Analysis of Sports Consumption Behavior and Design of Marketing Strategy from the Perspective of Physical Health Management

Huaping Zhang[1], Yue Dong[1*]

Abstract

Most individuals are willing to purchase suitable sporting goods and/or services to support physical health management (PHM). PHM-oriented research of sports consumption behavior (SCB) aids in objectively grasping and understanding the impact of PHM demand and awareness on SCB and in providing consumers with products/services through scientific marketing. Existing research lacks appropriate methods for clustering high-dimensional data on sports consumption and ignores the continuity of PHM and the time aspect of key consumption behaviors. This study investigates the PHM-oriented SCB and analyses the appropriate marketing tactics to address the issues. First, the authors presented the flowchart for the associated data analysis on PHM-oriented sports consumption and clustered the multivariate sports consumption data. Next, a method for fine-grained correlation analysis was outlined for sports consumption data, and an algorithm for selecting the ideal results of canonical correlation analysis (CCA) with kernels was provided. Finally, the proposed algorithm's complexity was evaluated. The authors analyzed the relationships between PHM demand and attitude and SCB, developed differential analysis results, and evaluated marketing methods.

Keywords: physical health management (PHM); sports consumption behavior (SCB); clustering; canonical correlation analysis (CCA)

1. Introduction

Improving the nation's physical health is a long-term and challenging undertaking of social governance. Physical health management (PHM) of the population includes post-test intervention, monitoring, and data analysis in addition to physical health tests (Islam, Dias, & Huda, 2021; Tchetchik, Kaplan, & Blass, 2021; Yang & Ming, 2021). Like people's physical health resources, sports consumption behavior (SCB) resources are widely dispersed (Balachander & Paulraj, 2021; Wang & Lv, 2019). In-depth PHM-focused research on SCB helps to objectively grasp and comprehend the impact of PHM demand and awareness on SCB and to provide consumers with products/services via scientific marketing.

Numerous researchers have analyzed the advertising strategies for the sports service industry in the context of consumer upgrade (He et al., 2001; Kearney & Cole, 2003; Martynyuk et al., 2021; Waris & Hameed, 2020; Wells & Macdonald, 1999; Zheng & Xing, 2020). Delpla et al. (2020) investigated how to build China's sports service business by detecting consumer demand, extending the consumer market, and applying consumption regulation. Additionally, they provided an innovative operation mechanism and optimized the service contents. Under increased consumer demand, the mechanism enhances the connection and integration between the sports service business and other essential sports industries. Due to typical issues of national growth, such as economic slowdown and industry reorganization, many citizens are hesitant to purchase sports-related goods and services (Amasyali & El-Gohary, 2021; Cibinskiene, Dumciuviene, & Andrijauskiene, 2020; Fujiwara, 2020; Liu et al., 2020; Niu, 2017; Silva et al., 2020; Tian et al., 2017).

Taking into account the small size of the sports industry, the imperfection of the capital market, and the dearth of relevant talent in China, Minghong and Xinke (2012) propose specific measures for further adjusting the sports structure to satisfy the residents' demand for sports consumption and to inject new vitality into the development of the sports industry. Tseng et al. (2020) constructed a structural feature analysis model for sports consumption demand, compared the consumption behaviors of residents of different regions, genders, and ages, and found that participatory sports have a significant potential for marginal consumption and a strong propensity to purchase. Therefore, the planning for the development of the sports sector should take into account all variables comprehensively and create an overall balanced plan with the appropriate emphasis. Chen (2015) developed a structural equation model (SEM) for physical sports consumption, analyzed the impact of consumption motivation, health awareness, and quality perception on consumers' subjective consumer behavior (SCB), and proposed a marketing strategy optimization scheme to maximize the double mediating effects of scientific content and experience perception on consumption.

[1] School of Management and Economics, North China University of Water Resources and Electric Power, Zhengzhou 450046, China
[*]Corresponding author's email: X20211070717@stu.ncwu.edu.cn

The PHM-focused SCB is susceptible to a variety of internal and external stimuli. Physical health resources and sports consumption statistics are derived from diverse sources, have vastly distinct structures, and have a complicated relationship (Acharya & Adhikari, 2021; Beusker, Stoy, & Pollalis, 2012; Hu, 2014; Lemas et al., 2021; Singh et al., 2021; Tian, 2014). Domestic and international scholars have analyzed and summarized the characteristics, motivations, decision-making process, and influencing variables of SCB, modeled the factors influencing SCB, and provided hypotheses on these elements. But few efficient clustering approaches are available for high-dimensional sports consumption data, and the continuity of PHM and the temporal sequence of necessary consumption behavior receive less consideration. This study investigates the PHM-oriented SCB and analyses the appropriate marketing tactics to address the issues. The second section shows the flowchart for the associated data analysis on PHM-oriented sports consumption and concludes the clustering of multivariate sports consumption data. The third section describes the correlation analysis of sports consumption data at a fine-grained level. Section 4 provides an approach for selecting the optimal results of canonical correlation analysis (CCA) with kernels, while Section 5 examines the algorithm's complexity. The links between PHM demand and attitude and SCB were explored via experiments, differential analysis results were obtained, and marketing strategies were evaluated.

2. Clustering of Multivariate Sports Consumption Data

The data relating to PHM-focused sports consumption contain numerous variables. This work provides a fine-grained correlation analysis method capable of picking the ideal findings for batch processing massive data on sports consumption and correlation analysis of multivariate data.

This work aims to determine the relationships between the external components of the SCB of PHM-oriented consumers. Our method's input data consist of multivariate consumption data geared towards PHM, i.e., consumption data of sports products and related sports services. Figure 1 depicts the data analysis flowchart for PHM-oriented sports consumption. Observably, the initial stage involves clustering the multivariate consumption data.

Suppose the consumption behavior dataset A contains t variables, each corresponding to the M-day consumption data $A=\{a_{i1}, a_{i2}, \ldots, a_M\}$ of a sports product/service, reflecting the unique consumption behavior of this sports product/service. PHM is a step-by-step process that usually lasts for a period. Hence, it is wrong to cluster all the consumption data in each unit period into a group.

