Investigation Report on Tongzhou Marathon

Buzuoha Re, Ding Dang, Fu Tengfei, Zheng Yibo, Yu Zhichuan

Beijing Wuzi University, Beijing, China

ABSTRACT

In the context of the digital economy era, sports events have become an important vehicle for promoting national fitness. This study focused on the participants, spectators and staff of the Tongzhou Grand Canal Half Marathon. Through questionnaire surveys, it comprehensively utilized various analytical methods such as K-prototype clustering, structural equation model, Ordered Logit model and Apriori association rules, Empirically test the factors influencing satisfaction and the characteristics of service demands throughout the event. The study found that digital transformation can effectively identify and construct user profiles of "high satisfaction participants", "medium satisfaction actors", and "low satisfaction callers", precisely depicting group characteristics; Event information, venue facilities, and on-site services are key factors affecting the satisfaction of participants, while problems such as crowded starting areas, unclear signs, and untimely information updates significantly lower the experience evaluation; Different groups show different combinations of preferences for finisher services. Further analysis suggests that the optimization of the event experience should focus on the refinement of processes, the precision of services, and the balance of commercial sponsorship. Based on this, countermeasures are proposed in terms of optimizing event organization, strengthening data-driven decision-making, and improving spectator order, providing practical references for the high-quality development of sports events in the context of national fitness.

KEYWORDS: *Tongzhou Grand Canal Half Marathon, Satisfaction survey, Event optimization, Adaboost model, K-prototype model, Cross-analysis, SEM structural equation model, Ordered logit model, Apriori association rule algorithm.*

Research Background

In recent years, the state has attached great importance to the construction of the Beijing Sub-center. The introduction of policies such as the "Opinions of The State Council on Supporting the High-Quality Development of the Beijing Sub-Center" has provided clear direction guidance and strong support for the development of sports in Tongzhou District. In this context, Tongzhou District actively responds to the national policy call, with the core goal of promoting the deep integration of sports with culture, tourism, business and other industries, and holds various sports events as an important measure to stimulate the vitality of the city's sports economy [1].

From the perspective of regional development foundation, Tongzhou, as the sub-center of Beijing City, has significant advantages in hosting diverse sports events. On the one hand, it has well-developed sports venues and facilities, which can meet the needs of hosting events of different scales and types; On the other hand, Tongzhou, with its superior geographical location and relatively convenient transportation, and relying on its unique canal cultural resources [1, 2] has

---

[1] 金耀飞.副中心迎来体育超级周末[N].北京城市副中心报, 2025-09-12(001).
[2] 金耀飞.副中心水上赛事经济热出了圈[N]. 北京城市副中心报, 2025-08-13(001).

successfully hosted many sports events with regional characteristics and gradually accumulated experience in event organization and operation.

From the perspective of social demand, with the continuous development of China's social economy and the continuous improvement of people's living standards, public health awareness has significantly increased, and the attention and enthusiasm for participation in physical exercise and sports events have grown increasingly high [3]. This extensive social demand is not only reflected in the daily fitness activities of the general public, but also extends to the diverse demands for watching and participating in various sports events, providing a broad market space for the development of sports events in Tongzhou District. According to statistics, the number of sports events, the number of participating spectators and the area of sports venues in Tongzhou District have all shown an increasing trend year by year from 2020 to 2024, confirming the feasibility and necessity of the development of regional sports events.

However, during the rapid development of sports events in Tongzhou District, there are still many problems to be solved, such as the matching degree of event types and group demands, the quality of event service guarantees, and the maximization of the driving effect of events on the regional economy and society. For example, there are significant differences in cognitive channels, willingness to participate, and satisfaction evaluations of events among different age, income, and occupational groups, and there is still room for improvement in aspects such as time arrangement, transportation convenience, and finalization services in event organization.

Based on this, an in-depth analysis of the current situation of sports events in Tongzhou District, the identification of the characteristics of potential participants, and the precise grasp of satisfaction and demands at different stages are of great practical significance for promoting the high-quality development of regional sports events and facilitating the construction of the Beijing sub-center, and have become the core driving force for this study.

This study aims to systematically analyze the characteristics of the participating groups in sports events in Tongzhou, the influencing factors of satisfaction, and the preferences for service demands, providing data support and decision-making

references for the optimization and upgrading of events. This study obtained valid data of event-related groups through mixed sampling, combined with *Adaboost* model [ ], [4]*K-prototype* model, cross-analysis, structural equation model [ ], [5]*Ordered logit* model [ ], [6]*Apriori* association rule algorithm The analysis was conducted from dimensions such as group profiling, willingness to participate, satisfaction mechanism, and service association, and ultimately targeted optimization suggestions were formed.

