

学校代码: 10200

分 类 号: F830

学校名称: 东北师范大学

研究生学号: 2014101396



东北师范大学 硕士学位论文

大数据时代下零售行业客户分析模型研究

**Research on Customer Analysis and Analysis Model of Retail
Industry in the Era of Big Data**

作者: 张天劼

指导教师: 李春艳 教授

一级学科: 工商管理

二级学科: 企业管理

研究方向: 企业经营管理理论

学位类型: 学术硕士

东北师范大学学位评定委员会

2017年5月

独 创 性 声 明

本人郑重声明：所提交的学位论文是本人在导师指导下独立进行研究工作所取得的成果。据我所知，除了特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。对本人的研究做出重要贡献的个人和集体，均已在文中作了明确的说明。本声明的法律结果由本人承担。

学位论文作者签名： 张天劼 日期： 2017.6.7

学位论文使用授权书

本学位论文作者完全了解东北师范大学有关保留、使用学位论文的规定，即：东北师范大学有权保留并向国家有关部门或机构送交学位论文的复印件和电子版，允许论文被查阅和借阅。本人授权东北师范大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或其它复制手段保存、汇编本学位论文。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名： 张天劼

日

期： 2017.6.7

指导教师签名： 李 莹

日

期： _____

学位论文作者毕业后去向： _____

工作单位： _____ 电话： _____

通讯地址： _____ 邮编： _____

学校代码：10200
分 类 号：F830

学校名称：东北师范大学
研究生学号：2014101396



东北师范大学 硕士学位论文

大数据时代下零售行业客户分析模型研究

Research on Customer Analysis and Analysis Model of Retail Industry in the Era of Big Data

作者：张天劼

指导教师：李春艳教授

一级学科：工商管理

二级学科：企业管理

研究方向：企业经营管理理论

学位类型：学术硕士

东北师范大学学位评定委员会

2017 年 5 月

摘要

随着电子商务的兴起、大数据时代的到来，传统零售行业受到了很大的冲击和压力。由于电子商务在数字化上的先天优势，诸多零售电商巨头已经形成了完整的数据分析体系和业务应用链条，他们深谙大数据分析与应用之道，一方面不断扩充新产品线和产品类目，另一方面也深入挖掘消费者行为数据，精准满足消费者期望，不断侵蚀着传统线下零售市场份额。

传统零售商已经意识到数据分析的重要性，而随着信息技术与网络技术的不断普及，大数据思维逐渐深入人心。大数据由于其经济性、时效性的特点，与零售业务相结合，将逐渐应用于客户管理、业务模式以及风险控制等领域，掀起了零售业生产率增长的新浪潮。基于对互联网的逐步深入了解，零售行业逐渐明确大数据时代的发展方向，以大数据理念重塑零售业业务发展模式，注重大数据创新，提高企业管理水平。

面对大数据带来的机遇与挑战，现有的许多企业管理理论如客户生命周期理论，必须根据现实情况的变化，进行有针对性地更改。大数据驱动下的全生命周期价值客户是企业营销的核心人群，应根据不同客户生命周期和类型，应用不同的动态保持模型和策略。

本文在经典客户关系管理理论基础上，对零售业客户分析创新方法进行了研究，并结合零售企业的实际情况，进一步改良和拓展了传统客户分析方法，最终提出了一系列针对性的分析模型。

关键词：大数据；零售行业；客户分析模型；

Abstract

With the rise of e-commerce, the arrival of the era of big data, the traditional retail industry has been a great impact and pressure. Due to the inherent advantages of e-commerce in the digital, many retail giants have formed a complete data analysis system and business application chain, they know how big data analysis and application of the road, a continuous expansion of new product line and product category, on the other hand also dig the consumer behavior data, accurate to meet consumer expectations. Continue to erode the market share of traditional retail line.

Traditional retailers have been aware of the importance of data analysis, and with the popularity of information technology and network technology, big data thinking gradually popular. Big data because of the characteristics of economy, timeliness, combined with the retail business, will be gradually applied to customer management, business model, risk control and other fields, set off a new wave of productivity growth in the retail industry. Further understanding of the Internet based on the direction of development of the retail industry gradually clear the era of big data, big data to restore the idea of retail business development mode, pay attention to the big data innovation, improve enterprise management level.

