

零售客户精准分类研究——以唐山市卷烟零售户为例

孙炳华*, 孙丹丹*, 宁淑娟

唐山市路北区烟草专卖局(分公司), 河北 唐山

收稿日期: 2023年9月22日; 录用日期: 2023年10月8日; 发布日期: 2023年11月3日

摘要

为提升烟草行业诚信互助小组功能, 深入了解社群群体成员的基本情况, 以深挖社群运营价值, 建立更有效的小组社群化运行模式, 本文以唐山市卷烟零售客户为例, 根据烟草行业卷烟销售特点, 选出零售客户的15项基本属性指标, 包括人像指标、店铺属性、经营属性、行为特征四个维度, 并K-means聚类分析方法对零售户进行精准分类, 从类内误差和离散度两方面验证聚类结果。本文收集了730份有效零售户样本, 利用K-means聚类法划分为5类社群小组, 分类结果较好, 根据各聚类簇的特征, 提取零售客户的当前和潜在价值, 为烟草公司实施社群化管理服务提供依据。

关键词

客户精准分类, K-means聚类, 社群运营

A Study on Precise Classification of Retail Customers: Taking Tangshan Cigarette Retailers as an Example

Binghua Sun*, Dandan Sun*, Shujuan Ning

Tangshanlubei District Tobacco Monopoly Bureau (Branch), Tangshan Hebei

Received: Sep. 22nd, 2023; accepted: Oct. 8th, 2023; published: Nov. 3rd, 2023

Abstract

In order to enhance the function of the integrity mutual assistance group in the tobacco industry,

*第一作者。

gain a deep understanding of the basic situation of community members, explore the operational value of the community, and establish a more effective group community-based operation mode, this article takes Tangshan cigarette retail customers as an example, selects 15 basic attribute indicators for retail customers based on the sales characteristics of cigarettes in the tobacco industry, including four dimensions: portrait indicators, store attributes, business attributes, and behavioral characteristics, and K-means clustering analysis method is used to accurately classify retail households, verifying the clustering results from two aspects: intra class error and dispersion. This article collected 730 valid retail customer samples and divided them into 5 types of community groups using K-means clustering method. The classification results were good. Based on the characteristics of each cluster, the current and potential value of retail customers were extracted, providing a basis for tobacco companies to implement community-based management services.

Keywords

Accurate Customer Classification, K-Means Clustering, Community Operations

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来,随着卷烟市场化改革的不断推进,烟草公司持续推进诚信互助小组建设,优化小组功能,零售户经营卷烟的盈利水平也得到了较大提高,但卷烟零售户诚信互助小组建设也存在各种问题。比如烟草公司的运营服务模式难以创新突破,信息协同效率较低,小组参与活跃度不足等。在此背景下,社群运营理念受到了行业的关注。近年来,烟草公司对于社群运营以规范并提升小组建设能力为目标,采取多种方式,如创新培训模式,大力开展平台操作培训;小组活动分解到片到线到周,稳步实施;广泛征集客户建议需求,瞄准客户关注内容;更好搭建“交流台”。但长期存在建设与运营不匹配的问题。据此,为提升诚信互助小组的服务水平,本文借助社群化运营的理念,通过对零售客户进行精准分类,在此基础上根据分类结果采取不同的服务策略。

2. 文献综述

在客户分类指标的选择中,学者们大多基于客户细分理论,即五维度细分,地理细分、人口细分、行为细分、心理细分和价格-利益细分[1]。王萍[2]、闫相斌[3]等用客户特征变量和购买记录变量来分析客户的购买行为;尹朝阳[4]、李苏建[5]等指出卷烟客户分类要考虑客户销售行为和基本属性等因素;蒋国瑞[6]、万映红[7]等也选择客户行为的频数属性和基本属性对客户进行分类研究。客户分类是客户分析的基础,对分类后的数据进行挖掘更具针对性。学者们对于聚类算法应用于客户细分的探索和实践已有较多案例。如 P-median 聚类、K-means、SOM、系统聚类、模糊 C 均值(FCM)聚类等[8]。Chen S.等提出利用模糊层次分析法来评价 B2C 电商公司的客户价值[9]。叶志龙等改进 RFM 模型从当前价值、潜在价值、忠诚度三个方面建立适用于线上企业的客户价值评价及分类指标体系[10]。

