

# Crawler Data Visualization in the Tourism Industry Based on Python

Ao Wang \*

School of History Culture and Tourism, Sichuan Normal University, Chengdu, China

\* Corresponding author: 1910700518@mail.sit.edu.cn

**Abstract.** Data visualization is being widely used in various fields, such as medicine, biology, economics, and other disciplines. In the tourism industry, whether the display of tourism resources, the planning of tourism routes, the feedback of tourists' experience, and so on, all of them become clearer and more intuitive through data visualization. In particular, Python-based data visualization has many advantages in the tourism industry. This paper mainly analyzes the current situation and advantages of Python-based data visualization in the tourism industry, and puts forward corresponding improvement measures and suggestions, aiming to help tourism practitioners and tourists use data visualization more conveniently and efficiently. It is found that data visualization plays an important role in displaying tourist attractions, planning tourist routes, and enhancing the experience of tourism services. However, the problems of data visualization in tourism, such as insufficient interactivity and single visualization style, require improvement measures. For example, methods such as optimizing algorithms and introducing interactive visualization tools can solve these problems. The purpose of this study is to promote a wider and more effective application of data visualization in the tourism industry so that tourism-related industries can be more efficient and intuitive when using data visualization tools.

**Keywords:** Python, Data visualization, Web crawler, Tourism industry.

## 1. Introduction

Data visualization refers to the presentation of relatively obscure data visually and interactively, to graphically and intuitively express the information and laws contained in the data. Essentially, any method that can demonstrate the principles, laws, and logic of things with the help of graphics is called data visualization. Data visualization has become a commonly used technology in the fields of medicine, biology, economy, and so on, and it is also widely used in the tourism industry. Whether it is to show tourism resources, plan tourism routes, or statistics on visitor consumption of tourist attractions and other information, data visualization makes these data clearer and easier to understand. With the rapid development of information technology, the application of data visualization in the tourism industry is increasing and has become an indispensable part of the tourism industry.

Data visualization using Python for tourism has many advantages, which can give tourism practitioners and tourists more intuitive access to tourism information, save decision-making time, and provide help for practitioners to predict future trends, as well as enhance the tourism experience of tourists. Sun et al. used Python's requests library to process the tourism data in Xinjiang, visualized the text data, and converted it into charts in the form of maps, so that the data information related to the tourist attractions in Xinjiang is presented in the form of maps, which makes it convenient for tourists to formulate personalized journeys more suitable for themselves according to the attraction tips displayed on the maps, thus enhancing the tourists' tourism experience [1].

Tourist behavior is recorded by a variety of software, and data-based visualization facilitates better travel planning for social network users. However, current research has found that although it is possible to present and analyze tourism-related data in Python, there are still some shortcomings in applying data visualization to the tourism industry. The limitations of Python's algorithms and the singularity of visualization styles have caused inconvenience to the use of data visualization in the tourism industry.

This paper aims to study the application of data visualization in the tourism industry. It introduces the tools of data visualization and the current situation of data visualization in the tourism industry.

The paper will analyze its advantages and demonstrate the specific use of data visualization through different perspectives. Finally, suggestions will be provided on the problems and shortcomings of data visualization in the tourism industry. The purpose of this paper is to study the convenience and problems brought by data visualization to the tourism industry, and to give corresponding countermeasures and suggestions so that data visualization can be more widely and effectively used in the tourism industry in the future. The goal is to improve the efficiency and intuitiveness of data presentation tools for tourism-related industries in the future.

## **2. Related Concepts**

### **2.1. Introduction to Python**

Python is a programming language that is both simple and powerful, focusing on how to solve problems rather than syntax and structure. It is an interpreted, compiled, and object-oriented scripting language. The unique syntactic structure of the Python language is highly readable, making it easier for users to use and allowing people to focus on solving problems rather than learning various tedious syntax rules [2]. Using the Python language to create a web crawler and to process and store the data fully utilizes the advantages of the role of the Python language. In the tourism industry, tourism practitioners can use Python's tripartite library to visualize and present tourism-related data and display tourist attractions and tourist information more intuitively. Practitioners can also utilize Python's predictive models to forecast future tourism trends and thus provide better tourism services.

