BILINGUAL PUBLISHING CO.
Pioneer of Global Academics Since 1984

Review of Educational Theory
https://ojs.bilpublishing.com/index.php/ret

ARTICLE

Campus Economic Analysis Based on K-Means Clustering and Hotspot Mining

Xiuzhang Yang[1,2] Shuai Wu[2] Huan Xia[2,3*] Yuanbo Li[1] Xin Li[1]

1. Guizhou Vocational College of Economics and Business, 558022, China
2. School of Information, Guizhou University of Finance and Economics, Guizhou, 550025, China
3. Key Laboratory ofEconomics System Simulation, Guizhou University of Finance and Economics, Guizhou,550025, China

ARTICLE INFO

Article history
Received: 13 April 2020
Revised: 17 April 2020
Accepted: 24 April 2020
Published Online: 30 April 2020

Keywords:
Machine learning
K-Means clustering
Data mining
Consumer behavior
Campus economy
Economic regionalization

ABSTRACT

With the advent of the era of big data and the development and construction of smart campuses, the campus is gradually moving towards digitalization, networking and informationization. The campus card is an important part of the construction of a smart campus, and the massive data it generates can indirectly reflect the living conditions of students at school. In the face of the campus card, how to quickly and accurately obtain the information required by users from the massive data sets has become an urgent problem that needs to be solved. This paper proposes a data mining algorithm based on K-Means clustering and time series. It analyzes the consumption data of a college student's card to deeply mine and analyze the daily life consumer behavior habits of students, and to make an accurate judgment on the specific life consumer behavior. The algorithm proposed in this paper provides a practical reference for the construction of smart campuses in universities, and has important theoretical and application values.

1. Introduction

With the advent of the era of big data and the development and construction of smart campuses, the campus is gradually moving towards digitalization, networking and informationization. In order to achieve the digital construction of the campus, the campus card, as the basic project and an important part of the smart campus, is widely used in universities. At present, universities have established a com-

**Corresponding Author:*
Huan Xia (1981-), male, Hunan Province, China, Ph.D., professor, main research direction: computer simulation, big data analysis.
About the Author:
Xiuzhang Yang (1991-), male, Guizhou Province, China, master's degree, main research direction: Web data mining,Image Processing;
Shuai Wu (1994-), male,Jiangsu Province, China, master's degree, main research direction: machine learning;
Yue-qi Dou (1995), female, Shanxi Province, China, master's degree, main research direction: machine learning;
Nan Cong (1997), female, Shandong Province, China, master's degree, main research direction: machine learning.
Fund Project:
The authors acknowledge Science and Technology Project of Guizhou Province of China(Grant QKHJC[2019]1403) and (Grant QKHJC[2019]1041). Guizhou Province Colleges and Universities Top Technology Talent Support Program(Grant QJHKY[2016]068).

DOI: https://doi.org/10.30564/ret.v3i2.1837

prehensive campus card system [1]. The data described by campus card in different places, times and atmospheres have different values [2]. Therefore, mining the consumption data of the campus card can roughly reflect the students' daily consumption behavior habits, calculate the core components of the campus economy, and provide a practical reference for the optimization construction of smart campuses in universities.

Traditional consumer behavior surveys are mainly based on questionnaire surveys. The shortcomings of this survey method are mainly: the amount of data is small, and there is no continuity. It can't make effective predictions. The sorted data has the disadvantages of one-sidedness and poor timeliness. Aiming at the shortcomings of the questionnaire survey, this paper proposes a hot data mining algorithm based on K-Means clustering, which deeply mines the consumption data of the campus card, integrates the scattered consumption data into specific categories through K-Means clustering, and deep mine the potential value of consumption data. This algorithm deeply explores the student's consumption habits and library borrowing habits. It has certain guiding significance for the construction of university libraries and life services, and provides a practical reference for the better construction of university smart campuses. The algorithm the application of this method has important theoretical and practical value to the research of campus economy.

2. Current research situation at home and abroad

Campus economy, as an important part of economic regionalization, is a manifestation of the development of economic regionalization, and has become the focus of research by many scholars at home and abroad. Xu Juan [3] used questionnaires to investigate the views of students in higher vocational colleges; Zhang Chunhua [4] and others used 448 questionnaires combined with a factor analysis algorithm to analyze the influencing factors of college students 'online gaming consumption behavior; Liu Shangjun [5] and others passed questionnaires Investigate and combine Excal and Spss to analyze the consumption behavior of liquid milk from the perspective of product marketing 4p; Ren Jinhua [6] based on the principle of information entropy to construct an index to quantify the law of student consumption activities.

