.084
X3	0.748
X4	-0.266
X5	0.291

Figure 5-1-4: Significance level of each factor

We continue to use stepwise regression, where we keep introducing more important variables or eliminating unimportant ones to find the main influencing factors:

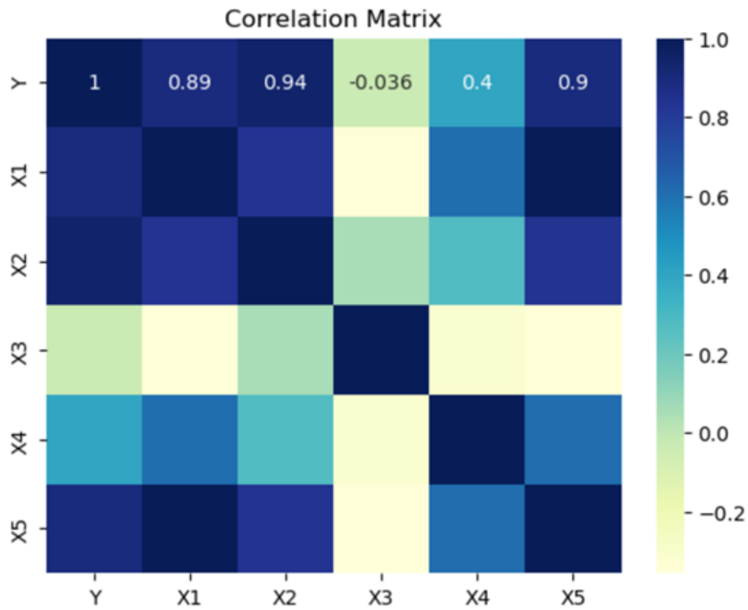

Figure 5-1-5: Results of stepwise regression analysis

By analyzing the data, we finally conclude that the number of charging stations and power consumption per kilometer are the key factors for the popularity of new energy electric vehicles. This suggests that these two factors largely influence the popularity of new energy electric vehicles. Specifically, a higher number of charging stations provides more convenient charging services for EVs, while lower power consumption per kilometer means higher energy efficiency and more economical cost of use.

These two factors have positively contributed to the development of new energy electric vehicles in China. The increase in the number of charging stations has increased user acceptance of new energy vehicles, solved the problem of charging convenience, and further boosted consumers' desire to purchase vehicles. At the same time, low power consumption means more environmentally friendly and economical vehicle operation, which is in line with society's pursuit of sustainable development and environmentally friendly travel.

5.2 Question 2: Development projections

5.2.1 Modeling: Time-series models

In order to predict the development of new energy electric vehicles in China in the next decade, we first need to analyze the development of new energy electric vehicles in the past decade. The first question has already provided a basic analysis of the various factors affecting the development in the past ten years. Based on this, the second question continues to use the data collected in the first question and retains the two most important factors for time series analysis.

By plotting a bump chart, the selected multiple categories are compared in successive dimensions and ranked according to the size of the data streams. Year was used as the x-axis, sale volume as size, charging piles number as ranking, and new energy vehicle energy efficiency as streams. as in Figure 3-2-1, we can visualize that each of the major factors increases as the and from this we can roughly infer that we can use the time series model to analyze.

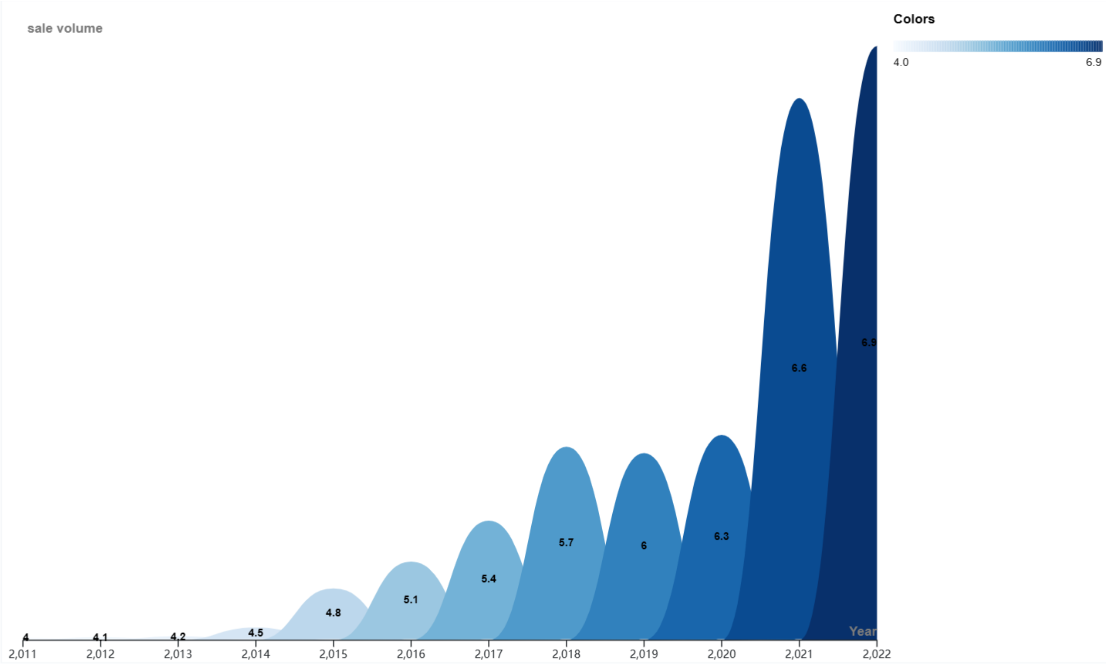

Figure 5-2-1: Bump chart of factors

Principle of TSLM Time Series Model:

The principle of the TSLM time series model is based on the Autoregressive Model, which assumes that observations at the current moment are correlated with observations at past moments. Specifically, the TSLM utilizes inputs from historical time steps as features to predict observations at the next time step. [1]

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 x_{2,t} + \varepsilon_t$$

Based on least squares estimation:

$$\sum_{t=1}^{T} \epsilon_t^2 = \sum_{t=1}^{T} \{(y_t\} - \beta_0 - \beta_1 X_{1,t} - \beta_2)^2$$

An autoregressive AR(p) is simultaneously fitted to the independent variables X1 and X2, and the model is constructed using the observations at the historical time step. The goal of the least squares method is to minimize the squared difference between the observations and the model predictions in order to obtain optimal model parameters.