已有的文献为本文进行零售客户的精准分类提供了有益的借鉴,但仍然存在着一些不足。一是,行业内的文献多从经营维度来对零售客户进行分类,未考虑客户本身的行为特征;二是,部分文献对零售客户分类结果的应用多从投放角度考虑,未考虑客户服务。为此,本文针对诚信互助小组的运营目标设

计问卷,对问卷数据进行预处理后保留有效数据,并利用 K-means 聚类算法进一步细分客户。

3. 研究方法

3.1. 客户标签库建立 - 问卷指标的设计

在 2018 年全国网建(浙江)现场会上,国家局提出:“要将小组建设作为一项基础性、系统性、长期性工作,反复抓,抓反复,使之真正成为密切客我关系、提升渠道掌控力的长效举措。”社区运营化的特征是以共同的兴趣、爱好、利益吸引聚集,做有态度的内容,做圈层化互动,以实现互帮互助,共建共赢。对于现存的诚信互助小组活动参与度与活力不够、组长的作用不突出问题,如何挖掘客户真正需求及潜在价值是关键,根据客户细分理论,一般从外在属性、内在属性、消费行为分类等方面考虑。本文从人像指标、店铺属性、经营属性、行为特征四个维度设计零售客户问卷,通过四个维度的描述,分类并挖掘客户价值,进而指导社区小组运营,引导互助互利共赢。

具体问卷指标体系见表 1。

Table 1. Index system of cigarette retail customer questionnaire

表 1. 卷烟零售客户问卷指标体系

一级指标	二级指标	问题	选项
人像指标	地理位置、性别、年龄、学历背景、年收入	1.请选择您所在的城市	唐山;其他
		2.您的性别	男;女
		3.您的年龄	20~29岁;30~39岁;40~49岁;50~59岁;60~69岁;69岁以上
		4.您的学历背景	高中及以下;专科及以上
		5.您的年收入区间	5万以下(不含5万);5~10万(不含10万);10~30万
店铺属性	市场类型、业态类型、前柜数量、背柜数量、经营面积	6.请选择您店铺的市场类型	城网;农网
		7.请选择您的业态类型	烟酒店;便利店;食杂店;超市;娱乐服务业
		8.请输入您的背柜数量	[填空题]
		9.请输入您的前柜数量	[填空题]
经营属性	有品吸区、经营偏好、终端类型、扫码系统情况	10.请填写您的卷烟经营面积	[填空题]
		11.请选择您所处的卷烟经营档位区间	30~27档;26~16档;15~8档;7~1档
		12.您的店铺是否有品吸区	是;否
		13.您的店铺终端类型	普通终端;加盟终端;合作终端;一般现代终端
		14.您的烟草扫码系统使用情况	未安装扫码系统;扫码;全品种扫码
行为特征	兴趣爱好	15.您的经营偏好(多选)	卷烟;酒类;茶类;其他
		16.您的兴趣爱好	运动;文娱;阅读;游戏;旅行;社交;公益;其他

①人像指标:主要从地理位置、性别、年龄段、学历背景、年收入进行衡量,地理位置用于筛选特定样本,性别等指标可直接反应出烟草零售客户的内在属性。②店铺属性:主要从市场类型、业态类型、前柜数量、背柜数量、经营面积指标进行衡量,反映出当前店铺的基本情况。③经营属性:主要从有品

吸区、经营偏好、终端类型、扫描系统情况进行衡量，反应出零售户对卷烟销售的经营偏好。④行为特征：评价指标主要为兴趣爱好，是反映小组成员个性化特征的重要指标。

指标体系中背柜数量、前柜数量、卷烟经营面积为定量指标，其余均为定性指标。本次问卷共收回 1117 份，有效问卷 730 份，有效率 65.4%。

3.2. 基于 K-means 聚类的客户分类方法

本文选择 K-means 聚类分类法，K-means 是一种快速聚类法，算法能够处理大数据，具有高效率的优点，且模型的解释性强，应用广泛。具体实现过程为：事先指定分类数 k ，算法开始随机从数据集选择 k 个样本点作为初始聚类点，然后遍历整个数据集的各个样本，将每个样本归到离它最近的初始点所在的类中，之后以各个类的记录的均值中心点取代之前的初始聚类点，如此不断迭代运算，直到目标函数 J 的取值不再发生变化为止，实现收敛，循环迭代的次数可以人为设定。