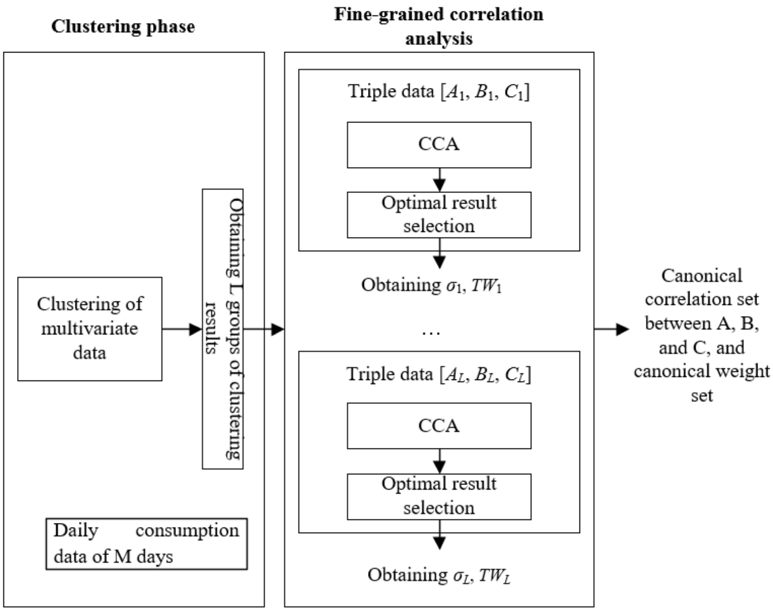

Figure 1. Flow chart

To cluster the PHM-oriented multivariate sports consumption data, it is necessary to first cluster the consumption data A_i on each variable in the form of a daily consumption curve, producing L groups of clustering results $us=\{us_1, us_2, \ldots, us_L\}$, where $us_i \in us$ is the clustering results of the consumption data on the i-th variable. Next, the difference between the daily consumption curves that always exist in all $us_{ij} \in us$ is solved through a set intersection and allocated to the same cluster, where us_{ij} is a cluster of us_i. In this way, the final clustering results can be obtained as US={US_1, US_2, …, US_L}. Each $US_k \in US$ contains the daily consumption data of M_k days and satisfies $\Sigma^L_{k=1}M_k=M$.

Based on the clustering results from the US sports consumption data, the M-day daily data in the entire multivariate time series are divided into L classes. Then, A_i, B_i, and C_i are obtained for US_k, which correspond to each variable and are taken as the inputs of fine-grained correlation analysis.

3. Fine-Grained Correlation Analysis of Sports Consumption Data

Like principal component analysis (PCA) and independent component analysis (ICA), CCA can statistically analyze the linear dependence between multi-dimensional vectors. Let [A_i, B_i] be the single-day data involving sports product and service consumption variables. The linear relationship between A_i and B_i can be identified through CCA. Let us denote the pair of projection weight matrices corresponding to A_i and B_i as g_i and h_i, and the transposition of matrix g_i as g^T_i. Then, g_i={g_{i1}, g_{i2}, …g_{it}} and h_i={h_{i1}, h_{i2}, …h_{iw}} can be defined as canonical weights. The canonical components, i.e., projections v_i and u_i, can be obtained by mapping [A_i, B_i] to the canonical space:

$$v_i = g_i^T A_i = g_{i1}a_{i1} + g_{i2}a_{i2} + \ldots + g_{it}a_{it}$$
$$u_i = h_i^T B_i = h_{i1}b_{i1} + h_{i2}b_{i2} + \ldots + h_{iw}b_{iw} \quad (1)$$

The goal of CCA is to maximize the correlation between v_i and u_i:

$$\sigma_i = \max_{g_i,h_i} \frac{D_{v_i-u_i}}{\sqrt{D_{v_i-v_i}}\sqrt{D_{u_i-u_i}}} = \max_{g_i,h_i} \frac{g_i^T D_{A_i-B_i} h_i}{\sqrt{g_i^T D_{A_i-A_i} g_i}\sqrt{h_i^T D_{B_i-B_i} h_i}} \quad (2)$$

Let D_{vi-ui} be the covariance matrix of v_i and u_i; D_{vi-vi} is the autocorrelation matrix of v_i; D_{ui-ui} is the autocorrelation matrix of u_i. To maximize the correlation coefficient σ_i between v_i and u_i, the numerator should be maximized without changing the denominator. Hence, formula (2) can be rewritten as:

$$\max_{g_i,h_i} g_i^T D_{A_i-B_i} h_i, \text{ s.t. } g_i^T D_{A_i-A_i} g_i = 1, h_i^T D_{B_i-B_i} h_i = 1 \quad (3)$$

Lagrange's equation can solve the target function:

$$\delta = g_i^T D_{A_i-B_i} h_i - \frac{\mu_{A_i}}{2}\left(g_i^T D_{A_i-A_i} g_i - 1\right) - \frac{\mu_{B_i}}{2}\left(h_i^T D_{B_i-B_i} h_i - 1\right) \quad (4)$$

Let us denote the correlation coefficient σ_i between v_i and u_i as an eigenvalue μ_{Ai}. The above formula can be expressed as a generalized eigenvalue problem:

$$\begin{pmatrix} D^{-1}_{v_i-v_i} & 0 \\ 0 & D^{-1}_{u_i-u_i} \end{pmatrix} \begin{pmatrix} 0 & D^{-1}_{v_i-u_i} \\ D^{-1}_{u_i-v_i} & 0 \end{pmatrix} \begin{pmatrix} g \\ h \end{pmatrix} = \mu_{A_i} \begin{pmatrix} g \\ h \end{pmatrix} \quad (5)$$

When μ_{Ai} reaches the eigenvalue, the first pair of canonical weights can be solved: $g^{(1)}_I$ and $h^{(1)}_i$. The second pair of canonical weights $g^{(2)}_I$ and $h^{(2)}_i$ can be solved by the optimization problem below:

$$\max_{g_i,h_i} g_i^T D_{A_i-B_i} h_i, \text{ s.t.}$$
$$g_i^{(2)T} D_{A_i-A_i} g_i^{(2)} = 1, h_i^{(2)T} D_{A_i-B_i} h_i^{(2)} = 1;$$
$$g_i^{(2)T} D_{A_i-A_i} g_i^{(1)} = 1, h_i^{(2)T} D_{A_i-B_i} h_i^{(1)} = 0 \quad (6)$$

Similarly, the daily data on sports product consumption, sports service consumption, and PHM are denoted as [A_i, B_i, C_i]. Then, the CCA can find a group of projection weight matrices g_i, h_i, and d_i, ensuring the canonical components' maximization. Let ω_i be the canonical component of the third variable C_i; d_i={d_{i1}, d_{i2}, ...d_{ie}} be the canonical weight of this component. Then, the ω_i value can be calculated by:

$$\omega_i = d_i^T C_i = d_{i1}c_{i1} + d_{i2}c_{i2} + \ldots + d_{ie}c_{ie} \quad (7)$$

For the projections v_i, u_i, and ω_i, which are obtained by mapping [A_i, B_i, C_i] to the canonical space, their maximum correlation can be calculated by:

$$\sigma_i = \max_{g_i,h_i,d_i} \frac{D_{v_i-u_i-\omega_i}}{\sqrt{D_{v_i-v_i}}\sqrt{D_{u_i-u_i}}\sqrt{D_{\omega_i-\omega_i}}} \quad (8)$$