1. Design of the research plan

1.1 Research Protocol and Questionnaire Collection The research was carried out in three phases: The first phase was questionnaire design. Based on the details and schedule of the event services published on the official account of Tongma, a multi-dimensional questionnaire was designed, covering basic information of the participants (age, gender, occupation, etc.), satisfaction with the core dimensions of the event (organization, track, supply, safety, volunteer service, schedule), and open-ended questions were added to collect suggestions for improvement; The second phase is data collection, which involves obtaining sample data through field research to ensure the validity of the sample; The third stage is data analysis, using models such as cluster analysis, Adaboost, and K-prototype to mine data patterns, ensure objective and accurate results, and provide a scientific basis for optimizing event quality.

The questionnaire collection was carried out at the event venue on November 10th, and there were two major problems in the implementation: First, the participants had low interest and resistance to the questionnaire due to their concern for the event; Some participants filled out the questionnaire perfunctorily, which affected the quality of the data.

In response to these issues, we promptly adjusted our strategy: First, we optimized the questionnaire presentation and added raffle rewards to enhance its appeal; Second, we strengthened communication and guidance, clearly informing the significance of the questionnaire and the feedback mechanism. After the adjustment, a large amount of valid data was successfully obtained, laying the foundation for subsequent analysis.

1.2 Data Preprocessing After the questionnaires were collected, we first carried out data organization and

[3]
蒋熙雯.文体旅融合发展背景下全民健身赛事促进城市活力提升的策略研究[C]//中国体育科学学会.第五届全民健身学大会论文摘要集——墙报交流（三）.北京体育大学:, 2024:248-249. DOI:10.26914/c.cnkihy.2024.067092.

[4] 胡欢武.机器学习基础[M].电子工业出版社:201904:384.
[5] 刘源,刘红云.结构方程模型应用[M].北京师范大学出版社: 202009:208.
[6] Hilbe M J .Logistic Regression Models[M].Taylor and Francis;CRC Press:2011-03-23:

screening to ensure quality. A total of 354 questionnaires were distributed this time. Some questionnaires had abnormal situations such as too short filling time or highly consistent answers, which could affect the objectivity of the data. Therefore, invalid questionnaires were excluded and 324 valid questionnaires were finally determined.

The valid questionnaires were further cleaned, including the elimination of duplicate data and the correction of incorrect values, to lay the foundation for subsequent analysis. In terms of descriptive statistics, the proportion of respondents covered volunteers, staff, contestants and audience in sequence; Sixty-nine percent had experience in similar events, with 178 men and 146 men respectively. The age group was concentrated between 18 and 45 years old, with more people earning between 9,000 and 15,000 yuan per month. The occupations were mainly business practitioners and freelancers. The sample matched the target group well.

Based on this, we began this study. First, a rough profile of the participants in the marathon was unearthed, providing a scientific basis for the organizers' subsequent publicity efforts. Then, in accordance with the questionnaire design, the data analysis section was divided into three parts: pre-race access analysis, in-race satisfaction survey, and post-race demand analysis to quantify the real needs of the participants.

2. Potential audience exploration for the event
2.1. Potential Audience mining for sports events based on Adaboost and K-prototype models

At present, sports events in Tongzhou are facing problems such as ambiguous characteristics of participating groups and insufficient targeted publicity and mobilization. Accurately identifying the differences in participation preferences and satisfaction among different groups is the key to optimizing event services and increasing the participation rate of residents. Group mining through Adaboost and K-prototype models aims to identify the core factors influencing residents' participation in sports events, divide differentiated groups and build profiles to provide a scientific basis for event organizers to formulate precise publicity strategies and improve service shortcomings, thereby promoting the high-quality development of regional sports events.

