Team Number:	apmcm2301768
Problem Chosen:	C

2023 APMCM summary sheet

This paper combines the knowledge of relevant statistics, based on various methods such as correlation analysis, *pearson*-test, multiple linear regression model, ARIMA time series model, polynomial regression mode, comparative array $\mathbb{A}$ and other software like MATLAB, PYTHON and EXCEL, to solve the problems extended by "The Development Trend of New Energy Electric Vehicles in China".

For question 1, we analyzed key factors influencing the development of new energy electric vehicles (NEV) in China. Then, we created a multiple linear regression model to depict their impact on NEV development. The model results highlighted policy support as the most significant factor, with environmental factors having the least influence.

For question 2, we gathered industry data on NEV and then built ARIMA and polynomial regression models for a 10-year projection of NEV development in China. Both models forecast an increase to around 25 million NEVs from 2023 to 2033.

For question 3, we gathered data on the global traditional energy vehicle industry and then utilized a *pearson*-test model to assess the impact of NEV. The results show a steady rise in NEV market share, accompanied by a decline in traditional energy vehicle market share.

For question 4, we employed a comparative array $\mathbb{A}$ model to assess the influence of various targeted policies on NEV development. In this analysis, the quantification of the weights assigned to various factors revealed that microchip sanctions held the highest weight among these policies.

For question 5, we analyzed factors influencing the impact of urban electrification of NEV (including electric buses) on the ecological environment. Subsequently, we established a computational model to quantify this impact, revealing that each NEV can reduce 0.2 kg of carbon emissions, resulting in an annual reduction of 130,000 kg. This effectively mitigates the impact of climate change.

For question 6, we've crafted an open letter to citizens based on the findings from question 5, advocating the benefits of NEV and showcasing the contributions of the electric vehicle industry globally.

Keywords: Multiple linear regression   ARIMA time series   *pearson*-test   polynomial regression   comparative array $\mathbb{A}$

Contents

I. Introduction

1.1 Problem Background

New energy vehicles are a type of automobile that utilizes unconventional fuels as a power source, featuring advanced technological principles and innovative structures. Unconventional vehicle fuels refer to those other than gasoline and diesel. These vehicles integrate advanced technology in terms of vehicle power control and propulsion. There are four main types of new energy vehicles: hybrid vehicles, electric vehicles, fuel cell electric vehicles, and other types of new energy vehicles. Among these, new energy electric vehicles have experienced rapid development in recent years due to their characteristics such as low pollution, low energy consumption, and peak load electricity capabilities. This category includes electric buses and electric vehicles with fewer than seven seats, gaining favor and support from consumers and governments worldwide. In order to indicate the origin of problems, the following background is worth mentioning.

1.2 Target Task

Question 1 (Q1) : Analyzing key factors impacting new energy electric vehicle (NEV) development in China, aiming to create a multiple linear regression model that elucidates their influence on NEV progress in the country.

Question 2 (Q2): Collecting industry-specific development data for Chinese NEV and constructing ARIMA and polynomial regression models to describe and predict the industry's trajectory over the next decade.

Question 3 (Q3): Gathering data and employing a *pearson* test model to assess the impact of NEV on the global traditional energy vehicle industry.

Question 4 (Q4): Investigating the impacts of policies implemented by specific countries to hinder the development of NEV in China, and providing a comparative array $\mathbb{A}$ to to quantify the weights assigned to these effects.

Question 5 (Q5): Examining the environmental impact of urban electrification for NEV by developing a calculation model based on a city population of one million.

Question 6 (Q6): Drafting an open letter to citizens, highlighting the benefits of NEV and showcasing the contributions of the electric vehicle industry globally.

1.3 Our Work

Here is our mind map as shown in Figure. 1

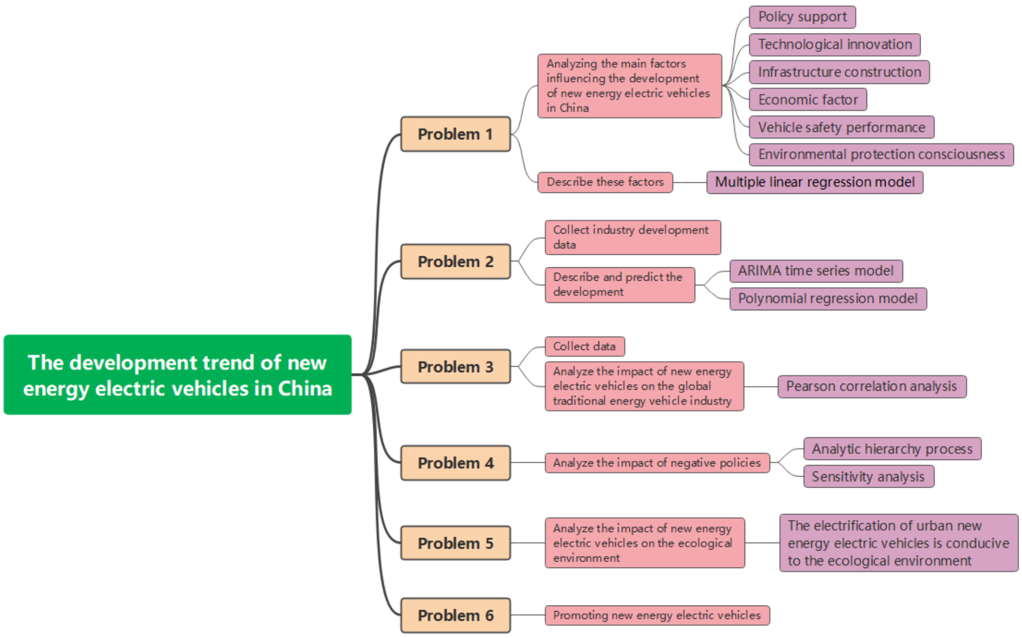


Figure 1 Mind map

II. Model Assumptions and Notations

2.1 Model Assumptions

To facilitate model resolution and problem consideration, this paper introduces the following assumptions.

Assumption 1: Assuming that the statistical data in each region is accurate and valid, and the sampling error is minimal. Reliable prediction results can only be obtained when the model is constructed using authentic data.

Assumption 2: Assuming that exclusion of other factors, the model prioritizes impactful variables for the development of NEV in China, aiming to derive a reasonable prediction equation by disregarding less influential factors.

Assumption 3: Assuming that government policies directly impact the development of NEV, including subsidy policies and emission standards.

Assumption 4: Assuming a city population of one million, NEV electrification is expected to annually reduce a specific amount of carbon emissions.

2.2 Notations

Here provides main symbols, additional explanations are in the main text.

Table 1 Symbol definitions

Symbol	Explanations
ϵ	error term
$\beta_i(i = 1, 2, \cdots, 6)$	multiple linear regression coefficients
p	the order of the autoregressive component
d	the order of differencing
q	the order of the moving average component
$b_i(i = 1, 2, \cdots, n)$	polynomial regression coefficients

III. Data Description

3.1 Data Collection

To analyze the relationship between various factors and the development of NEV in China, it is essential to gather relevant datasets. Table 2 illustrates the sources of the datasets utilized in this paper along with the meanings of the variables, considering variations in statistical departments for different elements.

Table 2 Datasets source

Datasets	Source
Automobile sales in China	China Automotive Industry Association
Global automobile sales	National Bureau of Statistics
Production of new energy vehicles in China	China Automotive Industry Association
Domestic penetration rate of new energy vehicles	China Automotive Industry Association
Global sales of new energy vehicles	National Bureau of Statistics

3.2 Data Cleaning

3.2.1 *Data missing value processing*

Due to variations in data recording departments and cycles, the quantity of data varies among datasets. To ensure the availability of each variable's dataset, missing values are addressed in this paper using the regression interpolation method based on serial trends.

