

Project 2

Visual Rebirth: How We Turned China's New Energy Vehicle Sales Data from Mess to Magic?

Author: Yidi Wang, Yeying Xing, Sui Zhao, Fanyu Liu

Team ID: T6

Date: 2024-12-13

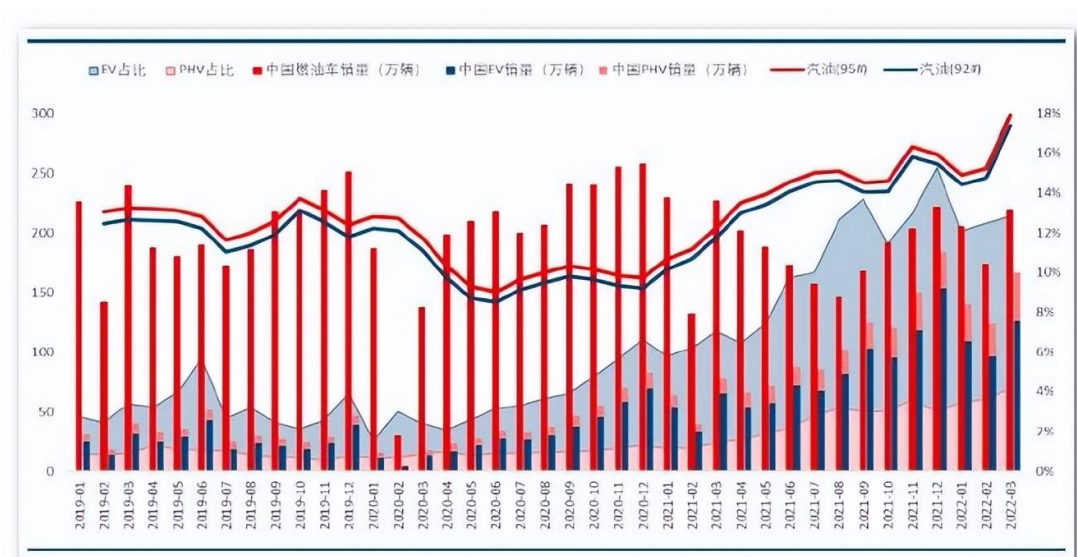
Abstract

This project aims to **analyze and improve** an information visualization that shows the sales trend of new energy vehicles and traditional fuel vehicles in China from 2019 to March 2022 in the context of gasoline price fluctuations. The selected visualization is relatively complex, including **stacked area charts, bar charts, and line charts**.

At the beginning of the project, we explained the reasons for the **theme selection** and **reproduce the original chart** using Python. Then, we deeply analyzed the information the visualization tries to convey and **the story it presents**. We also provided ideas for reading it. In addition, we evaluated **the strength and weakness of the original chart**. To address the identified shortcomings, we modified the visualization step by step.

In the end, we generated a new visualization that shows the development of China's new energy vehicle market more clearly. Notably, the improved visulization uncovers **new stories and insights** that are difficult to observe in the original visualization.

Source of visualized charts: https://m.thepaper.cn/newsDetail_forward_18584297



图说：2019年1月至2022年3月中国纯电、插混车及燃油车销量情况

来源：： Marklines，中汽协，Wind，中信建投

数据 @ 环球零碳

[Figure 1. The original chart]

1. Introduction

1.1. Background

New energy vehicles are powered by alternative fuels instead of traditional fuels like gasoline. They mainly include **pure electric vehicles (EVs)** and **plug-in hybrid electric vehicles (PHVs)**. Since 2019, China has steeped up its efforts to promote electric vehicles.

From 2019 to 2022, **gasoline prices rose** overall, leading to higher operating costs for traditional gasoline vehicles. In addition, China has continuously adjusted the **subsidy policy for new energy vehicles** to encourage the production and promotion of high-quality new energy vehicles in recent years. This, in turn, prompted consumers to switch to new energy vehicles.

