Big Data Analysis in Consumer Behavior: Tourism, Household Energy and Mobile Payment

Kunliang Xiao*

School of Mathematics, Shandong University, Jinan, China

*Corresponding author: 202200091131@mail.sdu.edu.cn

Abstract. As a matter of fact, with the improvement of people's living standard, research on consumer behavior is increasingly emerging. Among them, the use of big data to study consumer behavior has become an important part of today's consumer research. With this in mind, through the study of big data in three specific aspects (i.e., tourism consumption, household energy consumption, as well as mobile payment consumption), this article compares the models of big data in studying consumer behavior. According to the analysis, it reflects the high efficiency, speed, and convenience of big data. In addition, the implementation of big data analysis is feasible in the consumer behavior analysis. At the same time, the current limitations for the state-of-art applications and models are analyzed and demonstrated. In the meantime, the future prospects and improvement suggestions are proposed accordingly. Overall, these results shed light on guiding further exploration of consumer analysis in terms of the big data techniques.

Keywords: Big data; consumer behavior; tourism; household energy; mobile payment.

1. Introduction

After capitalism completed the industrial revolution, the economy developed rapidly, especially the commodity economy. The originally relatively stable market problems have become increasingly acute [1-3]. Major companies and groups are in tit-for-tat confrontation, and consumers have gradually become the focus of corporate competition. At the same time, research on consumer behavior has also made initial progress. From the end of the 19th century to the 1930s, major breakthroughs were made in the study of consumer behavior, and theories about the characteristics of consumers' psychological activities and behavioral rules began to emerge. By the beginning of the 20th century, the United States had greatly improved labor productivity thanks to the dividends of the Industrial Revolution. The production capabilities of major enterprises and groups had made a qualitative leap. Market demand could no longer meet the products produced, and the market was oversupplied, which led to a crisis among enterprises and a sharper competition. In such an environment, companies began to pay attention to consumer groups, try their best to meet consumer needs, and began to promote products. Various sales methods and advertising techniques began to show their talents on the corporate competitive stage. At the same time, some scholars began to theoretically study the methods by which companies satisfy consumers and consumers' choices about companies, forming the prototype of consumer behavior [4].

The study of consumer behavior is of great significance, which is reflected in three aspects: enterprises enhance marketing effects, consumers engage in scientific consumption, and improve macroeconomic decision-making levels. According to the research on consumer behavior, enterprises can grasp the psychological reactions and behavioral rules of consumers, which is beneficial to enterprises in providing services to consumers through market. The essence is to use various marketing methods and diverse promotion methods, which stimulate consumers' desire to buy and allow consumers to realize their purchasing behavior. Studying consumer behavior can enable consumers to rationally understand their own consumption characteristics and patterns, which is conducive to consumers' reasonable consumption and optimization of consumers' consumption behavior. At the same time, consumers’ own personality characteristics, interests and preferences, attitudes towards life, cognitive concepts, etc [5]. will all have an impact on consumers’ behavioral patterns and psychological characteristics. Therefore, consumer behavior research is more conducive to improving the quality of social consumption and the quality of life for consumers. Studying consumer behavior can better control and regulate the market and improve the level of macroeconomic decision-making [6]. For example, when the consumption preferences of most consumers generally change, the market will be in a situation of supply exceeding demand, resulting in many consumers being unable to meet their own consumption needs. The government can intervene in research on consumer behavior and regulate corporate production [7].

This article studies the application of big data in consumer behavior and adopts the following structure. Section 2 gives a basic description of the development of big data in consumer behavior. Sections 3, 4, and 5 introduce the application of big data in studying consumer behavior from three aspects: tourism consumption, household energy consumption, and mobile payment consumption respectively. Section 6 discusses the shortcomings of this study and suggestions for future researches, and Section 7 summarizes the entire article. The research in this article analyzes the application of big data in consumer behavior, provides ideas for studying consumer behavior, and also provides a research direction for big data applications.