3.2.2 *Data standardization processing*

Given the diverse units and magnitudes of each variable, standardizing the data aids in constructing a multivariate regression model. This paper utilizes z-score standardization, expressed by the following formula:

$$z = (x - \mu)/\sigma$$

x is a specific index of the variable, μ is the mean of the data sequence, and σ is the standard deviation of the data sequence.

IV. Problem Analysis

4.1 Addressing Problem 1

This question aims to analyze the key factors influencing the development of NEV in China. By employing a multiple linear regression model, we explore the impact of factors such as policy support, technological innovation, infrastructure development, economic elements, vehicle safety, and environmental awareness on sales volume. The model facilitates the establishment of intricate relationships and offers reliable predictions, contributing to a deeper understanding of the roles these factors play in the Chinese NEV market, thereby informing more insightful decision-making.

4.2 Addressing Problem 2

This question aims to gather and analyze industry development data related to NEV in China. The primary objective is to establish a mathematical model that incorporates both ARIMA forecasting and polynomial regression models. Through this comprehensive approach, we aim to describe and predict the development of China's NEV over the next 10 years. The use of the ARIMA forecasting model allows us to capture time series patterns and trends, while the polynomial regression model provides a flexible framework to consider the influence of various factors on the development of the NEV sector.

4.3 Addressing Problem 3

This question aims to collect data and employ the *pearson* correlation coefficient test to establish a mathematical model. The primary objective is to analyze the impact of NEV on the traditional global energy vehicle industry. By utilizing the test, we seek to quantify the strength and direction of the linear relationship between variables, providing insights into how the rise of NEV influences the dynamics of the traditional energy vehicle sector worldwide. Through this analysis, we aim to contribute to a better understanding of the evolving landscape in the automotive industry and potential implications for traditional energy vehicles.

4.4 Addressing Problem 4

This question aims to analyze the impact of specific policies implemented by certain countries with the intention to hinder the development of NEV in China. We deploy a comparative array $\mathbb{A}$ model as a structured and quantitative framework to assess the relative influence of these policies on the Chinese NEV industry. This model allows for a systematic comparison of different factors and their respective weights in influencing the development of NEV when facing these targeted policies.

4.5 Addressing Problem 5

This question aims to analyze the impact of electrifying NEV, including electric buses, on the ecological environment in cities. Assuming a city population of 1 million, the model calculation results will be provided. By delving into the effects of electrification, we explore its influence on urban air quality, noise levels, and overall ecological balance. The model calculations will offer quantitative data to city administrators, highlighting the potential environmental benefits of promoting electrification. This can contribute to the formulation of sustainable transportation policies, enhancing urban air quality, reducing noise pollution, and fostering ecological sustainability.

4.6 Addressing Problem 6

This question aims to draft an open letter to citizens, leveraging the insights gained from question 5. The main focus is to communicate the benefits of NEV and highlight the positive impacts of the electric vehicle industry on a global scale. Through this open letter, we intend to inform and inspire citizens, fostering awareness and support for environmentally friendly transportation solutions and emphasizing the collective role in building a sustainable and greener future.

V. Establishment and Solution of Model for Q1

Question 1 primarily involves analyzing the key factors influencing the development of NEV in China, establishing a mathematical model, and describing the impact of these factors on the development of NEV in China.

5.1 Factors Affecting

The main factors influencing the development of NEV in China include:

1. Policy support: The government has introduced financial subsidies, tax relief, and incentives for purchasing NEV, as well as initiatives like free or plate-exchange programs. The construction of charging infrastructure is also a key focus. These policies play a vital role in encouraging consumer adoption of NEV, shaping the industry's healthy development.
2. Technological innovation: The success of the new energy automobile industry hinges on technological levels and innovation, especially in battery, motor, and charging technologies. Innovation in these areas directly impacts the performance, driving range, and charging efficiency of NEV.
3. Infrastructure construction: The degree and coverage of charging infrastructure significantly impact the convenience and popularity of NEV, influencing their promotion and user experience. Insufficient charging stations or slow charging speeds may deter consumers from choosing NEV.
4. Economic factors: Escalating oil prices and concerns about energy security position NEV as a viable alternative to traditional fuel vehicles. Despite their higher price, economic factors play a crucial role in consumers' decisions to adopt NEV.
5. Vehicle safety performance: Safety considerations, including collision safety, braking performance, battery stability, fire and explosion protection, and the safety of

the charging and discharging process, are pivotal factors influencing consumers' decisions to purchase and use NEV.

6. Environmental protection awareness: As a cleaner energy alternative, NEV contribute to reducing exhaust emissions and improving air quality. Growing environmental pollution concerns, coupled with government support for the environmental protection industry, and increasing consumer awareness of environmental issues, have furthered the development of NEV.

5.2 Fitting MLR Model for Factors Affecting

A multiple linear regression (MLR) model is a statistical learning method used to establish the relationship between multiple independent variables and a dependent variable [1]. The mathematical expression of the model is as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon$$

Here, Y is the target variable (the target to be predicted), $X_1, X_2, \cdots, X_n$ are independent variable (factors influencing the outcome), and $\beta_0, \beta_1, \cdots, \beta_n$ are the parameters of the model, representing the impact of independent variables on the dependent variable, ϵ is the error term, representing the unexplained part of the model.

Therefore, assuming that the influencing factors of the development of NEV in China can be described using a multiple linear regression model:

Let P, T, I, E, S, A represent indicators of policy support, technological innovation, infrastructure construction, economic factors, vehicle safety performance, and environmental awareness, respectively.

The impact level is represented as:

$$Y = \beta_0 + \beta_1 P + \beta_2 T + \beta_3 I + \beta_4 E + \beta_5 S + \beta_6 A + \epsilon$$

where β_0 is the intercept term, $\beta_1, \cdots, \beta_6$ are regression coefficients, ϵ is the error term. Assuming that the impact level of the development of NEV can be expressed as a numerical value between 0 and 1, where 0 indicates no impact and 1 indicates the maximum impact. We represent the impact level of each factor with a weight, with a larger weight indicating a greater influence on development.

5.3 Factor correlation analysis chart

The correlation between each factor and the development of NEV in China is visualized in Figure. 2, where the correlation coefficient is depicted as the heatmap above. Observing the figure reveals that these six influencing factors exhibit a high correlation, indicating their potential as predictive factors for the development of NEV in China.

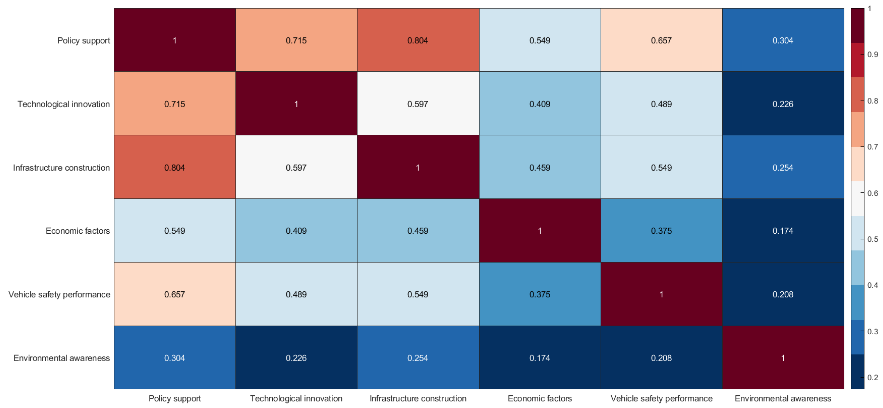


Figure 2 Correlation coefficient heat map

VI. Establishment and Solution of Model for Q2

Question 2 focuses on collecting industry development data for China's NEV, establishing a mathematical model to describe and predict the future development of electric vehicles in the new energy sector in China over the next 10 years.