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t$$

5.2.2 Model solving

The normal formulas indicate that the fitted values are calculated as follows:

$$\hat{y} = X\hat{\beta} = X(X'X)^{-1}X'y = Hy$$

Calculation of cross-tabulation statistics：

$$CV = \frac{1}{T}\sum_{t=1}^{T}\{[e_t\}/(1-h_t)]^2$$

Calculate the variance of the residuals:

$$\hat{\sigma}_e^2 = \frac{1}{T-k-1}(y - X\hat{\beta})'(y - X\hat{\beta})$$

Assuming that the errors are normally distributed, the 95% prediction interval is calculated as follows:

$$\hat{y} \pm 1.96\hat{\sigma}_e\sqrt{1 + x^*(X'X)^{-1}(x^*)'}$$

5.2.3 Drawing conclusions

By fitting the data through python programming, we generated the fitted data based on the TSLM model and obtained the fitting function as follows:

$$Y = 0.643X_{1,t} + 0.356X_{2,t}$$

In order to visualize the fitting effect, we made a comparison plot containing the original data and the fitted data, as shown in Figure 5-2-2, which shows that the fitting effect is very effective.

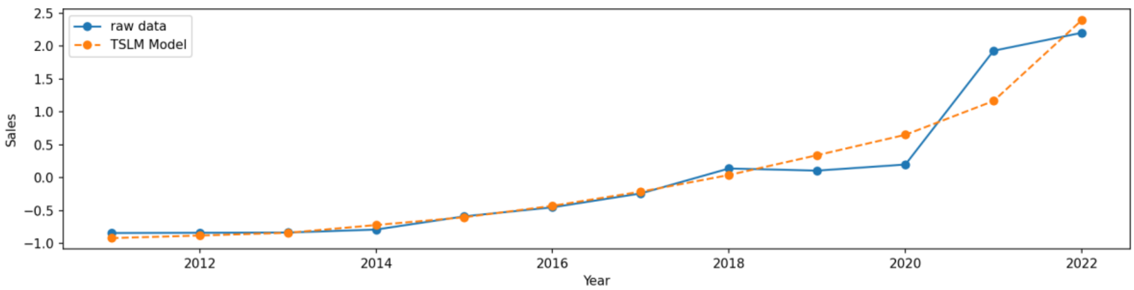

Figure 5-2-2 TSLM model and Actual data

Cross-validation:

The mean square error (MSE) for cross-validation is: 0.39120223361298332.

The mean square error is the average of the squares of the differences between the predicted and true values and measures the predictive accuracy of the model, with smaller values indicating that the model's predictions are closer to the true values.

Therefore, the mean square error of 0.3912 is relatively small, indicating that the model performs better in cross validation in terms of prediction.

prediction intervals:

```
[[0.82543829 3.30885801]
 [1.71637005 4.41565699]
 [2.5835069  5.54625087]]
```

As a result, we get a trend chart of projected EV vehicle sales values for the next ten years, as shown in Figure 5-2-3:

Figure 5-2-3 TSLM model fit curve to the original data

The line graph of the projected values for the next ten years visualizes that the production of electric vehicles in China will grow exponentially in the latter decade and will explode in 2030.

5.3 Question 3: Impact projections

5.3.1 Modeling: Grey correlation analysis models

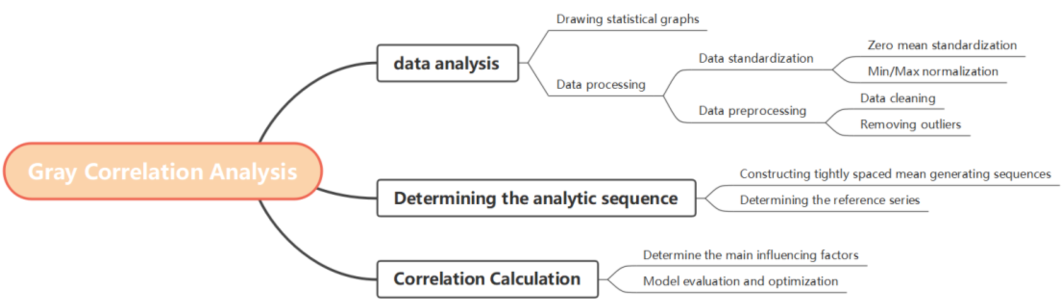

Figure 5-3-1 Gray Correlation Analysis Model

Data analysis:

From Global EV Data, collected by the International Energy Agency (IEA), we have data on global sales of new energy electric vehicles and conventional energy vehicles. Since all sales data are in units of (vehicle), we do not need to perform quantitative processing.

After preprocessing the data, plotting Parallel coordinates as shown in Figure 3-3-1, displaying multiple continuous time dimensions as axes, and connecting each row of new energy vehicles and fuel vehicles as lines on the axes to make the corresponding data more visualizable.