The covariance matrix $D_{vi-ui-\omega i}$ of v_i, u_i, and ω_i can be expressed as:

$$D_{v_i-u_i-\omega_i} = \begin{pmatrix} D_{v_i-v_i} & D_{v_i-u_i} & D_{v_i-\omega_i} \\ D_{u_i-v_i} & D_{u_i-u_i} & D_{u_i-\omega_i} \\ D_{\omega_i-v_i} & D_{\omega_i-u_i} & D_{\omega_i-\omega_i} \end{pmatrix} \quad (9)$$

From the properties of covariance, D_{vi-ui} is equivalent to that D_{ui-vi} and $D_{vi-ui-\omega i}$ are symmetrical, and the three covariance analyses of the diagonal are the autocorrelation matrices v_i, u_i, and ω_i, respectively. The maximization of the correlation between the three variables can be expressed as:

$$\max_{g_i,h_i,d_i} g_i^T D_{A_i-B_i} h_i + g_i^T D_{A_i-C_i} d_i + h_i^T D_{A_i-C_i} d_i,$$
$$s.t. \quad g_i^T D_{A_i-A_i} g_i = 1, h_i^T D_{B_i-B_i} h_i = 1, D_{C_i-C_i} d_i = 1 \quad (10)$$

The corresponding Lagrange's equation can be expressed as:

$$\delta = g_i^T D_{A_i-B_i} h_i + g_i^T D_{A_i-C_i} d_i + h_i^T D_{B_i-C_i} d_i - \frac{\mu_{A_i}}{2}\left(g_i^T D_{A_i-A_i} g_i - 1\right) - \frac{\mu_{B_i}}{2}\left(h_i^T D_{B_i-B_i} h_i - 1\right) - \frac{\mu_{C_i}}{2}\left(d_i^T D_{C_i-C_i} d_i - 1\right) \quad (11)$$

Generalized eigenvalues can characterize the above formula:

$$\begin{pmatrix} D^{-1}_{v_i v_i} & 0 & 0 \\ 0 & D^{-1}_{u_i u_i} & 0 \\ 0 & 0 & D^{-1}_{\omega_i \omega_i} \end{pmatrix} \begin{pmatrix} 0 & D_{v_i u_i} & D_{v_i \omega_i} \\ D_{u_i v_i} & 0 & D_{u_i \omega_i} \\ D_{\omega_i v_i} & D_{\omega_i u_i} & 0 \end{pmatrix} \begin{pmatrix} g \\ h \\ d \end{pmatrix} = \mu_{A_i} \begin{pmatrix} g \\ h \\ d \end{pmatrix} \quad (12)$$

The above analysis shows that the correlation coefficients of the three-variate consumption behavior dataset form a 3×3 symmetric matrix. Similarly, multiple canonical components and canonical weights can be obtained through CCA of multivariate data.

Since the traditional CCA is not good at capturing the nonlinearity between data, this paper introduces kernel functions (e.g., polynomial or Gaussian kernel function) to the CCA, enabling it to discover the nonlinear relationship. Based on the definition of the covariance matrix, D_{AiAi}, and D_{AiBi} can be written as $D_{AiAi}=A^T_iA_i$ and $D_{AiBi}=A^T_iB_i$, respectively; projection weight matrices g_i and h_i can be rewritten as $g_i=A^T_i\beta_i$ and $h_i=B^T_i\alpha_i$, respectively. Let $LP_{Ai}=A^T_iA_i$ and $LP_{Bi}=A^T_iB_i$ be the kernel projections of A_i and B_i, respectively; v_i and u_i be the canonical weights of LP_{Ai}, and LP_{Bi} projected to the canonical space via g_i and h_i, respectively. Then, the objective function of the CCA with kernels can be updated as follows:

$$\sigma_i = \max_{\beta_i,\alpha_i} \frac{\beta_i^T A_i A_i^T B_i B_i^T \alpha_i}{\sqrt{\beta_i^T A_i A_i^T A_i A_i^T \beta_i}\sqrt{\alpha_i^T B_i B_i^T B_i B_i^T \alpha_i}} = \max_{\beta_i,\alpha_i} \frac{\beta_i^T LP_{A_i} LP_{B_i}\alpha_i}{\sqrt{\beta_i^T LP_{A_i}^2 \beta_i}\sqrt{\alpha_i^T LP_{B_i}^2 \alpha_i}} \quad (13)$$

The maximization of correlation can be expressed as:

$$\max_{\beta_i,\alpha_i} \ \beta_i^T LP_{A_i} LP_{B_i}\alpha_i, s.t.\ \beta_i^T LP_{A_i}^2 \beta_i = 1,\ \alpha_i^T LP_{B_i}^2 \alpha_i = 1 \quad (14)$$

This paper constructs a regularized CCA with kernels using reversible kernel functions. The proposed method can control the trivial learning issues in correlation analysis that arise from overfitting. When the data dimension satisfies ξ<min{t, w}, the method helps to solve the generalized eigenvalues. Let us denote the regularization parameter as φ. Then, the objective function of the regularized CCA with kernels can be expressed as:

$$\sigma_i = \max_{g_i,h_i} \frac{g_i^T D_{A_i-B_i} h_i}{\sqrt{g_i^T D_{A_i-A_i} g_i + \phi\|g\|^2}\sqrt{h_i^T D_{B_i-B_i} h_i + \phi\|h\|^2}} \quad (15)$$

The above formula can be rewritten as:

$$\sigma_i = \max_{\beta_i,\alpha_i} \frac{\beta_i^T LP_{A_i} LP_{B_i}\alpha_i}{\sqrt{\beta_i^T LP_{A_i}^2 \beta_i + \phi\|g\|^2}\sqrt{\alpha_i^T LP_{B_i}^2 \alpha_i + \phi\|h\|^2}} = \max_{\beta_i,\alpha_i} \frac{\beta_i^T LP_{A_i} LP_{B_i}\alpha_i}{\sqrt{\beta_i^T LP_{A_i}^2 \beta_i + \phi\beta_i^T LP_{A_i}^2 \beta_i}\sqrt{\alpha_i^T LP_{B_i}^2 \alpha_i + \phi\alpha_i^T LP_{B_i}^2 \alpha_i}} \quad (16)$$

The maximization of correlation can be expressed as:

$$\max_{\beta_i,\alpha_i} \ \beta_i^T LP_{A_i} LP_{B_i}\alpha_i,$$
$$s.t.\ \ \beta_i^T LP_{A_i}^2 \beta_i + \phi\beta_i^T LP_{A_i}^2 \beta_i = 1,\ \alpha_i^T LP_{B_i}^2 \alpha_i + \phi\alpha_i^T LP_{B_i}^2 \alpha_i = 1 \quad (17)$$

With the regularized CCA with kernels, the maximization of the correlation of three-variate data can be expressed as:

$$\max_{\beta_i,\alpha_i} \ \beta_i^T LP_{A_i} LP_{B_i}\alpha_i + \beta_i^T LP_{A_i} LP_{C_i}\eta_i + \alpha_i^T LP_{B_i} LP_{C_i}\eta_i,$$
$$s.t.\ \ \beta_i^T LP_{A_i}^2 \beta_i + \phi\beta_i^T LP_{A_i}^2 \beta_i = 1,\ \alpha_i^T LP_{B_i}^2 \alpha_i + \phi\alpha_i^T LP_{B_i}^2 \alpha_i = 1,$$
$$\eta_i^T LP_{C_i}^2 \eta_i + \phi\eta_i^T LP_{C_i}^2 \eta_i = 1 \quad (18)$$

where $L_{Ci}=C_iC^T_i$ is the kernel projection of data C_i; $d_i=C^T_i\psi_i$ is the corresponding weight matrix d_i.

