

基于K-prototype聚类的职业院校学生校园消费行为分析

——以宣城职业技术学院为例

庞波¹, 吴以兵², 汪青华³

(宣城职业技术学院 1 党委办公室; 2 学生处; 3 总务处, 安徽 宣城 242000)

摘 要: 职业院校数字化转型环境下, 数据挖掘技术成为促进职业院校教学管理精准化、决策科学化的重要途径。本文基于K-prototype聚类算法, 以宣城职业技术学院学生相关校园消费行为数据为切入点, 探究职业院校学生消费行为与学生主体的交互影响, 建立了四类学生校园消费行为类别, 分析不同类别学生群体的基本情况、学习情况、身心健康情况的具体特征及相关性。研究得出: 性别年龄因素与职业院校消费行为存在相关性; 校园月均消费金额较低的学生自主学习能力较弱; 校园日均消费次数较高的学生更容易产生相关心理问题。职业院校学生校园消费管理行为分析为新形势下学校数字治理和学生未来职业发展提供决策参考。

关键词: 数据挖掘; 职业院校; 学生校园消费; K-prototype 聚类

中图分类号: G641

文献标识码: A

文章编号: 1674-5337 (2023) 03-0025-12

一、研究背景

当前, 职业院校学生管理趋向信息化、数字化转变。世界数字教育大会职业教育数字化转型发展平行论坛指出, 提升数字化治理能力是转型与重塑职业教育新生态的重要主题。^[1]在学习贯彻党的二十大精神服务国家数字化战略行动中, 教育部教育管理信息中心党委书记、主任李建聪提出“提升教育数据治理水平。深度挖掘数据在驱动教育数字化发展的价值”^[2]。可以说, 随着教育数据治理逐步推进, 数据日渐成为人们分析教育、反思教学的一部分, 而如何挖掘数字治理中教育“数据化”呈现的学生问题, 以个性化助力学生发展成为关键^[3]。通过多年职业院校智慧校园建设, 目前大部分院校已具备基本信息化平台及各类应用

系统, 积累了大量数据资源, 是支撑和推进职业院校学生相关工作决策部署的有力基础。

近年来, 职业院校学生消费观念和消费行为的培养受到学校管理者和一线教师的重视, 同时, 在数据挖掘技术支持下也引发了教育研究者的关注。建立职业院校学生消费行为模型成为提高学生校园学习和生活质量, 支持学校贫困资助, 及实现个性化教育管理的重要途径。基于此, 本研究以大数据技术为依托, 以职业院校学生校园消费行为为研究中心, 通过关联学生主体基本信息、学习、身心健康层面相关数据, 分析不同消费行为学生主体特征, 精准引导职业院校学生树立正确消费观, 形成健康合理消费计划, 同时, 为职业院校学生思想教育和学院相关管理决策提供依据。

基金项目: 安徽省2019年职教教学学会教育科研规划一般课题“大数据技术在职业教育的应用”(Azcyj052)。

作者简介: 庞波 (1971—), 男, 宣城职业技术学院党委副书记、副教授。

二、文献综述

对于相关学生校园消费行为的已有研究主要包含以下几个部分：其一，针对“校园一卡通”系统挖掘学生消费相关数据，主要通过K-means聚类算法探究职业院校学生的消费行为偏好和消费水平。有学者在K-means聚类基础上，利用FP-growth关联规则算法设计学生消费行为分析系统^[4]，也有利用Apriori关联算法^[5]、主成分分析法^[6]判别学生消费水平，也有通过校园卡中学生图书馆图书借阅次数、生活习惯、饮食习惯构建学生校园行为画像^[7]。其二，通过实证研究分析学生消费行为。如通过网络数据资源或手机APP分析学生线上和线下消费行为和存在的问题^[8]；利用问卷调查中职学生消费行为和类型^[9]或进行归类比较^[10]。其三，对学生消费行为的影响因素进行探究。有学者提到学生自我财务管理行为受到学校所处环境影响^[11]，有学者通过实证研究指出学生金钱管理受其生活管理能力的影响^[12]，也有研究指出家庭、社会、学校、学生主体会对大学生消费行为产生影响^[13]。其四，从校园贷视角研究当代大学生消费观、消费现状及不良消费的危害^[14-15]。

纵观相关研究可以发现，我国对于学生消费行为的研究较为全面，且一直受到学者关注。由于职业院校学生本身的特殊性，其消费行为特征不局限于单一因素和单一类别，往往存在

多方面因素的共同作用。从校园大数据中获取学生真实消费数据来分析和预测学生不同层面行为，对学生发展和学院管理决策都具有一定意义和价值。

三、研究设计

（一）数据采集与处理

1. 数据采集。本文数据依托于宣城职业技术学院数字化校园数据库。基于该学院管理模式和信息设备，学生校园消费通过校园一卡通系统完成。其中，餐饮消费主要在食堂进行，学生生活类必需品由校内门面房提供，校园线下消费均可通过一卡通刷卡完成，宿舍水电费由一卡通线上平台缴费。基于此，采集宣城职业技术学院学生在2022年第二学期10月、11月校园食堂消费金额及次数为学生校内餐饮消费情况数据；获取校园一卡通系统中门面房消费、水电费等金额及次数（除去食堂消费部分）为学生校园日常生活消费情况数据。

同时采集此次校园消费行为数据对应的学生信息，以学生性别、年龄作为学生基本信息。学生学习层面，通过采集教务系统学生学业成绩班级排名与图书馆图书借阅次数分别评价其学习结果和自主学习能力；学生身心健康层面，采集学生出入宿舍测温设备系统中数据异常情况作为衡量学生身体健康状况指标，从学生校园信息管理系统中学生心理平台获取学生心理健康情况数据。具体如表1所示。

表1 职业院校学生校园消费行为相关数据采集维度及指标

维度	数据采集指标	维度	数据采集指标
校园消费行为	餐饮消费金额	基本信息	年龄
	餐饮消费次数		性别
	日常生活消费金额	学习层面	学业成绩班级排名(百分比)
	日常生活消费次数		图书馆图书借阅次数
学生		身心健康层面	体温异常次数
			心理健康情况