Faced with the opportunities and challenges brought by big data, many existing enterprise management theories, such as the theory of customer life cycle, must be changed according to the actual situation. The life cycle value customer driven by big data is the core of enterprise marketing, and different dynamic maintenance models and strategies should be applied according to different customer life cycle and types.

Based on the classical theory of customer relationship management, customer analysis of retail innovation is studied, combined with the actual situation of retail enterprises, to further improve and expand the traditional customer analysis method, finally put forward a series of models for analysis.

Key words: Big data; Retail industry; Customer analysis model;

目录

摘要.....	2
目录.....	4
1 绪论.....	6
1.1 研究背景.....	6
1.2 研究意义.....	6
1.3 研究方法.....	7
1.4 研究框架.....	7
2 零售业客户管理理论及相关研究文献综述.....	9
2.1 零售业客户管理理论综述.....	9
2.2 零售业客户分析的文献综述.....	13
2.3 当前客户分析理论研究面临的挑战.....	16
3 零售业客户分析方法.....	17
3.1 传统分析方法.....	17
3.1.1 直接聚类法.....	17
3.1.2 最短距离聚类法.....	18
3.1.3 最远距离聚类法.....	18
3.2 大数据分析.....	18
3.2.1 大数据分析的数据源构成.....	18
3.2.2 大数据分析的核心数据维度.....	19
3.2.3 大数据分析的核心指标.....	19
4 零售业大数据客户分析模型设计.....	25
4.1 客户价值模型（RFM）.....	25
4.2 客户生命周期模型（NESS）.....	26
4.3 客户粘性模型（RFD）.....	27
4.4 客户身份模型（CIM）.....	28
5 北京朝阳大悦城客户分析模型应用的案例分析.....	29

5.1 案例背景..... 29

5.2 模型建立及应用..... 29

5.3 应用总结..... 33

6 结论..... 35

6.1 主要结论..... 35

6.2 研究不足..... 35

参考文献..... 36

致谢.....40

表索引

表 1 会员特征指标表20

表 2 会员购买分析指标21

表 3 会员购物行为分析指标22

图索引

图 1 RFM 模型数据分布图30

图 2 NESS 模型数据分布图31

图 3 店铺客流等级分布图33

1 绪论

1.1 研究背景

最近几年，大数据被越来越多地应用到各个领域。一般来讲，大数据是个定性的概念，是指快速增长且难以被普通数据分析方法所处理的数据集。大数据具有数据量大、结构复杂、实时性高、价值密度低等特点。

在“新常态”经济背景和电子商务浪潮的冲击下，我国传统零售业的发展面临着巨大的转型压力，其发展离不开新思维、新技术的支持。在信息与网络技术不断创新和普及下，大数据思维兴起并逐渐深入人心。大数据具有宏观性、全覆盖性、潜在的经济性及时效性特点，与其零售业务相结合，并逐渐应用于零售行业的客户管理、业务模式以及风险控制等领域，将会掀起了零售业生产率增长的新浪潮。基于零售行业对互联网的深入了解，逐渐明确和发现了大数据时代的自身发展方向，即以大数据理念重塑零售业业务发展模式，注重大数据创新，提升价值竞争力，提高企业管理水平。

面对大数据带来的机遇与挑战，现有的许多企业管理理论如客户生命周期理论，必须根据现实情况的变化，进行有针对性地更改。大数据驱动下的全生命周期价值客户是企业营销的核心人群，应根据不同客户生命周期和类型，应用不同的动态保持模型和策略。

1.2 研究意义

数据分析体系是企业数据分析的方法论和核心工具，用来指导数据分析的思路，引导企业的数据分析师及决策者进行一个完整的数据分析。数据分析体系是在剔除具体数据信息之后的数据分析方法库和模型库，企业没有分析体系会令数据分析工作缺乏整体思路，难以保证质量。

在大数据时代，企业采集数据的渠道及手段获得了前所未有的拓展，企业有能力掌握并量化更多的内外部环境信息，进而对企业管理的理论及实践工具

提供了新的视角。

面对海量数据、多维度数据、线上线下数据，简单机械地套用传统的数据模型反而会缩小视野、降低敏锐度。本次研究旨在改良和拓展传统客户分析方法，建设大数据思维下的客户分析模型，对零售企业的技术组织、人才队伍建设、各类分析模型应用具有全面的指导意义。