4. Selection of Optimal Results

The computing would be unnecessarily laborious to process the massive associated data of sports consumption if every group of canonical correlation coefficients and the corresponding canonical components and weights were analyzed. The regularized CCA with kernels can recognize the correlations between most data but cannot ensure the optimal mean of the maximum correlations. To solve the problem, this paper proposes an algorithm to select the optimal results of the regularized CCA with kernels.

Firstly, a CCA is performed on each pair of $[A_i, B_i]$, where $A_i∈A$ and $B_i∈B$. Suppose each correlation analysis only produces a group of canonical correlation coefficients σ_i, canonical components $TC_i=[v_i, u_i]$, and canonical weights $TW_i=[g_i, h_i]$. Based on the correlation between u_i and u_{NE}, the data B_{MS} in B, which is the most similar to a new data B_{NE} not belonging to B, can be searched for by:

$$B_{MS} = \underset{B_i}{argmax} CC(u_i, u_{NE}) = \underset{B_i}{argmax} CC(u_i, h_i^T u_{NE}) \quad (19)$$

Let us denote correlation coefficient $\sigma_{vi\text{-}vNE}$ as $CC(u_i, u_{new})$ and the projection of B_{NE} via h_i to the space of u_i as u_{NE}. The k-nearest neighbors (k-NN) algorithm makes it possible to find l B_{MS} most similar to B_{NE} and the most similar data A_{MS} in A. The mean of l A_{MS} can be calculated by:

$$A_{NE} = \frac{1}{l}\left(A_{MS}^{(1)} + A_{MS}^{(2)} + \ldots + A_{MS}^{(l)}\right) \quad (20)$$

Let θ_1, θ_2, …, θ_l be the weights of similarity. Then, the weighted mean of l A_{MS} can be calculated by:

$$A_{NE} = \theta_1 A_{MS}^{(1)} + \theta_2 A_{MS}^{(2)} + \ldots + \theta_l A_{MS}^{(l)} \quad (21)$$

If l=1, then:

$$A_{NE} = A_{MS} \quad (22)$$

Based on the known canonical components and weights, the result can be verified through the inverse process. For $[A_i, B_i]$, the r groups of σ_i, TC_i, and TW_i can be expressed as:

$$\sigma_i = \left\{\sigma_i^{(1)}, \sigma_i^{(2)}, \ldots, \sigma_i^{(r)}\right\},$$
$$TC_i = \left\{\left[v_i^{(1)}, u_i^{(1)}\right], \left[v_i^{(2)}, u_i^{(2)}\right], \ldots, \left[v_i^{(r)}, u_i^{(r)}\right]\right\},$$
$$TW_i = \left\{\left[g_i^{(1)}, h_i^{(1)}\right], \left[g_i^{(2)}, h_i^{(2)}\right], \ldots, \left[g_i^{(r)}, h_i^{(r)}\right]\right\}。 \quad (23)$$

Let $B_j∈B$ and i≠j. For each group of $\sigma_i^{(\tau)}[v^{(\tau)}_i, u^{(\tau)}_i]$ and $[g^{(\tau)}_i, h^{(\tau)}_i]$, it is possible to find the most similar $B'^{(\tau)}_i$ to B_i:

$$B_i'^{(\tau)} = \underset{B_j}{argmax} CC\left(u_i^{(\tau)}, h_i^{(\tau)T} B_j\right), \quad (24)$$

Suppose $B'^{(\tau)}_i$ differs very slightly from B_i, and the corresponding $A'^{(\tau)}_i$ and A_i are approximately equal. The hypothetical data $A'^{(\tau)}_i$ can be described as $A^{*'(\tau)}_i$. Then,

$$A_i^{*\prime(\tau)} = A_i \quad (25)$$

The accuracy of the hypothetical data $A^{*'(\tau)}_i$ against the true data $X'^{(t)}_iA'^{(\tau)}_i$ can be evaluated by the correlation coefficients capable of demonstrating the validity of canonical components and weights. Then, the selection mechanism of optimal results can be expressed as:

$$h_\tau = \underset{h_i^{(\tau)}}{argmax} CC\left(A_i'^{(\tau)}, A_i^{*\prime(\tau)}\right) = \underset{h_i^{(\tau)}}{argmax} CC\left(A_i'^{(\tau)}, A_i\right), \quad (26)$$

where, $A'^{(\tau)}_i$ is from $[A'^{(\tau)}_i, B'^{(\tau)}_i]$ ($B'^{(\tau)}_i$ can be obtained by formula (24); A_i is from $[A_i, B_i]$.

Similarly, the selection mechanism of optimal results can be derived for the triple $[A_i, B_i, C_i]$. The τ-th group of CCA results are denoted as $\sigma_i^{(\tau)}$, $[v^{(\tau)}_i, u^{(\tau)}_i, q^{(\tau)}_i]$ and $[g^{(\tau)}_i, h^{(\tau)}_i, d^{(\tau)}_i]$. Based on the mean correlation coefficient, the most similar $B'^{(\tau)}_i$ and $B'^{(\tau)}_i$ can be found for B_i and C_i:

$$B_i'^{(\tau)}, C_i'^{(\tau)} = \underset{B_j, C_j}{argmax}\left[CC\left(u_i^{(\tau)}, h_i^{(\tau)T} B_j\right) + CC\left(q_i^{(\tau)}, d_i^{(\tau)T} C_j\right)\right] \quad (27)$$

where, $B_j \in B$; $C_j \in A$; $i \neq j$. Formula (26) can be converted into:

$$h_i, d_i = arg \underset{h_i^{(\tau)}, d_i^{(\tau)}}{max} CC\left(A_i'^{(\tau)}, A_i^{*\prime(\tau)}\right) \quad (28)$$