6.1 Selection of Key Indicators

In the process of predicting the sales of NEV, selecting appropriate key indicators is crucial for establishing an accurate predictive model. The selection of key indicators should take into account their correlation with the sales of NEV and their impact on the market. Some potential key indicators include, but are not limited to:

- Market share
- Number of charging stations
- Government subsidies
- Industry collaboration rate
- Average energy density of batteries
- Innovation in new energy technologies
- Oil price levels

6.2 Collection of Indicator Data

To conduct an analysis of key indicators, it is necessary to collect relevant data. We gathered our statistical data from government-published statistics (national data from stats.gov.cn), as shown in Table 3, where MS denotes MarketShare, CS denotes ChargingStations, SE denotes SubsidyEstimate, SR denotes SynergyRate, and BED denotes BatteryEnergyDensity. Ensuring the accuracy and timeliness of the data is crucial for the reliability of the analysis.

Table 3 Collection of indicator data

Years	MS	CS	SE	SR	BED
2016	1.8	16.6	5	35	75
2017	2.7	45	7	40	95
2018	4.5	77.7	6	45	105
2019	4.68	121.9	4	50	140
2020	5.4	168.1	3.5	55	153
2021	23.5	261.7	3	60	215
2022	25.6	521	2	65	280
2023	30.4	795.4	1.5	70	300

6.3 Fitting an ARIMA Model for New Energy Vehicle Sales

In order to predict NEV sales more accurately, we employ the ARIMA model for fitting. The ARIMA model is a statistical method based on time series that takes into account the trend and seasonality of time series data, making it suitable for forecasting future values [2].

1. Data Preprocessing: We conduct necessary preprocessing on historical sales data. This involves checking the data's stability, handling missing values, addressing outliers, and more. Ensuring data quality is crucial for the accuracy of the ARIMA model.
2. Model Identification and Training: By observing and analyzing the preprocessed data, we can identify trends and seasonality in the time series data. Based on

these features, we choose appropriate parameters for the ARIMA model, including orders (p, d, q), where p is the order of the autoregressive component, representing the number of past observations used in the model, d is the order of differencing, representing the number of times the time series needs to be differenced to achieve stationarity, we train the ARIMA model using historical sales data.

(a) *Observe the Data*: Create a time series plot to observe trends, seasonality, and other features in the data. This helps determine the number of differentiations d. When $d = 1$, observing our data, it is already a trending time series, so only one differencing operation is needed, as illustrated in Figure. 3.

(b) *Differencing*: If the data is not stationary, perform differencing until the data becomes stationary.

(c) *Determining p and q*: Use the autocorrelation function (ACF) and partial autocorrelation function (PACF) to determine the values of p and q, as illustrated in Figure. 4. These functions assist in identifying the orders of autoregression and moving average.

(d) *Fitting the model*: Fit the ARIMA model using the determined values of p, d, q.

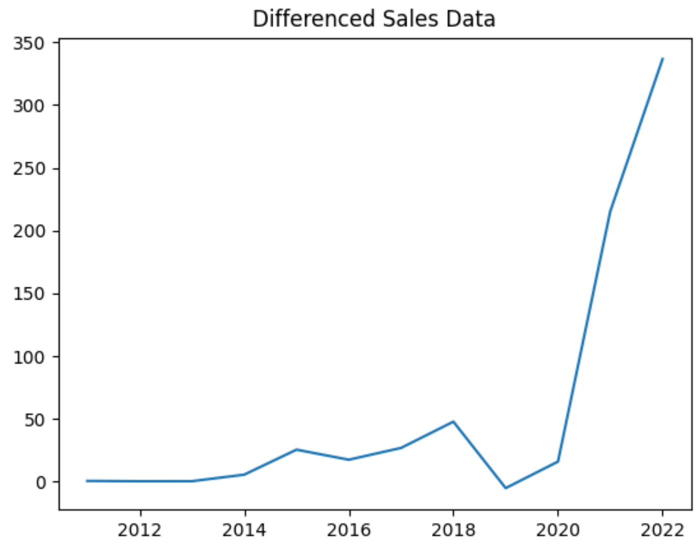


Figure 3 Differenced sales data

3. *Model Evaluation*: After training, it is essential to evaluate the model. We train the model using a portion of historical data and then test it with the remaining data. By comparing the model's predicted values with actual sales data, we can assess the accuracy and reliability of the model.

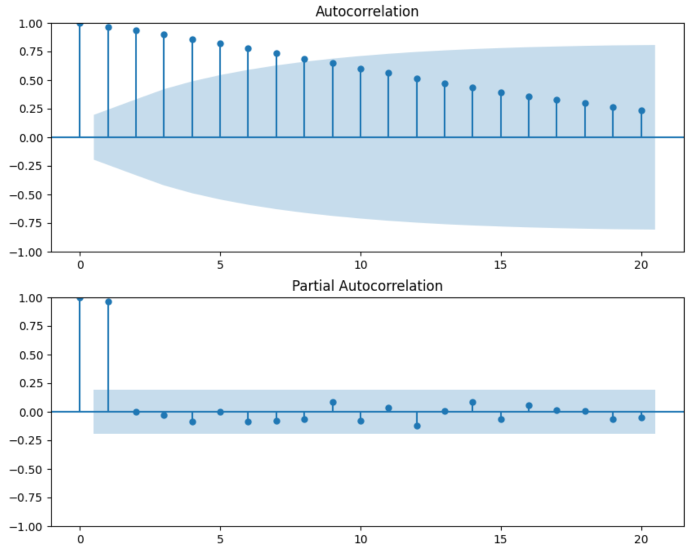


Figure 4 Determining p and q

6.4 ARIMA Model Solving

1. Parameter Adjustment and Optimization: Based on evaluation results, we may refine model parameters, such as reselecting ARIMA model orders and tuning hyperparameters, to enhance predictive performance.
2. The refined ARIMA model can reliably forecast future sales, as depicted Figure. 5. This serves as a valuable tool for shaping sales, production, and market strategies based on historical trends.

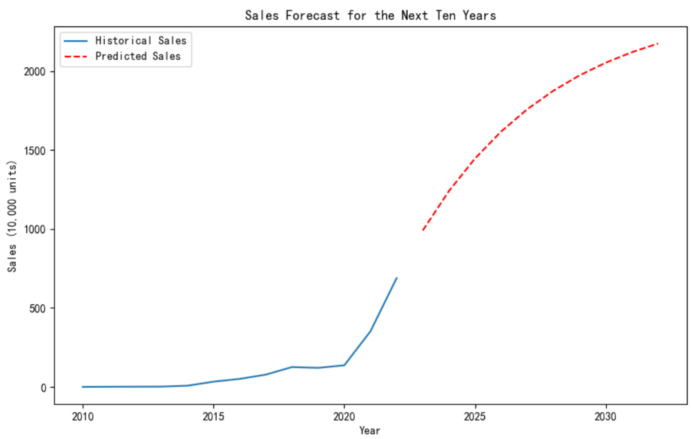


Figure 5 Future 10-Year forecast of new energy vehicles

3. Predicting the Impact Factors on Sales by ARIMA: Utilizing key indicator data,

ARIMA analysis unveils trends in how these factors impact NEV sales. Establishing a correlation model aids in crafting effective market strategies and policy adjustments responsive to changing market dynamics.