本文基于社群运营的五类典型人群“教父型”、“圈子型”、“学习型”、“服务型”、“产品型”，将分类类别定为 5。

K-means 分类实施步骤如下：

第一步：首先筛选唐山市样本，并对数据进行预处理。考虑数据合理性，卷烟经营面积通常小于 70 m^2 ，前柜数量通常小于 13，背柜数量通常小于 8。共剔除无效数据 387 个，剩余有效数据 730 个。

第二步：将三个定量数据背柜数量、前柜数量、卷烟经营面积进行标准化处理，标准化公式为 $y_i = \frac{x_i - \bar{x}}{s}$ ， \bar{x} 为序列的均值， s 为方差，新序列 y_1, y_2, \dots, y_n 的均值为 0，方差为 1，且无量纲。

第三步：将处理后的 15 个指标带入 K-means 分类模型，分类设置为 5，迭代 10 次，得出分类结果。

第四步：验证模型有效性。由于 K-means 属无监督聚类，没有比较直接的聚类评估方法，面可从类内的紧密程度和类间的离散程度来评估聚类效果。当类内误差越小，类间离散度越大表明各类差异性较大，分类效果好。下表显示，各类别类内误差在 30~65 之间，而离散度为 73.6%，结果较好。具体见表 2 所示。

Table 2. Clustering effect indicators

表 2. 聚类效果指标

	第一类	第二类	第三类	第四类	第五类
均方误差	37.632	49.869	50.78	46.33	61.86
离散度	73.6%				

3.3. 客户特征提取

对客户特征的描述是分类研究的最终目标，客户特征的描述需结合市场实际。社群化运营的关键是小组成员的粘度。围绕“圈层、兴趣、内容、利益”展开研究，“圈层”能反应小组成员的社会交际匹配度，健康的圈层能更大激发出小组成员的有效互动；“兴趣”是维持小组长期共存的重要纽带；“内容”是可以反应小组成员销售行为特征，也是小组培育的出发点；“利益”互帮互助、互利共赢维系的纽带，能体现小组成员经营性特征及方向。

本文运用描述统计法提取分类特征，对各类零售用户的人像指标、店铺属性、经营属性、行为特征四个维度指标进行分析对比。由于数据多为定性指标，适合众数、频率分析来提取关键信息，进行纵向、横向对比分析。

4. 基于唐山市零售客户的实证分析

4.1. 零售客户分类结果分析

本文基于调查问卷数据,以730个唐山市零售客户为例,通过对数据进行预处理,采用K-means方法完成零售客户的分类,共分成五类,第一类116个、第二类96个、第三类166个、第四类174个,第五类178个。

提取各类数据的众数,其中背柜数量、前柜数量、卷烟经营面积为均值,结果见表3所示。

Table 3. Characteristics of cigarette retail customer classification results

表 3. 卷烟零售客户分类结果特征

类别	第一类	第二类	第三类	第四类	第五类	
人像指标	性别	男性	男性	男性	女性	男性
	年龄	40~49岁	40~49岁	40~49岁	40~49岁	40~49岁
	学历背景	高中及以下	高中及以下	高中及以下	高中及以下	高中及以下
	年收入区间	5~10万 (不含10万)	5~10万 (不含10万)	5万以下 (不含5万)	5万以下 (不含5万)	5万以下 (不含5万)
	市场类型	城网	城网	城网	农网	农网
店铺属性	业态类型	烟酒店	烟酒店	便利店	便利店	便利店
	背柜数量	0.55	0.78	~0.34	~0.25	~0.39
	前柜数量	0.33	0.35	~0.25	~0.21	~0.31
	卷烟经营面积	1.10	0.96	~0.37	~0.36	~0.54
	卷烟经营档位区间	26~16档	26~16档	26~16档	26~16档	26~16档
经营属性	有品吸区	否	是	否	否	否
	系统使用情况	未安装扫码系统	未安装扫码系统	未安装扫码系统	未安装扫码系统	未安装扫码系统
	店铺终端类型	普通终端	普通终端	普通终端	普通终端	普通终端
	经营偏好	卷烟、酒类	卷烟、酒类	卷烟	卷烟、其他	卷烟
	行为特征	兴趣爱好	运动	旅行	阅读	其他

从人像指标看,第一类、第二类客户多为中等收入,第四类用户多为女性群体;从店铺属性看,第四类、第五类用户主要为农网类型,第三类、第四类、第五类多为便利店业态,第一类、第二类客户店铺背柜、前柜及经营面积相对更大更多;从经营属性看,第二类客户更多设有品吸区,第一类、第二类客户的经营偏好偏向卷烟类和酒类。而第四类客户偏向卷烟和其他类,其他主要为日用百货;从行为特征看,第一类客户和第五类客户更多爱好运动,第二类客户爱好旅行,第三类客户爱好阅读,第四类客户爱好其他。