The study was analyzed through a three-step process of feature importance inquiry, K-prototype clustering, and resident profiling as follows:

2.1.1. Variable Selection and Adaboost model Feature Importance exploration

2.1.1.1 Variable Selection: Cluster variables were screened from three dimensions: residents' basic personal information, participation status, and event satisfaction, including gender (0 = male). 1= female), age (0= under 18 to 5= over 56), occupation (0= student to 5= freelancer), income (0= under 3000 to 4= over 15000), participation frequency (1=1 to 3=≥3), participation status (0= audience to 3= Staff, and event planning, time, features, environment, facilities, transportation, medical care, service, food and beverage satisfaction (1 = very dissatisfied to 5 = very satisfied), variable details are shown in Table 2-1:

Table 2-1 Variable Selection

Dimensions	Variables	Variable Content	Dimensions	Variables	Variable Content
Personal Basic Information	Gender	0- male; 1- female	Event satisfactio	Planning	1- very dissatisfied;
	Age	0-18 under; 1-18-25 years old; 2-26-35 years old; 3-36-45 years old; 4-46-55 years old; 5 to 56 years old and above		Time	2- fairly dissatisfied;
					3- Average;
	Career	0- Student at school; 1- Sports-related industry personnel; 2- Government agencies; 3- Public institutions; 4- Business units; 5- Freelancers		Features	4- Relatively satisfied;
	Income	0-3000 or less; 1-3000 to 5000; 2-5000 to 9000; 3-9000 to 15000; 4- 15,000 or more		Environme	5- Very satisfied

			n	nt
Participation status	Frequency of participation	1-1 times; 2-2 times; 3- More than or equal to 3 times	Facilities	
	Participation Identity	0- Audience 1- contestants; 2- Volunteer; 3- Staff	Traffic	
			Medical service	
-			Catering	

2.1.1.2 Adaboost Feature Importance Analysis with "whether participated in the Tongzhou event" (0 = not participated, 1 = participated) as the dependent variable and the above 15 variables as the independent variables, the Adaboost ensemble learning algorithm was used to analyze the feature importance. The results showed that "event time arrangement" was the most important feature, followed by "convenience of event venue transportation" and "event medical emergency services"; "Occupation" and "frequency of participation" had minimal impact on participation behavior, with feature importance approaching zero. Therefore, these two variables were excluded, and the remaining 13 variables were retained for subsequent clustering.

2.1.2. K-Prototype Algorithm Clustering Analysis

For the 13 retained variables containing categorization data, the K-prototype clustering algorithm was used to classify the samples. Eventually, three groups were obtained, with proportions of 25.62%, 35.49%, and 38.89% respectively, and the groups were evenly distributed. Analysis of variance showed that there were significant differences ($p<0.05$) among the three groups in nine satisfaction indicators such as event planning, time, and characteristics, with a contour coefficient of 0.283, verifying the rationality of the clustering model.

Table 2-2 Mean Group Satisfaction

	Participants in Satisfaction	Low satisfaction participants	High satisfaction participants	F	p
Overall planning and process of the event	2.90±1.31	2.17±1.00	4.08±0.88	102.298	0.000*
The schedule of the event	3.19±1.28	2.11±0.92	4.03±0.92	105.178	0.000*
The event's publicity and promotion showcase its characteristics	3.17±1.33	2.13±1.00	3.87±0.87	82.185	0.000*
Overall environmental hygiene conditions of the event venue	2.98±1.41	2.23±0.94	4.13±0.85	99.725	0.0*
The completeness of facilities within the site	2.99±1.22	2.17±1.01	4.03±0.88	101.135	0.000*
Accessibility of the event venue	2.89±1.33	2.04±0.91	4.09±0.88	121.822	0.000*
Service attitude of event staff	3.22±1.41	2.03±0.91	4.00±0.92	103.62	0.000*
Catering services provided during the event	2.93±1.30	2.12±1.01	4.24±0.77	134.208	0.000*
Medical emergency services provided during the event	3.17±1.38	1.96±0.89	3.92±0.95	104.462	0.000*

Based on the average satisfaction values of each group (Table 2-2) and the chi-square analysis results (Table 2-3), the three groups were named "Participants with medium satisfaction", "Participants with low satisfaction", and "Participants with high Satisfaction".