2. Basic Descriptions

With a new round of information technology and technological revolution, mankind has entered the era of big data. The data generated by people every day, such as shopping information, browsing history, etc., are being widely collected and saved. Big data has penetrated into every aspect of people's lives and continues to provide people with more convenient and efficient services. IBM proposed the 5V characteristics of big data, namely Volume, Velocity, Variety, Value, and Veracity, which fundamentally revealed the nature of the rapid development of big data in this era and pointed out the direction for in-depth research on big data in the future. In the current research, big data is divided into three types, namely structured data, semi-structured data and unstructured data, among which unstructured data occupies the main part of big data. Unstructured data refers to data that records user operation behaviors and processes. The difference from structured data is that it is real-time and diverse. The main difficulty in current big data processing is not how to obtain a large amount of data, but how to professionally process the data, that is, to extract valuable data from a large amount of low-quality data to obtain better analysis decisions. At present, big data is widely used, such as prediction of crime rate, prediction of the spread of the new corona virus, improvement of urban traffic conditions, etc. [8].

Big data is becoming a powerful weapon for analysis and research in the new era. There are three main models for big data research on consumer behavior: U&A model (Usage and Attitude Research), which mainly studies consumers' attitudes and usage habits towards goods. It is widely used in the research of disposable consumer goods such as food and beverages, fast food and snacks. The AIDMA model was proposed by American advertising scientist Lewis in 1898. It describes the five stages that consumers go through from understanding to actual purchase: Attention, Interest, Desire, Memory, and Action. The next is consumer user value analysis model, which mainly judges the user's economic value based on the difference between the enterprise resources consumed by the user and the value created for the enterprise. Meanwhile, the user's market value is calculated based on the user's praise and loyalty. It divides users into 4 categories: high market value and high economic value; low market value and high economic value; high market value and low economic value; low market value and low economic value. Then enterprises can use different marketing methods for different categories [9].

3. Applications in Tourism

As people's living standards continue to improve, the tourism industry has gradually prospered, and data analysis has penetrated into all aspects of the tourism industry. The rise of big data in recent years has injected vitality into the tourism industry, making the tourism industry that originally developed steadily become more prosperous. The rapid development of the tourism industry has promoted research on tourism consumer behavior. Studying tourism consumer behavior is conducive to evaluating and discussing consumer behavior, rationally planning consumers' travel methods, and alleviating overtourism and the number of tourists in tourist destinations. At the same time, one creates a harmonious tourism atmosphere for tourist destinations, provide consumers with high-quality tourism services, and avoid the pain points of overcrowded tourism.

Big data plays an important role in studying tourism consumer behavior from three aspects: tourists' own consumption behavior, tourism market consumption situation, and green and low-carbon tourism consumption. In terms of tourists' own consumption behavior, big data analyzes tourists from three characteristics: basic attributes of tourists, tourist origin, tourist behavior and preferences. Describing the basic attributes of different tourists, including gender, age, education, occupation, etc., from different angles, one can have different understandings of tourism consumers. Analyzing tourist source areas can determine where tourists gather from, and analyze the extent to find which tourist attractions attract tourists from different regions. For example, an attraction is more likely to attract tourists from areas with a completely different climate, scenery, and culture. Due to this difference and people's general curiosity, tourists from other regions are more likely to visit this attraction. In terms of tourist behavior and preferences, different tourists have different preferences for things, but the things that a tourist like are related. For example, tourists who are interested in humanities and history are more likely to go to historical sites, local folk customs and other cultural places, but they rarely go to scenic spots such as mountains, flowing water, famous mountains and ancient trees. By grasping consumers' consumption preferences, one can better carry out precise marketing and help consumers locate travel methods that suit them without being disgusted.

Regarding the consumption situation in the tourism market, big data can divide the tourism market into regions and count the number of consumers and consumption levels in each region. Appropriately reduce consumer prices in areas with higher consumption levels to provide consumers with reasonable and affordable consumer services. Big data can also compare consumer behavior in different regions, formulate different marketing strategies to meet different consumer behaviors, make marketing more precise, and improve the quality and efficiency of marketing. At the same time, big data can connect the central consumption system to other consumption subsystems, collect the consumption status of each subsystem in real time, and make predictions and analyzes of future consumer behavior.