1.3 研究方法

（1）理论联系实际方法

大数据时代下传统零售企业的数据分析方法，既是理论问题，又是现实问题。因此，本论文研究坚持理论联系实际方法，坚持从理论与实际相结合的视角提出问题、分析问题和解决问题。

（2）文献研究法

通过收集、查阅相关文献资料，了解、梳理和掌握零售业客户分析领域已取得的理论成果和研究进展，理解零售业数据分析的研究现状与未来研究方向，为后面的研究提供理论依据。

（3）跨学科研究法

大数据时代下零售企业的数据分析模型研究，同时涉及企业管理学、统计学的理论和知识。因此，本论文研究将采用跨学科研究方法进行综合研究。

（4）模拟法(模型方法)

客户分析模型研究是本文的主要目的，本文将依照客户数据分析的主要理论，分析和研究不同的数据维度与度量，最终创设或优化相应的分析模型。

1.4 研究框架

本文共分 6 章。第一章为绪论。第二章主要对零售业客户管理理论及相关研究进行了综述，主要有零售业的客户生命周期理论、零售业的客户分析理论，并对当前理论研究面临的挑战做了简要分析。第三章总结了零售行业客户分析方法，包括传统分析方法和大数据思维下新的分析方法，并重点论述了后者。在第四章，根据相关理论综述和分析方法总结，设计了几个零售业大数据客户

分析模型。第五章以北京朝阳大悦城为案例，介绍并分析了上述模型在实际企业数据分析与经营管理中的应用情况。第六章是本文的结论部分。

2 零售业客户管理理论及相关研究文献综述

2.1 零售业客户管理理论综述

2.1.1 零售业的内涵及特征

零售业 (retail industry) 是指通过商品交易行为将生产和服务各行业生产的产品和劳务提供给消费者的商业活动。角度不同对零售业的定义亦有所差异,但本质特征是一致的。主流的零售业定义分为两种:一种是商品营销学角度的定义,一种是企业管理学角度的定义。前者认为,零售业是制造商通过批发商、中间商等将商品和服务出售给消费者的一种营销业务。后者认为,零售业是包括所有把有形商品和无形服务销售给普通公众的服务企业的集合,不仅包括有店铺零售商也包括无店铺零售商,它们不改变商品的形式,只是链接从制造商到消费者的服务行业。

在信息时代,传统零售业发生了深刻的革命。网络技术打破了零售市场时空界限。传统零售企业经营成功的选址首要因素已经变得不再那么重要。任何零售商只要通过一定的努力,都可以将目标市场扩展到全国乃至全世界,零售业的竞争将更趋激烈。必须更多地依靠经营管理的创新才能占有市场份额。

2.1.2 零售业的客户生命周期理论

客户生命周期可以从企业和客户两个角度理解。从企业角度出发,是指企业欲对某一客户进行开发开始,到实现为客户提供服务完毕的时间段。从客户角度出发,是指一个客户从对为其服务的企业了解、购买其产品和服务开始,到终止双边交易关系为止时间段。可见,客户生命周期描述的是客户关系从一种状态或一个阶段,向另一种状态或另一个阶段转换的期间。无论哪个角度,客户生命周期可以是一次完结的,也可以是重复多次的。产品和服务的性质、质量不同,客户生命周期亦不相同。客户生命周期的延长和重复,客户生命周期管理——阶段管理、过程管理、生命周期价值管理、客户忠诚度及期限管理

等对零售企业至关重要，而产品性质、品牌、区域、便捷、质量等是影响客户生命周期的重要变量。

客户的生命周期理论是企业产品生命周期理论的一种演变，但对商业企业来讲，产品的生命周期远不如客户的生命周期更为重要。在互联网、大数据、云计算时代，利用信息技术以及互联网技术等做客户生命周期管理，协调企业与顾客间在销售服务上的交互，向客户提供创新式的个性化服务，成为现代零售业企业的核心竞争力。

2.1.3 零售业的客户分析理论

客户分析是指客户的需求分析。这种客户需求分析是根据客户的各种信息和数据来分解的客户需要。客户分析包括分析客户特征，评估客户价值，以针对性地制订相应的营销策略与资源配置计划。

客户分析包括诸多内容：客户特征、客户意见、客户咨询、客户商业行为、客户满意度、客户忠诚度、客户注意力、客户信誉度、客户购买力、客户流动等等。通过对客户的舆情收集、多元信息分析，客户价值评价、改进产品、服务质量和便捷性，以便留住，吸引客户，延长客户生命周期。