Table 2-3 Results of Chi-square Analysis

class	Name	In the satisfaction participation group	Low satisfaction participants	High-satisfaction participants	Total	Caffe	p
Gender	male	46 (55.42)	56 (48.70)	76(60.32)	178 (54.94)	3.291	0.193
	female	37 (44.58)	59 (51.30)	50 (39.68)	146 (45.06)		
Age	Under 18	1 (1.97)	2 (2.74)	1 (2.06)	4 (7.41)	33.843	0.000*
	18-25 years old	25 (24.10)	18 (13.91)	26 (18.25)	69 (18.21)		
	26-35 years old	11 (10.84)	37 (27.83)	30 (1.43)	78 (20.99)		
	36-45 years old	15 (14.46)	30 (23.48)	41 (28.57)	85 (23.15)		
	46-55 years old	11 (13.25)	17 (14.78)	17 (13.49)	45 (13.89)		
	Over 56 years old	9 (9.28)	11 (10.26)	13 (9.70)	33 (16.36)		
Monthly income	Under 3000	16 (19.28)	18 (15.65)	18 (14.29)	52 (16.05)	40.407	0.000*
	3000-5000	15 (18.07)	13 (11.30)	5 (3.97)	33 (10.19)		
	5000-9000	26 (31.33)	14 (12.17)	20 (15.87)	60 (18.52)		
	9000-15000	12 (14.46)	56 (48.70)	59 (46.83)	127 (39.20)		
	More than 15,000	14 (16.87)	14 (12.17)	24 (19.05)	52 (16.05)		
Entry status	Audience	16 (19.28)	16 (13.91)	16 (12.70)	48 (14.81)	33.43	0.000*
	Contestants	16 (19.28)	16 (13.91)	20 (15.87)	52 (16.05)		
	Volunteers	23 (27.71)	45 (35.13)	51 (40.48)	119 (36.73)		
	Staff	10 (12.05)	33 (28.70)	34 (26.98)	77 (23.77)		

2.1.3. Satisfaction analysis based on resident profiling

High satisfaction participants: Predominantly male, aged 36 to 45, mostly high-income groups, mainly volunteers, with high satisfaction in dimensions such as event organization, venue facilities, and services. Satisfaction actors in the event: Predominantly male participants, aged 18 to 25, belonging to the middle-income group, with moderate satisfaction in various dimensions of the event. Low satisfaction to be called: Predominantly female, aged 18 to 25, middle-income group, mostly staff, with low satisfaction in terms of event organization, venue facilities, services, etc.

3. Marathon event satisfaction analysis

3.1. Pre-race satisfaction analysis

3.1.1 Based on cross-analysis of event information acquisition preferences To clarify the differences in channels through which different groups obtain information about the Tongzhou Marathon event, using participation identity, age, monthly income, occupation as grouping variables and information acquisition channels as test variables, significant association dimensions were screened through Pearson chi-square test, and then cross-analysis was conducted. The results are as follows.

3.1.1.1 Chi-square test After the test, the significance of gender with all information acquisition channels was greater than 0.05, the degree of event awareness was significantly associated only with television and radio, and the other channels were not significantly associated, so these two grouping variables were excluded. Participation status was significantly associated with television broadcasts, Internet news sites, social media platforms, recommendations from friends or family; Age is significantly associated with television broadcasts, Internet news sites, social media platforms, event websites, and referrals from friends or family; Monthly income is significantly associated with television broadcasts and Internet news sites; Occupation was significantly associated with television and radio, newspapers or magazines, and these dimensions were included in the subsequent cross-analysis.

Table 3-1 Pearson Chi-square Test Table

	Television broadcasts	Internet news sites	Social media platforms	The official website of the event	Recommendations from friends or family	Live promotion or advertising	Newspapers or magazines
Participation identity			0.045	0.988	0.008	0.195	0.809
Age		0.015	0.009	0.028	0.001	0.683	0.296
Monthly income			0.691	0.382	0.183	0.280	0.490
Event awareness	0.017	0.432	0.204	0.403	0.529	0.887	0.576
Gender	0.901	0.552	0.128	0.701	0.789	0.496	0.229
Occupation	0.035	0.508	0.591	0.198	0.149	0.147	0.031

3.1.1.2 Cross-analysis
3.1.1.2.1 Cross-analysis of participation identity and understanding channels

Table 3-2 Cross-analysis table of Participation identity and access channels

	Internet news sites		Social media platforms		Recommendations from friends or family		Television broadcasts	
	Unselected	Select	Not selected	Select	Unselected	Select	Unselected	Select
Audience	32	16	21	27	23	25	27	21
Contestants	28	24	19	33	24	28	33	19
Volunteers	86	33	58	61	67	52	86	33
Staff	59	18	34	43	39	38	59	18

According to Table 3-2 data: Viewers prefer to obtain information through social media platforms, recommendations from friends or family; Contestants use social media platforms as their main channel; Volunteers and staff also favored social media platforms, with 51.34% and 59.65% of the choices respectively.