可见第一类零售户主要特征为城网、中等收入、卷烟经营店面较大、爱好运动;第二类零售户较第一类为城网、中等收入、设置有品吸区、爱好旅行;第三类零售户为城网、低收入群、小型便利店铺、爱好阅读;第四类零售户为农网、女性、低收入群、经营卷烟及其他零售品,爱好特别;第五类零售户

农网、低收入群、小型卷烟便利店铺、爱好运动。

4.2. 各类客户特征分析及服务建议

为进一步挖掘各类客户的特点，本文利用频率分析，分析各类客户的兴趣爱好结构特征，以更精准定位零售户。具体见图 1 所示。

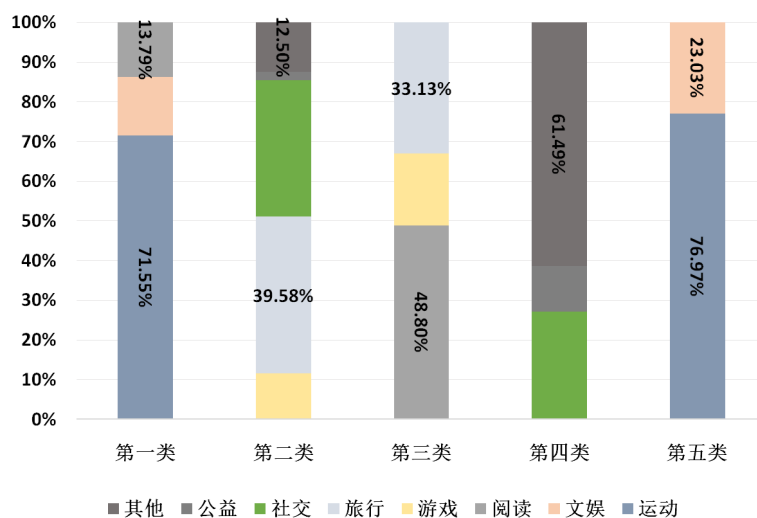


Figure 1. Interest and hobby characteristics of retailers by category

图 1. 各类别零售户的兴趣爱好特征

可见，第一类零售户除运动外，对文娱和阅读同样感兴趣，第二类人群更偏向旅行及社交；第三类零售户爱好阅读和旅行，第四类零售户爱好其他及社交，第五类零售户爱好运动及文娱。

基于社群运营的典型特征，结合分类结果分析，将五类零售用户分别定义为“产品型”、“教父型”、“学习型”、“兴趣型”、“圈子型”，并针对性提出烟草服务建议。

第一类零售户主要为经营城网的较大烟酒店铺，合作方式为非加盟及合作，终端类型多为普通及一般现代终端，客户经营偏好为烟酒类，近一半零售户设有吸烟品区，爱好运动的中等收入群。这类社群的特点为产品较好，烟类品种丰富，且服务达到一定规模，对于产品型社群，产品承载了功能属性和情感，此类社群的关键是维持较高的参与度，提升小组成员的尊重和认同感。建议从打造新品宣传平台、以共同促进新品上柜的目标，推广优秀产品，聚合社群，刺激营销。运动类爱好特征推荐饮料类服务搭配。

第二类零售户主要为经营城网的烟酒店，店铺规模中等，属非加盟及合作，终端类型多为普通及一般现代终端，部分有扫码或全品种扫码系统，一半多零售户设置有吸烟品区，多为 49 岁以下爱好旅游及社交的中等收入群体。此类零售户定位为教父型社群，这类社群潜在价值高。此类社群对群组运营者要求高，小组长需打造自身人格魅力，专业长处。对于烟草行业建议推行精准的运营规则及奖励机制，建议打造客户经验交流平台，取长补短，共同成长，搭建营销人员现场培训平台，从卷烟陈列，重点品牌、新品陈列、店面维护技巧开展现场示范式培训，打造卷烟推介技巧现场指导平台，提升客户卷烟推介能力。通过各类商务合作增强信息分享及交流。

第三类零售户主要为城网偏小型商铺，非烟酒、非娱乐业态类型，无设置吸烟品区，多爱好阅读及旅行。此类群体定位为学习型社群，学习型社群特定是内容为主，用户大多热衷学习知识，因共同需求形成某一特定人群。对于卷烟零售户学习社群，此类群体较为特定，对于产品的热衷需进一步加强为对