1.2. Reasons for Selection

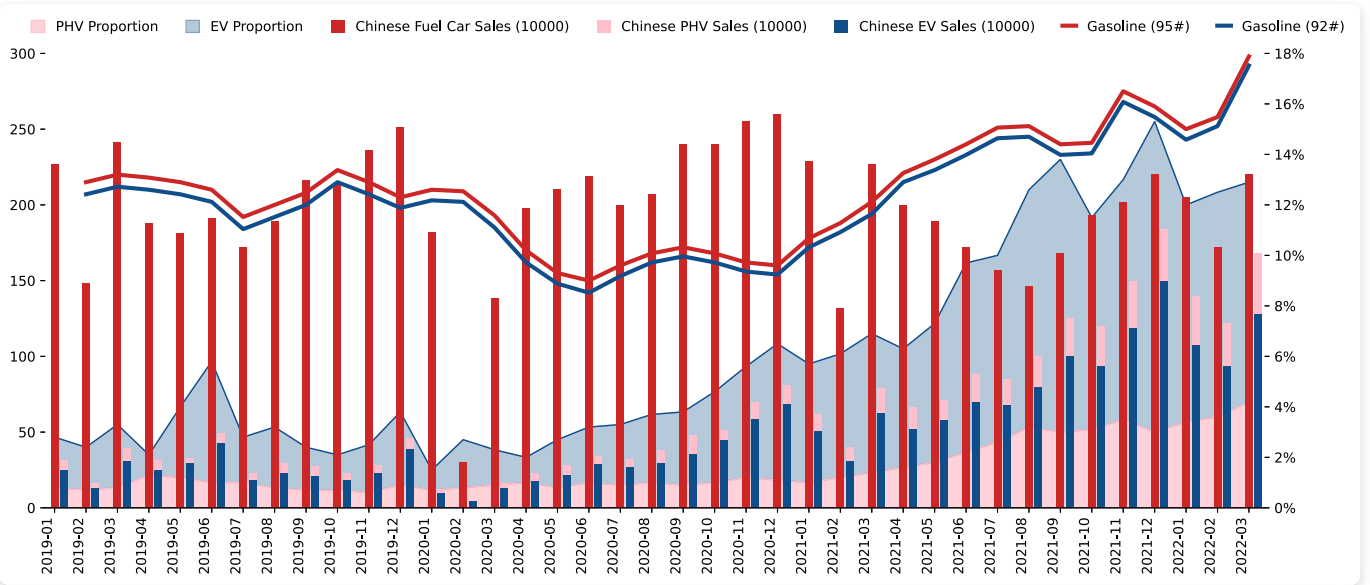
We want to study the changes in sales of traditional fuel vehicles and new energy vehicles in China. We also interested in the factors that affecting it. This chart comprehensively shows **the changes in gasoline prices, sales of traditional vehicles and new energy vehicles, and market penetration** in recent years, which attracted our attention. Thus, we select it as our topic. Generally, we have the following questions:

- What is the **overall trend** of traditional vehicles and new energy vehicles in recent years?
- Does the annual sales of new energy vehicles show **cyclical changes**?
- What **factors** affect the sales of new energy vehicles? How to further promote the development of new energy vehicles based on these factors?

To address these questions, we modified the original chart step by step for better visualization and analysis.

1.3. Stories in the Chart

We replicated the original chart and translated the Chinese text to English in the legend.



[Figure 2. The reproduced chart]

As shown in figure 2, the chart combines three forms of presentation: **bar chart**, **line chart**, and **stacked area chart**.

- **The bar chart shows the sales volume of various types of vehicles**, with the left y-axis (in units of 10,000 vehicles). The red bar on the left represents the sales of fuel vehicles. The upper part of the stacked bar on the right represents PHV sales, while the lower part represents EV sales.
- **The stacked area chart depicts the market share of new energy vehicles**, aligned with the percentage scale on the right y-axis. The upper part of the stacked area chart represents the EV proportion, while the lower part corresponds to the PHV proportion.
- **The two lines above the chart represent the fitted trend lines of gasoline price movements**. The trend line does not correspond to the y-axis on either side, and their heights have no practical significance. The red line above corresponds to 95# gasoline, and the black line below corresponds to 92# gasoline.