3.1.1.2.2 Cross-analysis of age and channels of understanding

Table 3-3 Cross-analysis table of Age and Access channels

	Television broadcast		Internet news sites		Social media platforms		The official website of sports events		Recommendations from friends or family	
	Unselected	Select	Unselected	Select	Unselected	Select	Unselected	Select	Unselected	Select
Below 18	6	18	8	16	3	21	8	16	2	22
18-25	39	20	36	23	20	39	33	26	31	28
26-35	51	17	49	19	30	38	27	41	38	30
36-45	56	19	53	22	41	34	24	51	43	32
46-55	28	17	29	16	19	26	21	24	17	28
Over 56	32	21	37	16	24	29	15	38	27	26

According to Table 33 data: The group under 18 prefers social media platforms, recommendations from friends or family; For the 18-25 age group, social media platforms are dominant; The 26-35 and 36-45 age groups are more inclined to official websites of sports events, with 60.87% and 68.92% of the choices respectively; The 46-55 age group mostly obtained information through social media platforms, event websites, and recommendations from friends or family; People over 56 prefer official websites of sports events.

3.1.1.2.3 Cross-analysis of monthly income and access channels

Table 3: Cross-analysis Table of Income and Channels of Understanding in April

	Internet news sites		Television broadcasts	
	Unselected	Select	Unselected	Select
Under 3000	33	19	32	20
3000-5000	15	18	17	16
5000-9000	28	32	27	33
9000-15000	107	20	101	26
More than 15,000	29	23	35	17

According to the data in Table 3-4,the proportion of those under 3,000 yuan, 3,000-5,000 yuan, and 5,000-9,000 yuan who obtain information through television broadcasts and Internet news websites varies slightly; The 9,000-15,000 yuan group prefers television broadcasts; The group above 15,000 yuan is more inclined to Internet news sites.

3.1.1.2.4 Cross-analysis of occupations and channels of understanding

Table 3-5 Cross-analysis Table of occupations and Access Channels

	Television broadcast		Newspaper or magazine	
	Unselected	Select	Not selected	Select
Students at school	19	19	17	21
Sports-related professionals	16	14	6	24
Government agencies	33	18	18	33
Public institutions	34	23	12	45
Corporate units	58	20	20	58
Freelancer	52	18	13	57

According to the data in Table 3-5, the proportion of students in school choosing television broadcasts and newspapers or magazines is comparable; People in sports-related industries, government agencies, public institutions, enterprises, and freelancers all prefer newspapers or magazines, with the proportions being 80.00%, 64.71%, 78.95%, 74.36%, and 81.16% respectively.

3.1.2 Analysis of the factors influencing the intention to sign up for sports events based on the structural equation model

3.2. Analyze the factors influencing satisfaction during the event based on the Ordered Logit model

Based on the survey data of participants during the Tongzhou Grand Canal Half Marathon, this study aims to identify the key factors influencing participants' willingness to participate or watch again in the future. By constructing the Ordered Logit model, the study found that several variables in event organization, facilities, and participant background characteristics had a significant impact on willingness, providing empirical evidence for event organizers to optimize experiences and provide precise services in the middle of the event.

3.2.1 Model Settings and Variables This study uses the Ordered Logit model for analysis. The dependent variable was "Your willingness to participate in or watch the Tongzhou District Marathon event in the coming year", which was a level five ordered variable (1= very reluctant, 5= very willing). The independent variables cover three dimensions of satisfaction and demographic characteristics:

Event organization satisfaction: including the overall planning, timing and promotion of the event;

Event facility satisfaction: including venue environment, facility completeness and transportation convenience;

Event service satisfaction: including staff service, catering and medical emergency services;

Participant characteristics: including gender, age, monthly income, occupation and participation status, etc.

3.2.2 Model results Compared with the likelihood ratio test of the analysis model showed a significant overall fit ($\chi^2(14) = 25.483$, $p = 0.030$), indicating that the independent variables have significant explanatory power for the willingness to participate. The following are the variables that reached the significant level ($p < 0.05$) and their impact analysis:

The impact of the event organization dimension: Event timing (OR=1.196, p=0.032) has a significant positive effect on participation intention. For each satisfaction level increase in timing, participation intention increases by approximately 19.6%. However, the negative coefficient for event promotion (OR=0.787, p=0.004) indicates that in the current mid-stage of the event, the higher the satisfaction with promotion, the lower the willingness to participate. This may reflect the gap between the on-site experience and the pre-event promotion commitment, which is worthy of in-depth examination by the organizers.