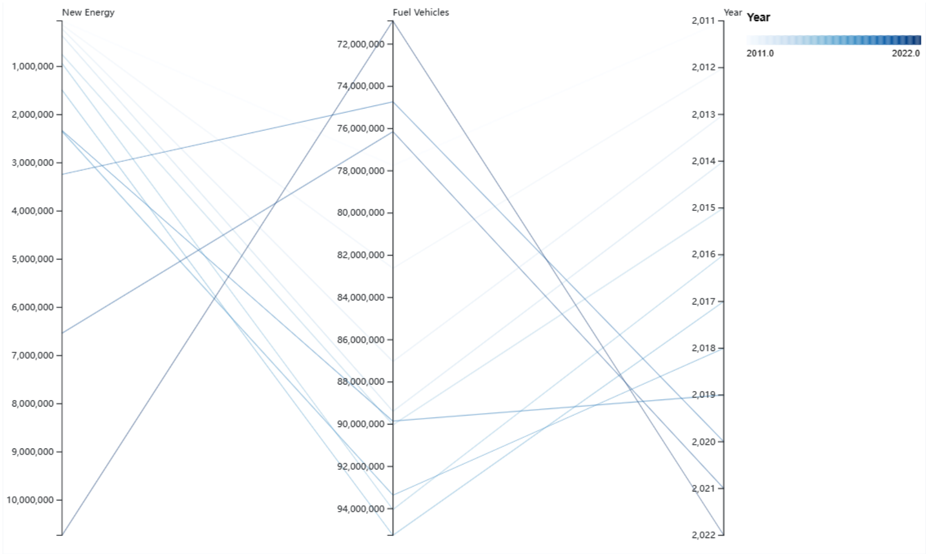

Figure 5-3-1 Parallel coordinates of sales volume

Determine the analysis sequence:

First of all, it is necessary to clearly define the reference series and the comparison series. In this study, we will take the sales volume of new energy electric vehicles as the independent variable to study its impact on the development of the global traditional energy automobile industry. As the reference sequence, we choose the global sales of traditional energy vehicles to reflect the overall development of the traditional energy vehicle industry. Since the dependent variable is time-varying, we select the year as the indicator for assessment, which is denoted by X0 and Xi (i =1, 2, ..., 20) are denoted. Separately, we use the sequences thus determined as follows: [2]

$$\begin{cases} parent\ series\text{: Global sales of conventional energy vehicles} \\ \text{Sub series: Global sales of new energy electric vehicles} \end{cases}$$

Comparing the child and parent series and combining the following formula, we can calculate the correlation coefficients of each indicator of the child series with the parent series

$$Parent\ Sequence\text{：}\ x_0 = [x_0(1), x_0(2), \cdots, x_0(n)]T$$

$$Subsequence\text{:：}\ x_1 = [x_1(1), x_1(2), \cdots, x_1(n)]T$$

In order to account for both the overall trend of the sequence and the characteristics of localized changes:

$$a = min_i min_k \left| x^0(k) - x_i(k) \right|$$

$$b = max_i max_k |x^0(k) - x_i(k)l|$$

With the above formula, we can calculate the two-level minimum difference (a) and the two-level maximum difference (b). By considering the maximum and minimum values together, we can more fully characterize the sequence and the trend of change, and thus accurately calculate the degree of correlation. Using this method, the relationship between the overall trend and local characteristics can be balanced to a certain extent, thus improving the accuracy and reliability of the gray correlation model. Therefore, we substitute it into the final correlation coefficient calculation formula:

$$Y((\mathrm{x_0(k), x_i(k)}) = \frac{a + \rho b}{|\mathrm{x_0(k) - x_i(k)}|}$$

Substitute the discrimination factor β = 0.5 to find the value of the correlation coefficient $Y(x_0(k), x_i(k))$.

5.3.2 Model solving

The GCC matrix is calculated by Matlab gray correlation analysis:

Table 5-3-1 GCC matrix

0.8593	0.8275	0.8007	0.7876	0.7862	0.7647	0.7611	0.7763	0.7965	0.9043	0.9189	1

The gray correlation is obtained by averaging the GCC matrix. With the help of Matlab calculations, we obtained that the gray correlation between the sales of new energy electric vehicles and the sales of traditional energy vehicles is 0.8319.

Strength of association: The correlation coefficient ranges from -1 to 1. A value of 1 indicates a perfect positive correlation, meaning that when one variable increases, the other increases in equal proportion. We arrive at a correlation coefficient of 0.8319 close to 1, indicating a strong positive correlation between global sales of conventional energy vehicles and sales of new energy electric vehicles.

On the premise that the correlation between global new energy electric vehicles and the traditional energy automotive industry is particularly strong, we continue to comprehensively analyze the impact of new energy electric vehicles through swot analysis.

3.3.3 Drawing conclusions

SWOT analysis results:

Table 5-3-2 SWOT Analysis

Strengths	Weaknesses
Proven Technology: The traditional energy automotive sector boasts extensive experience in reliable internal combustion engine technology and production processes. **Large Supply Chain:** Established traditional automakers benefit	**Environmental Concerns:** The rise of new energy vehicles has raised questions about the environmental friendliness of traditional automobiles, potentially leading to increased government regulations. **Oil Price Volatility:**

from a robust and stable supply chain network, enabling efficient large-scale production and cost reduction. **Brand Recognition:** Certain traditional automakers hold significant market share and widespread brand recognition, providing them a competitive edge in the market.	The traditional automotive industry is directly impacted by fluctuations in international oil prices, which may result in sales and profit instability. **Technology Upgrade Cycles:** Traditional automotive manufacturers may face periodic technology upgrades to meet market demands for new technologies.
Opportunities	**Threats**
Transition to New Energy: Traditional automakers have the opportunity to transition to the new energy vehicle market by developing technologies such as electric vehicles to meet market demands. **Integration of Smart Technologies:** Integrating smart technologies, such as autonomous driving and connectivity, can open up new markets for traditional automakers. **Integration of Renewable Energy:** Leveraging renewable energy sources, such as solar and wind power, offers traditional automakers new power alternatives.	**Intensified Competition:** The increasing competition in the new energy vehicle market may challenge the market share of the traditional automotive industry. **Government Policies:** Policies encouraging new energy vehicles by the government pose a potential threat to the traditional automotive industry, such as emission restrictions and subsidies. **Shift in Market Demand:** Growing consumer demands for environmental friendliness and new technologies could lead to a decline in market demand for traditional automobiles.

