Data Visualization Analysis of JD Mobile Phone Based on Python and ECharts and Bootstrap

Luqiao Fan[(✉)], Banxiang Duan, and Jie Gao

Guangdong Polytechnic of Science and Technology, Computer Engineering and Technical College, Guangzhou 510640, China
536714979@qq.com

Abstract. In order to more intuitively understand and mine the information behind mobile phone sales data, ECharts is used to visualize mobile phone sales data. After getting the sales data of three popular current mobile phones in JD website in 2022 through Python crawler technology, use ECharts to present the data in the form of bar chart, pie chart, word cloud chart, etc. Built a visual website interface through Bootstrap, and made use of the rich interaction functions of ECharts. Enable users to quickly access the information they need. In large data processing activities, ECharts can adapt to the visualization goal of high speed and low cost.

Keywords: big data analysis · data visualization · crawler · Python · ECharts · Bootstrap

1 Introduction

With the rapid development of science and technology, smart phones are being updated and iterated at a quick speed. There are various mobile phone brands at home and abroad, and which brand of mobile phone is more popular, which configuration of mobile phone is more favored by consumers, and what functions of mobile phone users value more are worth us to explore the information. In addition, sales companies also hope to provide more valuable strategic support for their future development and market competition through data analysis and mining of a large number of shopping data generated by users. This paper selected three popular current mobile phones, analyzed the sales data information and user comments information of the mobile phones, and built a visual website interface of sales data with Bootstrap, which will help consumers better understand their personal needs, and sales companies can have a more intuitive understanding of the market demand, so as to promote the sales of products.

In the era of large data, the exponential growth of various resources and data information is dazzling, and the most effective way for researchers to obtain information from the outside world comes from the visualization system [1]. When data is displayed in a visual graph, analysts can generally quickly obtain the hidden information behind it. Data visualization is to use computer graphics, image processing and other technologies to convert data into images and interact with each other, so as to express information and trends more intuitively [2]. In this paper, relevant data of three popular current

mobile phones obtained from JD website are used to visualize mobile phone data of multiple dimensions through ECharts, so as to conduct more in-depth observation and analysis [3, 4].

2 Research Overview

2.1 Introduction to Related Technologies

Python is the most popular programming language today, involving a wide range of application fields. Python is concise, easy to read, extensible, and has many class libraries. At present, it is brilliant in the field of big data, artificial intelligence, data analysis and so on, also known as the "glue language". Pandas is an open-source, third-party Python library built on top of Numpy and Matplotlib, and is known as one of the "big three" of data analysis (NumPy, Matplotlib, and Pandas). Pandas has become a must-have high-level tool for Python data analysis, and it aims to be a powerful, flexible data analysis tool that can support any programming language [5]. BeautifulSoup can parse and extract data from web pages. It is not the Python standard library, so you need to install it separately. It supports not only CSS selectors, but also the HTML parser from the Python standard library, and the XML parser from LXML. It realizes the usual way of navigating and searching documents, saves a lot of work time, and improves the efficiency of the development project.

ECharts is an open source and free visual chart tool owned by Baidu. It is a pure JavaScript chart library, which can run smoothly on PC and mobile devices. ECharts is not only compatible with most current browsers, but also relies on the lightweight Canvas class library ZRender at the bottom, providing intuitive, vivid, interactive and highly personalized data visualization charts. In addition, ECharts' innovative drag-and-drop weight calculation, data view, range roaming and other features greatly enhance the user experience, giving users the ability to mine and integrate data. ECharts is a data visualization tool with interactive graphical user interface, which is a deep data interactive visualization tool [6].

Bootstrap is an HTML, CSS, JavaScript based front-end framework released by Twitter for rapidly developing Web applications and websites. At present, the latest version of Bootstrap is Bootstrap5. With the Sass variable, a large number of mixins, responsive raster system, extensible prefabricated components, and powerful plug-in system based on jQuery, a prototype can be quickly developed or the whole App can be built.

2.2 Overall Design

Visual display: read the required data from CSV, and use ECharts visualization tool to display the data in the form of bar charts, pie charts, and word clouds.

Graph interaction: data visualization includes two core elements: visual presentation and interaction. Interaction not only alleviates the contradiction between limited visualization space and data overload, but also actively involves users in the process of mental model construction, helping users understand data and discover patterns. This article will use the mouse hover prompt box, button switch, link jump and other methods to achieve chart interaction.

3   System Implementation

This system mainly performs data visualization on Jingdong mobile phone sales data. The system is divided into three modules, which are the data crawler module, data cleaning module and data visualization module. First, we run the Python crawler module to save the crawled JD mobile phone sales data as CSV format files, and then use Pandas to clean and process the data. Finally, we build a visual website interface through Bootstrap, and use the rich interactive functions of ECharts to enable users to quickly obtain the information they need. Display interactive data visualization interface.