The impact of the event facilities dimension: The ease of transportation at the event venue (OR=1.218, p=0.016) was the most influential of all significant variables. For every unit increase in accessibility, the willingness to participate increased by approximately 21.8%, highlighting the importance of smooth logistical support to the experience of participants during the event.

The impact of participant background characteristics: Participation identity (OR=1.171, p=0.042), monthly income (OR=1.124, p=0.038), and occupation (OR=1.124, p=0.037) all have a significant positive impact on willingness. This suggests that groups with a specific identity (such as multiple participants), higher income, or a

specific occupation have a stronger willingness to continue participating in the future, which helps organizers identify and maintain their core audience.

3.3. Analyze the finisher service needs of participants based on the Apriori association rule algorithm

Based on the Apriori association rule model, we aim to dig out the intrinsic connections between different finisher service demands from the feedback of marathon participants. By analyzing patterns such as "how likely are participants to also pay attention to transportation connections when they focus on result inquiries?", the model helps us identify key patterns in the race service mix. This is of great practical guidance for event organizers, as it can precisely guide the optimal allocation of service resources, such as bundling closely related services for improvement, thereby more efficiently enhancing overall participant satisfaction and maximizing service benefits

3.3.1 Model Selection to mine frequent itemsets and association rules from large amounts of data. For example, in the retail industry, by analyzing the records of consumers' purchases of goods, one can discover association rules such as "the probability that a customer who buys bread also buys milk is very high." The mining of such associations is very helpful for understanding the underlying patterns within the data and can reveal the co-occurrence patterns among different elements. The algorithm has relatively low requirements for data types and is suitable for handling transactional data. Whether it's binary attribute data (such as whether a customer bought a certain item, represented by 0 and 1) or multi-valued attribute data (such as the customer's rating of the item), the Apriori algorithm can be used for analysis. The association rules generated by the Apriori algorithm are highly interpretable compared to some complex machine learning algorithms, such as deep neural networks. It presents the association directly in the form of "if A, then B", and people can easily understand the meaning of the rule.

3.3.2 Model Building We put the data into the spssau Tongzhou Grand Canal Half Marathon Satisfaction Survey Analysis Report in the format required by SPSSau, mainly in two columns (the first column is the serial number, the second column is the name of the finisher service demand, and the same product name can be in the same row). Identify frequent itemsets by using the Apriori property in an iterative manner to identify the frequently occurring itemsets in the dataset. Generate association rules, based on the frequent itemsets, generate association rules that satisfy the minimum support and minimum confidence.

Support refers to the frequency with which an item set appears in the dataset. For item A, support indicates the proportion of records in the dataset that contain A. Support can help determine how common a rule is. Let the total number of transactions in dataset D be n, and the Support (X) of item set X:

$$Support(X) = \frac{\left|\{t \in D \mid X \subseteq t\}\right|}{n}$$

Here, $\left|\{t \in DX \leq t\}\right|$ is the number of transactions that contain item set X.

Confidence is the probability that item B also appears in the records where item A appears. For rule A → B, confidence indicates the probability of B occurring while A is satisfied. A high level of confidence indicates higher reliability of the rule. Confidence level for association rule $X \to Y$ (where X and Y are item sets and $X \cap Y = \textcircled{7}$) $Confidence(X \to Y)$:

$$nf(X \to Y) = \frac{Support(X \cup Y)}{Sopport(X)} \text{Youdaoplaceholder0 (formula 7)}$$

Lift refers to whether the appearance of item A has an impact on the appearance of item B. It determines the extent to which the appearance of A boosts the occurrence of B by comparing the product of the confidence of the records containing A and the support of B with the product of the support of B. If the boost is greater than 1, A has a positive effect on B; otherwise, it has a negative effect. Association rule $X \to Y$ $Lift(X \to Y)$:

$$Lift(X \Rightarrow Y) = \frac{Confidence(X \Rightarrow Y)}{Support(Y)} \text{ (Formula 8)}$$

When the lift is greater than 1, it indicates a positive correlation between X and Y, that is, the presence of X increases the likelihood of the presence of Y, which helps to determine valuable association rules.

3.3.3 Result Analysis According to the results in Table 3-6 below, nearly 20% of the participants were concerned about the accurate way of score inquiry, and about 18% were concerned about transportation connection. When people were concerned about the exact way to check their scores, there was a 45.9% chance that they were also

concerned about transportation connections. An increase of 2.58 indicates a strong positive correlation between the two.