5. Algorithm Complexity Analysis

For a group of $[A_i, B_i, C_i]$, where $A=\{a_{i1}, a_{i2}, \ldots a_{it}\} \in R^{o\times t}$, $B=\{b_{i1}, b_{i2}, \ldots b_{iw}\} \in R^{o\times t}$, and $C=\{c_{i1}, c_{i2}, \ldots c_{ie}\} \in R^{o\times e}$ are daily consumption data, the time complexity of the CCA on $[A_i, B_i, C_i]$ is $FZ(on^2)+FZ(on^3)$, with $n=\max\{t, w, e\}$. The time complexity of the covariance matrix and matrix multiplication, inversion, and eigenvalue decomposition are denoted as $FZ(on^2)$ and $FZ(on^3)$, respectively. The time complexity of each run of the optimal result selection algorithm is represented as $FZ(r(Q-1))$, where $r=\min\{t, w, e\}$, and Q is the number of A_i in A. Then, the time complexity of the CCA with kernels, which directly regularize the three-variate consumption dataset A, B, and C can be calculated by:

$$M \cdot \left[FZ(on^2) + FZ(on^3) + FZ\big(r(Q-1)\big)\right] = FZ(rQ^2 + (n^3 + on^2 - r)Q) \quad (29)$$

The greater the Q value, the higher the time complexity. An additional multivariate data clustering algorithm was introduced to lower the time complexity.

Let S be the number of iterations before convergence; CN_k is the number of clusters after the clustering of the k-th variable; $FZ(Q \cdot CN_k \cdot S)$ be the time complexity of our algorithm in processing single-variate consumption data; $FZ(\Sigma^T_{k=1} Q \cdot CN_k \cdot S)$ be the time complexity of our algorithm running for T times; $\Pi^T_{k=1} Q_k$ be the time complexity of the search for common clusters through a set intersection. Then, the time complexity of the multivariate data clustering algorithm can be calculated by:

$$FZ(\textstyle\sum_{k=1}^{T} Q \cdot CN_k \cdot S + \prod_{k=1}^{T} Q_k) \quad (30)$$

The sum of daily consumption curves corresponding to each variable in each cluster can be assumed as follows:

$$\textstyle\sum_{k=1}^{L} Q_k = Q \quad (31)$$

Then, the time complexity of our algorithm can be calculated by:

$$\begin{aligned} & FZ\left(\sum_{k=1}^{T} Q \cdot CN_k \cdot S + \prod_{k=1}^{T} L_k\right) + \sum_{k=1}^{L} Q_k \left[FZ(on^2) + FZ(on^3) + FZ\big(r(Q_k - 1)\big)\right] \\ & = FZ\left(\sum_{k=1}^{T} Q \cdot CN_k \cdot S + \prod_{k=1}^{T} L_k\right) + \sum_{k=1}^{L} FZ\left(rQ_k^2 + (n^3 + on^2 - r)Q_k\right) \\ & = FZ\left(\textstyle\prod_{k=1}^{T} L_k + r\sum_{k=1}^{L} Q_k^2 + \prod_{k=1}^{t} L_k \cdot S \cdot Q + (n^3 + on^2 - r)Q\right) \end{aligned} \quad (32)$$

When the PHM-oriented associated data of sports consumption contain a huge amount Q of single-day data, and the number of iterations S is small, the following inequality can be obtained:

$$\textstyle\sum_{k=1}^{L} L_k + r\sum_{k=1}^{L} Q_k^2 + \sum_{k=1}^{t} L_k \cdot S \cdot Q < r \cdot Q^2 \quad (33)$$

In this case, the multivariate data clustering algorithm can optimize the time complexity of the proposed algorithm.

6. Experiments and Results Analysis

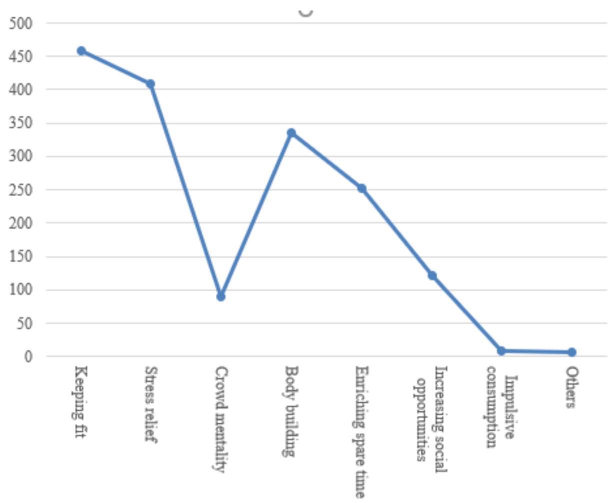

Figure 2. Motivations of sports consumption

Table 1

Regional sports consumption motivations

Motivation	Penetration rate	Response	
		N	Response rate
Keeping fit	94.90%	458	26.40%
Stress relief	83.30%	409	24.20%
Crowd mentality	15.20%	65	3.90%
Body building	73.40%	332	20.70%
Enriching spare time	51.60%	256	16.20%
Increasing social opportunities	25.30%	127	7.60%
Impulsive consumption	1.30%	6	0.50%
Others	1.30%	6	0.50%
Total	349.3%	1659	100%

Using the wjx.cn-developed app, the official questionnaire was distributed randomly. From November 20 to November 27, 2020, 2,156 valid and 144 invalid responses were collected. Valuable samples on the reasons for sports consumption were subjected to descriptive analysis. There are 1,157 (53.66%) males and 999 (46.33%) females among the 2,156 samples. Figure 2 depicts the sports consumption motives uncovered by the investigation. The motivations for regional sports consumption are listed in Table 1. It can be extrapolated that regional sports consumption incentives are becoming increasingly different. 94.90 %, 83.30 %, 73.40 %, and 51.60 % of exercisers are motivated by maintaining fitness, relieving stress, growing muscle, and enhancing leisure time, respectively. Therefore, sports consumption can partially satisfy consumer demands to maintain physical health, relieve job stress, and encourage social activities.

Figure 3 depicts the consumption levels of sports products and services, whereas Figure 4 depicts sports consumption in various regions. When a region's monthly mean sports consumption is below 300 yuan, more people purchase sports items than sports services; when the monthly mean sports consumption of a region is above 300 yuan, more people purchase sports services than sports products. Overall, sports product consumption is higher than sports service consumption. Figure 5 highlights the levels of sports consumption among various income groups. The monthly income of different socioeconomic groups affects the amount of sports consumed. For the group with a monthly income below 4,000 yuan, the monthly mean sports consumption is typically less than 100 yuan. For the group with a monthly income between 4,000 and 5,000 yuan, the monthly mean sports consumption is typically less than 200 yuan.