This system selected three flagship mobile phones from three brands, namely Apple's iPhone 13, Huawei's Mate 40 Pro and Xiaomi's Mi 12. By analyzing these three flagship mobile phones, the processed data is divided into purchase time information and user review information, and according to the distribution of purchase time information, monthly sales statistics histogram and daily sales statistics histogram are made, respectively; according to the user review information, the word frequency statistics pie chart of the number of user comments and the word cloud chart of the user comment content are made respectively.

3.1   Monthly Sales Statistics bar Chart

Figure 1 is a monthly sales bar chart of iPhone 13 sales in 2022 by month. The monthly sales chart of the iPhone 13, which was released on September 15, 2021, shows that sales continued to be enthusiastic in the months after the new phone was released. In November, due to the Singles' Day event, there was another sales peak, through the interactive chart, showing the purchase volume of 5616 units in November.

Figure 2 is a monthly sales histogram of Huawei Mate 40 Pro's mobile phone sales in 2022 by month. The Huawei Mate 40 Pro was launched on October 30, 2021, and as can be seen from its monthly sales histogram, it has a sales peak in the two or three months after the new launch of the phone. In November, due to the Double 11 event, there was another sales peak, and through the interaction of the chart, the number of purchases in November was 3488 units.

Figure 3 is a monthly sales histogram of Xiaomi Mi 12's mobile phone sales by month in 2022. The Xiaomi Mi 12 was launched on December 28, 2021, and as can be seen from its monthly sales histogram, there was a sales peak within one month after the new launch of the phone. In November, due to the Double 11 event, there was another sales peak, and through the interaction of the charts, the number of purchases in November was 9852 units.

3.2   Daily Sales Statistics bar Chart

Figures 4, 5 and 6 shows the bar charts of the daily sales of Huawei iPhone 13, Mate 40 Pro and Xiaomi 12 respectively. Through the situation of consumers buying mobile phones at different times of the day, it can be clearly seen that the consumption peak of different brands and models of mobile phones is not the same: the consumption peak of iPhone 13 is about 20 points. The Mate 40 Pro had two peaks, at 10 a.m. and 0 p.m. The consumption peak of the Mi 12 appeared at around 22:00. By analyzing the
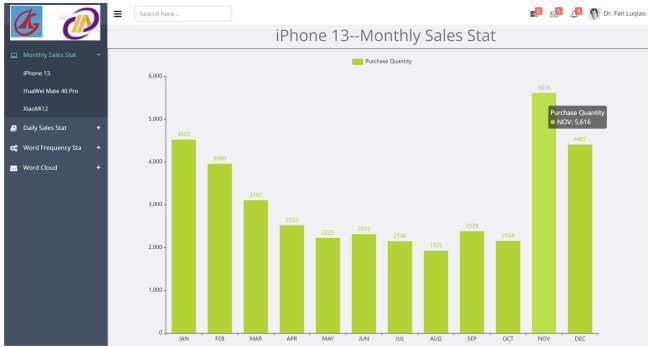
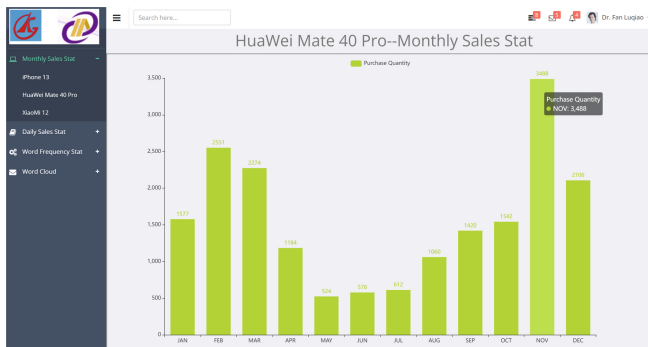**Fig. 1.** Bar chart of monthly sales of iPhone 13

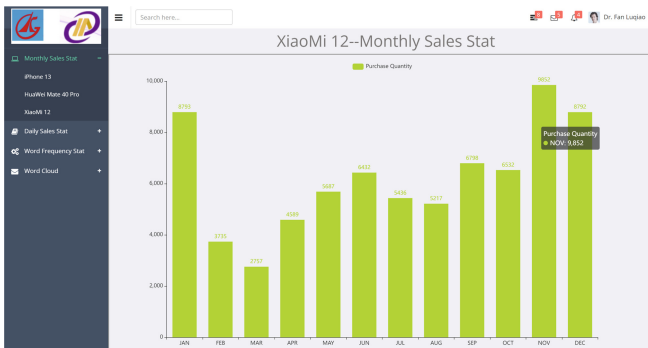**Fig. 2.** Bar chart of monthly sales of Huawei Mate 40 Pro

Fig. 3. Bar chart of monthly sales of Xiaomi 12

consumption peak in a day, we can roughly analyze the living habits of the consumer groups for different models of mobile phones, so as to make a rough portrait of the customers of each model of mobile phones.