Based on the Apriori association rule algorithm, the service requirements for participants to finish the race were analyzed, and the results are shown in Table 3-6 below:

Table 3-6 Summary of Association Rules

Item (A)	Subparagraph (B)	Previous item (A) support	Sub term (B) support	Total support	Confidence level	Boosting degree
Accurate ways to check grades	Convenient transportation links	0.196	0.178	0.09	0.459	2.58
Convenient transportation access	Accurate way to check your grades	0.178	0.196	0.09	0.506	2.58
The accurate way to check grades	Meaningful finisher mementos	0.196	0.208	0.121	0.617	2.968
Meaningful finisher mementos	Accurate way to check your results	0.208	0.196	0.121	0.582	2.968
Accurate ways to check grades	A suitable changing / shower environment	0.196	0.211	0.12	0.612	2.902
A suitable changing / shower environment	Accurate way to check grades	0.211	0.196	0.12	0.569	2.902
Convenient transportation links	Meaningful finisher mementos	0.178	0.208	0.109	0.612	2.944
Meaningful finisher mementos	Convenient transportation links	0.208	0.178	0.109	0.524	2.944
Convenient transportation access	A suitable changing / shower environment	0.178	0.211	0.104	0.584	2.769
A suitable changing / shower environment	Convenient transportation links	0.211	0.178	0.104	0.493	2.769
Meaningful finisher souvenirs	A suitable changing / shower environment	0.208	0.211	0.134	0.644	3.053
A suitable changing / shower environment	Meaningful finisher mementos	0.211	0.208	0.134	0.635	3.053
Accurate way to check your results	Meaningful finisher mementos and a suitable changing / shower environment	0.196	0.134	0.069	0.352	2.627
Meaningful finisher mementos	Accurate way to check results, appropriate changing / shower environment	0.208	0.12	0.069	0.332	2.764
A suitable changing / shower environment	Accurate way to check results, meaningful finisher mementos	0.211	0.121	0.069	0.327	2.703
Accurate way to check results, meaningful finisher mementos	A suitable dressing / shower ring habitat	0.121	0.211	0.069	0.57	2.703
Accurate way to check grades	Meaningful finisher mementos	0 12	0 208	0 069	0 575	2 764
A suitable changing / shower environment Meaningful finisher mementos, a suitable changing / shower environment	Accurate way to check your results	0.134	0.196	0.069	0.515	2.627

Convenient transportation links	Meaningful finisher mementos, suitable changing / shower facilities	0.178	0.134	0.059	0.331	2.474
Meaningful finisher mementos	Convenient transportation links and a suitable changing / shower environment	0.208	0.104	0.059	0.284	2.727
A suitable changing / shower environment	Meaningful finisher mementos and convenient transportation links	0.211	0.109	0.059	0.28	2.565
Meaningful finisher mementos, convenient transportation	A suitable dressing / shower ring habitat	0.109	0.211	0.059	0.541	2.565
Convenient transportation access, suitable changing / shower environment	Meaningful finisher souvenirs	0.104	0.208	0.059	0.567	2.727
Meaningful finisher mementos, suitable changing / shower environment	Convenient transportation links	0.134	0.178	0.059	0.44	2.474

About 21 per cent of people are concerned about meaningful finisher mementos, and nearly 20 per cent are concerned about ways to check results. When people focus on finisher mementos, there is a 58.2% chance they will also focus on the way they check their results. The increase was 2.96, indicating a strong correlation between the two.

About 21 per cent were concerned about the changing/shower environment, and nearly 20 per cent were concerned about the way of checking grades. When people focus on the changing/shower environment, there is a 57.9% chance they also focus on the way they check their grades. The increase was 2.96, indicating a strong correlation between the two. The higher the boost, the stronger the association rule.

It should be noted that the same degree of boost between meaningful finisher mementos and accurate result inquiry methods, and between appropriate changing/shower environments and accurate result inquiry methods, whether the former is pushed to the latter or the latter to the former, indicates that the association strength between them is equal. These closely related combinations of services are of great reference value for event organizers when planning finisher services.