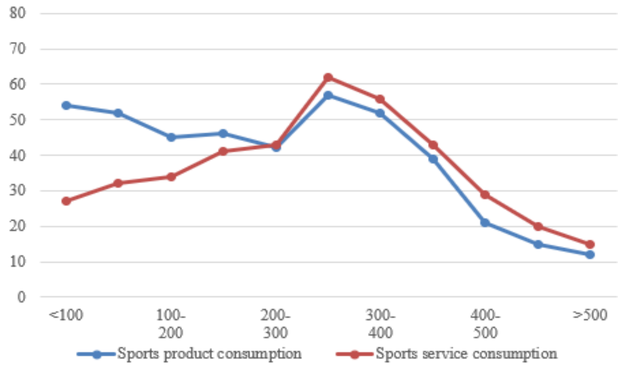

Figure 3. Consumption levels of sports products and services

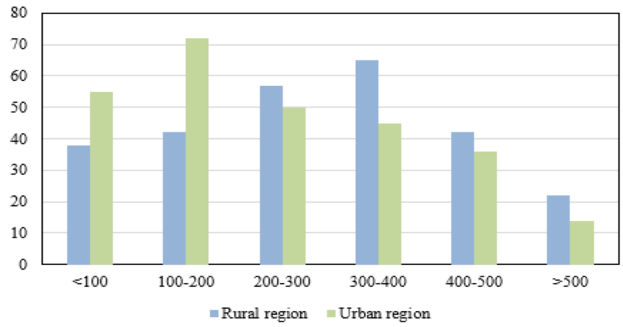

Figure 4. Sports consumption levels of different regions

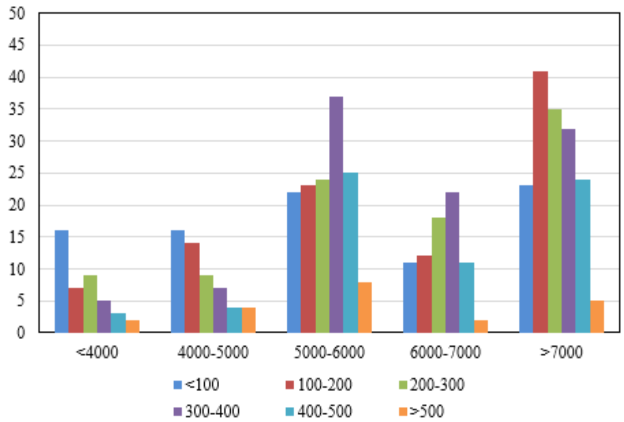

Figure 5. Sports consumption levels of different income groups

Table 2

Regional SCB and marketing directions

Consumption behavior	Marketing direction	Sample size	Min	Max	Mean	Mode	Standard deviation
PHM consumption	Intelligent physical monitoring equipment	600	3	6	4.35	6	0.951
	Sports data management software	600	4	6	4.21	6	0.624
	Physical test information platform	600	4	6	4.32	6	0.738
Sports product consumption	Sports equipment	600	3	6	4.63	6	0.653
	Sports facilities	600	5	6	3.76	6	0.979
	Sports apparels	600	2	6	4.53	5	0.605
	Sports newspapers and books	600	2	6	4.07	5	0.982
Sports service consumption	Sports fitness service	600	3	6	4.62	6	0.876
	Sports advertising service	600	2	6	4.29	6	0.943
	Sports entertainment service	600	1	6	4.37	6	0.839
	Sports health care	600	2	6	3.56	6	1.276

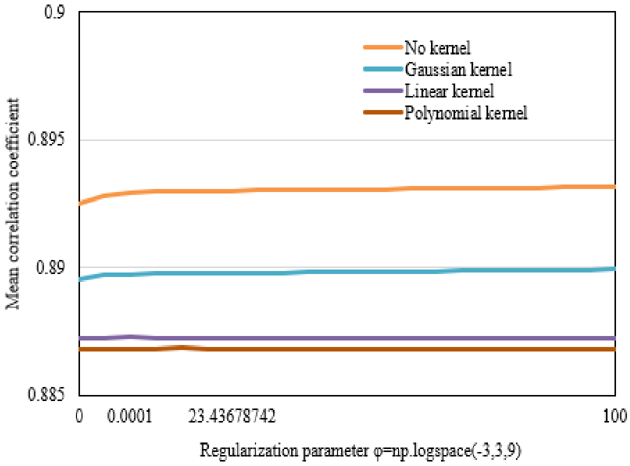

Figure 6. Mean correlation coefficient curves with different kernel functions

On SPSS21.0, descriptive statistical analysis was performed on the three types of sports consumption data. Table 2 summarizes the minimum, maximum, mean, median, and standard deviation for each type of sports consumption for each marketing approach. Note that the mean scores and standard deviations of sports health care, sports facilities, and sports newspapers and books were lower than those of the other marketing directions. Consequently, consumers hold vastly divergent sentiments and generally low familiarity with these three sports consumption categories. The relatively high mean scores and modest standard deviations for sports equipment, fitness services, and apparel indicate that consumers have similar views toward and high public recognition of these three sports consumption categories. The region has a relatively robust understanding of PHM and sports consumption. In the framework of the national fitness program, the attitude toward sports consumption is generally optimistic.

This work employs regularized or non-regularized CCA or CCA with kernels during data processing. Figure 6 compares the mean correlation coefficient curves for the absence of a kernel, the linear kernel, the polynomial kernel, and the Gaussian kernel. The more significant the positive correlation between variables, the closer the correlation coefficient is to 1. In Figure 6, the no-kernel CCA had the highest performance: the curve grew initially and tended to become steady until the regularization coefficient reached 23.43. In light of this, our trials chose the regularized CCA without kernels for correlation analysis of huge consumption data. Table 3 displays the associations between PHM demand and attitude and SCB. According to Table 3, the correlation coefficients of PHM demand and PHM attitude with sports product/service consumption were, at the 0.05 level of significance, 0.697 and 0.726, respectively. Significantly positive correlations indicate that the enormous sports data collected in the region mirror regional sports consumers' positive attitudes and behavior. There is a positive association between PHM demand, PHM attitude, sports preference, sports health concept, etc., and SCB, but the strength of the correlation varies. Consequently, the attitudes of various groups towards SCB differ significantly.