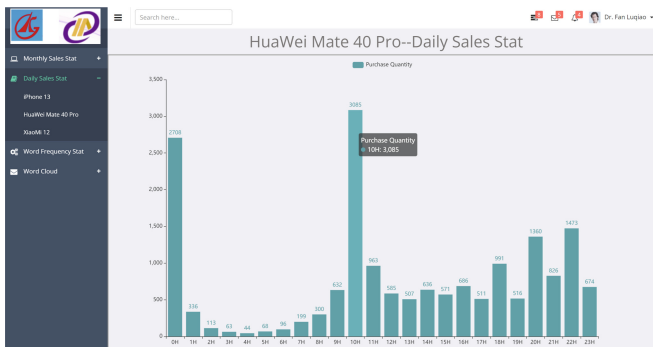**Fig. 4.** Bar chart of daily sales of iPhone 13
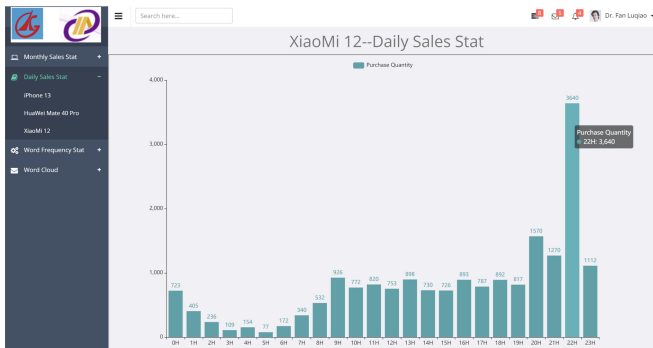

Fig. 5. Bar chart of daily sales of Huawei Mate 40 Pro

Fig. 6. Bar chart of daily sales of Xiaomi 12

3.3 Word Frequency Statistics Pie Chart

To make a pie chart of word frequency, we will use Jieba (Segmentation Library). Stuttering word segmentation library is an excellent Python third-party Chinese word segmentation library, which has been firmly ranked C in the field of Chinese word
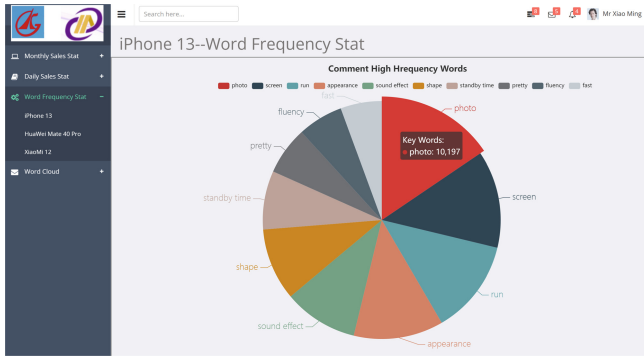
Fig. 7. Pie chart of word frequency statistics of iPhone 13

segmentation. Since each character in Chinese text is written continuously, each phrase in Chinese text is obtained by a special method, which is called word segmentation. Word segmentation is often needed in the field of Natural Language processing (NLP). The word segmentation principle of stuttering word segmentation library: Using a Chinese word library, determine the association probability between Chinese characters, the composition of words with large probability between Chinese characters, form the word segmentation results. In addition to tokenization, users can also add custom phrases. The stuttering thesaurus supports three segmentation modes: exact mode, full mode, and search engine mode.

According to the collected mobile phone data, the stuttering word segmentation library is used to segment the user's comment content, and the most frequent keywords in the user's comments of each mobile phone are analyzed to generate the pie chart shown in Fig. 7–9, and then the advantages of each mobile phone valued by users are analyzed. By observing Fig. 7, we can see that in the reviews of iPhone13, we can intuitively see the number and proportion of the top ten keywords appearing in the reviews of iPhone13, and we can see that the number of "photo", "screen" and "run" ranks the top three. Therefore, it is inferred that iPhone 13 does better in the camera function, the screen and the running speed of the system, which can be used as the advantage selling point of iPhone 13. When the user moves the mouse pointer to the "take a picture" sector of the pie chart, the value of the "take a picture" sector is displayed dynamically for 10197 times, indicating that 10197 users are most concerned about the feature of "take a picture". Through these three pie charts, the main selling points of these three models of mobile phones can be analyzed. Figure 8 shows the word frequency statistics of Huawei Mate 40 Pro, and similar results can be obtained. The word frequency statistics of MI 12 can be seen by observing Fig. 9, and similar results can be obtained, which will not be repeated here.