We got **the following stories** in the chart:

- Gasoline prices declined in 2020 due to the pandemic but began to **rise significantly** from 2021, with an even sharper upward trend in 2022.
- The market share of new energy vehicles has been increasing with fluctuations. The share of EVs was higher than that of PHVs and was **growing at a faster rate**.
- Sales of traditional fuel vehicles were **negatively correlated** with gasoline prices. Although rising fuel prices have led to a decrease in fuel vehicle sales, their sales remained higher than those of new energy vehicles.
- The sales of new energy vehicles tended to rise with increasing gasoline prices, although there was a **lag effect**.

1.4. The Effectiveness of the Visualization

- **Data-Ink Ratio:**

The chart contains no redundant elements, such as unnecessary borders or gridlines.

- **Lie Factor:**

The visual changes in the visualization accurately reflect the actual changes in the data. The chart does not truncate the y-axis, and the data change ratio matches the visual change ratio, resulting in a lie factor close to 1.

- **Amount of Information:**

The chart combines multiple chart types and presents monthly data from 2019 to March 2022, spanning a substantial time period and delivering a wealth of valuable insights.

2. Problem Statement

- **The amount of information in the chart is too large**, and includes **multiple types of charts**, which can cause cognitive load on readers.
- The **double y-axes** result in poor data readability, and readers may confuse which data corresponds to which y-axis.
- The chart lacks a **title** and **y-axis labels**. Charts should clearly inform readers about the subject and what the axes represent.

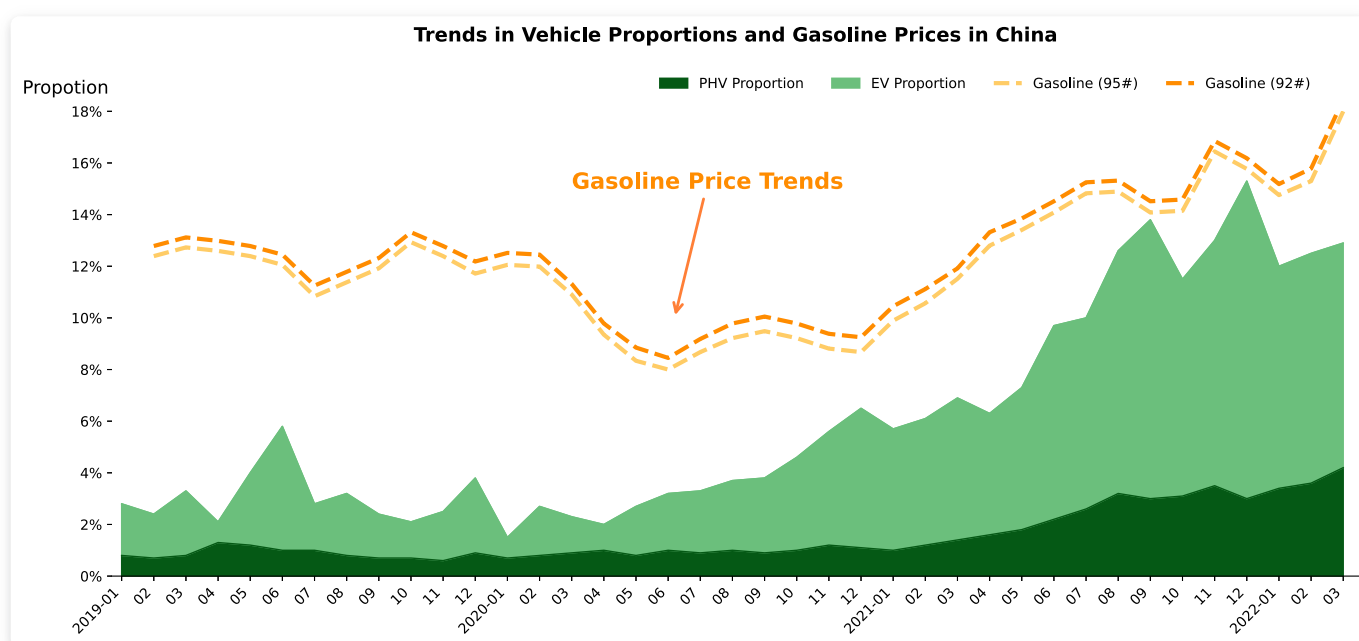
- **Time-series data should be presented using line charts**, but the original chart uses bar chart, which is not conducive to observing trends.
- **Visual variables** (such as color, shape, and shading) **should only be used for data variations**. Different data should be represented by different colors. However, in the graph, gasoline and new energy vehicles both use the same red and blue color scheme, making it difficult for readers to distinguish between them.
- The x-axis text labels in the chart are **rotated** 90 degrees, which makes them hard to read.
- The **trend lines** for the two oil prices are fitted and do not correspond to the y-axes on either side, and their heights have no meaning. We cannot calculate its **lie factor**. This can also **mislead** readers.
- The bars in the original chart are too narrow and should be widened to achieve the **golden ratio**.
- The chart **lacks descriptive text** and does not guide readers to focus on important information, which may cause them to spend too much time on irrelevant details.
- The **x-axis scale labels are redundant**, and unnecessary year information should be removed to improve the **data-ink ratio**.