5.4 Question 4: Policy impact analysis

5.4.1 Modeling: Difference-in-differences models

Principles of double-difference modeling: [3]

The difference-in-differences model (DID has been used primarily in sociology to assess policy effects. The model if based on a counterfactual framework to assess the change in the observed factor y under two scenarios, one in which the policy occurs and one in which it does not. If an exogenous policy shock divides the sample into two groups - the Treat group, which is subject to the policy intervention, and the Control group, which is not subject to the policy intervention, and there is no significant difference in y between the Treat and Control groups before the policy shock, then we can view the Control group's change in y before and after the occurrence of the policy as the change in y of the Treat group that is not subject to the policy. change in the Control group before and after the policy occurs as the situation in the Treat group when it was not exposed to the policy shock (a counterfactual result). By comparing the change in y in the Treat group (D0) with the change in y in the Control group (D1), we can obtain the actual effect of the policy shock ((DD=D0-D1)

The specific formula is as follows:

$$Y_{it} = \alpha + \delta D_i + \tau T_i + \beta(D_i * T_i) + \epsilon_{it}$$

Y_{it} is the outcome variable, D_i is the policy group dummy, T_i *are* the policy time dummy, D_i, T_i the interaction terms, δ, τ and β are the coefficients before each term, ϵ_{it} is the random error term.

Calculate β:

$$\beta = [E(Y|D = 1, T = 1) - E(Y|D = 1, T = 0)] - [E(Y|D = 0, T = 1) - E(Y|D = 0, T = 0)]$$

Finally, the fitting parameters of the double difference method model, such as coefficients, standard errors, R^2, etc., were calculated to determine the effectiveness of the model fit.

5.4.2 Model solving

Collect policy data:

The European Union held a summit in Brussels at the end of June 2023 to discuss strategic adjustments to China, the European Commission insisted on "de-risking" - "economic security strategy", to reduce the dependence on China, from this meeting, Europe and the United States adopted a series of measures to resist the development of China's new energy electric vehicles, including "anti-dumping" policy.

Collect data on export volumes:

Collect data on China's exports of new energy electric vehicles to the EU member states of Belgium and Thailand, process the data, and visualize the data to plot the table below:

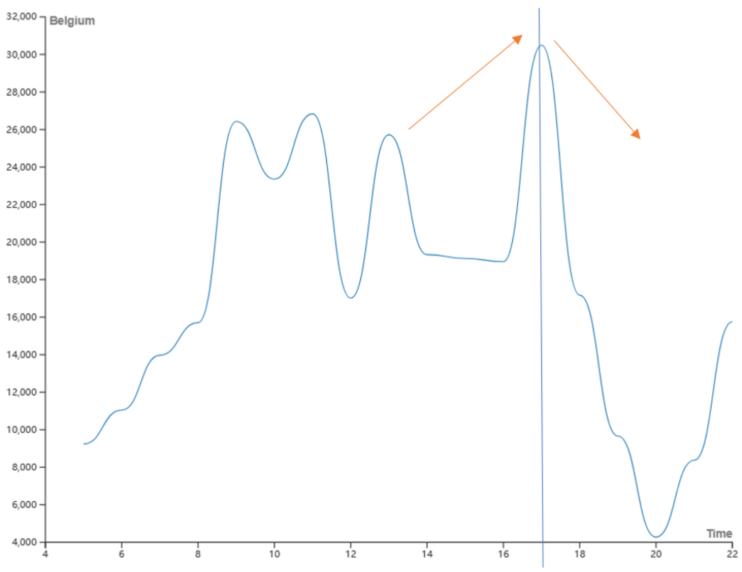

Figure 5-4-1 Line chart of export data

From the figure we can clearly see that the data was analyzed after the policy was enacted using SPSS to produce the results of the DID model.

5.4.3 Drawing conclusions

DID model results:

Table 5-4-1 DID model results

Results		Coefficient	Standard error	t	P
	Control	9823.692			
Before	Treated	19786.385			
	Diff（T-C）	9962.692	2105.517	4.732	0.000***
	Control	14039.6			
After	Treated	11047.8			
	Diff（T-C）	-2991.8	2257.915	-1.325	0.195
Diff-in-Diff		-12954.492	3994.937	-3.243	0.003***

Note: ***, **, * represent 1%, 5%, and 10% significance levels, respectively.

Visualization of the table contents, plotting Figure 5-4-2, provides a more intuitive view of the positive and negative effects of the policy.