The web crawler is a program that automatically extracts web pages for a search engine to download from the World Wide Web and is an important component of a search engine [3]. Traditional crawlers start with the URL of one or several initial web pages, obtain the URLs on the initial web page, and in the process of crawling the web page, keep extracting new URLs from the current page to put into the queue until certain stopping conditions of the system are met. In the tourism industry, tourism practitioners can use crawler technology to crawl data and information from tourism websites, airline websites, and tourism-related posts published by tourists on social platforms, and use data visualization to present their data, analyze and understand tourists' demand bias, to improve the industry.

### **2.2. Introduction to Tourism**

Tourism is a comprehensive industry involving tourism resources and related facilities that is exclusively or primarily engaged in entertaining and hosting tourists. It mainly includes hotels, restaurants, tourist transportation, scenic spot infrastructure services, retail, health, and entertainment services. Modern tourism has developed into a comprehensive industry that integrates the six elements of tourism, such as "food, accommodation, transportation, visit, shopping, and entertainment", and can satisfy people's spiritual needs. Among them, the scenic spot is the core element of tourism and is the main component and center link of tourism products and the tourism industry chain. Scenic area passenger ropeways and sightseeing bus services are important tourism facilities and services that enhance tourists' experience and enrich the mode of tourism. Tourism transportation is the "link" of the whole tourism industry, providing tourism services for tourists. Hotels are an important tourism infrastructure, providing accommodation, catering, leisure, and entertainment services for tourists. Tourist groups organized by travel agencies are one of the important sources of hotels, tourist transportation, and attractions. With the continuous development of the Internet, the composition of tourists is mostly casual. Tourists collect information on tourism resources on the Internet and plan the travel routes by themselves. This means that the role of travel agencies is gradually weakened, and the social service function of data visualization is gradually enhanced in the tourism industry [4].

### **2.3. Current Status of Data Visualization Use in The Tourism Industry**

With the increasing travel demand, most young people are opting for personalized travel and free travel arrangements. People usually choose the Internet to find tourism resources. However, much travel information on the Internet suffers from hidden promotions, lack of innovation, and vague prices, which makes it difficult for tourists to distinguish whether an attraction is worth visiting. Travel apps have sprung up, providing great convenience for people's travel planning. Common functions of these apps include tourist attraction recommendations, travel route planning, local specialties, and travel experience sharing. Through data analysis and visualization, they can provide a reference for future tourists' travel planning.

Generally, the relevant APP acquires a large amount of tourism information through crawler technology, cleans and screens the information, and selects the information that users need. Then, through the interface provided by the electronic map, the acquired information is displayed on the map, which helps users plan their trips and further improves tourists' sense of travel experience and satisfaction. A large number of tourism resources are provided on the Internet. The popular tourism website contains photographs of many tourist attractions, tourist series produced by others, the order of various tourist locations, and ratings of different tourist attractions. Users can access tourist information and visualization data from this site. The map shows the location of all tourist attractions. Each location has a sign displayed on the map. When the user clicks on these tourist signs, links to related descriptions and books are displayed. The color of the logo changes according to the popularity of the tourist attraction [5].

## **3. Methods of Using Data Visualization in The Tourism Industry**

### **3.1. Tourist Attractions' Use of Data Visualization**

Visualization is widely used in the tourism industry for tourism planning and destination promotion. Through technologies such as Geographic Information Systems (GIS) and Virtual Reality (VR), tourism practitioners can present information about a destination to potential tourists in the form of graphics and animations [6]. In this way, tourists can have a more intuitive understanding of the destination's attractions, culture, transportation, and other related information, so that they can make more informed travel decisions. At the same time, tourism organizations can also use visualization technology to design attractive promotional materials, such as beautiful maps, 3D models, and videos, to attract more tourists to their destinations.

Visualization is also valuable for forecasting and analyzing tourism activities. By collecting and analyzing large amounts of data, such as tourists' historical behaviors, preferences, and social media data, tourism practitioners can use visualization tools to reveal patterns and trends hidden in the data. The results of these analyses can help them better understand market needs, improve products and services, and develop more precise marketing strategies. In addition, visualization can help the travel industry predict foot traffic, optimize route planning and resource allocation, and improve operational efficiency.

Visualization technology also plays an important role in tourism studies. Through virtual reality technology, students and practitioners can immerse themselves in different tourism scenarios, such as monuments, natural landscapes, or cultural activities. This immersive experience can enhance learning and improve understanding of the characteristics and culture of the destination. At the same time, visualization can also help tourism practitioners receive online training through interactive diagrams, images, and videos so that they can learn and acquire professional knowledge more interestingly.