In recent years, domestic and foreign scholars have dedicated themselves to researching various fields through data analysis and machine learning technology. Liu Jian [7] applied machine learning and data analysis to tumor gene research, and proposed a double clustering algorithm based on regularized principal component analysis of double hypergraphs; Li Jianwei [8] analyzed online learning risks and made corresponding predictions through logistic regression ; Ye Yu [9] and others performed greening visibility data analysis based on machine learning to achieve large-scale analysis and obtain high-precision results; An Qiangqiang [10] and others based on machine learning and applied K nearest neighbor algorithm to improve the communication network The structured task processing efficiency and the clarity of the information feature portrayal; Qi Xuedan [11] based on machine learning and random forest algorithm to analyze the classification of Alzheimer's disease in depth.

At present, the research on campus economy mainly uses the traditional survey methods such as questionnaire surveys and visits to analyze the core distribution and composition of campus economy. Such survey methods have the disadvantages of contingency, small data sets, and low authenticity, reflecting the subjectivity and one-sidedness of the data, and cannot truly reflect the underlying core value of the data. The big data analysis system based on machine learning has the advantages of massive data, diverse data, fast speed, and high value, and can objectively and truly reflect the underlying core value of data.

Aiming at the shortcomings of the questionnaire survey and the advantages of big data analysis, this paper proposes a data analysis algorithm based on K-Means clustering and hotspot mining for campus economy, combined with data visualization technology to show the college students 'consumption habits, and deeply explore the college student 's consumption Habits, the core component of the campus economy is calculated, which can effectively optimize the allocation of resources for smart colleges and universities at the same time, and better fit the daily habits of college students.

3. Campus Economic Research Framework Model

3.1 Algorithm Framework Model

As a special consumer group, college students' consumption behavior has always attracted much attention. The purpose of this paper is to analyze and research the data of college student's campus card consumption data and library borrowing books, to deeply dig out the student's consumption habits, and to better analyze the college student's consumption based on K-Means clustering algorithm Behavioral habits, calculate the core components of the campus economy, and build the system

DOI: https://doi.org/10.30564/ret.v3i2.1837

framework diagram shown in Figure 1, the main steps are as follows:

(1) Collect relevant data of campus card and save it in csv file.

(2) Collecting data for data preprocessing, including: data integration, data cleaning, and outlier processing.

(3) Perform multi-dimensional data mining analysis on campus card consumption and library borrowing, and combine data visualization technology to present the analysis results in the form of charts.

(4) Relying on book borrowing data and campus card consumption data, combined with WordCloud word cloud analysis to predict students' habits and rules of library consumption behavior.

(5) Based on the K-Means clustering algorithm in Python environment to predict and analyze student consumption.

(6) Combined with hidden information such as consumption habits, a set of reference indicators for school management and decision-making on students' consumption behavior habits.

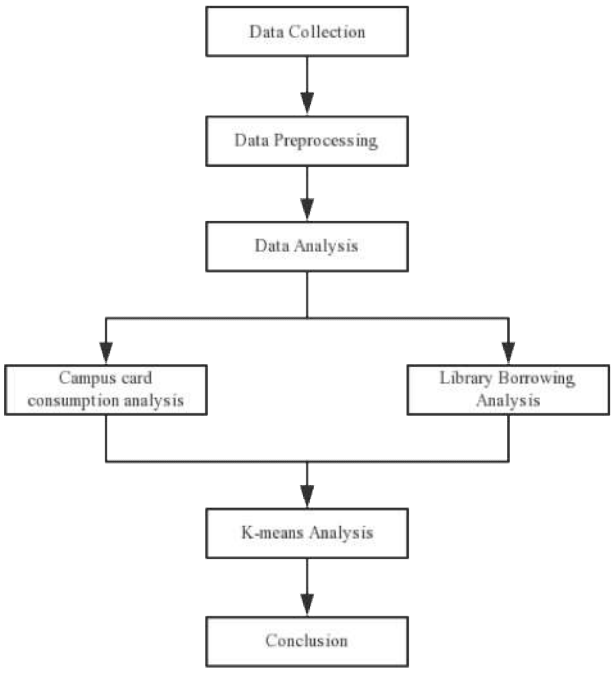

Figure 1. Campus Economic Analysis Framework

3.2 Data Preprocessing

The purpose of this article is to analyze the card data of colleges and universities, so there is no need to crawl the data, and directly call the historical consumption data in the school database. However, the initial data has problems such as inconsistent data, high dimensions, and redundancy, which affect the results of data analysis. In order to improve the quality of data analysis, this paper uses common data preprocessing methods to perform data preprocessing on the collected data, mainly including: data integration and data cleaning. In order to facilitate the comparative analysis of data on daily consumption behavior of college students, cluster analysis of book borrowing and K-Means cluster analysis of daily behavior habits of various students.