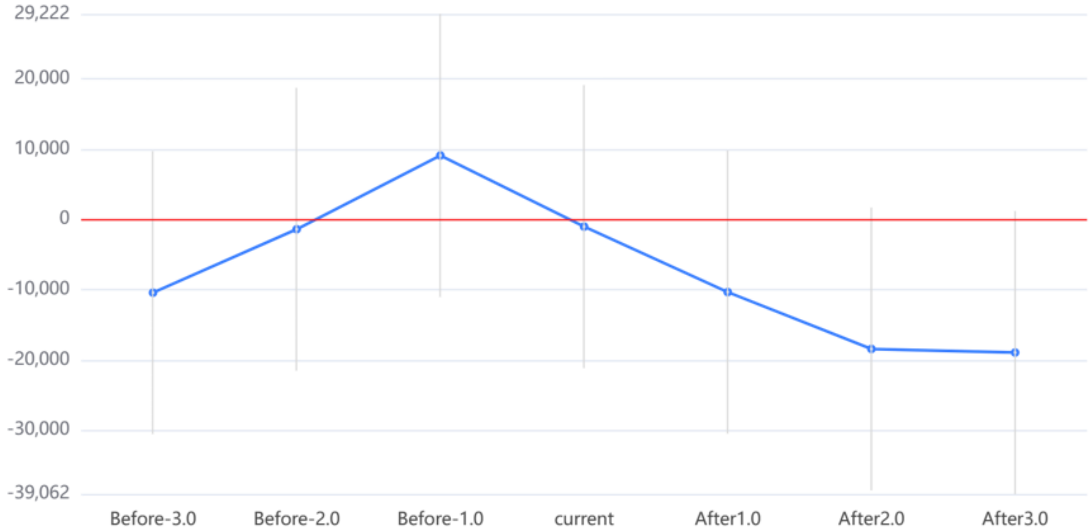

Figure 5-4-2 Map of policy effects

If the 95% confidence interval (the dashed part of the figure) contains a value of 0, it means that the difference between the treatment group and the control group is not obvious, and it is generally hoped that the situation occurs before the event (Before). If the 95% confidence interval is exceeded, it means that the policy effect is obvious, and it is generally hoped that the situation occurs after the event (After), and its above and below the value of 0 represents the positive and negative effects of the policy.

From the figure it can be seen that the event occurred after the March dotted line does not contain the value of 0, and below the value of 0, it can be obtained that the policy for China's exports of new energy vehicles to Belgium has an inhibitory effect, which leads to the conclusion that this policy for the development of China's new energy automobile industry has a negative impact.

The findings suggest a substantial influence of the EU's measures on curbing the export volume of Chinese new energy vehicles to Belgium, an EU member state. This underscores the significance of international policies in shaping the trajectory of China's new energy electric vehicle industry.

While the EU's anti-dumping policy appears to have a restraining effect on Chinese exports, it is crucial to acknowledge the resilience and adaptability of China's new energy vehicle sector. Despite facing impediments in certain markets, the industry as a whole continues to exhibit positive growth trends. This resilience is attributed to the diversified partnerships and collaborations that China has established with numerous other countries, contributing to the overall development and expansion of the new energy vehicle industry.

5.5 Question 5: Simulation and regression models

5.5.1 Modeling:

Building upon the foundation laid out in the China Automotive Low Carbon Action Plan 2022, our analysis is rooted in the comprehensive Life Cycle Assessment (LCA) methodology. This approach utilizes the China Automotive Life Cycle Database (CALCD) alongside the Automotive Life Cycle Assessment Model (CALCM) and the Automotive Life Cycle Assessment Tool (OBS). The study is further reinforced by data integration with the China Industrial Carbon Emission Information System (CICES).

The research and analysis, we are able to get the full life cycle carbon emissions of gasoline-fueled vehicles and electric vehicles in 2021 are $257.3gCO_2e/t*km$ and $131.8gCO_2e/t*km$, respectively, so we can simply consider the relationship between the carbon content of the whole city's cars and the number of fuel-fueled vehicles with the number of electric vehicles as: [4] [5]

$$C = C_f + C_e$$
$$C_f = 257.3O_f$$
$$C_e = 131.8O_e$$

Here, C, C_f, and C_e represent the city-wide carbon emissions from automobiles, fuel vehicles, and electric vehicles, respectively. Additionally, O_f and O_e signify the number of fuel vehicles and electric vehicles in the entire city, respectively.

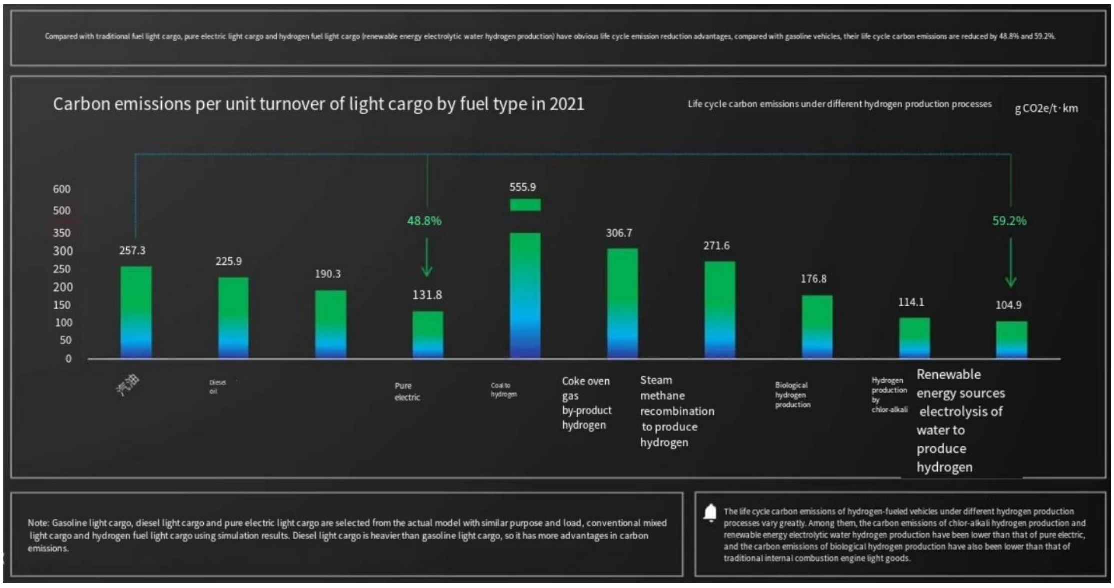

Figure 5-5-1 China Automotive Low Carbon Action Plan 2022

In Figure 5-5-1, we illustrate the China Automotive Low Carbon Action Plan 2022, emphasizing the considerable reduction (48.8%) in the entire life cycle carbon emissions of electric cars compared to gasoline cars. This substantial advantage positions electric vehicles as ecologically superior in terms of carbon balance.