目前我国零售企业对客户的分析手段较为简单，方法也不够系统和完善。一般的统计方法可以在一定程度上得出分析结果，但由于同企业发展阶段、规模、水准不同，经营的产品和服务不同，企业间存在很大的不平衡性，难以满足客户分析的特殊需求。客户分析理论强调制定专业、专项、系统、合理的客户分析方案，为挖掘潜在客户、培养忠诚客户提供决策支持的客户信息基础。现代信息技术则为这种大数据的客户分析提供了技术支撑。

2.1.4 零售业的客户关系管理（CRM）理论

关于客户关系管理（CRM）的概念界定，是一个不断演进和升华的过程。该理论的大体演进脉络如下：1980 年，全球第一家信息技术研究和分析的公司——高德纳（Gartner Group，成立于 1979 年）首次提出专门收集客户与公司联系的所有信息的“接触管理”（Contact Management）的概念，即，并让自己成为每一位用户的一站式信息技术服务公司。1985 年，巴巴拉·本德·杰克逊提出的关系营销概念丰富了市场营销的理论研究；1990 年，客户-企业关系营销概念演

变为客户关怀 (Customer care) 理念, 这种客户关怀多以电话服务中心为载体并成为客户信息资料来源渠道。1990 年, 美国高德纳公司 (Gartner Group) 提出企业资源计划 (ERP), 即建立在信息技术基础上, 以系统化的管理思想, 亦称供应链的管理思想。其应用范围从制造业扩展到各个行业乃至政府、学校的事业部门。ERP 通过融合数据库技术、客户服务器结构、图形用户界面、计算机辅助开发工具、可移植的开放系统等, 对企业资源进行了有效的集成, 强调对供应链进行整体管理。ERP 系统是将企业的资源流进行全面一体化管理的管理信息系统, 包括物流系统, 资金流系统, 技术流系统、人力流系统、信息流系统、工作流系统等的一体化管理。ERP 系统还包括以下主要的补充管理功能: 供应链、分销、运输、市场占有、客户服务、生产库存、质量监控, 测试创新, 业务流程, 产品数据、项目实施、设备维护、执行程序 (Manufacturing Executive System, MES)、法规标准、过程控制等补充功能。1999 年, 客户关系管理 (Customer Relationship Management) CRM 概念诞生。由高德纳团队 (Gartner Group) 提出的 CRM 理论认为: 在整个客户生命期中, CRM 都以客户为中心, 这意味着企业各类业务的运行过程及目的都要集中于满足客户的需要上。CRM 应用软件简化了与客户交流的各种渠道, 把直面沟通、电话联络、Web 访问、在线交流、线下体验等协调为一体, 由此, 企业掌握了按客户喜好与之进行交流的所有渠道, 便于企业寻找业务拓展所需的新市场, 便于开发与客户沟通的新渠道, 培育和提高了客户价值及与赢利性、平稳性密切相连的客户满意度和忠实度。IBM (国际商业机器公司团队) 则认为: 客户关系管理应贯穿于商业运营的全过程。通过将客户关系分类——关系管理、流程管理和接入管理, 识别、挑选、获取、发展和保持最优的客户关系。

在 ERP 系统中, 客户只是供应链中的一环, 单独提出 CRM 概念的原因, 一是 ERP 并没有给出良好的解决办法。在 ERP 的实际应用中人们发现, 针对客户在销售中的重要性及客户关系管理的多样性和复杂性, ERP 系统并没有很好地实现对供应链下游 (客户端) 的管理, 而客户关系管理虽然只是一个环节但却具有全方位的管理地位, 能够赋予企业最大化的源于客户的收益率; 二是到 90 年代末期, 随着互联网应用的普及, 客户信息处理技术如数据库、商业智能、

知识发现等技术迅猛发展，使得 CRM 市场一直处于一种爆炸性增长的态势。

客户关系管理系统既是一套原则制度，一个与客户关系相关联的包括市场营销、客户服务、信息化技术支持的商业流程，也是一套企业管理的软件和技术。由此可以看出，客户关系管理是一种理论与实践相结合的新的理论形态。