For green and low-carbon tourism consumption behavior, big data can collect the consumption accommodation level of different tourist locations and test the green consumption level of the tourist location. Implement consumer behavior constraints in areas with low levels of green consumption. For example, one can reduce or ban consumers’ consumption of flammable and explosive items, highly polluting items, and difficult-to-degrade items, improve consumers’ green consumption levels, and regulate consumer behavior from a macro perspective. Currently, the most popular theoretical models using big data to study tourism consumer behavior include the distribution center model, push-pull theory, optimal arousal theory, leisure incentive method and tourism career ladder model, among which push-pull theory is the most widely used [11]. Big data can also use UGC (user-generated content) to study tourism consumer behavior, collect tourists’ online travel notes, share pictures and videos to understand tourists’ personal preferences and consumption characteristics, thereby helping tourist attractions develop targeted tourism services to meet their needs. In the establishment of a tourism consumer behavior model, scholars Li and Cao analyzed the characteristics of tourists in the four continents of southern Xinjiang, including Aksu Prefecture, Kizilsu Kirgiz Autonomous Prefecture, Kashgar Prefecture, and Hotan Prefecture [11]. They used Train Browser and NLPIR to crawl, preprocess and mine tourist data from 500 online travel notes from online travel notes websites such as Mafengwo.com and Qunar.com [10]. They used a variety of analysis methods, mainly word frequency co-occurrence analysis, to obtain the per capita consumption level of tourists (seen from Fig. 1). Among them, per capita consumption ranges from 200 to 10,000 yuan, and is generally concentrated between 1,000 and 3,999 yuan (57.4%) [1]. According to the content of the travel notes, two reasons for low consumption are deduced: First, the distance from the destination is long, the number of days to visit is small, and the consumption is not high; the consumption is not high due to local relatives and friends or low-cost group tours [1].

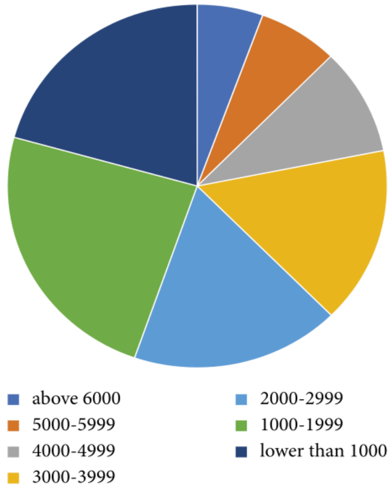

Fig. 1 Capita consumption.

4. Applications in Household Energy Consumption

As emerging information and digital technologies enter the energy industry, the energy system is gradually digitized, and big data has penetrated into all aspects of the energy industry. Energy plays a pivotal role in people's lives. The United Nations proposed in 2019 that energy in the new century must be reliable, sustainable, affordable, modern, and accessible to everyone. This is the seventh sustainable development goal. In sustainable development, the most important thing is to plan energy efficiently, ensure low energy supply costs, and minimize social and environmental costs. The most critical and effective way to improve energy planning and reconstruct energy composition is to emphasize the energy demand forecast model, because expected demand is exactly the definition of energy cost and can minimize energy planning.

In the energy demand prediction model, the most critical point is the prediction of household energy demand, because household energy is an important component of energy and has great planning potential in energy composition. The energy consumption structure is also affected by the new energy policies of relevant national departments. Incorporating energy consumption into the analysis of household consumption in more detail is of great significance to the analysis of household energy. However, household energy depends on household consumption structure. Household consumption structure is affected by people's lifestyle. At the same time, household income and expenditure also affect people's consumption and lifestyle. Therefore, the study of household energy consumer behavior is of great significance for the prediction of household energy.

Household energy consumption includes household direct energy and household indirect energy. Household indirect energy is explained as the energy contained in the goods and services purchased by households. These energies will not be used continuously like direct energy, but they actually far exceed household direct energy consumption. This is an important factor in household energy consumption. In addition, from the analysis of the supply and demand chain, households are the final destination of energy. Energy is completed from production to intermediate transportation, and finally flows into households. Households become the final users of energy. Therefore, in order to find out the decisive factors of changes in household energy demand, it is necessary to conduct a comprehensive analysis of direct and indirect household energy, so as to better guide the production department and transportation department to meet household energy needs and better improve the household energy consumption structure.