注:学业成绩为学生在校期间平均学年成绩班级排名;图书馆图书借阅次数为学生在校期间图书借阅次数;其余为2022年10月、11月数据

2. 数据处理。研究数据处理的过程分为两部分，一部分是对学生消费行为数据的处理。由于数据采集跨度为两个月，则后续将以月均餐饮消费金额和月均日常生活消费金额（10月、11月金额之和除以2）、日均餐饮消费次数和日均日常生活消费次数（10月、11月消费次数之和除以61）四项职业院校学生校园消费行为数值型数据进一步建模分析。另一部分是对学生相关数据的处理，由于获取了学生心理健康数据，因此需对学生学号进行转码加密。学习层面中新生缺失学年成绩专业排名数据，为保证相关消费行为数据的全面性和多样性，将其予以保留，同时，为保证数据质量和建模有效性，对校园一卡通系统中消费记录不全、基本信息缺失或不明确的数据样本进行剔除，经数据清

洗最终得到40552份样本数据。

3. 数据描述。综上得到职业院校学生校园消费行为数值型数据及分类型数据如下表2、表3所示。数值型数据（表2）中包含学生校园相关消费金额和次数的均值和中位数，其中，学生餐饮消费为食堂内消费情况，不包括学生外卖消费。分类型数据（表3）中共包含36个分类项，性别为2类；根据数据分布情况，将学生年龄在18岁以下、21岁以上单独分类，中间各年龄独成一类，共6类；平均学业成绩班级排名数据以每10%分类，共11类；图书借阅次数在0~5次，共6类；月均体温异常次数在5次及以上独成一类，5次以下各为一类，共6类；心理健康情况分类来源于心理健康平台采集结果，共5类。

表2 职业院校学生校园消费行为数值型数据表

数值维度	数值指标	均值	中位数
校园消费行为	月均日常生活消费金额(CB1)	249.00	200.65
	月均餐饮消费金额(CB2)	255.73	211.73
	日均日常生活消费次数(CB3)	1.42	1.08
	日均餐饮消费次数(CB4)	2.80	2.75

表3 职业院校校园消费行为主体分类型数据表

维度	数量	分类项	频数(百分比)	维度	数量	分类项	频数(百分比)
性别	2	男生	14349(35.38%)	在校期间图书馆图书借阅次数	6	0次	39512(97.44%)
		女生	26203(64.62%)			1次	532(1.31%)
年龄	6	未满18岁	3922(9.67%)			2次	284(0.70%)
		18岁	1652(4.07%)			3次	140(0.35%)
		19岁	10125(24.97%)			4次	40(0.10%)
		20岁	12517(30.87%)			5次	44(0.11%)
		21岁	8108(19.99%)			无异常	37759(93.11%)
		21岁以上	4228(10.43%)	体温异常次数	6	1次	2505(6.18%)
平均学业成绩班级排名	11	NULL	19088(47.07%)			2次	216(0.53%)
		前10%	1157(2.9%)			3次	44(0.11%)
		10%—20%	1832(4.52%)			4次	12(0.03%)
		20%—30%	2355(5.81%)			5次及以上	16(0.04%)
		30%—40%	2834(6.99%)	心理健康情况	5	正常	37402(92.23%)

续表

维度	数量	分类项	频数(百分比)	维度	数量	分类项	频数(百分比)
		40%—50%	2920(7.20%)			可能异常尚未约谈	1657(4.09%)
		50%—60%	2832(6.98%)			可能存在心理异常	196(0.48%)
		60%—70%	2687(6.63%)			在校适应异常	732(1.81%)
		70%—80%	2300(5.67%)			约谈后未发现异常	565(1.39%)
		80%—90%	1738(4.29%)				
		90%—100%	809(1.99%)				

(二) 研究方法

聚类算法是教育数据挖掘(EDM)中采用的算法之一,也是一种收集和呈现相似对象集群(组)的数据处理方法。教育数据聚类能将教育系统产生的主体数据进行组合^[14],通过算法将来自教育系统中的原始数据转化为对教育研究和实践更有意义的信息^[15],挖掘不同学习者的学习行为及教育教学活动中的教学风格,以此推进个性化学习和精准性教学。聚类算法中主流的包括K-means、K-medoids、Clara算法等^[16],主要应用于处理数值型数据。本研究中,由于获取的数据包含消费金额和频次等数值型数据,及学生性别、心理健康情况等分类型数据,因此采用聚类算法中K-prototype算法对相关数据进行聚类分析,得到不同类别的学生校园消费行为特征分布情况,如图1所示。

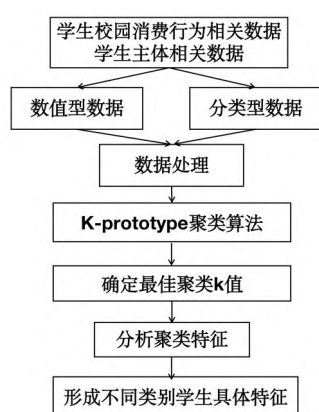


图1 基于K-prototype聚类算法学生消费行为分析框架

K-prototype算法是K-means算法和K-modes算法思想的继承^[17],其运用了描述数据簇的原型和混合属性数据间的相异性度量计算公式,

是处理混合属性(同时包括数值型和分类属性)聚类的典型算法^[18]。聚类过程中,若一个含有n个样本的混合属性数据集为 $X=X_1, X_2, \dots, X_n$,每个样本包含p个数值型属性, m-p个分类型属性,第i个样本为 $X_i=X_{i1}, X_{i2}, \dots, X_{ip}, X_{i(p+1)}, X_{i(p+2)}, \dots, X_{in}$ 。随机选取初始聚类个数为k, k个聚类中心表示为 $V=V_1, V_2, \dots, V_n$,其中, $V_j=V_{j1}, V_{j2}, \dots, V_{jp}, V_{j(p+1)}, V_{j(p+2)}, \dots, V_{jm}$ 。