4. Comparative Analysis of Campus Economic Behavior

Students' daily consumption patterns are diversified, and various factors can lead to differences among consumers. This article will make a comparative analysis of the campus card data of students in each major of the 14-level Information Institute, and dig deep into the factors that affect the differences in consumption between majors. This article will conduct a multi-dimensional detailed analysis based on the historical data of campus card consumption, and conduct a comparative study and analysis of student consumption from multiple angles. It mainly includes: a comparative analysis of the consumption of students in various majors, a comparative analysis of the consumption of students in each canteen, daily consumption behavior and consumption Trend analysis, from the results of the analysis to obtain student consumption information of practical value.

4.1 Professional Comparative Analysis

The purpose of this article is to analyze and study the differences between the consumption behaviors of students in various majors. Based on the visual analysis of Echarts, a statistical chart of the frequency of consumption in each major shown in Figure 3 is drawn.

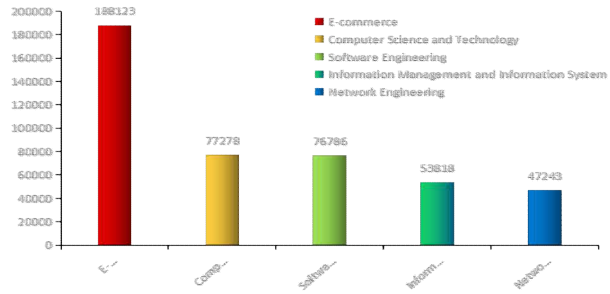

Figure 2. Statistics of Consumption Times by Major

Figure 2 shows the total number of times students spend in each major. The highest number of e-commerce purchases is 188,123. The lowest number of network

DOI: https://doi.org/10.30564/ret.v3i2.1837

engineering purchases is 47,243. Because the number of majors is not uniform, the data analysis is one-sided. Based on the number of students in each major, the statistical analysis shown in Table 1 is performed. According to Table 1, it can be seen that the average consumption times of the students in the five major information management and information systems are the largest, about 1794 times per person; the average consumption times of the students in the software engineering department are the lowest, average 1280 times per person; the consumption times of the other three majors are basically flat, All around 1500 times.

Table 1. Per capita consumption statistics by specialty

Professional Title	Total consumption	Number of people	Per Capita
E-commerce	188123	121	1555
Computer Science and Technology	77278	50	1546
Software Engineering	76786	60	1280
Information Management and Information System	53818	30	1794
Network Engineering	47243	30	1575

From the number of consumption, it can be predicted and analyzed that the students of the Information Institute have a higher dining rate in the cafeteria and are accustomed to eating in the cafeteria. Information management and information system students have the highest dining rate in the cafeteria. Software engineering students may be accustomed to ordering takeaways or going to off-campus stores for meals.

4.2 Canteen Comparative Analysis

College students generally have a strong subjective consciousness and a sense of consumption. This article aims to deeply explore the potential consumption behavior habits of school students. The data analysis is based on the four canteens of a university as a consumption point. Based on Echarts data visualization software, the pie-like statistics shown in Figure 5 Illustration. It can be seen from Figure 5 that the consumption proportion of each canteen is different: among them, the C1 canteen has the highest proportion, which is 36.6%; the second, which is basically the same, is the C3 canteen, which is 35.3%; the C4 canteen, which is 20.80%; C2 canteens accounted for the lowest proportion, only 7.3%. It can be roughly speculated that the students in this university are generally used to go to C1 and C3 canteens, which are the main consumption places of the school.