Currently, the models used to use big data to study household energy consumer behavior include multiple discrete continuous models, carbon emission coefficient models and quantile regression models. The carbon emission coefficient method is the most commonly used method in the study of carbon emissions from energy consumption, that is, multiplying energy consumption by the carbon emission coefficient to obtain the carbon emissions of the corresponding energy [11]. Quantile regression was first proposed by Koenker and Bassett. It is a modeling method that estimates the linear relationship between a set of regressors X and the quantiles of the explained variables Y. It is more robust to non-normal distributions [12]. The multiple discrete continuous model uses the saturation parameter to represent the utility structure in which marginal utility decreases as consumption increases, which shows better adaptability than the structure using only the logarithmic function. In their study of carbon emissions from household energy consumption, scholars such as Hu and Guo selected 2,455 household samples as observation objects and considered eight indicators: lnELE, lnincome, EDU, lnsquare, Appliance, Uice, Knowledge, and EsaveTV as given in Table. 1.

Table 1. Indicator.

lnELE	Per capita carbon emissions from household electricity consumption (logarithm)
lnincome	Household per capita income (logarithm)
EDU	Education level of family members
lnsquare	Per capita living area (logarithm)
Appliance	Total household appliances
Uice	Time of using the refrigerator
Knowledge	Ladder electricity price policy
EsaveTV	Way to turn off the TV

The carbon emissions of the sample are calculated using the carbon emission coefficient method, and then the heterogeneous impact of various factors on household electricity consumption is further explored through the quantile regression model [11]. The basic attributes and properties of 8 indicators were obtained. Since the VIF values of each variable are less than 5, it can be seen that there is no collinear relationship between the variables (seen from Table 2).

Table 2. Descriptions of the varible.

variables	Mean	Std.Dev.	Min	Max	VIF
lnELE	6.13	0.83	3.29	9.29	-
lnincome	9.60	1.16	4.20	15.20	1.28
EDU	0.15	0.35	0.00	1.00	1.15
lnsquare	3.70	0.96	1.39	6.46	1.04
Appliance	5.34	2.49	0.00	17.00	1.57
Uice	3.65	2.03	0.00	5.00	1.32
Knowledge	0.40	0.49	0.00	1.00	1.13
EsaveTV	0.17	0.37	0.00	1.00	1.00

5. Applications in Mobile Payment

With the rapid development of network technology and mobile phones, the number of mobile payment tools has shown a steady growth trend. Mobile payment is widely used in people's lives. Due to its rapid, efficient, convenient and safe, mobile payment is accepted and used by more and more people. The research content of mobile payment consumption behavior is the personal behavior and usage characteristics of consumers when using mobile payment tools. It studies the time, place and usage method of users when they use mobile payment, and then concludes the behavioral rule of users. This rule can be used to formulate targeted marketing strategies to meet the diverse needs of users. For example, if the data finds that the number and amount of payments made by users for a certain type of product are significant, then the manufacturer can promote more such products to users. This not only meets the needs of users for their own consumption, but also enables producers to obtain more benefits.

The models used to study mobile payment consumption behavior mainly include data-driven models and knowledge-driven models [12]. The data-driven model mainly analyzes the user's payment behavior by collecting user data when using mobile payment tools, such as payment time and location, payment amount, number of payments, etc. It focuses on the data itself, processes the data such as averaging, variance, etc., and uses mathematical tools such as interpolation and fitting to obtain the characteristics contained in the data column itself, and then predicts the user's consumption behavior and formulates the user's profile in advance. Sales strategy. The knowledge-driven model is to further explore and explain the obtained user data, and fundamentally explain the reasons why users produce this behavior. For example, if it is found that a user has recently purchased significantly more baby products, and using a data-driven model can only infer from the data itself that the baby products are showing a growth trend, then the only way is to adopt a marketing strategy to increase the promotion of baby products. Using the knowledge-driven model, one can deduce based on the event that the user has a child at home, and based on existing knowledge, one can deduce that there are pregnant women at home. Therefore, while increasing the sales of baby products, one should also promote maternity products.

Data-driven models can be divided into two categories. One type is to conduct statistical analysis on the data sequence itself, to obtain the patterns between the data, and to use statistical methods to analyze user behavior using the patterns. The other type is to orient the data source and the result to correspond, and find the relationship between the data source and the result through methods such as interpolation and fitting. Furthermore, machine learning methods can also be used to allow the machine to conduct in-depth learning on a part of the data, find out the relationship between the data source and the result, and then test the other part of the data to gradually obtain a higher matching relationship and make decisions on user behavior. predict. In a study of Indian consumer mobile payments by scholars such as Pushp Patil and Kuttimani Tamilmani, they collected 445 online questionnaires, 210 paper questionnaires, and a total of 655 usable response. They used confirmatory factor analysis (CFA) testing to analyze the measurement model by examining the convergent validity, discriminant validity, and internal consistency of the construct. They ultimately identified three significant positive predictors: performance expectations, behavioral intentions, and grievance adjustment. After analyzing it, the results were obtained (seen from Fig. 2) [12].