品牌的认同,但存在壁垒,需激发成员的“上进心”。建议以活动形式激发社群热情及参与度,增强社群小组成员的归属感,通过打造经营技巧,法律法规培训平台,提升客户技能,增强诚信意识,提升客户盈利,通过行业政策宣传、行业发展形势宣传平台,增强客户信心。服务推荐搭配食品、零食类营销。

第四类零售户主要为农网、业态类型有较多的娱乐服务业,基本为无吸烟品区,经营偏好为其他,女性群体,爱好社交、公益及其他。此类群体特征突出,定位为兴趣型社群,是指拥有特殊兴趣爱好的用户群体,群用户比较容易活跃,用户持续稳定,转化效率高。建议提升社群的情怀,营销模式的多样性,挖掘兴趣中的利他因素,结合自身特点特长,形成终极目标,在建群之初就大力宣传。如从日用百货入手,建立服务。建议搭建新兴的线上线下烟草互动聊天群,让大家更多的相聚交流、娱乐,利用活动推广产品来赚取利润。

第五类零售户主要为农网销售,超市业态,无吸烟品区,群体偏40岁以上中年,爱好为运动及文娱。定位为圈子型社群,此类社群有固定的专业和分享性,群内专业人士交流精密,用户粘度强,在烟草行业的特点是这类零售户相对经验丰富。此类社群运营关键是让一些小众圈子成长为大圈子是关键,建议通过小组信息沟通平台、小组例会等方式加强小组长的沟通联络,同时做好培训指导,采取“线上+线下”模式,组织人员集中学习小组活动规范,带动整体积极性,服务建议推荐水果、食品、饮料类。

5. 结论

本文首先建立烟草社群零售户的问卷指标,从人像指标、店铺属性、经营属性、行为特征四个维度建立指标体系,共16个二级指标。问卷回收1117份,文本研究唐山市零售户特征,利用K-means聚类将730个零售户划分为五类,模型有效。通过数据的挖掘,实现了零售户的精准分类,诚信互相小组可分为五类典型社群,即“产品型”、“教父型”、“学习型”、“兴趣型”、“圈子型”。研究发现产品型社群主要为城网烟草规模较大的零售户,教父型主要为城网经营中等规模烟酒店的零售户,学习型主要为城网小型超市便利商铺的零售户,兴趣型主要为农网女性经营的零售户,圈子型主要为农网中年经营超市的卷烟零售户。各类客户特征突出。

基于零售客户不同的优势特点,本文提出了针对性的管理服务建议,有利于实现对于零售客户群体的个性化服务,有利于优化企业资源配置,对于培育卷烟品牌,推动卷烟发展水平有积极作用。

参考文献

- [1] 夏维力,王青松. 基于客户价值的客户细分及保持策略研究[J]. 管理科学, 2006, 19(4): 35-38.
- [2] 江涛,邹平. 专卖体制下烟草行业统一分销模式设计[J]. 烟草科技, 2010(3): 63-65.
- [3] 王萍. 运用数据挖掘技术预测客户购买倾向——方法与实证研究[J]. 情报科学, 2005, 23(5): 738-741.
- [4] 闫相斌,李一军,叶强. 基于购买行为的客户分类方法研究[J]. 计算机集成制造系统, 2005, 11(12): 1769-1774.
- [5] 尹朝阳. 台议卷烟客户的分类[J]. 商场现代化, 2006(8): 82-83.
- [6] 李苏建,刘伟民. 模拟退火K均值聚类算法及其应用研究[J]. 微计算机信息, 2008, 24(21): 182-184.
- [7] 蒋国瑞,王晓谊. 基于数据挖掘的电信家庭电话分类研究[J]. 科技管理研究, 2010(12): 111-114.
- [8] Ma, Y., Ding, J. and Hong, W. (2010) Delivering Customer Value Based on Service Process: The Example Teso.com. *International Business Research*, 3, 131-135. <https://doi.org/10.5539/ibr.v3n2p131>
- [9] Chen, S., Jian, T. and Yang, H. (2011) A Fuzzy AHP Approach for Evaluating Customer Value of b2c Companies. *Journal of Computers*, 6, 112-116.
- [10] 叶志龙,黄章树. 基于新三维客户细分模型的线上会员客户价值研究[J]. 统计与信息论坛, 2016(5): 96-101.