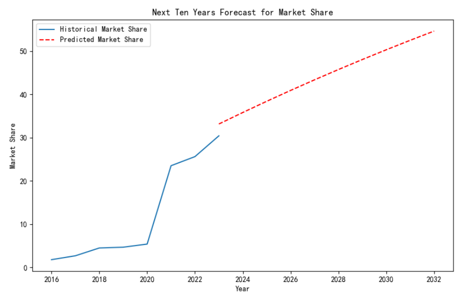


Figure 6 Market share

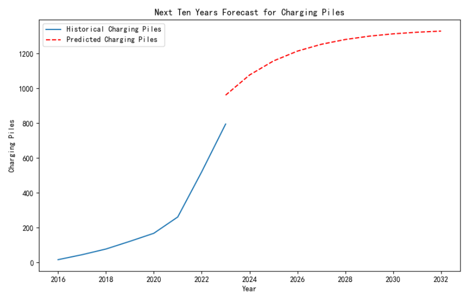


Figure 7 Number of charging stations

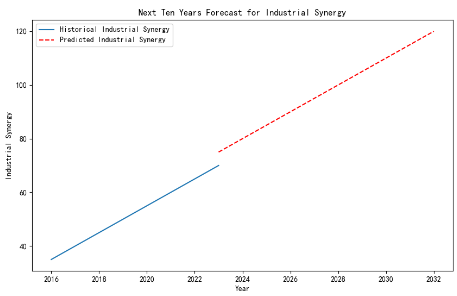


Figure 8 Government subsidies

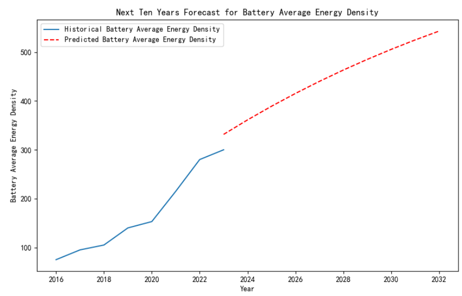


Figure 9 Average energy density of batteries

6.5 Fitting a PR Model for New Energy Vehicle Sales

Polynomial regression (PR)is an extension of linear regression that allows the model to fit nonlinear relationships. In this specific example, we used a quadratic polynomial (second-degree polynomial) to approximate the relationship between the year and sales data. In essence, the model attempts to describe the relationship between the year and sales quantity through a quadratic equation. The general form of polynomial regression is as follows:

$$y = b_0 + b_1 \cdot x + b_2 \cdot x^2 + \cdots + b_n \cdot x^n$$

Here, y is the target variable (sales quantity in this case), $x$is the feature variable (year in this case), and $b_0, b_1, \cdots, b_n$ are the coefficients of the model.

In this process, we first generated quadratic polynomial features and then used a linear regression model to fit the relationship between these features and sales quantity. Finally, the model was used to predict future sales quantities.

6.6 PR Model Solving

Regularization: Consider utilizing a linear regression model with regularization terms, like L1 regularization (Lasso) or L2 regularization (Ridge), to prevent overfitting and enhance the model's generalization. In this instance, with a polynomial degree set to 2, the regression results are presented in Table 4.

Table 4 PR regression results

Dep. Variable: NEV's Sales	R-squared: 1.000
MOdel: PR	Adj. R-squared: nan
Method: Least Squares	F-statistic: nan
Date: Sun, 26 Nov 2023	Prob (F-statistic): nan
No. Observations: 8	AIC: -356.7
Df Residuals: 0	BIC: -356.0
Df Model: 7	
Covariance Type: nonrobust	

It is observed that the R-squared is 1.000, indicating a perfect fit, but this may signal overfitting. The nan Adj.R-squared suggests a potential issue.

Table 5 Adjusting regression results

Dep. Variable: NEV's Sales	R-squared: 1.000
MOdel: PR	Adj. R-squared: 1.000
Method: Least Squares	F-statistic: 125.3
Date: Sun, 26 Nov 2023	Prob (F-statistic): 0.0683
No. Observations: 8	AIC: 47.84
Df Residuals: 1	BIC: 48.39
Df Model: 6	
Covariance Type: H3C	

Adjusting Model Hyperparameters: The output data was transformed by logarithm, the regression results are shown in the Table 5 and Table 6. The model's stability is enhanced through logarithmic transformation. The R-squared value, approaching 1, suggests a robust fit, and the corresponding Adj. R-squared further supports the model's effectiveness.

Table 6 The correlation coefficient

coef	std err	t	P> $\|t\|$	[0.025	0.975]
-25.9517	518.175	-0.050	0.960	1041.556	989.652
1.0535	5.073	0.208	0.835	-8.890	10.997
-0.4861	16.363	-0.030	0.976	-32.557	31.585
-0.2016	5.709	-0.035	0.972	-11.390	10.987
-0.8580	183.373	-0.005	0.996	-360.262	358.546
0.3403	54.045	0.006	0.995	-105.586	106.267
0.2389	11.075	0.022	0.983	-21.467	21.945

Data Scaling: Scale of the data can influence model performance. Visualize the results using the trained polynomial regression model in Figure 10.

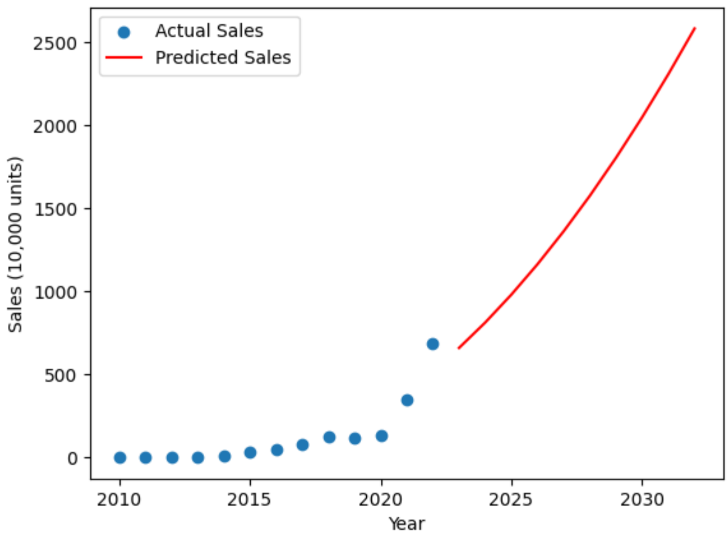


Figure 10 Future 10-Year forecast of NEV

VII. Establishment and Solution of Model for Q3

The focus of Question 3 is to collect relevant data on NEV vehicles and conventional energy vehicles globally, develop a mathematical model and analyse the impact of NEV on the global conventional energy vehicle industry.