Table 3

Correlations of PHM demand and PHM attitude with SCB

Variable	PHM demand	PHM attitude	Sports preference	Sports health concept	Sports product consumption	Sports service consumption	Mean	Standard deviation
PHM demand	1						4.3752	0.59812
PHM attitude	0.691	1					4.2981	0.96437
Sports preference	0.525	0.423	1				4.5847	0.54893
Sports health concept	0.653	0.472	0.671	1			4.6489	0.59246
Sports product consumption	0.697	0.589	0.426	0.492	1		4.5156	0.59195
Sports service consumption	0.726	0.627	0.603	0.645	0.543	1	4.0192	0.73432

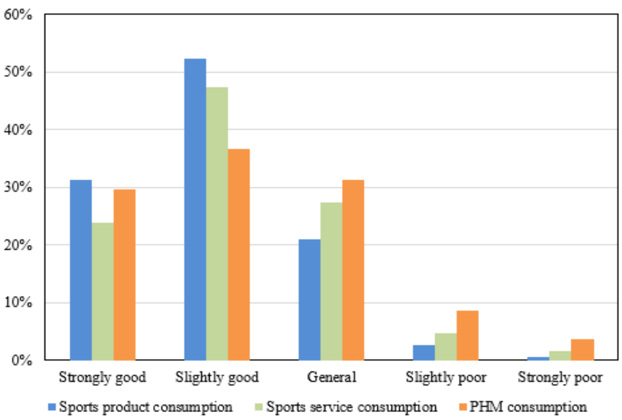

Figure 7. The convenience of different types of sports consumption

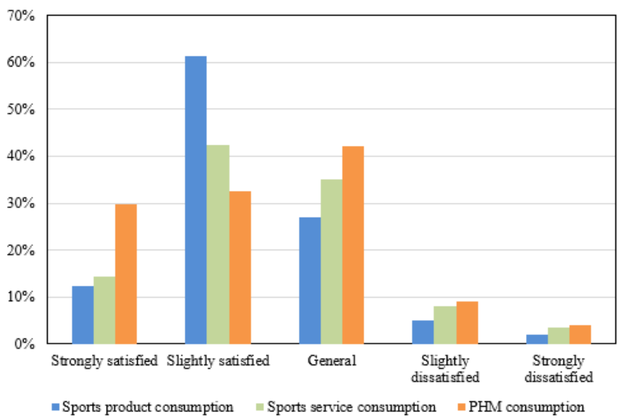

Figure 8. Satisfaction with different types of sports consumption

Figure 7 illustrates the convenience of various sports consumption methods. It can be noticed that sports product consumption accounted for a significant amount of strongly good and good convenience. In contrast, PHM consumption accounted for a significant portion of general and poor convenience. Physical sports products can be purchased online and offline. However, the consumption of PHM requires several supplementary facilities or software/hardware. Without venues or equipment, sports services cannot be purchased. Consequently, the three forms of sports consumption differ in terms of convenience.

The happiness with various sorts of sports consumption is depicted in Figure 8. It may be inferred that most customers were moderately or highly satisfied with their consumption of sports products but somewhat or strongly dissatisfied with their consumption of PHM. Overall, regional consumers purchase sports products above the use of sports services and PHM. Figure 9 examines the propensity to purchase additional sports consumption of various categories. Similar to the satisfaction results, customers in the region are more likely to acquire sports products in the future than sports services or PHM.

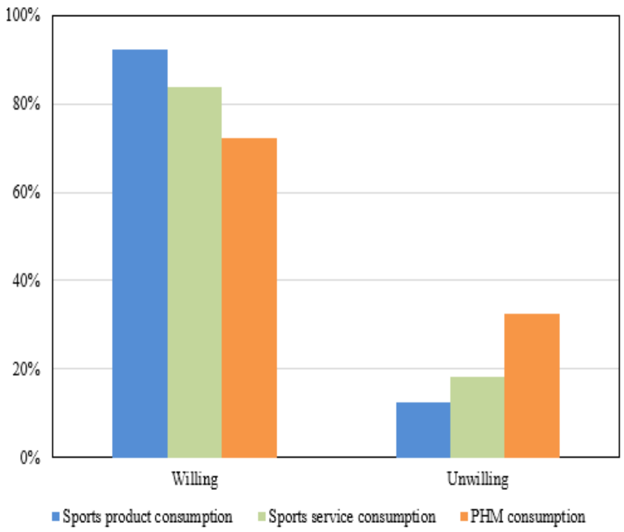

Figure 9. Willingness to further purchase different types of sports consumption

Several conclusions can be drawn from the above experimental results: (1) sports consumption demand promotes sports consumption and PHM consumption intentions; (2) economic foundation promotes sports consumption intention; (3) sports consumption satisfaction promotes sports consumption and PHM consumption intentions; (4) physical exercise habit promotes sports consumption and PHM consumption intentions; (5) physical exercise atmosphere promotes sports consumption and PHM consumption intentions.

7. Conclusions

This article investigates the PHM-focused SCB and investigates the pertinent marketing tactics. After elucidating the flow of associated data analysis on PHM-oriented sports consumption, the authors completed the clustering of multivariate sports consumption data, described a fine-grained correlation analysis approach for sports consumption data, and developed an algorithm for selecting the optimal results of CCA with kernels. Next, a survey was conducted on regional motives for sports consumption. A difference analysis was performed on the sports consumption levels of groups with distinct consumption kinds, geographies, and income levels. The regional SCB and marketing directions were summarized, the mean correlation coefficient curves were plotted under different kernel functions, and the correlations of PHM demand and PHM attitude with SCB were analyzed, revealing the convenience, satisfaction, and willingness to purchase again for various sports consumption types.

The information gathered regarding the demand for sports products and services is vitally significant to sports product and service providers. To become more competitive in the market, businesses must scan and study the current status and various impactors of SCB and develop relevant marketing strategies.

This study emphasizes the SCB characteristics, behavioral orientation, and impacting behavioral elements confronting PHM. A thorough comprehension of these concerns facilitates the identification of future sports customers, the strengthening of consumer loyalty, and the stimulation of sports consumption demand. Then, the rapid development of sports consumption, sports economics, and the sports industry is possible.

This investigation yields a multitude of outcomes for the SCB encountering the PHM. Nonetheless, there are certain restrictions. For instance, the authors did not compare consumer groups whose consumption behaviors and exercise habits may differ. In addition, the method of statistical analysis must be significantly enhanced to handle and evaluate data more effectively.

Acknowledgments

This paper is supported by Colleges and Universities' Young Key Teachers Fostering Program of Henan Province, China (Grant No.: 2019GGJS102) and the Program for Science & Technology Innovation Talents (Humanities and Social Sciences) in Universities of Henan Province, China (Grant No.: 2021-CX-006).

References

Acharya, B., & Adhikari, S. (2021). Household energy consumption and adaptation behavior during crisis: Evidence from Indian economic blockade on Nepal. Energy Policy, 148, 111998. https://doi.org/10.1016/j.enpol.2020.111998

Amasyali, K., & El-Gohary, N. (2021). Machine learning for occupant-behavior-sensitive cooling energy consumption prediction in office buildings. Renewable and Sustainable Energy Reviews, 142, 110714. https://doi.org/10.1016/j.rser.2021.110714

Balachander, K., & Paulraj, D. (2021). Building Energy Time Series Data Mining for Behavior Analytics and Forecasting Energy consumption. KSII Transactions on Internet and Information Systems, 15(6), 1957-1981. https://doi.org/10.3837/tiis.2021.06.001

Beusker, E., Stoy, C., & Pollalis, S. N. (2012). Estimation model and benchmarks for heating energy consumption of schools and sport facilities in Germany. Building and Environment, 49, 324-335. https://doi.org/10.1016/j.buildenv.2011.08.006