3. Modification

For all modified charts, we deleted redundant year sticks that highlight the start of each year, moved y label to the top of y-axis, noted the units on the legend and coordinates, and added a bold title to each chart.

The following is our **explanation and analysis of the three modified images**.

Step 1



[Figure 3. The modified stack area and trend lines chart]

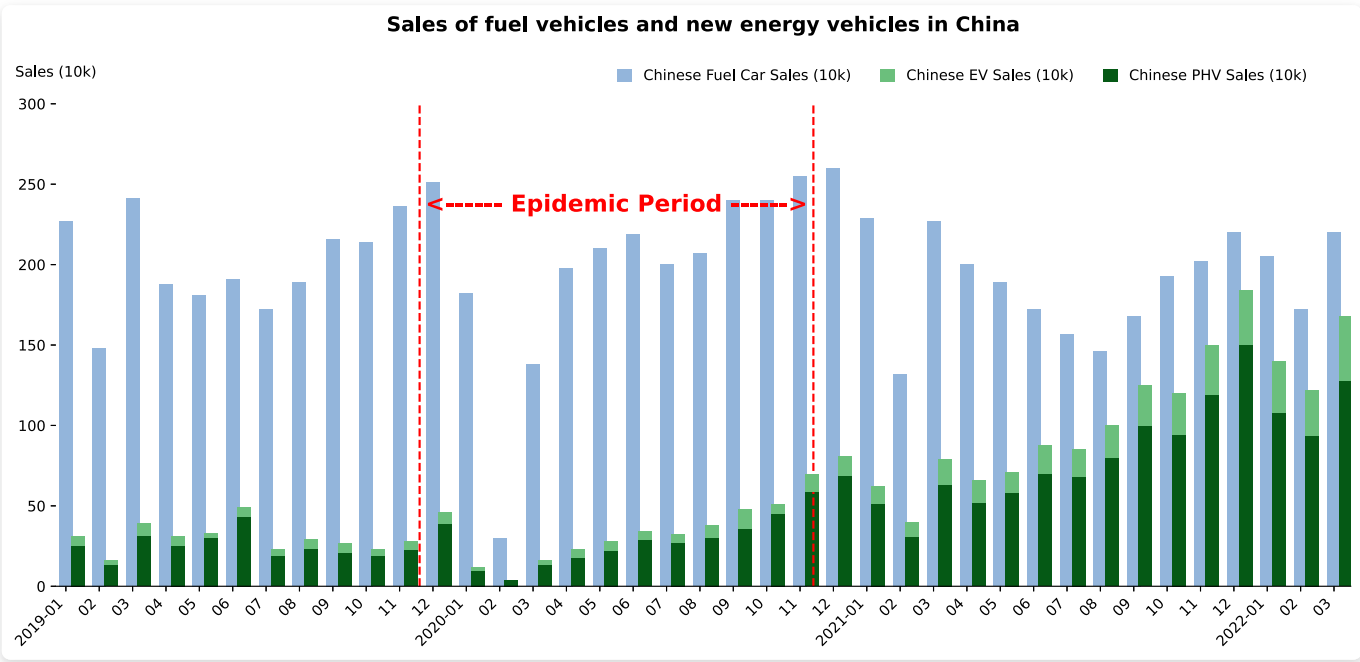
Firstly, we separated the stacked areas and gasoline price trend lines from the original chart. We **changed the color** of stacked areas chart from 'blue/pink' to 'light green/deep green' which is more suitable for the theme of new energy vehicles. We **used dashed lines to represent trend lines**, highlighting that they **do not** have corresponding y-axis values.