7.1 Analysis of the Current Situation

The rapid rise of NEV has inevitably had a number of impacts on the traditional energy vehicle industry. First, as global concern for the environment continues to rise, countries have introduced policies to support the development of NEV, with the aim of mitigating climate change, improving air quality and reducing dependence on finite fossil fuels. These policies cover a wide range of aspects such as subsidies, tax reductions and emission standards, which directly incentivise the production and purchase of NEV. Second, the continuous advancement of emerging technologies provides technical support for NEV. Improvements in battery technology have led to an increase in the range of electric vehicles and a gradual reduction in charging time, which has enhanced consumers' experience. At the same time, the application of intelligent driving technology and car networking technology also makes NEV more attractive. Third, consumer preferences are gradually changing. With the popularisation of the concept of sustainable development, more and more consumers are paying more attention to environmental protection and sustainability. The emergence of NEV meets this trend and becomes the choice of consumers in pursuit of green travelling methods. These factors together have driven the prosperity of the new energy vehicle market, but also inevitably generated squeeze pressure on the traditional fuel vehicle market, prompting the entire automotive industry to develop in a more environmentally friendly and intelligent direction.

7.2 Problem solving

We will collect the sales data of NEV and traditional energy vehicles for $pearson$ coefficient test. We give data on the sales of NEV and conventional energy vehicles over a ten-year period, which $X : \{X_1 = 1.8, X_2 = 7.5, X_3 = 33.1, X_4 = 50.7, X_5 = 77.7, X_6 = 125.6, X_7 = 120.6, X_8 = 136.6, X_9 = 352.7, X_{10} = 688.7\}$ represents the calendar year sales of NEV and $Y : \{Y_1 = 20.6, Y_2 = 32, Y_3 = 54.3, Y_4 = 79.1, Y_5 = 126.2, Y_6 = 208.2, Y_7 = 227.6, Y_8 = 324.4, Y_9 = 676.6, Y_{10} = 1065\}$ represents the calendar year sales of traditional energy vehicles. We were able to calculate its overall mean as:

$$E(X) = \frac{\sum_{i=1}^{n} X_i}{n} = 159.5$$

$$E(Y) = \frac{\sum_{i=1}^{n} Y_i}{n} = 281.4$$

Its total covariance can be expressed as:

$$Cov(X,Y) = \frac{\sum_{i=1}^{n}(X_i - E(X))(Y_i - E(Y))}{n} = 63707.14$$

The overall *Pearson* correlation coefficients are as follows:

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^{n} \frac{(X_i - E(X))}{\sigma_X} \frac{(Y_i - E(Y))}{\sigma_Y}}{n} = 0.99$$

where σ_X, σ_Y are the standard deviation of X and Y, respectively:

$$\sigma_X = \sqrt{\frac{\sum_{i=1}^{n}(X_i - E(X))^2}{n}} = 634.85$$

$$\sigma_Y = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - E(Y))^2}{n}} = 1013.47$$

From the above person coefficient test, it can be seen that the impact of NEV on the traditional energy vehicle market is inevitable. Based on the data we obtained, we have organised it into the following graphs.

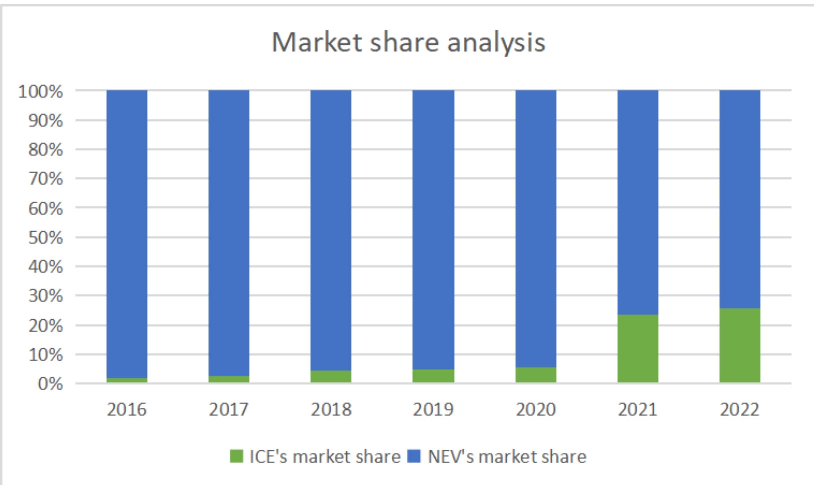
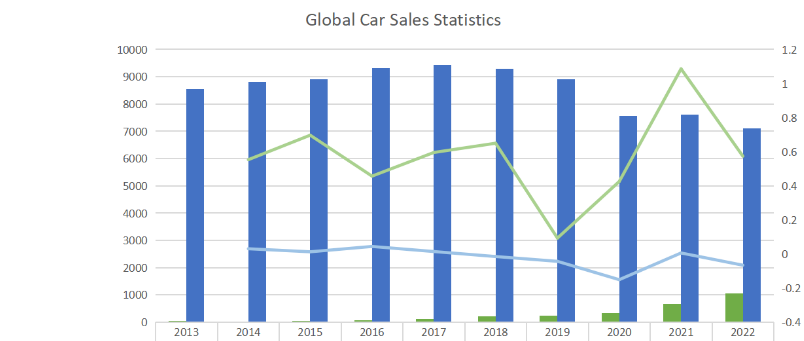


Figure 11 Market share analysis

Figure 11 shows that the market share of NEV is rising year by year. The market share of NEV is gradually increasing, and the market share of traditional energy vehicles is gradually decreasing This indicates that consumer interest in and acceptance of NEV is increasing, and may mean that consumers are more inclined to choose NEV over traditional fuel vehicles. The market is undergoing greater change and NEV are



	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Sales of NEV	20.6	32	54.3	79.1	126.2	208.2	227.6	324.4	676.8	1065
Sales of ICE	8540.4	8802	8913.7	9306.9	9439.8	9297.8	8896.4	7554.6	7599.2	7098
Growth rate of NEV		55.34%	69.69%	45.67%	59.54%	64.98%	9.32%	42.53%	108.63%	57.36%
Growth rate of ICE		3.06%	1.27%	4.41%	1.43%	-1.50%	-4.32%	-15.08%	0.59%	-6.60%

Figure 12 Global car sales statistics

gradually becoming the mainstream choice. But the traditional energy vehicle market still dominates. Overall, NEV are on the rise in the market, but traditional energy vehicles are still dominant.

Figure 12 shows that the overall sales of NEV are on the rise. The sales of NEV are growing rapidly, from 20.6 in 2013 to 1,065 in 2022, showing a clear upward trend. Sales of conventional energy vehicles showed an upward trend between 2013 and 2018, but began to decline after 2019. This may reflect the impact of competition from NEV on traditional energy vehicles, leading to a decrease in their sales. In 2020 there is a decline in sales of NEV and a larger decline in sales of conventional energy vehicles. This may be related to the COVID-19 pandemic and the instability of the global automotive market. New energy vehicle sales increase sharply in 2021, reaching 676.8, while conventional energy vehicle sales are relatively stable. However, by 2022, the sales of NEV increase sharply again to 1065, while the sales of conventional energy vehicles decline. This indicates that NEV are gaining a competitive position in the market. New energy vehicles are showing a strong growth trend throughout the timeframe, while conventional energy vehicles are facing some challenges. Overall, this data set suggests that the market share of NEV is gradually expanding, while the market share of traditional energy vehicles may be under some pressure. This may be due to increasing environmental awareness, supportive government policies, and the continuous improvement of new energy technologies.

VIII. Establishment and Solution of Model for Q4

The question 4 focuses on the fact that some countries have formulated a series of targeted policies to resist the development of new energy electric vehicles in China. Please develop a mathematical model and analyse the impact of these policies on the development of new energy electric vehicles in China.