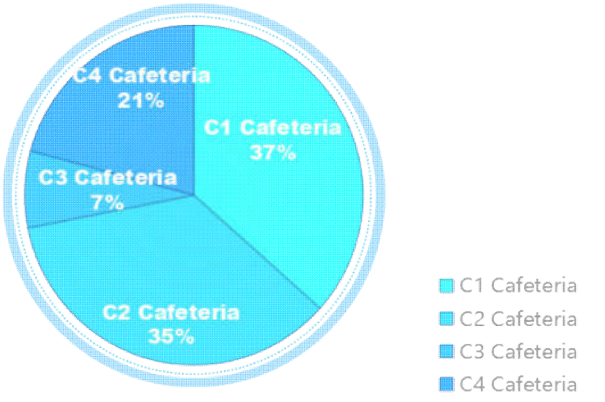

Figure 3. Consumption Distribution by Canteen

The potential self-condition factors of the canteen may cause the consumption behavior of school students to lean towards the canteen. The purpose of this article is to deeply explore the potential factors that affect the behavioral habits of students in the school. Based on the geographical location of each canteen, the price of meals, and the taste of the meals, the situation table shown in Table 2 is used to analyze the potential factors. Combining with the analysis of the dining distribution of each canteen in Figure 3, it can be seen that the price of meals is not the main factor affecting the dining habits of students. The main factors are the geographical location of the canteen and the taste of the meals.

Table 2. Basic evaluation indicators of each canteen

Cafeteria	Dormitory distance	Classroom distance	Meal prices	Meals taste
C1 Cafeteria	Relatively close	Relatively far	Cheap	Delicious
C2 Cafeteria	Near	Near	Expensive	Unpalatable
C3 Cafeteria	Near	Near	Expensive	Delicious
C4 Cafeteria	Relatively far	Far	Moderate	General

4.3 Comparative Analysis of Consumption

With the continuous development of the Internet, online consumption, mainly online shopping and takeaway, has replaced physical consumption on campus. In this paper, the statistical analysis is made on the number of campus students' consumption on the 14th level of the School of Information. Consumption is still the main consumption method of school students.

DOI: https://doi.org/10.30564/ret.v3i2.1837

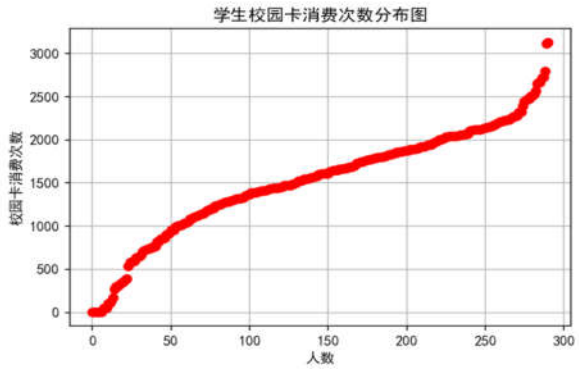

Figure 4. Distribution of campus card consumption

This article analyzes the time of three meals at the peak time of consumption 24 hours a day, and draws a peak time line chart shown in Figure 6 based on Echarts. The horizontal axis is the 24-hour time axis within one day, and the vertical axis is the number of consumers at that point in time. It can be seen from FIG.5 that there are three peak meal periods in the line chart, which are 6 to 8 am in the morning; 10 to 12 noon and 4 to 7 pm, corresponding to morning, lunch and dinner. It can be roughly inferred from the high peak that some students do not go to the cafeteria for breakfast, especially breakfast.

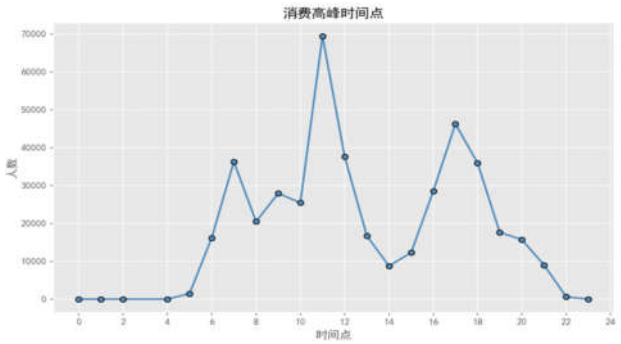

Figure 5. Peak consumption time distribution

With the continuous development of society, people's living standards continue to improve, and the consumption of students has also developed in various ways. Most students are not self-sufficient, and the main source of finance is mainly provided by their parents. It can roughly analyze the household economic situation from the student's consumption level at school, in order to select the targeted poverty alleviation target. Rely on Python to draw a scatter plot of the consumption amount shown in Figure 7. The horizontal axis represents the number of people, the vertical axis represents the total amount of student consumption at school, and the blue scattered points represent the amount of consumption. According to Figure 6, it can be seen that the overall consumption level of students is generally around 6000 to 10,000 yuan; 50 people are above 16,000 yuan, it can be roughly estimated that these 50 people have a good family situation; there are 2100 people who spend less than 2,000 yuan, and these 2100 people can Candidates for targeted poverty alleviation.