Consequently, when evaluating the ecological impact of electric vehicles in a city with a population of 1 million, our focus is on the carbon emissions saved by replacing all fuel vehicles with electric vehicles, serving as a pivotal indicator of impact.

5.5.2 Model solving

We firstly collect the population data of more than 20 cities in China as well as the ownership of new energy vehicles and traditional fuel vehicles, and plot the graphs to find out the relationship between the ownership of new energy vehicles and the population:

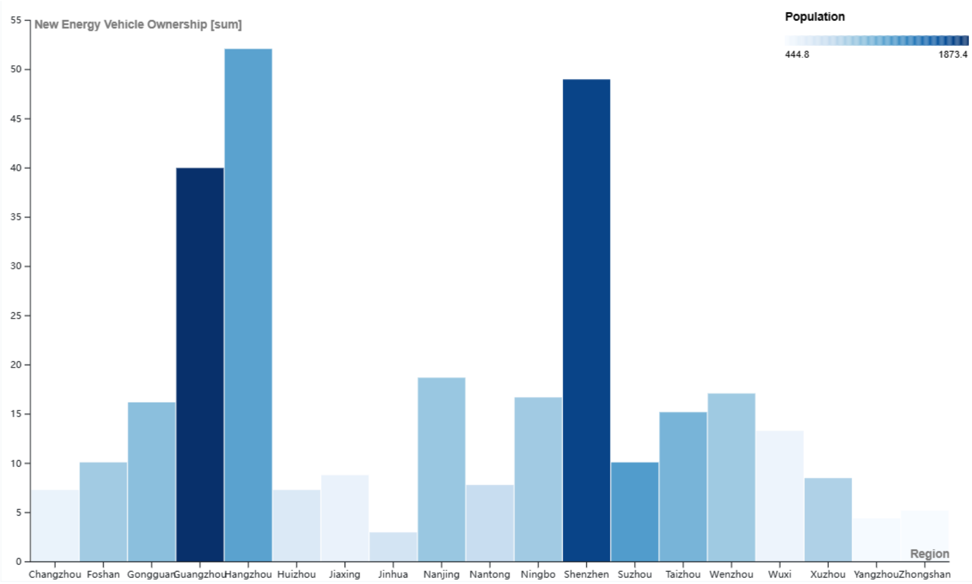

Figure 5-5-2 New energy vehicles ownership and population's connection

Also, we plot the correspondence between conventional car ownership and population:

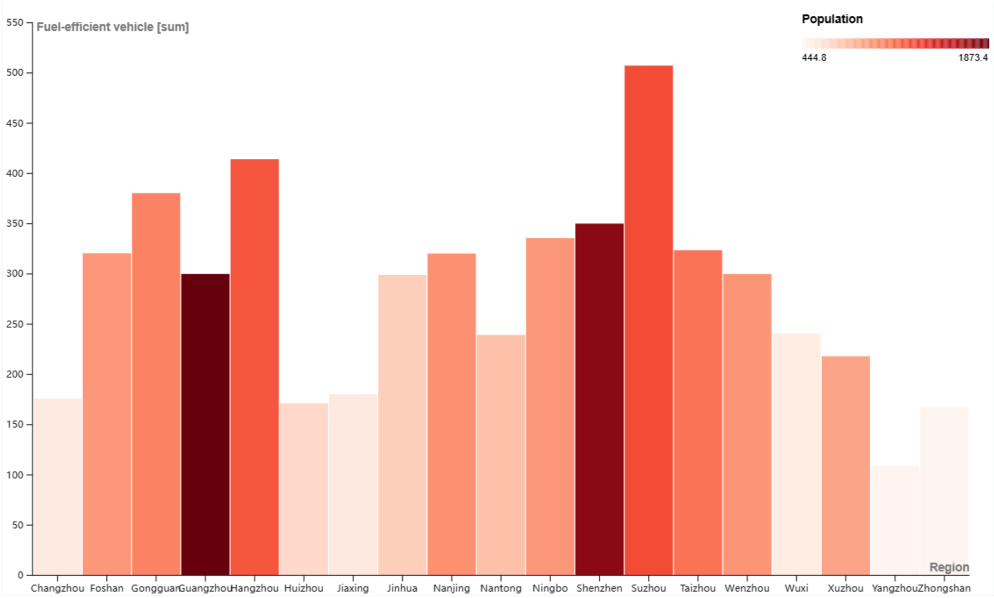

Figure 5-5-3 Conventional car ownership and population's connection

In the preliminary analysis of the graphs, we found that the population data of Shenzhen and Guangzhou are too large, which is too different from the predicted population of 1 million, so we exclude these two points as outliers and do not carry out the subsequent calculations.

We analyzed these data after processing them and fitted the ownership of both types of vehicles to the population of several Chinese cities in 2021 with the help of excel fitting function, and the fitting process is as follows:

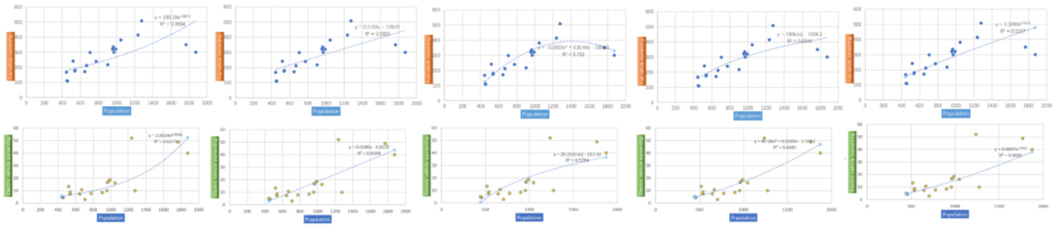

Figure 5-5-4 Fitting process diagram

5.5.3 Drawing conclusions

We selected the model with the best goodness-of-fit as the final model:

$$O_f = 78.602e^{0.0014x}$$

$$O_e = 0.0007n^{1.4432}$$

Bringing the data in, we end up with:

$$\Delta C = C - 131.8O_e = 89.87493373$$

Therefore, the total carbon reduction is about 89.87 when fuel vehicles are fully replaced by electric vehicles.