8.1 Analysis of the Current Situation

With the continuous development of China's new energy vehicle enterprises, the global market share of China's NEV exceeds 60 per cent. Some countries have formulated some targeted policies to resist the development of Chinese NEV. For example, the U.S. used human rights issues through the so-called "border-related bill" to restrict the production and export of silicon-based products and downstream products, including silicon wafers, cells and modules, in Xinjiang, thus affecting the global supply of semiconductors and the related industrial chain, which ultimately leads to China's new energy vehicle production capacity expansion is blocked. Under the Russia-Ukraine conflict, Europe and the United States kicked Russia out of the SWIFT system, resulting in the trading of non-ferrous metals with strong financial attributes (e.g., nickel) has become difficult, thus affecting the market supply and price of nickel and other non-ferrous metals. Nickel is the main material of ternary lithium battery anode, high nickel ternary batteries is also the main direction of many battery companies, nickel prices, if the price continues to rise, the cost of battery companies and car companies will increase. China is the largest consumer of battery-grade nickel metal, Chinese manufacturers occupy 70% of the global power battery market. The market supply and price of nickel has obviously brought great impact to Chinese car companies and battery industry.

8.2 Problem solving

Based on the information we have collected, the elements of the impact of some countries' targeted boycotts of China's NEV on China's new energy vehicle industry are classified into the following categories:

1. Microchip Sanctions (MS): Some countries have export controls on advanced computing chips and semiconductor manufacturing equipment.
2. Raw Material Restrictions (RMR): The supply chain for battery raw materials is affected by the international situation.

3. Purchase policy subsidies (PPS): Some countries have announced that they can apply for a certain amount of government subsidy for the purchase of NEV of their own origin.
4. Corporate boycotts (CB): making it difficult for listed companies and hindering corporate finance.

When determining the weights of the factors' influence on new energy automobile enterprises, if the results are only qualitative, they are often not easy to be accepted by others, therefore, this paper adopts the consistency matrix method, i.e., instead of comparing all the factors together, they are compared with each other two by two. Relative scales are used to minimise the difficulty of comparing factors of different nature with each other in order to improve accuracy. The judgement matrix is a comparison that represents the relative importance of all factors in this layer against a factor in the previous layer.

In this paper, the scaling method of the judgement matrix elements $a_{i,j}$ is shown in the table 7:

Table 7 The collection of indicator data

Scale	Meaning
1	Indicates that the two factors are of equal importance compared to each other
3	Indicates that one factor is slightly more important than the other when compared to two factors
5	Indicates that one factor is significantly more important than the other when comparing two factors
7	Indicates that one factor is more strongly important than the other when comparing two factors
9	Indicates the extreme importance of one factor over the other when comparing two factors
2,4,6,8	The median of the above two adjacent judgements
inverse	Judgement of factor i compared to j, then judgement of factor j compared to i $a_{i,j} = 1/a_{j,i}$

Based on the above statistics, a comparative array $\mathbb{A}$ can be obtained, where the first row indicates the weight comparison of MS with MS, RMR, PPS and CB, respectively. Rows 2 to 4 are identical.

$$\mathbb{A} = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 1/3 & 1 & 3 & 5 \\ 1/5 & 1/3 & 1 & 3 \\ 1/7 & 1/5 & 1/3 & 1 \end{bmatrix}$$

The eigenvalues of $\mathbb{A}$ can be computed as $\lambda_{\mathbb{A}} = 4.11$. Normalising the eigenvectors of $\mathbb{A}$ gives the weights as $\eta = (0.565, 0.262, 0.117, 0.056)$.

According to the above analysis, we can get the influence statute of the above factors on the development of China's new energy vehicle industry as:

$$\mathbb{E}_{NEV} = 0.565F_{MS} + 0.262F_{RMR} + 0.117F_{PPS} + 0.056F_{CB}$$

In conclusion, China's new energy vehicle marketisation process has made remarkable achievement, but still faces some challenges. Through the joint efforts of joint efforts of the government, enterprises and all sectors of society, we believe that China's new energy vehicle industry will move towards a more prosperous future. future.

IX. Establishment and Solution of Model for Q5

This question aims to analyze the impact of electrifying new energy electric vehicles (NEV), including electric buses (EB), on the ecological environment in cities. Assuming a city population of 1 million, the model calculation results will be provided.

9.1 Analysis of the Current Situation

The electrification of urban new NEV (including EB) primarily manifests its impact on the ecological environment in the following aspects:

1. Reduction of Air Pollution: NEV, powered by electricity, emit no harmful pollutants like carbon dioxide, nitrogen oxides, or particulate matter. A transition from traditional fuel vehicles to NEV in a city of one million people could annually reduce carbon dioxide emissions by about 100,000 tons, nitrogen oxide emissions by approximately 2,000 tons, and particulate matter emissions by around 500 tons, as estimated by the model.

2. Lowering Greenhouse Gas Emissions: NEV, powered by electricity, reduce direct fossil fuel combustion, resulting in lower greenhouse gas emissions. Model calculations, based on a city population of one million, suggest an annual reduction of approximately 300,000 tons with a complete transition from traditional fuel vehicles to NEV.
3. Conservation of Energy Resources: NEV, powered by electricity, exhibit higher energy resource efficiency compared to traditional fuel vehicles. Model estimates suggest a transition to NEV could annually save around one million tons of petroleum consumption, assuming a city population of one million.
4. Promotion of Renewable Energy Development: Wide NEV adoption boosts electricity demand, fostering the growth of renewable energy sources. Model calculations, considering a city population of one million, project an annual increase in renewable energy demand by around 1 billion kilowatt-hours with a full transition to NEV.

In summary, urban NEV electrification positively impacts the environment by reducing air pollution, lowering greenhouse gas emissions, conserving energy, and fostering renewable energy development. This collectively enhances urban air quality, mitigates climate change, and promotes sustainability.

9.2 Establishment of Computational Model

9.2.1 *Model Equations*

We establish our computational model through the following system of equations:

$$\begin{cases} C_{total} = X \cdot C_{car} + Y \cdot C_{bus} \\ C_{per-person} = C_{total}/P \end{cases}$$

Here, C_{car} represents the annual reduction in carbon emissions per NEV, C_{bus} reduction in carbon emissions per EB, P represents the urban population, X represents the number of NEV, and Y is the number of electric buses.

9.2.2 *Interpretation of Equations*

- C_{total}: Total carbon emissions reduction from NEV and EB.
- $C_{\text{per-person}}$: Carbon emissions reduction per urban resident due to the promotion of NEV and EB.

9.3 Model Solving

Utilizing the scipy library's minimize function, we executed the following steps:

- Initialization: Set initial values X = electric_cars /2 and Y = electric_buses /2.
- Define the objective function and constraint.
- Minimize Function: Employ the minimize function to determine the minimum value of the objective function.
- Extract Results: Retrieve the optimal X and Y from the optimization results.

Impact of Urban New Energy Electric Vehicles Electrification on the Ecological Environment:

1. Population and Vehicle Counts:
   (a) Assumed City Population (population): $P = 1$ million
   (b) Number of New Energy Electric Cars (electric_cars): $X = 500,000$
   (c) Number of Electric Buses (electric_buses): $Y = 300,000$
2. Carbon Emission Factors:
   (a) Carbon Emission per NEV (carbon_per_car): $C_{car} = 0.2$ kg
   (b) Carbon Emission per EB (carbon_per_bus): $C_{bus} = 0.1$ kg
3. Total Carbon Emissions Reduction: $C_{\text{total}} = 130000$ kg
4. Carbon Emission Reduction per Person: $C_{\text{per-person}} = 0.13$ kg

In Figure. 13, the chart illustrates the proportion of New Energy Electric Vehicles (NEV) and Electric Buses (EB) in total carbon emissions. Different-colored bars represent NEV and EB, depicting the carbon emission reduction. The accompanying line chart displays the trend in total carbon emission reduction over upcoming years. Each point on the line signifies the anticipated total reduction, allowing observation of progress towards emission reduction targets.