K-prototype算法中混合型相异度为数值型相异度和分类型相异度分开求再相加所得。其中,数值型相异性度量基于欧式距离得到样本 X_i 与聚类中心 V_l 间的距离公式为:

$$d_{Num}(X_i, V_l) = \sum_{j=1}^p |X_{ij} - V_{lj}|^2$$

分类型相异性度量基于汉明距离得到的距离公式为:

$$d_{cat}(X_i, V_l) = \sum_{j=p+1}^m \delta(X_{ij}, V_{lj})$$

$$\delta(X_{ij}, V_{lj}) = \begin{cases} 0, & X_{ij} = V_{lj} \\ 1, & X_{ij} \neq V_{lj} \end{cases} \quad (p+1 \leq j \leq m)$$

混合类型数据的相异度度量计算公式为:

$$d(X_i, V_l) = d_{Num}(X_i, V_l) + \gamma d_{cat}(X_i, V_l)$$

K-prototype算法中两种类型数据间的权重采用权重系数 γ 进行调节^[19],一般来说,分类型数据在聚类分析中起到的作用比较小,即考虑将其权重设置为较小值,而数值型数据则可以设置较高权重,避免聚类结果值偏向分类型特征^[20]。本研究根据实际数据分类情况分别设

置六项分类型数据的权重系数。其算法步骤如

图2所示^[21]。

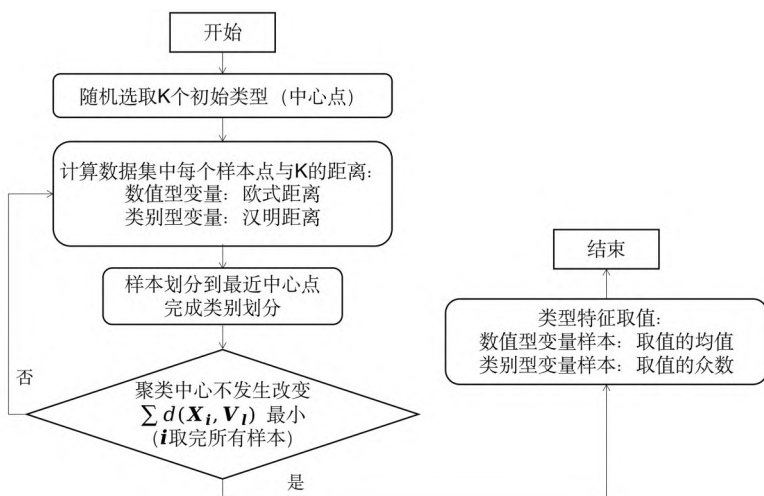


图2 K-prototype算法主要步骤流程图

本研究利用手肘法确定 K-prototype 聚类个数 K 值，其核心指标为样本聚类的误差平方和

(SSE)，
$$SSE = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2$$
 其中， C_i 为第 i 个

簇， p 为 C_i 的样本点， m_i 是 C_i 的质心。随着聚类数 k 的增大，样本划分趋向精细化，每簇聚合程度逐渐提高，SSE 逐渐变小。当 k 小于真实聚类值时，随着 k 值增加，每簇聚合程度增加，SSE 下降幅度较大；达到真实聚类值后，再增加 k 后，每簇聚合程度会迅速变小，SSE 下降幅度骤减后趋于平缓。本研究从 $k=2$ 开始取值，通过 python 实现 k 与 SSE 关系图如下图 3 所示，可以看出，当 $k=4$ 时，再增加 k 值后，SSE 下降趋势开始逐步减缓，由此确定真实聚类值 $k=4$ 。

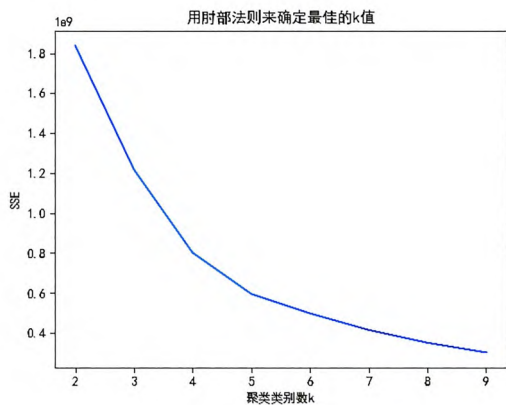


图3 聚类数 k 与 SSE 关系图

四、结果分析

(一) 职业院校学生校园消费行为聚类结果

利用 K-prototypes 聚类算法，以 $k=4$ 得到具有显著行为特点的四类职业院校学生校园消费行为模型。四簇包含的数据个数如表 4 所示，每类中学生人群分布较为均匀，表明整体聚类效果较好。其中，簇 3 类学生数量占比最高 (38.84%)，簇 1 类学生数量占比最低 (18.02%)。

表 4 学生校园消费行为聚类数据分布

类别编号	簇 1	簇 2	簇 3	簇 4
频数	7309	8955	15751	8537
(百分比)	(18.02%)	(22.08%)	(38.84%)	(21.05%)

根据聚类的方差得出各簇类差异性，对四簇中学生消费行为四项数值型变量的均值进行对比，其中，簇 3 类学生月均消费金额最低，日均消费频次最低，簇 4 类学生月均餐饮消费金额及日均餐饮消费频次最高，簇 1 类学生月均日常生活消费金额及日均日常生活消费次数最高，簇 2 类和簇 3 类学生的月均日常生活和餐饮消费水平相差较小，簇 1 类学生群体月均日常生活消费远高于餐饮消费。总体上，该职业院校学生

校园消费金额最高群体(簇4类)均值为937.2, 表5所示。
其消费水平显著高于簇3类学生(268.27), 如

表5 学生校园消费行为各类别方差分析差异对比结果

消费行为指标	簇1	簇2	簇3	簇4	F	p
	均值±标准差	均值±标准差	均值±标准差	均值±标准差		
CB1	439.30±137.94	202.65±138.91	116.27±90.72	410.78±201.01	13460.629	0.000**
CB2	173.83±118.81	246.23±154.95	152.00±127.95	526.42±131.94	15759.446	0.000**
CB3	2.45±0.92	0.78±0.66	0.60±0.56	1.09±0.91	12458.887	0.000**
CB4	2.69±0.88	3.20±0.69	1.82±0.82	4.38±0.97	17579.76	0.000**