Chen, J. (2015). Research on the Current Situation of Sports Consumption of Students from College of Oujiang. The Open Cybernetics & Systemics Journal, 9(1), 1968-1973. http://dx.doi.org/10.2174/1874110X01509011968

Cibinskiene, A., Dumciuviene, D., & Andrijauskiene, M. (2020). Energy consumption in public buildings: The determinants of occupants' behavior. Energies, 13(14), 3586. https://doi.org/10.3390/en13143586

Delpla, I., Legay, C., Proulx, F., & Rodriguez, M. J. (2020). Perception of tap water quality: Assessment of the factors modifying the links between satisfaction and water consumption behavior. Science of the total environment, 722, 137786. https://doi.org/10.1016/j.scitotenv.2020.137786

Fujiwara, M. (2020). The determinants of consumption behavior in fast fashion-From the point of view of sns, smartphone technology and consumption values. Journal of the Japan Research Association for Textile End-Uses, 61(4), 299-307. https://doi.org/10.11419/senshoshi.61.4_299

He, J.-P., Stein, A. D., Humphrey, H. E., Paneth, N., & Courval, J. M. (2001). Time Trends in Sport-Caught Great Lakes Fish Consumption and Serum Polychlorinated Biphenyl Levels among Michigan Anglers, 1973–1993. Environmental science & technology, 35(3), 435-440. https://doi.org/10.1021/es001067p

Hu, S. (2014). Urban women sports goods consumption market research and analysis. Journal of Chemical and Pharmaceutical Research, 6(7), 2759-2761. https://www.jocpr.com/articles/urban-women-sports-goods-consumption-market-research-and-analysis.pdf

Islam, M. T., Dias, P., & Huda, N. (2021). Young consumers'e-waste awareness, consumption, disposal, and recycling behavior: A case study of university students in Sydney, Australia. Journal of Cleaner Production, 282, 124490. https://doi.org/10.1016/j.jclepro.2020.124490

Kearney, J. P., & Cole, D. C. (2003). Great Lakes and inland sport fish consumption by licensed anglers in two Ontario communities. Journal of Great Lakes Research, 29(3), 460-478. https://doi.org/10.1016/S0380-1330(03)70451-1

Lemas, D. J., Loop, M. S., Duong, M., Schleffer, A., Collins, C., Bowden, J. A., Du, X., Patel, K., Ciesielski, A. L., & Ridge, Z. (2021). Estimating drug consumption during a college sporting event from wastewater using liquid chromatography mass spectrometry. Science of the total environment, 764, 143963. https://doi.org/10.1016/j.scitotenv.2020.143963

Liu, F., Lim, E. T., Li, H., Tan, C.-W., & Cyr, D. (2020). Disentangling utilitarian and hedonic consumption behavior in online shopping: An expectation disconfirmation perspective. Information & Management, 57(3), 103199. https://doi.org/10.1016/j.im.2019.103199

Martynyuk, O., Sugak, L., Stepova, H., Antoniuk, V., & Mileiko, I. (2021). Models of Distributed Systems Testing Based on Energy Consumption in Behavior. In 2nd International Workshop on Intelligent Information Technologies and Systems of Information Security, March 24–26, 2021, Khmelnytskyi, Ukraine (pp. 481-493). CEUR-WS.org. http://ceur-ws.org/Vol-2853/paper52.pdf

Minghong, Z., & Xinke, Y. (2012). An Empirical Analysis of the Influencing Factors of Sports Consumption: Based on the Data from 1990 to 2009 in China. In W. Hu (Ed.), Advances in Electric and Electronics (pp. 325-331). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-28744-2_41

Niu, G. (2017). Research on Sports Consumption Structure of First-tier City Residents in China. AGRO FOOD INDUSTRY HI-TECH, 28(3), 3489-3491. https://www.teknoscienze.com/tks_issue/vol_283/

Silva, W. M., Martínez, L. C., Plata-Rueda, A., Serrão, J. E., & Zanuncio, J. C. (2020). Respiration, predatory behavior and prey consumption by Podisus nigrispinus (Heteroptera: Pentatomidae) nymphs exposed to some insecticides. Chemosphere, 261, 127720. https://doi.org/10.1016/j.chemosphere.2020.127720

Singh, D., Zerriffi, H., Bailis, R., & LeMay, V. (2021). Forest, Farms and Fuelwood: Measuring changes in fuelwood collection and consumption behavior from a clean cooking intervention. Energy for Sustainable Development, 61, 196-205. https://doi.org/10.1016/j.esd.2021.02.002

Tchetchik, A., Kaplan, S., & Blass, V. (2021). Recycling and consumption reduction following the COVID-19 lockdown: The effect of threat and coping appraisal, past behavior and information. Resources, Conservation and Recycling, 167, 105370. https://doi.org/10.1016/j.resconrec.2020.105370

Tian, H., Zhou, M., Zhou, T., & Wang, W. (2017). An empirical study on the current situation and influencing factors of sports consumption in different regions. Agro Food Industry Hi-Tech, 28(1), 1117-1121. https://www.teknoscienze.com/tks_issue/vol_281/

Tian, M. (2014). Adaptive filtering method-based China sports consumption development research. BioTechnology: An Indian Journal, 10(9), 3442-3452. https://www.researchgate.net/publication/297140302

Tseng, M.-L., Sujanto, R. Y., Iranmanesh, M., Tan, K., & Chiu, A. S. (2020). Sustainable packaged food and beverage consumption transition in Indonesia: Persuasive communication to affect consumer behavior. Resources, Conservation and Recycling, 161, 104933. https://doi.org/10.1016/j.resconrec.2020.104933

Wang, J., & Lv, B. (2019). Big data analysis and research on consumption demand of sports fitness leisure activities. Cluster Computing, 22(2), 3573-3582. https://doi.org/10.1007/s10586-018-2207-y

Waris, I., & Hameed, I. (2020). Promoting environmentally sustainable consumption behavior: an empirical evaluation of purchase intention of energy-efficient appliances. Energy Efficiency, 13(8), 1653-1664. https://doi.org/10.1007/s12053-020-09901-4

Wells, S., & Macdonald, S. (1999). The relationship between alcohol consumption patterns and car, work, sports and home accidents for different age groups. Accident Analysis & Prevention, 31(6), 663-665. https://doi.org/10.1016/S0001-4575(99)00025-1

Yang, C., & Ming, H. (2021). Detection of sports energy consumption based on IoTs and cloud computing. Sustainable Energy Technologies and Assessments, 46, 101224. https://doi.org/10.1016/j.seta.2021.101224

Zheng, P., & Xing, J. (2020). Boundedness and large-time behavior for a two-dimensional quasilinear chemotaxis-growth system with indirect signal consumption. Zeitschrift für angewandte Mathematik und Physik, 71(3), 98. https://doi.org/10.1007/s00033-020-01320-w