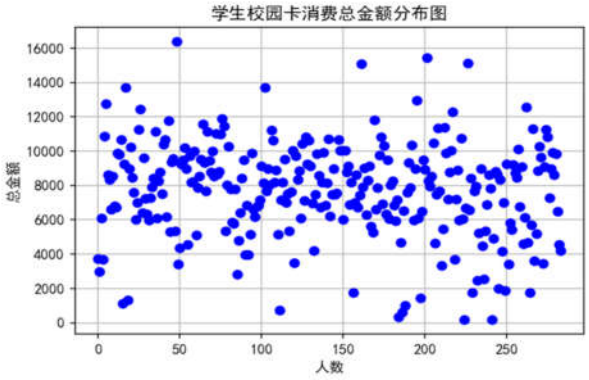

Figure 6. Scatter plot of total consumption

5. Library Borrowing Analysis

As one of the three pillars of school running, the university library is mainly responsible for the teaching and research tasks of the school, and it is an important place to improve the quality of students at school. The library, as the main place for students to study outside the school, not only provides a place for students to study, but also provides books and literature for students to read and research. Library book borrowing and student entry time can roughly reflect the main learning direction and content of the student during that time. This data mining cannot confirm the main learning direction of students. It has been recommended that libraries purchase this type of books for students to read and optimize the library book purchase mechanism. At the same time, it can continuously provide books for the main directions of students, further improving the students at school. Quality training.

WordCloud displays the "keywords" that appear frequently in the text from a visual angle. It combines Chinese word segmentation, word frequency statistics, and visualization into a text visualization technology. The word frequency is represented by the size of the "topic words". Browse users quickly discover core keywords.

Based on WordCloud and Jieba word segmentation module in 1 Python environment, the topic name mining is performed on the bibliographic name of the 14-level information college student campus card borrowing data set. By examining the word frequency, the high-frequency subject words in the school's borrowing bibliography are borrowed from a large number of multidimensional bor-

DOI: https://doi.org/10.30564/ret.v3i2.1837

rowings.

(1) From the words "English", "Advanced Mathematics", "Computer", "Economics", "Management" and other words, it can be roughly concluded that this type of discipline is the key learning content for students of the School of Information;

(2) From "C language", "Photoshop", "Java", "CS6", "HTML", "database", "office", "C ++", and "PHP", it can be roughly inferred that the students of this school of information are familiar with such technologies Interested, especially in "C language" technology;

(3) From the words "E-commerce", "College English", "Program Design", "Network Security", "Web Design", "Modeling Statistics", "Web Page" and "Probability Theory", we can roughly infer that the school of information is in school Students are interested in such courses, especially those who are most interested in "programming design";

(4) From the words "Practical", "Principle", "Analysis", "Detailed", "Breakthrough", "Practice", "Guidance" and other words, we can roughly speculate that the students of this school of information have higher demand for books on practical training;

(5) From the words "Life", "Prose", we can roughly infer that the students of this school of information;

(6) From "Dictionary", "Level four and six", "Vocabulary", "Language", "Dictionary", "Level four vocabulary" and other words, it can be roughly inferred that students of Information School focus on studying level four and six apart from the main subject English.

In summary, when it comes to borrowing books, students of the School of Information Science are mainly inclined to books related to this major, showing a desire for their professional knowledge. It is no longer limited to textbook knowledge, and is more willing to obtain more from outside the classroom related information. University libraries can take information colleges as an example. The focus is on purchasing books that conform to the discipline construction of the school, so that students can borrow and improve the relevant discipline skills and academic literacy.

6. K-Means Campus Economic Cluster Analysis

K-Means clustering is a distance-based iterative algorithm [12], which uses distance as an evaluation index of similarity, that is, the closer the two are, the higher the similarity is and the greater the possibility of clustering. This article is based on K-Means clustering analysis of the student consumption of the 14-level information college.