* $p < 0.05$ ** $p < 0.01$

(二) 职业院校学生校园消费行为相关性分析

为进一步探究本研究中学生性别、年龄、学习情况、身心健康情况与学生消费行为的相关性, 宏观上通过SPSS 26.0对相关数据进行相关性和多元方差分析。

1. 相关性分析。以年龄为变量, 利用相关性分析四项学生校园消费行为与年龄间的关系, 如表6所示, 可以看出, 学生校园消费行为四项数据与年龄均呈现显著水平, 且Pearson相关系数均为负数, 即消费行为与年龄间为显著负相关关系。

表6 学生校园消费行为与学生年龄相关性分析

年龄		CB1	CB2	CB3	CB4
	相关系数	-0.079**	-0.075**	-0.073**	-0.017**
	显著性	0.000	0.000	0.000	0.000

* $p < 0.05$ ** $p < 0.01$

同时, 通过学生校园消费行为与年龄的散点图和线性趋势可看出(如图4所示), 两两变量间均呈下降型线性趋势, 宏观上可看出, 随着学生年龄的增长, 校园消费金额降低, 校园消费次数降低。

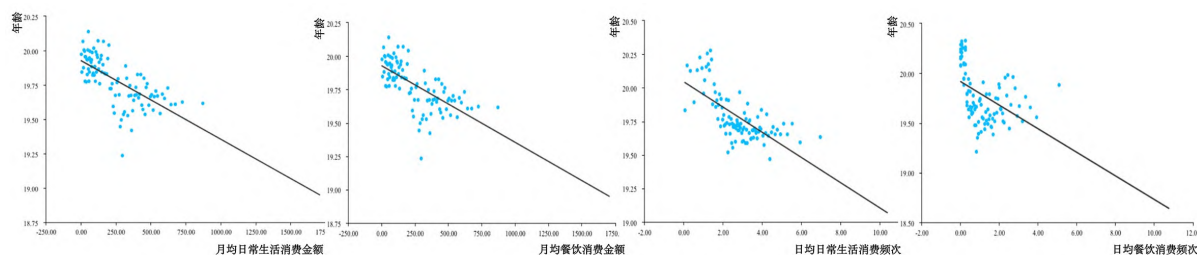


图4 职业院校学生校园消费行为与年龄关系散点图

2. 多元方差分析。考虑研究中为多个自变量与多个因变量的相互关系, 采用多元方差分析学生校园消费行为与学生主体的相关影响, 具体如表7所示, 可以看出, 在学生基本情况中, 性别与年龄对其校园消费行为的效应显著性为 $0.000 < 0.001$, 达到极显著水平。在学习情况中, 图书馆图书借阅次数对日均餐饮消费次数的显著性达到极显著水平, 对于月均餐饮消

费金额及日均日常生活消费次数的显著性均小于0.05, 也是显著的; 学生平均学业成绩排名对日均日常生活消费次数也达到显著水平($p = 0.023 < 0.05$)。在身心健康情况中, 体温异常次数对日均餐饮消费次数及日均日常生活消费次数的影响是显著的, 同时, 心理健康情况仅对日均餐饮消费次数的影响达到显著水平($p = 0.005 < 0.05$)。

表 7 职业院校学生校园消费行为与学生主体间效应检验

		df	均方	F	显著性	偏 Eta 平方
基本情况	性别	CB1	1	15098187.93	412.709	0.000
		CB2	1	17489527.58	468.493	0.000
		CB3	1	108.5	63.905	0.000
		CB4	1	267.888	167.941	0.000
	年龄	CB1	5	2971767.366	81.233	0.000
		CB2	5	3134278.317	83.958	0.000
		CB3	5	31.481	18.542	0.000
		CB4	5	285.392	178.915	0.000
学习情况	图书馆 图书 借阅 次数	CB1	5	137555.672	3.653	0.003
		CB2	5	121633.195	3.162	0.007
		CB3	5	5.086	2.99	0.011
		CB4	5	16.543	10.161	0.000
	平均 学业 成绩 排名	CB1	10	18402.61	0.489	0.899
		CB2	10	14437.27	0.375	0.958
		CB3	10	3.536	2.078	0.023
		CB4	10	2.136	1.312	0.217
身心健康情况	体温 异常 次数	CB1	5	50467.278	1.334	0.246
		CB2	5	59987.789	1.551	0.17
		CB3	5	4.946	2.898	0.013
		CB4	5	7.74	4.7	0.000
	心理 健康 情况	CB1	4	41928.7	1.108	0.351
		CB2	4	41366.44	1.07	0.37
		CB3	4	1.061	0.622	0.647
		CB4	4	6.164	3.743	0.005

表 8 四类校园消费型学生群体基本信息分布表

簇类		簇 1 类	簇 2 类	簇 3 类	簇 4 类
性别	男	3868(52.92%)	1876(20.95%)	4007(25.44%)	4598(53.86%)
	女	3441(47.08%)	7079(79.05%)	11744(74.56%)	3939(46.14%)
年龄分布	未满 18 岁	605(8.28%)	1209(13.05%)	1314(8.34%)	794(9.30%)
	18 岁	251(3.43%)	435(4.86%)	674(4.28%)	292(3.42%)
	19 岁	2884(39.46%)	2395(26.74%)	2918(18.53%)	1928(22.58%)
	20 岁	1682(23.01%)	2122(23.70%)	5213(33.10%)	3500(41.00%)
	21 岁	1361(18.62%)	1848(20.64%)	3403(21.60%)	1496(17.52%)
	21 岁以上	526(7.20%)	946(10.56%)	2229(14.15%)	527(6.17%)

（三）职业院校学生校园消费行为类别特征分析

1. 学生基本信息特征分析。四簇学生类别学生基本信息具体分布如表8所示，学生数据中男女所占比例为14349：26203，女生约为男生的1.8倍。

四类学生群体性别比例分布直方图如图5（a）所示。其中，簇2类学生群体中，女生数量远高于男生数量，约为男生数量的3.7倍；簇3类学生群体中，女生数量约为男生数量的2.9倍；簇1类学生群体和簇4类学生群体中，男女占比大致相同，且男生多于女生。根据学院学生男女比例及各簇类男女分布情况可看出，男生的校园消费金额整体上高于女生。另外，数据中学生年龄在15~28岁，其中，以18岁为界限，小于18岁为中职学生群体，18岁以上为高职和大专学生群体。四类中各年龄学生人数分