From the first new chart, we can better understand **the background of the original chart**. In the context of an overall upward trend in oil prices, the market penetration rate of energy vehicles is also increasing.

More Details:

In 2019, gasoline prices remained relatively stable. At the start of 2020, the gasoline price fell sharply due to the pandemic but rebounded with the economic recovery, albeit staying at relatively low levels. In 2021, rising demand combined with delayed supply led to a significant increase in gasoline prices. Since 2022, geopolitical conflicts between Russia and Ukraine, along with EU energy sanctions against Russia, have caused a significant increase in gasoline prices. As gasoline prices rose, the market share of new energy vehicles steadily increased. From 2019 to 2020, the growth of energy vehicles in market share was gradual. From 2021 to 2022, the market share of new energy vehicles increased significantly.

Step 2

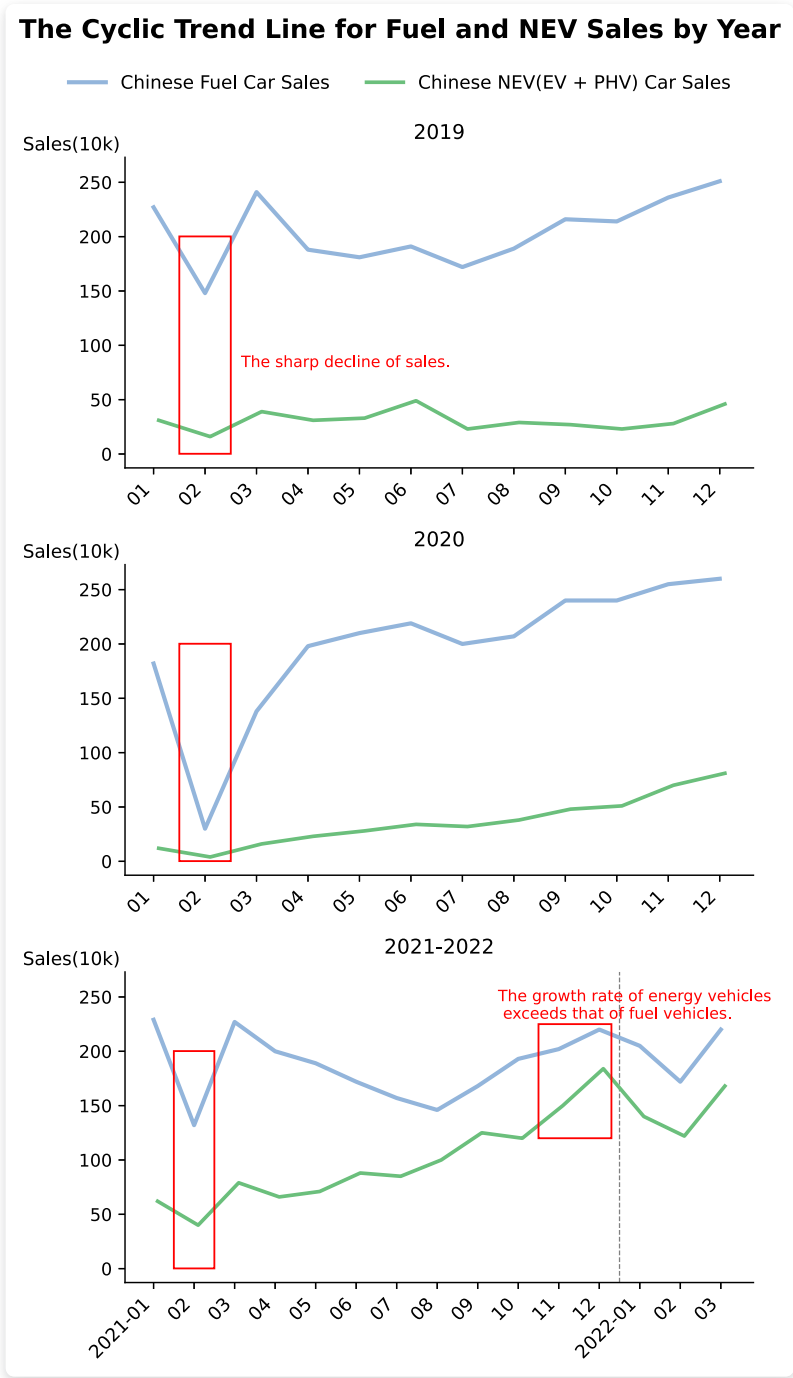


[Figure 4. The modified bar chart]

Secondly, we separated the bar chart and **changed the color** from 'red/ pink/ blue' to 'light blue/ light green/ deep green', and we **widened each bar** to reach the golden ratio. The pure bar chart **focuses on the sales volume value** for different vehicles. We can see a distinct difference between the end and the start. On the whole, though the sales of fuel cars almost always maintained around 180, **the sales of new energy cars had significantly increase from 40 to 150**. We also noticed an unusual decline in 2020 with total sales being less than 50, and believe that the '**epidemic period**' is the main reason, so we annotated it on the new chart.

Step 3

Thirdly, we split the complete image by year and drew **trend lines of sales volume** According to the bar chart (2022 only has three months) so we merged them with 2021.



[Figure 5. The cyclic line chart]

Here we found some **cyclical regularity** every year:

- The total sales of new energy vehicles have shown an **overall upward trend** every year, with the **upward trend increasing over the years.**
- The overall growth of new energy vehicles in 2021 **exceeded** that of traditional gasoline vehicles, possibly due to China's introduction of subsidy policies for new energy vehicles in 2021.
- The sales of both types of cars **decrease significantly every February**, possibly because this period falls during the Chinese Spring Festival holiday.

4. Conclusion

In conclusion, we separated the original into **stacked area + gasoline price trend lines, bar chart, and line chart**, revealing more detailed background of our project, the overall sales volume changes for different vehicle and the cyclical regularity on sales volume for each year respectively. Through the modification and further exploration, we found **more insights** beyond "the increase market penetration rate of energy vehicles as gasoline price increasing". **The epidemic, policies, Spring Festival, and other factors also have a significant impact on overall vehicle sales, and can be reflected in our visualization.**

Based on these factors, we can do the following things to **further promote the development of new energy vehicles**

- The government can **further optimize subsidy policies** for new energy vehicles and strengthen infrastructure development to promote the widespread adoption of new energy vehicles.
- Given the typical decline in car sales every February, companies can **launch promotional activities** for new energy vehicles during the Spring Festival to stimulate consumer demand.
- Society should **strengthen education and awareness of environmental protection**, encouraging more people to pay attention to and purchase new energy vehicles, thus promoting the realization of sustainable transportation.

The article and all the codes are now available on the public platform

Github:<https://github.com/zhaosuisui/Data-Visualization>