6.1 The Best K Value Analysis

K-Means is a method that minimizes the error between the sample and the square of the particle as the objective function. The sum of the distance errors formed by the squares of each type of cluster and the sample points within it is called distortions. For the mass points of each cluster, the lower the degree of distortion, the denser the sample members included; for the mass points of each cluster, the higher the degree of distortion, the sparser the sample members included. At the same time, the degree of distortion is negatively related to the categories of K-Means clustering. The more K-Means clustering categories, the worse the distortion; on the contrary, the fewer K-Means clustering categories, the better the distortion. However, there is a certain degree of discrimination data. When the critical point is reached, the degree of distortion will be greatly improved. After crossing the critical point, the degree of distortion will be worsened. Therefore, to a certain extent, this critical point can be roughly regarded as the best clustering performance point of K-Means clustering analysis.

Since the clustering algorithm cannot confirm the classification category in advance, the optimal K value is confirmed first. This article relies on the law of the elbow to determine the optimal k value, that is, to confirm the optimal k value according to the degree of improvement in the degree of distortion, and draw the optimal k value map shown in Figure 7.

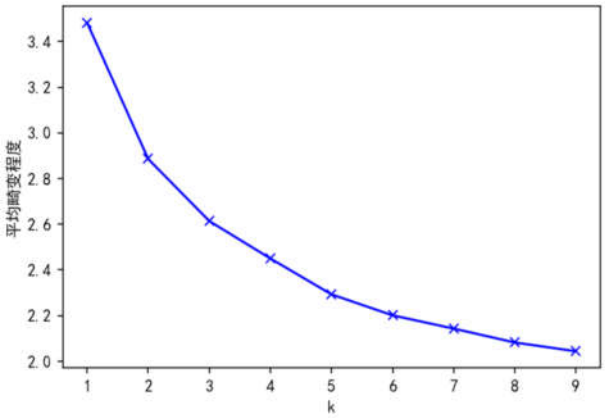

Figure 7. Optimal k value

The abscissa indicates the value of k, and the ordinate indicates the average degree of distortion. According to the principle of the elbow rule, when k = 2, the degree of distortion (y value) is greatly improved, and k = 2 can be roughly confirmed as the number of clusters. Therefore, the elbow has the highest curvature when k = 2, so for the K-Means clustering analysis of the 14-level information college student campus card consumption data set, the op-

DOI: https://doi.org/10.30564/ret.v3i2.1837

timal number of clusters should be k = 2. It is confirmed that the optimal k value is 2, that is, K-Means two-cluster analysis is the most suitable analysis method for the student campus card consumption situation data set of the level 4 information college.

6.2 Analysis of K-Means Clustering Consumption Behavior

K-Means two-cluster analysis is the most suitable analysis method to confirm the optimal k value as the data set of the campus card consumption situation of the 4th-level information college students. Therefore, the data set of 2014 student campus card usage in the School of Information Technology was divided into two categories based on K-Means clustering, and the K-Means two-cluster scatter distribution map shown in Figure 8 was drawn.

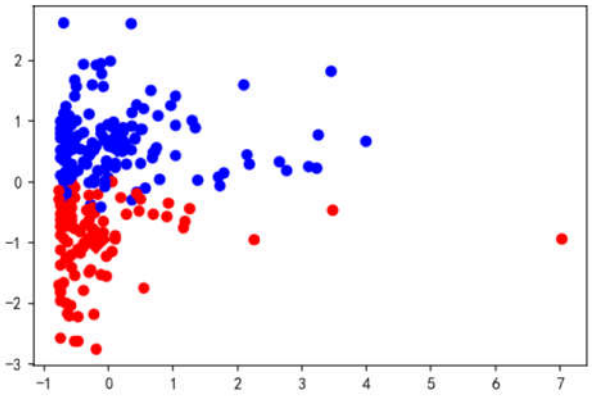

Figure 8. K-Means two cluster analysis

Figure 8 can analyze that the blue dots are generally located above the red scatter dots, and classify the consumer behavior of students into two categories. Combining the data summary information in Table 3, the student's consumption information can be roughly divided into two categories. By comparing the information of borrowing times and consumption times of the two types of students, it can be inferred that the living habits of the second type of students are more regular than those of the first type. The number of times borrowed from books can roughly predict that these students belong to the class that loves learning.