布情况如图5（b）所示，可以看出，中职学生群体中，簇3类学生居多；高职、大专学生群体中，校园较高消费和较低消费的大多为20岁；生活消费显著高于餐饮消费的学生群体大多为18~19岁年龄段学生，该阶段学生大多为新生群体。

2. 学习情况分析。四类群体具体学习情况分布如表9所示，在四类群体中，图书馆图书借阅次数最多为5次，最多分布于簇4类学生群体，图书馆图书借阅0次占比由高到低分别为簇3（98.22%），簇1（97.32%），簇2（96.84%），簇4（96.71%）。四类校园消费型学生群体图书馆图书借阅次数折线图如图6所示，簇2和簇4类学生自主学习能力强于簇1和簇3类学生，由此可见，校园消费水平较低的学生自主学习能力较弱，日常生活消费水平高于餐饮消费的学生群体在学习中表现出较弱积极主动性。

表9 四类校园消费型学生群体学习情况分布表

簇类	簇1类	簇2类	簇3类	簇4类
图书馆 图书 借阅 次数	0次	7113(97.32%)	8672(96.84%)	15471(98.22%)
	1次	100(1.37%)	134(1.50%)	157(1.00%)
	2次	54(0.74%)	91(1.02%)	68(0.43%)
	3次	26(0.36%)	31(0.35%)	42(0.27%)
	4次	5(0.07%)	16(0.18%)	7(0.04%)
	5次	11(0.15%)	11(0.12%)	6(0.04%)
	null	4461(61.03%)	4343(48.50%)	6009(38.15%)
平均学业 成绩排名	前10%	197(2.70%)	181(2.02%)	617(3.92%)
	10%—20%	265(3.63%)	487(5.44%)	728(4.62%)
	20%—30%	369(5.05%)	577(6.44%)	972(6.17%)
	30%—40%	441(6.03%)	712(7.95%)	1209(7.68%)
	40%—50%	510(6.98%)	612(6.83%)	1109(7.04%)
	50%—60%	398(5.45%)	497(5.55%)	1270(8.06%)
	60%—70%	272(3.72%)	431(4.81%)	1395(8.86%)
	70%—80%	195(2.67%)	584(6.52%)	1093(6.94%)
	80%—90%	127(1.74%)	343(3.83%)	925(5.87%)
	90%—100%	74(1.01%)	188(2.10%)	424(2.69%)