4. Conclusions and Recommendations

The investigation of this model at this critical point in the middle of the event shows that the future intentions of the participants are not determined by a single factor, but are influenced by the core aspects of the event operation and the characteristics of the participants themselves. Based on the analysis results, the following suggestions are made:

4.1. Provide precise services to different participants to enhance overall satisfaction

For low-satisfaction groups: For the "low-satisfaction individuals to be called" identified by the K-prototype model (predominantly female staff), enhance their event experience by optimizing work schedules, reducing on-site workloads, adding staff rest areas and exclusive benefits (such as upgraded work meals, post-event souvenirs); In response to the low proportion of female participants, special benefits for women were introduced (such as special MEDALS for the women's group and pre-race lectures on women's health) to enhance publicity and mobilization for female runners.

Focus on core participants: Target high-income, high-loyalty "high satisfaction participants of the event" as the core service recipients, open exclusive registration channels, set up VIP rest areas for the event, invite them to participate in the event preparation opinion collection, and strengthen the sense of belonging of the core group to the event.

4.2. Optimize the core links of event operation and address key pain points

Time and traffic optimization: Based on the Adaboost model that "event timing is a core feature affecting

participation", combined with the analysis results of the Ordered Logit model, further optimize event timing (avoid urban commuting peak), and work with the transportation department to dynamically plan track lockdowns using the "first lockout, first lift, fast lockout, fast lift" model, At the same time, the number of shuttle bus services was increased and temporary parking areas were added to address the traffic inconvenience reported by the participants.

Event process improvements: Continue the humanized mechanism of "sign up first, pay after winning" from 2025, further simplify the registration and collection process, open online competition package mailing service to reduce on-site queuing time for participants; To address congestion in the starting area, the race will start in zones based on the race results, and the track marking system will be optimized to avoid unclear route guidance.

4.3. Upgrade the promotion strategy to match the information access preferences of different groups of people

Channel targeting: Based on cross-analysis results, focus on social media (Douyin, wechat official account) for the 18-25 age group, and release event short videos and runner stories; For those over 36 years old and sports-related industry practitioners, professional event information will be released through official websites, newspapers and magazines; For the audience, use word-of-mouth recommendations from friends and offer incentives such as "Invite friends to watch the game" small gifts.

Avoid disconnection between promotion and experience: Before the event, accurately disclose the details of the event services (such as the distribution of supply stations and the locations of medical support points) through official channels. During the event, update the event dynamics in real time to ensure that the contents promised in the promotion, such as "canal culture experience and special supplies", are consistent with the on-site services.

4.4. Improve the service guarantee system to align with the finisher service requirements

Strengthen event support: On the basis of existing medical support, increase the number of mobile medical personnel on the track and extend the post-event medical service time; In addition to regular supplies, the supply station will add canal culture-themed food in line with the local characteristics of

Tongzhou to enhance the uniqueness of the supply experience.

4.5. Deepen the integration of brand culture and enhance the core appeal of the event

Cultural IP creation: Continuing the innovative design of the Lighthouse theme medal, launch differentiated cultural and creative products (such as canal theme competition uniforms, collectible event badges) each year in combination with the canal culture, and deeply bind the event with the Tongzhou Grand Canal culture.

Event integration with the city: Expand the scale of the marathon carnival, add Grand Canal folklore experiences, intangible cultural heritage displays, etc., to allow participants to experience the local culture of Tongzhou in their spare time from the event; At the same time, interactive activities such as the "Running Club Elite Ranking List" will be held to enhance the social attributes and team participation of the event.

References

[1] Jin Yaofei. Sports Super Weekend in the sub-center [N]. Beijing Sub-Center News, 2025-09-12(001).

[2] Jin Yaofei. The Water Sports economy in the sub-center has gone viral [N]. Beijing Sub-center News, 2025-08-13(001).

[3] Jiang Xiwen. Research on Strategies for Promoting Urban Vitality through Mass Fitness Events in the Context of Integrated Development of Culture, Sports and Tourism [C]// Chinese Society of Sports Science. Abstract Collection of Papers from the Fifth National Fitness Science Conference - Poster Exchange (Part 3). Beijing Sport University; , 2024: 248-249. DOI:10.26914 / Arthur c. nkihy. 2024.067092.

[4] Hu Huanwu. Fundamentals of Machine Learning [M]. Publishing House of Electronics Industry: 201904: 384.

[5] Liu Yuan, Liu Hongyun. Application of Structural Equation models [M]. Beijing Normal University Press: 202009: 208.

[6] Hilbe M J .Logistic Regression Models [M].Taylor and Francis; CRC Press: 2011-03-23