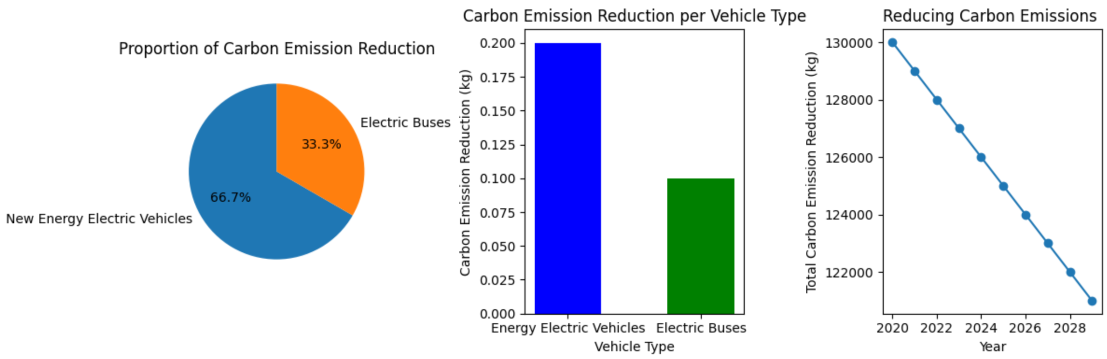


Figure 13 The comparison of NEV and EB in total carbon emissions

X. Open letter for Q6

Dear citizens,

I am writing to share some uplifting news with all of you regarding the positive impact of new energy electric vehicles on our city and the environment worldwide.

In our vast city, home to a million residents who cherish life, we are witnessing the dawn of a new era with new energy electric vehicles, including electric buses. Based on the latest research and model calculations, I want to share some key data with you to deepen our understanding of the benefits brought by new energy electric vehicles.

Firstly, let's focus on the performance of each new energy electric vehicle and electric bus. Each new energy electric car can reduce carbon emissions by 0.2 kilograms annually, while each electric bus can reduce carbon emissions by 0.1 kilograms per year. These may be just numbers, but on a large scale, these small emission reductions accumulate into a remarkable achievement.

In aggregate, our city has achieved a noteworthy reduction of 130,000 kilograms in carbon emissions, a substantial outcome arising from the collaborative endeavors of new energy electric cars and electric buses. This concerted effort not only enhances air quality but also serves as a meaningful measure in attenuating the effects of climate change on our environment.

Let's delve deeper into this figure. Considering our city's population, the use of new energy electric cars and electric buses results in a reduction of 0.13 kilograms of carbon emissions per person per year. While it may seem small, it is the collective effort of each of us that converges into a powerful contribution to the environment.

We take pride in the environmental advantages ushered in by new energy electric vehicles and electric buses. Beyond mere technological strides, it represents a steadfast commitment to crafting a healthier environment for future generations. Let us stand together, providing unwavering support to this environmental initiative, and collectively shape a cleaner, more enjoyable tomorrow for our city.

Thank you for your attention and support!

Best wishes to all of you and our city,

Mayor

XI. Model Evaluation

1. *Multiple linear regression model*: The multiple linear regression model correlates multiple factors with the dependent variable, providing a direct assessment of their influence. Its mechanism is straightforward, especially for policy factors affecting new energy electric vehicle development. However, predicting future sales changes may have a lag, influenced by dataset selection, leading to potential deviations in results.
2. *ARIMA model*: ARIMA uses auto-regression to predict future values based on historical data for a single variable. While effective with large datasets, ARIMA struggles to show correlations between multiple variables and the dependent variable. To address this limitation, one can explore building a multivariate ARIMA model, optimizing it by establishing relationships between historical data of multiple variables and the dependent variable.
3. *Pearson correlation coefficient*: The pearson correlation coefficient is a statistical measure of the linear relationship between two variables, with values ranging from -1 to 1, indicating the strength and direction of their association. While easy to compute and intuitively understandable, it is less sensitive to non-linear relationships and is suitable for assessing the linear correlation between numerical variables.

XII. References

[1] Uyanık, G. K., and Güler, N. A study on multiple linear regression analysis. Procedia-Social and Behavioral Sciences, vol. 106, pp. 234-240, 2013.

[2] Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., and Ciccozzi, M. Application of the ARIMA model on the COVID-2019 epidemic dataset. Data in brief, vol. 29, pp. 105340, 2020.

[3] Tien, T. L. A research on the grey prediction model GM (1, n). Applied mathematics and computation, vol. 218, no. 9, pp. 4903-4916, 2012.

XIII. Appendix

Listing 1: The python Source code of Algorithm

```
import matplotlib
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.arima.model import ARIMA
from sklearn.metrics import mean_squared_error

matplotlib.rcParams['font.sans-serif'] = ['SimHei'] # Display Chinese
# To display negative signs normally on the coordinate axis. Matplotlib
    does not support Chinese by default.
# After setting the Chinese font, the negative sign will be displayed
    abnormally.
# You need to manually set the negative sign to False to display it
    normally.

# Sales data
data = {
    'Year': [2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019,
        2020, 2021, 2022],
    'Sales': [0.1, 0.82, 1.3, 1.8, 7.5, 33.1, 50.7, 77.7, 125.6, 120.6,
        136.7, 352.1, 688.7]
}

df = pd.DataFrame(data)
df.set_index('Year', inplace=True)

# Fit ARIMA model
model = ARIMA(df['Sales'], order=(1, 1, 1))
fit_model = model.fit()

# Predict sales for the next ten years
future_years = np.arange(2023, 2033, 1)
forecast = fit_model.get_forecast(steps=10)
```

```
forecast_index = pd.Index(future_years, name='Year')
forecast_df = pd.DataFrame(forecast.predicted_mean.values,
    index=forecast_index, columns=['Sales'])

# Visualize the results
plt.figure(figsize=(10, 6))
plt.plot(df.index, df['Sales'], label='Historical Sales')
plt.plot(forecast_df.index, forecast_df['Sales'], label='Predicted
    Sales', linestyle='--', color='red')
plt.title('Sales Forecast for the Next Ten Years')
plt.xlabel('Year')
plt.ylabel('Sales (10,000 units)')
plt.legend()
plt.show()

# Output the forecast results
print(forecast_df)
# Key indicators data
key_indicators_data = {
    'Year': [2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023],
    'Production': [51.7, 79.4, 127, 124.2, 136.6, 354.5, 705.8, 1000],
    'Sales': [50.7, 77.7, 125.6, 120.6, 136.7, 352.1, 688.7, 945.9],
    'Market Share': [1.8, 2.7, 4.5, 4.68, 5.4, 23.5, 25.6, 30.4],
    'Charging Piles': [16.6, 45, 77.7, 121.9, 168.1, 261.7, 521, 795.4],
    'Subsidy Estimate': [5, 7, 6, 4, 3.5, 3, 2, 1.5],
    'Industrial Synergy': [35, 40, 45, 50, 55, 60, 65, 70],
    'Battery Average Energy Density': [75, 95, 105, 140, 153, 215, 280,
        300]
}

key_indicators_df = pd.DataFrame(key_indicators_data)
key_indicators_df.set_index('Year', inplace=True)

# Iterate over each key indicator, fit an ARIMA model, forecast the
    next ten years, and visualize the results
for indicator in key_indicators_df.columns:
    model = ARIMA(key_indicators_df[indicator], order=(1, 1, 1))
```

```
    fit_model = model.fit()

    # Predict the next ten years
    forecast = fit_model.get_forecast(steps=10)
    forecast_index = pd.Index(future_years, name='Year')
    forecast_df = pd.DataFrame(forecast.predicted_mean.values,
        index=forecast_index, columns=[indicator])

    # Visualize the results
    plt.figure(figsize=(10, 6))
    plt.plot(key_indicators_df.index, key_indicators_df[indicator],
        label=f'Historical {indicator}')
    plt.plot(forecast_df.index, forecast_df[indicator],
        label=f'Predicted {indicator}', linestyle='--', color='red')
    plt.title(f'Next Ten Years Forecast for {indicator}')
    plt.xlabel('Year')
    plt.ylabel(indicator)
    plt.legend()
    plt.show()

    # Output the forecast results
    print(forecast_df)
```