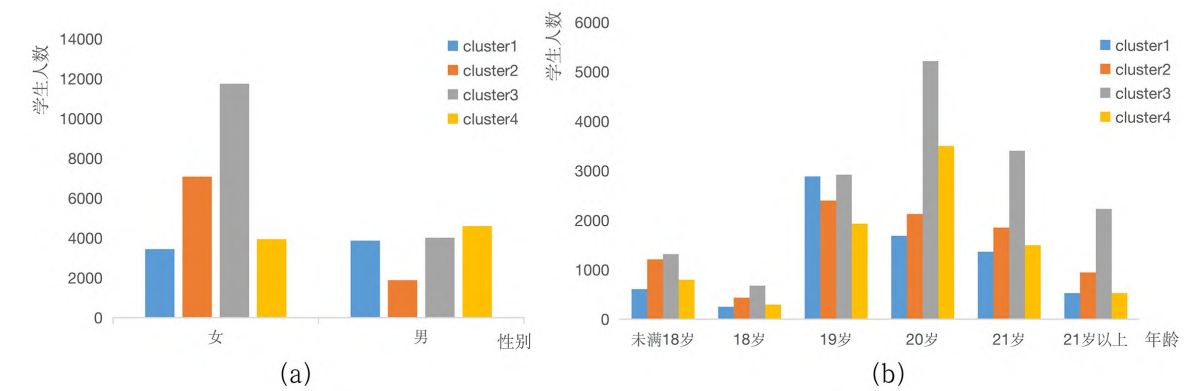


图5 四类校园消费型学生群体性别年龄分布直方图

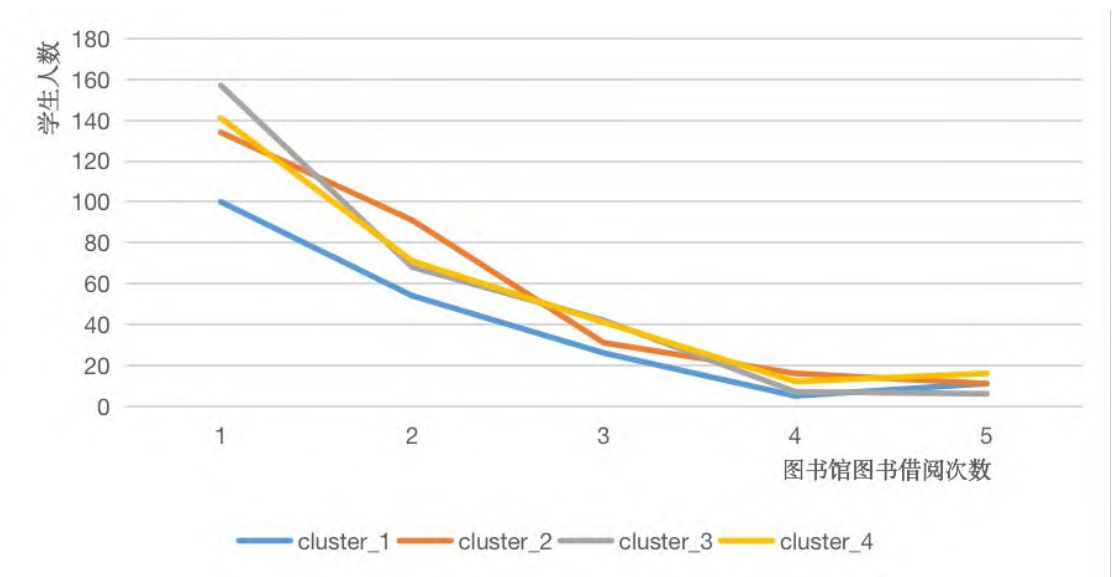


图6 四类校园消费型学生群体图书馆图书借阅次数折线图

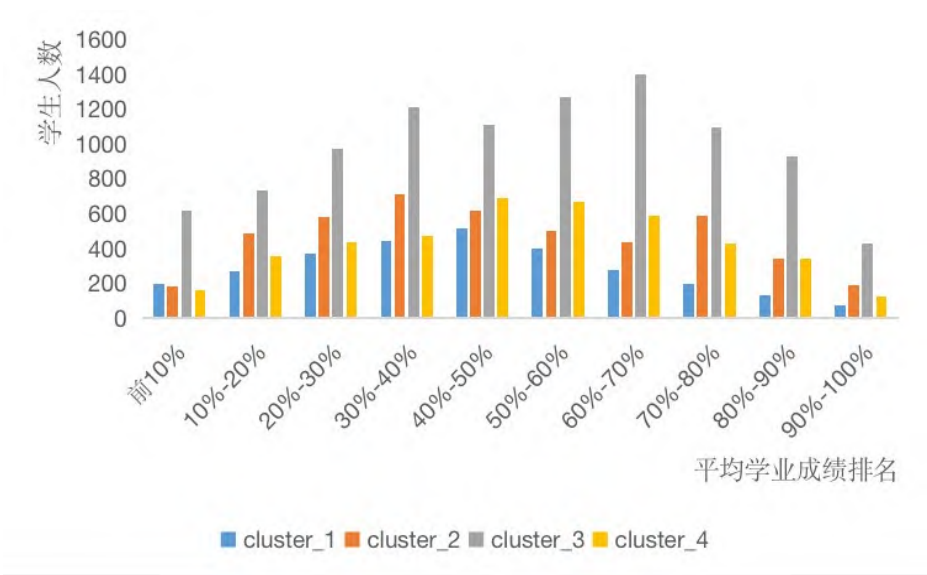


图7 四类校园消费型学生群体平均学业成绩排名直方图

表 10 四类校园消费型学生群体身心健康情况分布表

簇类		簇 1 类	簇 2 类	簇 3 类	簇 4 类
体温异常次数	无异常	6875(94.06%)	8391(93.70%)	14494(92.02%)	7999(93.70%)
	1 次	379(5.19%)	519(5.80%)	1120(7.11%)	487(5.70%)
	2 次	42(0.57%)	39(0.44%)	90(0.57%)	45(0.53%)
	3 次	11(0.15%)	6(0.07%)	22(0.14%)	5(0.06%)
	4 次	0(0.00%)	0(0.00%)	12(0.08%)	0(0.00%)
	5 次及以上	2(0.03%)	0(0.00%)	13(0.08%)	1(0.01%)
	正常	6666(91.20%)	8292(92.60%)	14585(92.60%)	7859(92.06%)
心理健康情况	可能异常尚未约谈	331(4.53%)	351(3.92%)	628(3.99%)	347(4.06%)
	可能存在心理异常	34(0.47%)	47(0.52%)	67(0.43%)	48(0.56%)
	在校适应异常	147(2.01%)	155(1.73%)	266(1.69%)	164(1.92%)
	约谈后未发现异常	131(1.79%)	110(1.23%)	205(1.30%)	119(1.39%)

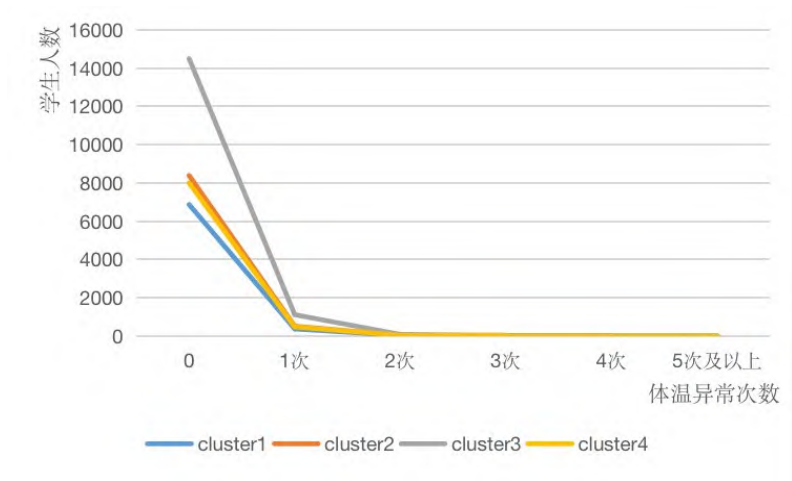


图 8 四类消费型学生群体健康异常次数折线图

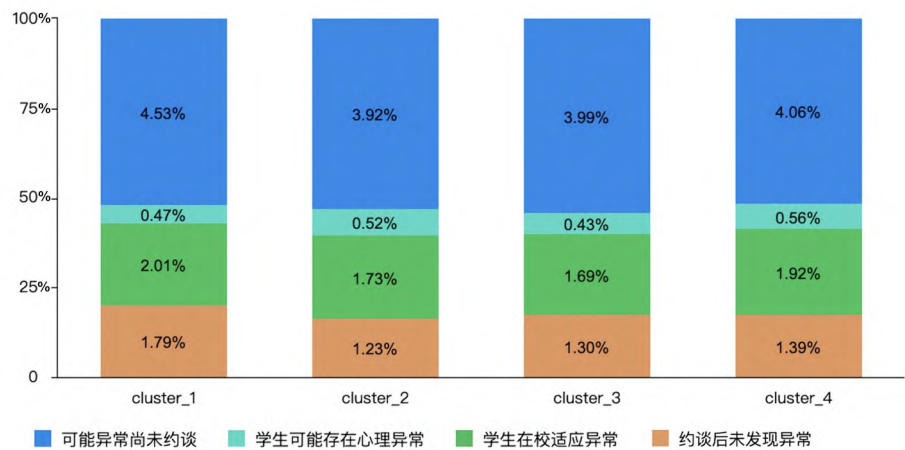


图 9 四类校园消费型学生群体心理情况占比堆积图

对四类学生消费行为中的学习情况进行分析,得到学科成绩排名直方图(图7)。根据已有排名数据,排名在班级前10%的学生在簇3类学生中占比最高(3.92%),簇4类中占比最低(1.90%);四类消费行为学生群体中中等生(班级排名位于40%—60%间)占比由高到低分别为,簇4类(15.88%)、簇3类(15.10%)、簇1类(12.42%)、簇2类(12.38%);排名在班级90%—100%的学生在簇3类中占比最高,为2.69%。由此宏观上得出,校园日均生活消费次数较高的学生群体中多数学业成绩在中等偏下水平。