Table 3. K-Means clustering information summary

Category	Book Borrowing	During class	Weekend	Rest mealtime	Breakfast	Lunch	Dinner	Transaction value	Consumption
I	27	945	229	475	183	289	226	6370	1174
II	41	1576	419	734	369	504	386	9425	1995

There are many factors affecting the use of campus card, and in addition to the subjective factors of students themselves, they are also subject to objective conditions. In this paper, the K-Means campus card usage frequency cluster analysis is performed according to time, and the statistical analysis of the campus frequency cluster at the time shown in figure 11. The horizontal axis represents the date, and the vertical axis represents the number of users.

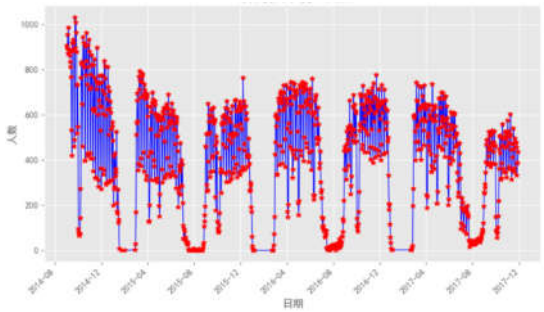

Figure 9. K-Means time distribution cluster analysis

Figure 9 shows that there are 7 places with more focus, indicating that campus card use was more frequent during this period. During the school year from March to July and September to December, the number of consumptions is expected to be more; from January to February and July to August each year, it is in the summer and winter vacations, and the consumption is less. And during the summer vacation, there are more consumers than during the winter vacation. It can be predicted that during the summer vacation, some students choose to stay at school or go home later. From 14 to 17 years, the number of use of campus card has been gradually reduced. It can be predicted that students are not familiar with the surrounding environment at the beginning of school. They are just entering high school when they are in high school. They continue to live in high school and eat in the cafeteria. With the familiarity with the surrounding environment, and the rise of takeaways and Alipay. Some students started paying with Alipay and ordering takeaways. The use of campus cards in the cafeteria has also decreased.

7. Conclusion

In today's society, Internet technology and big data technology continue to develop, and people are paying more and more attention to the value of data. Data is no longer simply stored as a journal. In order to better build a smart campus, and closer to the daily consumption habits of school students, optimize the allocation of resources for smart campus construction and better develop the campus economy. The K-Means clustering proposed in this paper combines hotspot mining analysis algorithm and data visualization technology to multi-dimensionally mine 14-level information college student campus card con-

DOI: https://doi.org/10.30564/ret.v3i2.1837

sumption related data. This article compares and analyzes the analysis of the consumption of students in various majors, the analysis of the consumption of students in the canteens, the analysis of daily consumption behaviors and consumption trends; the analysis of the school library construction in combination with the WordCloud technology; and finally the K-Means cluster analysis model Divide school students into two broad categories.

This article delves into the data related to the campus card usage of students at the 14-level School of Information, and deeply explores the relationships and connections between the data. The following suggestions are made for the campus economy:

(1) Compared with the price of the restaurant, we should pay more attention to the taste of the restaurant; we can combine the Internet to order food online, set up a cafeteria take-out function, and take out services.

(2) Consumption times of campus card for students in school are roughly distributed between 1000-2000. It can be estimated that most students' consumption is quite reasonable.

(3) The comparison of the peak consumption period shows that the daily peak meals are: breakfast (6 to 8), lunch (11 to 12), and dinner (5 to 7). Lunch is the largest and can be prepared during the meal preparation process; breakfast is the smallest and can be reduced as appropriate. Can be prepared 1 to half an hour in advance according to the peak period.

(4) Combined with the WordCloud technical analysis, it can be seen that when students borrow books, they are mainly inclined to the relevant books of this major, showing their desire for the professional knowledge they have learned. . University libraries can purchase related books in combination with the Chinese Library Classification.

(5) Through K-Means clustering, it can be found that students who consume more often borrow more books. It can be predicted that such students belong to a hard-working, regular life, and are active students.

(6) Through K-Means time clustering, we can find that the number of consumptions is more during the school period from March to July and September to December each year; January to February and July to August each year are in summer and winter vacations, and the number of consumption is relatively less. And during the summer vacation, there are more consumers than during the winter vacation. It can be predicted that during the summer vacation, some students choose to stay at school or go home later. Part of the window of a restaurant can be opened to facilitate students' daily dining while saving corresponding costs.