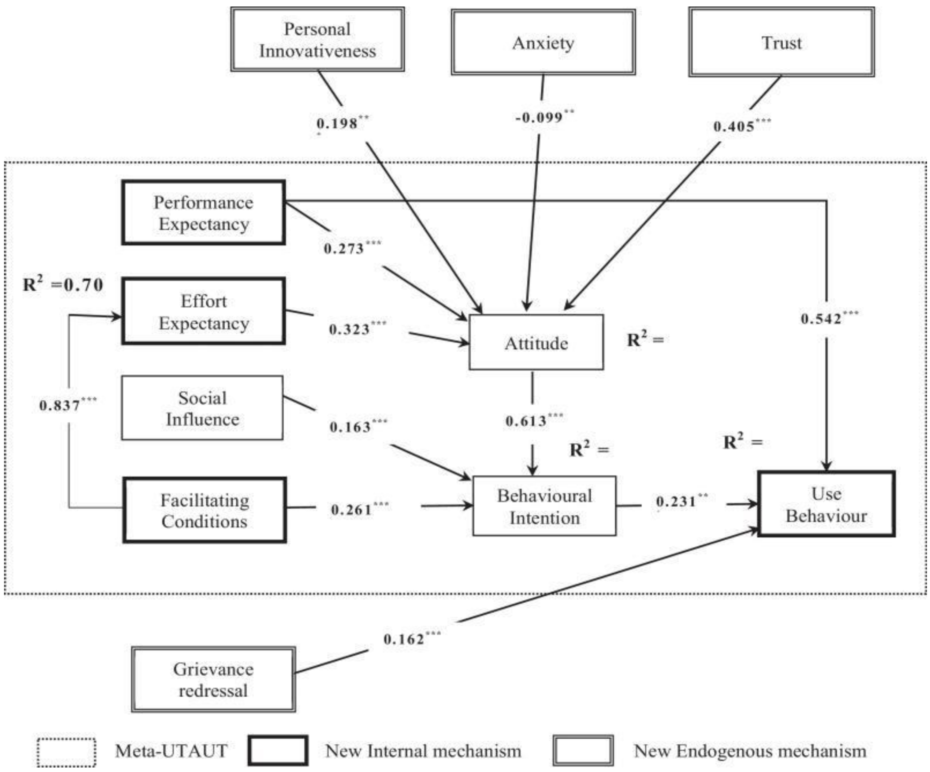

Fig. 2 Analysis results.

Facilitating conditions are the strongest predictor of effort expectation, with a path coefficient value of 0.837; predictive attitude is the strongest predictor of behavioral intention, with a path coefficient value of 0.613 [12]. Studying mobile payment consumption behavior has three implications for producers, consumers, and the government. For manufacturers, understanding consumer behavior can better promote precise sales to consumers, adapt to a more intense corporate competition environment, and obtain greater benefits. For consumers, understanding their own consumption behavior can help them understand mobile payments rationally and avoid falling into the vortex of mobile payment tools. They can also reasonably view merchants' mobile promotions and have a better experience when using mobile payment tools. For the government, studying mobile payment consumption behavior can identify the dependence of users in different regions on mobile payment tools, and govern and regulate the use of mobile payment tools in different regions, which is of great significance to regional planning and business district design [13].

6. Limitations & Future Outlooks

At present, big data research on consumer behavior still has many shortcomings, which are mainly reflected in three aspects: insufficient and inaccurate data sources, limitations of the research area, and failure to consider cross-disciplinary communication and comparison. On the one hand, the sources of data are insufficient and inaccurate because consumer data itself is private content, and consumers are unwilling to disclose it when conducting questionnaire surveys. When conducting information retrieval on major institutional platforms, the collected data is often not fully opened due to confidentiality measures. As a result, the full picture of the data cannot be analyzed, and only part of it can be intercepted to make inferences. On the other hand, there is a huge amount of consumer data, and errors and omissions will inevitably occur during storage, resulting in missing and inaccurate data. Moreover, the data collected from various websites often have certain tendencies. For example, some Internet writers and keyboard warriors post a large number of posts on the Internet and write articles to guide public opinion, which makes some consumers lose their way and fail to carry out the research. The operation of real consumption behavior causes the collected data to have a greater tendency and cannot reflect the real consumption behavior of a wide range of consumers.