3.4   Comment Information Word Cloud

Word cloud, as a new chart tool, can achieve good communication effect by visually highlighting the "keywords" that appear frequently in the text. The stammer library is also used to create a word cloud of mobile phone comments.
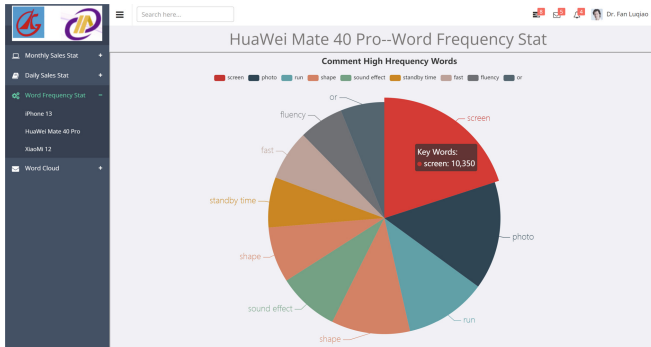
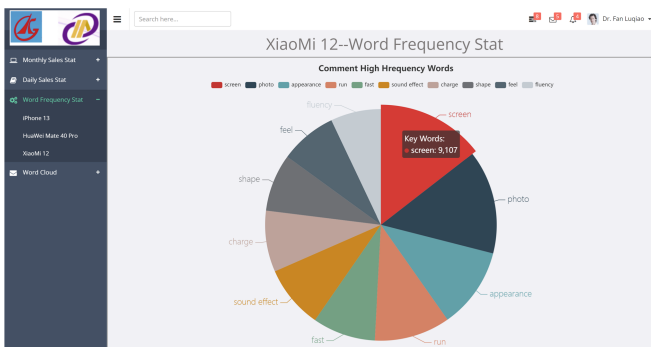**Fig. 8.** Pie chart of word frequency statistics of Huawei Mate 40 Pro

Fig. 9. Pie chart of word frequency statistics of Mi 12

Figure 10 is the word cloud of the iPhone 13. When you click the "iPhone 13" hyperlink in the "Word Cloud" in the left navigation bar, you can visually display the word cloud generated from user reviews of iPhone 13 in the center chart display area. The more frequent the word is, the larger the font size is displayed. Through the word cloud, the main high-frequency content in the comments can be intuitively seen. Each keyword is divided by green, blue, orange, purple, light green and brown colors, indicating words with more than 10000, 8000, 6000, 4000, 2000 and less than 2000 comments frequency, respectively. It can be seen that "clear", "very fast", "clear picture", "fast delivery" and "well-packed phone" are the top praise words for iPhone 13, which have great appeal to users.

Fig. 10. Word Cloud of iPhone 13

4   Conclusion

This paper focuses on the application of Python, Bootstrap and ECharts technology in mobile phone data visual analysis. Through the display and analysis of the visualization results of the sales data of three popular new mobile phones in 2022, it can help users better meet their own needs, provide more constructive opinions for the majority of users to buy mobile phones, and also provide certain decision support for the mobile phone sales industry. The word cloud map analysis of Huawei Mate 40 Pro and Xiaomi 12 is basically similar, which will not be repeated here.

References

1. Comba, Joao L. D..Data Visualization for the Understanding of COVID-19 [J]. Computing in Science and Engineering. 2020,22(11):81–86.
2. Pradipta Biswas, KamalPreet Singh Saluja, Somnath Arjun, et. al. COVID-19 Data Visualization through Automatic Phase Detection[J]. Digital Government: Research and Practice. 2020,1(10):1–8.
3. Xiaoquan Ou , Zuying Zhu, Junyan Chen, et al. A Study of Data Visualization of the Neo-coronary Pneumonia Epidemic[C]. In proceedings of 2nd IEEE Eurasia Conference on IOT, Communication and Engineering 2020. pp 315–317.
4. Juan Shen, Qiang Wang. The Trend of Research on Alzheimer's Disease based on Bibliometrics and Visualization using Echarts[C]. 2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI). pp 491–497.
5. Gao Yan.Crawling and Analysis Data of Big Data Post Based on Selenium Framework[J].Industrial control computer,2020, 33(02):109–111.
6. Fan Luqiao,Gao Jie, Duan Banxiang.Mobile phone sales data visualization system based on Python & ECharts[J].Computer programming skills & maintenance, 2022 (6):78- 81.