Visualization technologies can also provide value-added services to the travel experience. For example, travel apps can utilize augmented reality to provide navigation and narration features at attractions, making it easier for tourists to explore and learn about the destination. In addition, virtual tour guides can also be enabled by visualization technologies, allowing tourists to interact with virtual

tour guides via smartphones or VR devices to access more detailed tour information and stories. These innovative visualization applications provide a personalized and enriched tourism experience that enhances tourists' engagement and satisfaction.

### **3.2. Visitors' Use of Data Visualization**

Hotels utilize visualization technology to offer a range of services, including room selection, transportation navigation, personalized suggestions from AI travel assistants, management of tourism activities, and analysis of visitor feedback and social media. These services aim to assist travelers in planning their trips more effectively, understanding tourism resources, and enhancing their overall travel experience.

First, when it comes to the hotel and accommodation experience, by using visualization technology, hotels can show guests different types of rooms and suites, including the layout, furnishings, and decorative style. This can help guests better choose the right accommodation option and get a feel for the stay ahead of time. Second, in terms of transportation and navigation, visualization technology can provide real-time traffic information, route planning, and navigation features in mobile apps or navigation systems. Travelers can easily find the best routes, avoid congestion, and know the location and schedule of public transportation. Meanwhile, AI travel assistants can also be used for data visualization. AI travel assistant is an application that uses artificial intelligence technology to provide travelers with personalized, efficient, and safe travel services. By integrating various data sources and visualization tools, AI Travel Assistant can provide travelers with personalized suggestions and recommendations, including attractions, restaurants, shopping, and activities. This visual presentation of information helps travelers better plan their trips and discover new experiences that may be of interest to them. Then, regarding travel activity and event management, visualization tools help travelers manage and coordinate various travel activities and events, including booking management, resource allocation, and scheduling. Through intuitive interfaces and charts, tourists can better track and monitor the progress of activities. Finally, visualization technology can play a role in visitor feedback and social media analysis. Through visualization technology, tourists can provide real-time feedback on tour information on social media so that other tourists can understand what they think of the service and experience. These analytics can help other tourists save time on information collection and receive real-time information on the latest tourism resources to arrange and plan their travel trips more effectively.

## **4. Suggestion**

### **4.1. Advantages of Using Data Visualization in The Tourism Industry**

Data visualization enhances the decision-making efficiency of tourists through more intuitive data. Data visualization makes complex tourist attraction information, hotel information, and transportation and travel information clear, easy to understand, and compare, helping consumers make quick choices. It can provide intuitive images and charts to help tourists understand and filter the travel itinerary, find a more suitable way to travel and save consumers' decision-making time.

Data visualization can also help travel practitioners predict future trends in the industry. Through data visualization, travel companies, tourist attractions, etc. can see past and present trends in tourism consumption, tourists' preferences for travel destinations, and other information, and predict future developments. This helps industry players to prepare in advance and make improvements to tourist attractions and routes, etc., by keeping up with market preferences, thus promoting the efficiency and sustainable development of the tourism industry.

In addition, the application of data visualization can enhance the experience of tourism consumers. Data visualization can help tourism practitioners understand customers' needs and preferences for tour itineraries and excursions, etc., to provide more personalized services. Through data visualization, travel companies and tourist attractions can better understand tourists' feedback and satisfaction with their travel itineraries and thus improve the quality of tourism services.

Data visualization facilitates tourism planning. For tourism practitioners, data visualization allows them to better understand the needs and potential of tourist destinations and thus develop more effective planning for tourist attractions. For example, scenic spots can use maps and charts to show the tourist flow, distribution of attractions, and traffic conditions of a destination. For tourists, data visualization can be used to present information about tourist attractions, routes, and other information, as well as feedback from other tourists on their travel itineraries, so that they can make better reference to plan their travel itineraries.

#### **4.2. Shortcomings in The Use of Data Visualization in The Tourism Industry**

Python is slow in data processing for tourism data visualization. Currently, Python-based data visualization has the problem of slow speed when processing large amounts of tourism data. Since Python is an interpreted language, its execution efficiency is relatively low, especially when dealing with large-scale data such as tourist attractions or transportation itineraries. This provides inconvenience for real-time decision-making by tourism practitioners and tourists.