The limitation of research area mainly means that most research on consumer behavior is limited to a certain place. For example, when conducting a questionnaire survey, only the consumer behavior in one area is studied, and the results can only explain this area. However, these conclusions may not necessarily apply to consumer behavior across the country or even around the world. The limitation of the research area is mainly due to the cost of investigation and geographical restrictions. Investigating consumer behavior in different regions involves the impact of different cultural environments and is more expensive. At present, the use of big data to study consumer behavior is still in its initial stage. It only considers using the data itself to explore consumer behavior using mathematical models, and does not involve in-depth knowledge of economics and psychology. The research is more concerned with whether the construction of the model itself is appropriate, without considering the communication and comparison between disciplines, and is single-minded.

In future research on consumer behavior, more psychological suggestions should be used, formal research methods should be adopted, and more accurate, effective and detailed information from consumers should be collected through certification by national agencies. Expand the scope of research to include ethnic minorities who are easily overlooked in the field of consumer behavior research. It is also necessary to consider the communication and integration between interdisciplines. For example, big data obtained from the perspective of economics will also have important application value in the future.

7. Conclusion

To sum up, this article explores the application of big data in different areas of consumer behavior research. The full article explores the significance of big data on consumer behavior research from three aspects: tourism consumption, household energy consumption, and mobile payment consumption. Comparative analysis of big data models for studying consumer behavior reflects the efficiency, speed, and convenience of big data, and highlights the important value of big data for studying consumer behavior. However, the research in this article still suffers from insufficient and inaccurate data sources, limitations in the scope of the research area, and failure to consider the limitations of cross-disciplinary communication and comparison. In future research, more accurate, effective and detailed information should be collected, the scope of research should be expanded, and the integration of disciplines should be considered. The research in this article can provide reference for future research on consumer behavior using big data, and can also summarize the application of big data models.

References

[1] Alcalde J. Implementation of stable solutions to marriage problems. Journal of Economic Theory, 1996, 69(1): 240-254.

[2] Panasyuk M V, Pudovik E M, Sabirova M E. Problems of labor market of modern Russia in conditions of stable economic growth. Life Science Journal, 2014, 11(6s): 487-489.

[3] Sotomayor M. Three remarks on the many-to-many stable matching problem. Mathematical social sciences, 1999, 38(1): 55-70.

[4] Kuokkanen H, Sun W. Companies, meet ethical consumers: Strategic CSR management to impact consumer choice. Journal of Business Ethics, 2020, 166: 403-423.

[5] Li S, Kallas Z. Meta-analysis of consumers' willingness to pay for sustainable food products. Appetite, 2021, 163: 105239

[6] Zukin S, Maguire J S. Consumers and consumption. Annu. Rev. Sociol., 2004, 30: 173-197.

[7] Foxall G R. Foundations of consumer behaviour analysis. Marketing theory, 2001, 1(2): 165-199.

[8] Fan J, Han F, Liu H. Challenges of big data analysis. National science review, 2014, 1(2): 293-314.

[9] Matz S C, Netzer O. Using big data as a window into consumers’ psychology. Current opinion in behavioral sciences, 2017, 18: 7-12.

[10] Li J, Cao B. Study on Tourism Consumer Behavior and Countermeasures Based on Big Data. Computational Intelligence and Neuroscience, 2022, 12: 612051.

[11] Hu Z, Guo J. The influencing factors on consumer of family carbon emission. Operation and Management, 2022, 5: 98-106.

[12] Yang X. Mobile payment recognition based on knowledge base and data analysis. Chongqing University of Posts and Telecommunications Ranking, 2020.

[13] Patil P, Tamilmani K, Rana N P, et al. Understanding consumer adoption of mobile payment in India: Extending Meta-UTAUT model with personal innovativeness, anxiety, trust, and grievance redressal. International Journal of Information Management, 2020, 54: 102144.