3. 身心健康情况分析。四类校园消费型学生群体身心健康情况如表10所示,四类学生群体中出现健康问题的簇类学生由高到低为簇3(7.98%),簇4(6.30%),簇2(6.30%),簇1(5.94%)。由四类消费型学生群体健康异常次数折线图(图8)可以看出,月均消费金额少、日均消费次数低的学生群体出现健康异常次数较高,月均日常生活消费高于餐饮消费的学生群体出现健康异常次数最低。

关联四类校园消费型学生群体得到心理情况占比堆积图如图9所示,其中“学生可能存在心理异常”的占比由高到低分别为簇4类(0.56%),簇2类(0.52%),簇1类(0.47%),簇3类(0.43%)。“学生在校适应异常”的占比由高到低分别为簇1类(2.01%),簇4类(1.92%),簇2类(1.73%),簇3类(1.69%)。由此得出,校园消费较高群体学生比消费较低群体学生更可能存在心理问题;校园生活消费高于餐饮消费的学生群体更有可能存在在校适应异常的心理情况。

五、研究总结

整体上,本研究结果得到的职业院校学生校园消费结构合理,学生基本情况层面,男生校园消费高于女生,高年级学生消费管理能力强于低年级学生。学习层面,自主学习能力较弱的学生月均消费金额低,同时月均生活消费金额远高于餐饮消费。身心健康层面,出现体温异常次数高的学生日均消费次数低,学生存在在校适应异常或可能存在心理问题的学生日

均生活消费次数更高。

同时,针对四类校园消费型的学生群体应采用个性化的引导制定相应教育模式。对簇1类学生,其日常校园生活开支较高且男生居多,其可能在于水电费和在门面房的消费居高,学习主动性较弱,在校适应异常人数占比最多,应提早发现并关注这类学生的日常生活情况,开展校园户外活动,引导其形成良好生活作息习惯。对簇2类学生,其整体上自我消费管理能力较强,能基本保证一日三餐饮食规律,学习成绩较为优秀,应积极在校园内利用学生榜样作用,发扬积极消费观。对簇3类学生,其整体上校园消费水平较低,消费次数较低,可能部分为走读生或外卖消费情况,这类学生出现健康异常次数最多,学业成绩排名在班级前10%和后10%的学生较多,应引导其做好学习计划和日常生活计划,合理规划和利用时间。对簇4类学生,其校园消费水平最高,心理可能存在异常情况的占比最高,这类学生也需及时干预,精准到个人开展谈话,了解其消费行为和心理问题是否存在关联。总体上,职业院校学生主体特征复杂多样,班杜拉三元交互决定理论把个体自身主体因素、行为因素、环境因素作为相互决定的实体^[22],由此从三个层面对培养职业院校学生消费管理能力提出建议。

在学生主体层面,培养学生消费认知能力,加强校园消费观引导教育。由于职业院校学生生理和心理特征的差异性,学校管理工作应最大限度调动学生的积极主动性,对于学习管理能力较强的学生需积极挖掘其生活管理的潜能,定期开展多样化主题校园生活圈活动,关联学生专业技能,激发学生参与热情,创造条件鼓励学生主动自我发展;另外,积极推动校园消费观文化环境建设,潜移默化形成学生管理消费的意识,积极调整和反思不良消费行为,培养学生养成良好消费观念和消费计划。

在学生行为层面,融合学生校园生活与学习,建立个性化与多样化评价机制,完善学校管理工作制度体系。职业院校评价机制需根据不同年级学生特点多维度展开评价,及时增强学生自我效能感与成就感,让学生发现自身不足,同时,阶段性记录学生学习、生活情况,引导学生对不同阶段的差异进行反思和总结,积极合理作出调整和适当约束;避免只注重学

习培养而忽视生活的管理模式,完善校园消费场所管理,减少学生校园不良消费隐患;以学生真实数据支撑优化管理工作,进一步完善职业院校数据治理框架。

在环境层面,系紧家庭和学校教育联系的绳索,合力构建学生健康消费观。本文中职业院校学生校园消费行为也会受到外界环境(如学校和家庭环境)的作用,学生消费观念会“遗传或继承”原生家庭父母的消费习惯^[23],家庭教育中要给学生必要的自我管理空间和权利,学校环境间接影响学生消费管理能力的培养,辅导员应主动关注特殊家庭背景的学生,与学生家庭定期沟通形成教育合力,让学生在被理解和认可的同时严格要求自己,从口头上的服从转变为行动上的主动执行。

本研究数据仅来源于一所职业院校,研究结论可能存在一定局限性;消费行为相关指标的关联性是否完全等同于指标间的因果性还需进一步验证。下一步对于职业院校学生校园消费行为的研究将关注相关外界环境要素的作用,探究影响职业院校学生消费行为的更多成因。同时,还将在数字化校园中继续挖掘数据,探寻数据模型的理论支撑,建立职业院校学生自我管理行为模型,体现数据价值,以期提升职业院校数据治理工作实效,为职业院校学生科学规划学习、厘清生活目标提供指导方向。

参考文献:

- [1] 世界数字教育大会职业教育数字化转型发展平行论坛在京举行[EB/OL].http://www.moe.gov.cn/jyb_xwfb/xw_zt/moe_357/2023/2023_zt01/mtbd/202302/t20230216_1045077.html.
- [2] 学习贯彻党的二十大精神,服务国家教育数字化战略行动[EB/OL].http://www.moe.gov.cn/jyb_xwfb/xw_zt/moe_357/jjyzt_2022/2022_zt17/xxgc/xxgc_bt/202212/t20221226_1036169.html.
- [3] 李璇律,丁念金.数字治理的教育“数据化”危机:出场、运作与转化[J].中国电化教育,2022(9).
- [4] 冯拥武.基于数据挖掘技术的大学生消费行为分析系统的设计研究[D].兰州:西北民族大学,2020.
- [5] 徐凯.基于高校智慧校园平台大数据的学生行为研究[D].广州:广东工业大学,2021.
- [6] 范媛,蔡敏.基于校园卡消费数据的学生消费水平判别模型建立[J].计算技术与自动化,2021(2).
- [7] 张艳分,卢小清,刘禹,等.基于大数据平台的大学生校园行为探析[J].中国教育信息化,2019(1).
- [8] 梅磊.大数据背景下高职院校学生消费行为分析及其正面引导策略——基于四川文化传媒职业学院[J].中国教育信息化,2016(21).
- [9] 罗晓芳.中职学生的消费行为对幸福感的影响研究[D].成都:四川师范大学,2019.
- [10] 张淑萍.当代中职生消费理念的调查分析[J].才智,2014(23).
- [11] 吴伶.高职学生自我管理能力分析[J].职业,2012(33).
- [12] 李军.我国高职学生自我管理能力的影响因素及策略研究[D].南昌:南昌大学,2016.
- [13] 刘宝明.当代大学生消费行为现状分析及问题研究[D].大连:大连理工大学,2008.
- [14] Dutt A, Ismail M A, Herawan T. A Systematic Review on Educational Data Mining [J]. IEEE Access, 2017(5).
- [15] Romero C, Ventura S. Educational data mining: A review of the state of the art [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2010(6).
- [16] 贺玲,吴玲达,蔡益朝.数据挖掘中的聚类算法综述[J].计算机应用研究,2007(1).
- [17] 翟鸣宇,程建,王苏桐,等.基于K-prototype聚类的学生教育画像分析[J].大连理工大学学报(社会科学版),2021(6).
- [18] 李若楠.混合数据的多视图聚类算法研究[D].长春:东北师范大学,2022.
- [19] Guo Dong-wei, Chen Ying-jie, Chen Jing-wen. A k-prototypes algorithm based on adaptive determination of the initial centroids [C]. Proceedings of the 10th International Conference on Machine Learning and Computing, New York: ACM Press, 2018.
- [20] 贾子琪,宋玲.一种面向混合型数据聚类的k-prototypes聚类算法[J].小型微型计算机系统,2020(9).
- [21] 陈韩,王雷,蒋子云.基于K-prototypes的混合属性数据聚类算法[J].计算机应用,2010(8).
- [22] 周文霞,郭桂萍.自我效能感:概念、理论和应用[J].中国人民大学学报,2006(1).
- [23] 李静,沈书生,刘芳铭,等.大数据视阈下当代大学生消费行为新特征的调查研究[J].中国电化教育,2017(12).

(责任编辑:李永山)

Confrontation and Reconstruction: The Evolution of Interpersonal Ecology and Layer Differentiation in University Dormitories

Ren Haihua^{1,2}

(1 School of Marxism, Nantong Vocational University, Nantong, Jiangsu 226007; 2 Yangzhou University Base of Jiangsu Province Research Center for the Theoretical System of Socialism with Chinese Characteristics, Yangzhou, Jiangsu 225009)

Abstract: As an important part of the school, the dormitory is a special field that connects the school with the family and society. It has the basic function of environmental education and also serves as an important space for the basic social form, that is, the sub-social form. The "stable-conflict-stable" interpersonal ecology hidden within university dormitories determines the sub-social form of university dormitories. Based on a deep analysis of the original structure of university dormitories, this paper explores the spatiotemporal evolution trajectory of interpersonal ecology in university dormitories and explores the internal force that drives the dynamic adjustment of interpersonal ecology in university dormitories through layer differentiation using layer theory. The interest layer promotes layer breaking and achieves the first layer differentiation; the emotional layer brings psychological distance closer and achieves group emotional connection; the interest layer enhances communication dimensions and achieves group relationship harmony; and the value layer consolidates interpersonal ecology and achieves stability of interpersonal ecology. This further enhances the effectiveness of ideological and political education in universities under dormitory time and space.

Key words: universities; dormitories; interpersonal ecology; confrontation; layer differentiation

Analysis of Campus Consumption Behavior of Students in Vocational Colleges Based on K-prototype Clustering

—Taking Xuancheng Vocational and Technical College as an example

Pang Bo¹, Wu Yibing², Wang Qinghua³

(1 Party Committee Office, Xuancheng Vocational and Technical College, Xuancheng, Anhui 242000, China; 2 Student Affairs Division, Xuancheng Vocational and Technical College, Xuancheng, Anhui 242000, China; 3 General Affairs Division, Xuancheng Vocational and Technical College, Xuancheng, Anhui 242000, China)

Abstract: In the digital transformation environment of vocational colleges, data mining technology has become an important way to promote precise teaching management and scientific decision-making in vocational colleges. Based on the K-prototype clustering algorithm, this paper takes the data of campus consumption behavior related to students of Xuancheng Vocational and Technical College as the entry point to explore the interactive influence between students' consumption behavior and student subjects in vocational colleges. It establishes four categories of students' campus consumption behavior and analyzes the specific characteristics and correlations of basic conditions, study conditions, physical and mental health conditions of different categories of student groups in vocational colleges. The research concludes that gender-age factors are correlated with consumption behavior in vocational colleges; students with lower average monthly campus consumption have weaker independent learning ability; students with higher average daily campus consumption are more likely to have related psychological problems. The analysis of campus consumption management behaviors of students in vocational colleges provides decision-making reference for digital governance of schools and students' future career development under the new situation.

Key words: data mining; vocational colleges; student campus consumption; K-prototype clustering