Since reducing carbon content to reduce the urban greenhouse effect is of great significance, we can conclude that the electrification of urban new energy electric vehicles (including electric buses) has a great impact on the ecological environment, specifically in terms of the reduction of carbon content.

5.6 Question 6: Open Letter to the Public

Dear Citizens:

I am writing to share with you an important finding about the far-reaching impact of new urban energy electric vehicles on our ecosystem.

Through in-depth research and data analysis, we have selected the best-fit model with the following findings:

Final model:

$$O_f = 78.602e^{0.0014x}$$

Expected value model:

$$O_e = 0.0007n^{1.4432}$$

After substituting the data into the model calculations, we get the following:

$$\Delta C = C - 131.8O_e = 89.87493373$$

In other words, when traditional fuel vehicles are completely replaced by electric vehicles, the total carbon emission reduction is about 89.87. This is an encouraging figure that reflects the significant effect of new urban energy electric vehicles (including electric buses) in reducing carbon emissions.

Reducing carbon emissions to minimize the urban greenhouse effect is of great significance to our urban ecosystem. Therefore, we conclude that the electrification of urban new energy electric vehicles has had a great impact on the ecological environment, especially in terms of carbon emission reduction.

It is in this context that I would like to emphasize to you the multiple benefits of new energy electric vehicles and the positive environmental and social contributions of the electric vehicle industry around the world.

First of all, new energy electric vehicles have a significant role to play in improving urban air quality due to their low-pollution and low-energy consumption characteristics. By reducing tailpipe emissions, we can effectively protect the air around us and create a fresher air environment for our homes.

Secondly, the promotion of new energy electric vehicles will help to reduce the dependence on limited energy resources and leave a more sustainable ecological environment for our future generations. This is not only about our own quality of life, but also a precious gift to future generations.

Finally, as a global trend, the new energy electric vehicle industry is becoming a new engine for the national economy. Through technological innovation and industrial upgrading, our country has become one of the world's most recognized leaders in the field of new energy electric vehicles.

Therefore, I sincerely invite you to join us in supporting the promotion and application of new energy electric vehicles. Through everyone's efforts, we can collectively contribute to the creation of a cleaner and more livable urban environment.

Thank you for your time and let's work together for a better future!

VI. Future Work

1.Enhancing Predictive Models:

For Question 1, our future work involves refining the multiple linear regression model used to analyze the main factors affecting the annual sales of new energy electric vehicles in China. We will explore additional variables that might influence the industry and conduct a thorough sensitivity analysis to identify potential factors not considered initially. This will contribute to a more comprehensive understanding of the dynamics influencing the development of the electric vehicle sector.

2. Improving Forecasting Accuracy:

Building on the outcomes of Question 2, our next step is to enhance the time series model for predicting the sales of new energy electric vehicles in China over the next decade. We will investigate advanced forecasting techniques and consider incorporating more variables for a more accurate prediction. Continuous refinement of the model will ensure its reliability and usefulness for stakeholders in the electric vehicle industry.

3. Global Industry Impact Analysis:

Question 3 aims to analyze the impact of new energy electric vehicles on the global traditional energy automobile industry. Future work involves expanding the scope of the analysis by incorporating additional variables and considering regional variations. Moreover, we plan to delve deeper into the SWOT analysis to provide more nuanced insights into the opportunities and challenges faced by the traditional energy automobile industry globally.

4. Policy Impact Assessment:

Building on the findings of Question 4, we intend to broaden our analysis of the impact of international policies on China's new energy electric vehicle industry. This includes extending our investigation to encompass a broader range of countries and policies. By incorporating a comparative analysis, we aim to provide a more comprehensive understanding of how different policy interventions influence the development trajectory of China's electric vehicle industry.

5. Comprehensive Ecological Impact Study:

Looking ahead to Question 5, our future work involves expanding the ecological impact analysis of the electrification of new energy electric vehicles in urban settings. We plan to include more cities in our dataset, ensuring a representative sample, and refining our methodology for assessing carbon emissions reduction. Additionally, we will explore other environmental indicators to offer a holistic understanding of the ecological benefits associated with the widespread adoption of new energy electric vehicles in urban environments.

Continuing to advance our research along these lines will not only deepen our understanding of the various facets influencing the new energy electric vehicle industry in China but also contribute valuable insights for policymakers, industry stakeholders, and the general public. The evolving landscape of the electric vehicle sector demands ongoing research and analysis to inform sustainable and forward-thinking strategies.

VII. References

[1] Sean J. Taylor and Benjamin Letham (2018) Forecasting at Scale, The American Statistician, 72:1, 37-45, DOI: 10.1080/00031305.2017.1380080
[2] hou, M., & Dong, W. (2023). Grey correlation analysis of macro- and micro-scale properties of aeolian sand concrete under the salt freezing effect. Structures, 58, 105551. doi. 10.1016/j.istroc.2023.105551.
[3] Wenli Mao,Junxin Yan,Jianming Sun. Has the trade war promoted the return of U.S. manufacturing industry--an empirical study based on double difference analysis[J]. Northern Economy and Trade,2023(09):20-25+29.
[4] Li Chunbo. Theoretical research and application of carbon emission flow diagram for electric vehicles[D]. Hangzhou University of Electronic Science and Technology,2023.DOI:10.27075/d.cnki.ghzdc.2023.001601.
[5] Cheng Gang. Analysis of Influencing Factors of Automobile Ownership [D]. Hebei University of Economics and Trade,2021.DOI:10.27106/d.cnki.ghbju.2021.000410.