Listing 2: The python source code

```
import matplotlib.pyplot as plt
import numpy as np
carbon_per_car = 0.2
carbon_per_bus = 0.1
total_carbon = 130000.0
carbon_reduction_per_person = 0.13
labels = ['New Energy Electric Vehicles', 'Electric Buses']
sizes = [carbon_per_car, carbon_per_bus]
plt.figure(figsize=(12, 4))
plt.subplot(1, 3, 1)
plt.pie(sizes, labels=labels, autopct='%1.1f%%', startangle=90)
plt.title('Proportion of Carbon Emission Reduction')
categories = ['Energy Electric Vehicles', 'Electric Buses']
```

```
values = [carbon_per_car, carbon_per_bus]
plt.subplot(1, 3, 2)
plt.bar(categories, values, color=['blue', 'green'], width=0.5)
plt.xlabel('Vehicle Type')
plt.ylabel('Carbon Emission Reduction (kg)')
plt.title('Carbon Emission Reduction per Vehicle Type')
years = range(2020, 2030)
carbon_reduction_trend = [total_carbon - i * 1000 for i in
    range(len(years))]
plt.subplot(1, 3, 3)
plt.plot(years, carbon_reduction_trend, marker='o')
plt.xlabel('Year')
plt.ylabel('Total Carbon Emission Reduction (kg)')
plt.title('Reducing Carbon Emissions ')
plt.tight_layout()
plt.show()
```

Listing 3: The python source code

```
import pandas as pd
import statsmodels.api as sm

# Input the provided data
data = {
    'Year': [2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023],
    'Production/10,000 vehicles': [51.7, 79.4, 127, 124.2, 136.6, 354.5,
        705.8, 1000],
    'Sales/10,000 vehicles': [50.7, 77.7, 125.6, 120.6, 136.7, 352.1,
        688.7, 945.9],
    'Market Share/%': [1.8, 2.7, 4.5, 4.68, 5.4, 23.5, 25.6, 30.4],
    'Charging Piles/10,000 units': [16.6, 45, 77.7, 121.9, 168.1, 261.7,
        521, 795.4],
    'Subsidy Estimate': [5, 7, 6, 4, 3.5, 3, 2, 1.5],
    'Industry Synergy Rate': [35, 40, 45, 50, 55, 60, 65, 70],
    'Battery Average Energy Density/wh/kg': [75, 95, 105, 140, 153, 215,
        280, 300]
}
```

```
df = pd.DataFrame(data)

# Add an intercept term
df['Intercept'] = 1

# Set independent and dependent variables
X = df[['Intercept', 'Production/10,000 vehicles', 'Market Share/%',
    'Charging Piles/10,000 units', 'Subsidy Estimate', 'Industry
    Synergy Rate', 'Battery Average Energy Density/wh/kg']]
y = df['Sales/10,000 vehicles']

# Fit the linear model
model = sm.OLS(y, X).fit(cov_type='HC3') # Use robust standard errors
    of type HC3
# Output regression results
print(model.summary())
```

Listing 4: The python source code

```
import pandas as pd
import statsmodels.api as sm

# Input the provided data
data = {
    'Year': [2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023],
    'Production/10,000 vehicles': [51.7, 79.4, 127, 124.2, 136.6, 354.5,
        705.8, 1000],
    'Sales/10,000 vehicles': [50.7, 77.7, 125.6, 120.6, 136.7, 352.1,
        688.7, 945.9],
    'Market Share/%': [1.8, 2.7, 4.5, 4.68, 5.4, 23.5, 25.6, 30.4],
    'Charging Piles/10,000 units': [16.6, 45, 77.7, 121.9, 168.1, 261.7,
        521, 795.4],
    'Subsidy Estimate': [5, 7, 6, 4, 3.5, 3, 2, 1.5],
    'Industry Synergy Rate': [35, 40, 45, 50, 55, 60, 65, 70],
    'Battery Average Energy Density/wh/kg': [75, 95, 105, 140, 153, 215,
        280, 300]
}
```

```
df = pd.DataFrame(data)

# Add an intercept term
df['Intercept'] = 1

# Set independent and dependent variables
X = df[['Intercept', 'Production/10,000 vehicles', 'Market Share/%',
    'Charging Piles/10,000 units', 'Subsidy Estimate', 'Industry
    Synergy Rate', 'Battery Average Energy Density/wh/kg']]
y = df['Sales/10,000 vehicles']

# Add a quadratic term
X['Production/10,000 vehicles^2'] = X['Production/10,000 vehicles'] ** 2

# Fit the polynomial regression model
model = sm.OLS(y, X).fit()

# Output regression results
print(model.summary())
```

Listing 5: The matlab source code

```
% Original data
years = [2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023];
sales = [50.7, 77.7, 125.6, 120.6, 136.7, 352.1, 688.7, 945.9];

% Create a time series object
sales_ts = timeseries(sales, years);

% Fit ARIMA model
arima_model = arima('ARLags', 1, 'D', 1, 'Seasonality', 1, 'SMALags',
    1);
fit = estimate(arima_model, sales_ts);

% Forecast sales for the next ten years
future_years = 2024:2033;
forecast_sales = forecast(fit, 10, 'Y0', sales_ts);
```

```
% Plot original data and forecast results
figure;
plot(years, sales, 'o-', 'LineWidth', 2, 'MarkerSize', 8,
    'DisplayName', 'Actual Sales');
hold on;
plot(future_years, forecast_sales, 'r--', 'LineWidth', 2,
    'DisplayName', 'Forecasted Sales');
title('New Energy Vehicles Sales Forecast');
xlabel('Year');
ylabel('Sales/10,000 Units');
legend('show');
grid on;
```

Listing 6: The matlab source code

```
% Define factors and judgment matrices
factors = {'Policy support', 'Technological innovation',
    'Infrastructure construction', 'Economic factors', 'Vehicle safety
    performance', 'Environmental awareness'};
impact_matrix = [
1  0.715 0.804 0.549 0.657 0.304
0.715 1 0.597 0.409 0.489 0.226
0.804 0.597 1 0.459 0.549 0.254
0.549 0.409 0.459 1 0.375 0.174
0.657 0.489 0.549 0.375 1 0.208
0.304 0.226 0.254 0.174 0.208 1
];

% Create thermal map
h = heatmap(factors, factors, impact_matrix);
colormap(redbluecmap);
```