Campus economy, as an important part of regional economy, is a manifestation of regional economic development. The campus economy is geared towards students and relying on schools. With the continuous development of the Internet, it has become an important part of the regional economy. Based on the advantages of big data analysis, this paper proposes a data analysis algorithm based on K-Means clustering and hotspot mining. It also combines data visualization technology to show the college students 'consumption habits, deeply explore the college students' consumption habits, and calculate the core composition of the campus economy. Partly, it can effectively optimize the allocation of resources while constructing a smart campus for colleges and universities, and it is more in line with the daily habits of college students. It has important theoretical and application values for regional economic research.

Acknowledgment

The authors acknowledge Science and Technology Project of Guizhou Province of China, the project name is Time-varying simulation of circular diaphragm wall structure (Grant QKHJC[2019]1403), Application of Knowledge Map Construction Method of Guizhou Multi-source Geographic Data in Public Opinion(Grant QKHJC[2019]1041). Guizhou Province Colleges and Universities Top Technology Talent Support Program, the project name is Big data analysis and evaluation system for remote real-time monitoring of directional drilling rig (Grant QJH KY[2016]068).

References

[1] Chen Feng. Analysis and data mining of dining consumption behavior of college users based on campus card system [J]. China Education Informationization, 2014, (09): 47-49.

[2] Lu Sifeng, He Naiming. Analysis and guidance of college students' consumption behavior. Journal of Beijing Technology and Business University, 2003 (18): 62-65.

[3] Xu Juan. Investigation and Analysis of Students' Consumption View in Higher Vocational Colleges [J].Journal of Chongqing College of Electric Power, 2019,24 (01): 57-59.

[4] Zhang Chunhua, Wen Lu. An Empirical Study of Online Game Consumption Behavior and Its Influencing Factors——Based on the Differential Analysis of College Students' Gender and Education [J].Jiangsu Social Science, 2018 (06): 50-58.

[5] Liu Shangjun, Liu Jun, Jiang Zhihui.Analysis of the Consumption Behavior of Liquid Milk among College Students——Based on a Survey of Students in Tarim University [J].Journal of Tarim University, 2017,29 (03): 79-85.

 DOI: https://doi.org/10.30564/ret.v3i2.1837

[6] Ren Jinhua. Study on the Law of Student Campus Consumption Based on Card Data [D].Huazhong Normal University, 2018.
[7] Liu Jian. Research on tumor gene expression profile data analysis method based on machine learning [D]. China University of Mining and Technology, 2018.
[8] Li Jianwei, Su Zhanyu, Huang Yiru.Research on Risk Prediction of Online Learning Based on Big Data Learning Analysis [J].Modern Educational Technology, 2018,28 (08): 78-84.
[9] Ye Yu, Zhang Lingzhu, Yan Wentao, Zeng Wei.Humanistic perspective measurement framework for street greening quality: large-scale analysis based on Baidu streetscape data and machine learning [J]. Landscape Architecture, 2018,25 (08): 24- 29.
[10] An Qiangqiang, Li Zhaoxing, Zhang Feng, Zhang Yaqiong. Unstructured big data analysis algorithm of communication network based on machine learning [J].Electronic Design Engineering, 2018,26 (14): 53-56.
[11] Qi Xuedan. Classification of Alzheimer's disease based on brain MRI and machine learning [D].Shanxi Medical University, 2018.
[12] Anil K. Jain. Data clustering: 50 years beyond K-means [J]. Pattern Recognition Letters, 2009,31 (8).
[13] Wang Pengfei.Coupling Development of Beijing-Tianjin-Hebei Circulation Industry and Regional Economy: Theory and Empirical [J].Commercial Economic Research, 2019 (09): 153-155.
[14] Wu Shuai, Yang Xiuyi, Xia Huan, Tian Guijiang, Zhao Ziru, Liu Xifeng, Gao Tong. Maximum Likelihood Estimation Based on Normal Distribution [J]. South Agricultural Machinery, 2019, 50 (19): 20-21.
[15] Wu Shuai, Xia Huan, Yang Xiuyi, Yu Xiaomin, Zhao Ziru, Dou Yueqi. Discussion on the protection of ancient buildings based on sentiment analysis of Notre Dame fire [J]. Information Technology and Information Technology, 2019 (08): 154-159.
[16] [12]Wu Shuai, Xia Huan, Yang Xiuyi, Zhao Ziru, Dou Yueqi, Liu Xifeng. Analysis of O2O offline store crowdfunding expansion business model based on knowledge map [J]. Mall Modernization, 2019 (15): 9-10.

DOI: https://doi.org/10.30564/ret.v3i2.1837