VIII. Appendix

Draw heat maps by python

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# data
data = {
    'Year': [2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020,
2021, 2022],
    'Sale Volume': [0.9, 1.3, 1.8, 7.5, 32.9, 50.2, 76.8, 124.7, 120.6,
132.3, 350.7, 385.1],
    'Per Capita Income': [14551, 16510, 18311, 20167, 21966, 23821,
25974, 28228, 30733, 32189, 35128, 36883],
    'Charging Piles': [0.68, 1.17, 2.3, 3.96, 5.78, 20.4, 44.6, 77.7,
121.9, 168.1, 261.7, 521],
    'Average Oil Price': [7.63, 7.83, 7.73, 7.51, 6.07, 5.95, 6.42,
7.22, 6.81, 5.71, 6.83, 7.61],
    'Government Budget': [0.7, 1.3, 2.1, 4.2, 27.8, 212.7, 375.6, 471.1,
274.3, 152.9, 194.8, 199],
    'NEV Energy Efficiency': [4, 4.1, 4.2, 4.5, 4.8, 5.1, 5.4, 5.7, 6,
6.3, 6.6, 6.9]
}

# Create DataFrame
df = pd.DataFrame(data)

# Scatter matrix
sns.pairplot(df)
plt.show()

# Thermal map
correlation_matrix = df.corr()
plt.figure(figsize=(10, 8))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm',
linewidths=.5)
plt.title('Correlation Heatmap')
plt.show()
```

R language

```
# Load the required package
```

```
library(readxl)

# Read Excel data
data <- read_excel("C:/Users/....xlsx")

# De-dimensionalize
scaled_data <- scale(data[, c("Y", "X1", "X2", "X3", "X4", "X5")])

# Construct multiple linear regression model
model <- lm(Y ~ X1 + X2 + X3 + X4 + X5, data =
as.data.frame(scaled_data))

# View regression results
summary(model)

# Stepwise regression
# Merge the dependent variable Y and the de-dimensionalized independent
variable into a single data frame
 final_data <- cbind(data[, "Y"], data)

# Perform stepwise regression analysis
model <- lm(Y ~ ., data = final_data)

# Stepwise regression using all independent variables
final_model <- step(model)
```

Time series model forecasting

```
import pandas as pd
import numpy as np
import statsmodels.api as sm
import matplotlib.pyplot as plt
import warnings

warnings.filterwarnings("ignore", message="kurtosistest only valid for
n>=20", category=UserWarning)

data = pd.read_excel(r"C:/Users/....xlsx")
df = pd.DataFrame(data)

from statsmodels.tsa.ar_model import AutoReg

# Fitting an autoregressive model for X1
X1_model = AutoReg(df['X1'], lags=1).fit()
```

```
# Fitting an autoregressive model with X2
X2_model = AutoReg(df['X2'], lags=1).fit()

# Construction of the TSLM model
X = pd.DataFrame({
    'const': 1,
    'X1': df['X1'],
    'X2': df['X2']
})
model = sm.OLS(df['Y'], X).fit()

# Construct the independent variable matrix X
X = pd.DataFrame({
    'const': 1,
    'X1': df['X1'],
    'X2': df['X2']
})

# Construct the vector of dependent variables Y
Y = df['Y']

# Construction of the TSLM model
model = sm.OLS(Y, X).fit()

# Constructing regression formulas
formula = "Y = {:.3f} + {:.3f} * X1 + {:.3f} *
X2".format(model.params[0], model.params[1], model.params[2])
print("The regression equation for the TSLM model is: ")
print(formula)

# Data projected for the next decade
future_years = range(df['year'].max() + 1, df['year'].max() + 11)
pred_X1 = X1_model.predict(start=len(df), end=len(df) + 9)
pred_X2 = X2_model.predict(start=len(df), end=len(df) + 9)

future_X = pd.DataFrame({
    'const': 1,
    'X1': pred_X1,
    'X2': pred_X2
})
pred_values = model.predict(future_X)

# Report Display
```

```
fig, ax = plt.subplots(2, 1, figsize=(10, 8))

# Plot a line graph of the TSLM model with the actual data together.
ax[0].plot(df['year'], df['Y'], marker='o', linestyle='-', label='raw
data')
ax[0].plot(df['year'], model.fittedvalues, marker='o', linestyle='--',
label='TSLM Model')
ax[0].set_xlabel('Year')
ax[0].set_ylabel('Sales')
ax[0].set_title('TSLM model and actual data')
ax[0].legend()

# Mapping of data 10 years after projections
ax[1].plot(df['year'], df['Y'], marker='o', linestyle='-', label='raw
data')
ax[1].plot(future_years, pred_values, marker='o', linestyle='--',
label='Predicted results')
ax[1].set_xlabel('Year')
ax[1].set_ylabel('Sales')
ax[1].set_title('Forecasting sales figures ten years from now')
ax[1].legend()

plt.tight_layout()
plt.show()
```

Gray correlation analysis:

```
Matrix=input('input data:');
19
[s,n]=size(Matrix);
% s is the number of samples, n is the number of sequences
del = zeros(s,n-1);
for i = 1:n-1
del(:,i) = abs(Matrix(:,1) - Matrix(:,1+i));
% Serial difference
end
r=0.5;
% Resolution factor, normally 0.5
M = max(max(del));
% Get the maximum difference between the two levels
m = min(min(del));
% Get the minimum difference between the two levels
A = m + r * M;
B = r * M;
GCC = zeros(s,n-1);
for i = 1:n-1
```

```
GCC(:,i) = A ./ (B + del(:,i));
% Get the gray correlation coefficient of each sequence
end
GCD = mean(GCC);
% The gray correlation coefficients are averaged to obtain the gray
correlation degree
disp(GCD)
```