Python's visualization results lack interactivity. Python's visualization libraries (e.g., Matplotlib and Seaborn), although powerful, are relatively weak in terms of interactivity. The data related to the tourism industry is constantly updated, and it is necessary to be able to interact with the visualization results and adjust the parameters on the fly to obtain more accurate data analysis results. In addition, imperfections in the filtering and sorting of data can have an impact on the decision-making of tourism practitioners and tourists.

Python presents a more homogeneous style of data visualization. Python-based data visualization in the tourism industry usually uses common statistical charts, such as pie charts, bar charts, and line graphs. This single style of visualization makes the presentation of information lack innovation and personalization and fails to meet the diversified display of different data types and needs of the tourism industry.

#### **4.3. Improvements in The Use of Data Visualization in The Tourism Industry**

High-performance computing libraries as well as optimization algorithms can be introduced to improve data processing efficiency. To solve the problem of slow data processing, high-performance computing libraries such as NumPy, Pandas, etc. can be introduced when using Python. These libraries can make better use of CPU and memory resources, which can speed up data processing and analysis and improve data processing efficiency. More efficient data structures and algorithms can also be used to implement the data processing and mapping process to reduce the cost of time and increase the speed of data processing.

Development of interactive visualization tools. To enhance the interactivity of the visualization results, interactive visualization tools based on Python can be developed. For example, with the help of libraries such as Plotly, interactive maps, timelines, etc. can be realized, enabling tourism practitioners and tourists to adjust and modify the data in real-time according to their needs.

Use diverse visualization representations. To enrich the visualization style, try exploring other types of visualization charts. For example, different types of charts such as word clouds and heat maps in Python libraries to present tourism data. In addition, VR and AR technologies can be combined to provide tourists with a more immersive and personalized data visualization experience.

In addition, other domain technologies can be combined to improve the use of data visualization in tourism. When improving the data visualization tools in Python, other related domain technologies, such as GIS and AI, can also be combined to provide more comprehensive and accurate data analysis for tourism practitioners and tourists. For example, GIS technology is used to visualize geographic locations.

## 5. Conclusion

This paper analyzes and researches the use and advantages of data visualization in the tourism industry, analyzes and summarizes its shortcomings, and gives corresponding improvement measures. In this paper, it is found that the use of data visualization in the tourism industry involves a wide range of advantages over ordinary data presentation. For tourism practitioners, data visualization is used in the promotion of tourist attractions, the prediction of tourism activities, tourism study and the enhancement of tourism services. For tourists, data visualization makes it more intuitive and convenient to choose hotels and accommodations, transportation, travel planning, and feedback and sharing of travel experiences. However, at the same time, there are some problems in the use of data visualization in the tourism industry, such as insufficient interactivity, single visualization style, etc. According to these problems, this paper also provides some suggestions and improvement measures, such as optimizing algorithms, introducing interactive visualization tools and combining them with other fields of technology. Future research should explore additional computer languages and methods for visualizing tourism information. This exploration will broaden the application of data visualization in the tourism industry. Furthermore, integrating various industries with tourism can provide more intuitive and effective data presentations for both tourism practitioners and tourists, as well as individuals in related fields. Python-based data visualization will play a greater role in the tourism industry in the future, providing better and more intuitive data support for tourism practitioners and tourists, thus promoting the sustainable development of the tourism industry.

## References

- [1] Sun C., Wang Y. Python-based tourism data visualization applications. *Computer Programming Skills & Maintenance*, 2022 (11): 100 - 103.
- [2] Li X., Zhao Y., Xiao Z.L., et al. Analysis of data analysis and visualization design of tourism website based on Python crawler. *Computer Knowledge and Technology*, 2022, 18 (33): 58 - 60+70.
- [3] Zhang L., Sun Y.F. Visual analysis of Yuncheng tourism data based on Python. *Computer Era*, 2022 (10): 85 - 88. DOI: 10.16644/j.cnki.cn33 - 1094/tp.2022.10.021.
- [4] Xiao, H. Data Visualization on Tourism. In: Nguwi, Y.Y. (eds) *Tourism Analytics Before and After COVID-19*. Springer, Singapore. 2023.
- [5] Zhao Q. Data analysis and visualization of travel website based on Python crawler. *Electronic Design Engineering*, 2022, 30 (16): 152 - 155.
- [6] Solazzo, G., Maruccia, Y., Ndou, V. et al. How to exploit Big Social Data in the Covid-19 pandemic: the case of the Italian tourism industry. *Serv Bus* 2022, 16, 417 – 443.