2.1.5 客户行为理论

客户即为消费者，客户行为亦为消费者行为。对消费者行为的分析，有两大观测视角：

一是经济学，二是管理学。经济学是重点从效用角度研究消费者行为及消费者行为与需求定理的关系；管理学是重点从市场营销角度探讨消费者行为与企业决策和企业盈利关系，但二者分析基础具有相通性。

第一，客户行为理论强调消费者主权。消费者主权是指企业要根据消费者的需求进行生产，否则企业就会失去生存的基础。企业可以引导和创造新消费，新需求，进而开拓新市场，但这是以消费者认同为前提的。第二，客户行为理论强调消费效用最大化。因为消费者购买物品是为了效用，包括物美价廉、便利化、偏好、满足感、炫耀、时尚、地位身份标识等方面的需求。物品的效用越大，消费者愿意支付的价格越高。因此消费者行为追求的是效用最大化。消费者效用分为总效应与边际效应。前者是指一定时期内，在消费一种或几种商品过程中，消费者所获得的效用总和。后者是指一定时间内，消费者增加单位商品所引起的效用增加量，亦称为边际效用。按边际效用递减规律：消费者购买商品越多，边际效用越小，反之亦然。消费者购买商品支付价格以边际效用为标准。总效用与边际效用的关系表现为总效用是边际效用之和。当边际效用为零时，总效用达到最大；当边际效用为正数时，总效用是增加的；当边际效用为负数时，总效用减少。第三，客户行为理论强调消费者均衡。消费者把有限的货币收入用于购买何种商品、购买多少能达到效用最大，即为消费者均衡。实现消费者均衡的限制条件： $P_1X_1 + P_2X_2 + \dots + P_nX_n = M$ ，均衡条件： $MU_1 / P_1 = MU_2 / P_2 = \dots = MU_n / P_n$ 。第四，客户的消费行为受其替代效应和收入效应影响。当具有相同效用或类似效用的两种商品，一种商品价格下降，另一种商品价格不变或上涨，便会产生替代效应，消费者会购买性价比高的商品。当消费

者购买两种商品时，如果一种商品名义价格下降，便可产生货币收入购买力相对增强的效果。从而形成可购买更多商品的收入效应。

2.2 零售业客户分析的文献综述

从现有的研究文献看，关于零售业的专项研究，包括零售业的发展态势、转型创新、营销模式、竞争力、物理管理等方面的研究文献较多；关于客户管理的研究，多为电商、企业、地产、旅游、银行等行业的专项客户关系管理及精准营销的客户管理，有关客户资源管理的研究文献相当丰富，但专门的零售业的客户分析研究的文献不是很多。从近年的零售业客户分析的研究文献看，主要围绕企业竞争力提升、数据挖掘、客户满意度等维度展开。

2.2.1 零售企业竞争力与客户关系研究

基于客户关系管理的零售企业竞争力提升研究，王丽芳（2016）^[1]认为，我国零售企业与外资零售企业相比实力悬殊，发展面临巨大压力，而提高零售业企业竞争力的关键在于，"以顾客为中心"，与顾客建立持久的关系。近三十年，客户关系管理已成为众多商业零售企业提升竞争优势的利器。客户关系管理可以找准目标客户，培育客户忠诚度，发展连锁经营。具体可采取提高客户资料信息收集能力，加强客户信息的深层次挖掘，充分开发会员功能。丰富针对性的精准营销手段，提供网络增值服务，强化电子营销手段等对策，田秀政（2015）^[2]。

2.2.2 数据挖掘与客户关系研究

近年来运用大数据及大数据分析技术进行客户分析和管理的研究逐步深入。王永波（2016）的研究认为，大数据挖掘可按照一定标准将客户细分成不同的客户群体，一般来说，同一个客户群体表现为具有相似的行为，如购物偏好等。客户细分有助于企业针对性吸引客户，留住客户，进而获得最大利润。客户细分可将客户属性分为一般属性和交易属性，前者是针对不同状态的客户提供合适的产品或服务，称之为以客户状态为导向；后者主要是针对交易数据特点，通过客户购买树分析客户购买商品的集合构建、树型层次结构及距离定

义来度量客户相似度，客户聚类导致的商品聚类等消费倾向^[3]。蔡玖琳（2015）探讨了基于数据挖掘的零售业客户细分的方法，发现已有的零售业客户细分研究存在尚未形成完整的适用于零售业的客户细分指标体系，以及客户细分模型准确度偏低的两方面的不足。鉴于此在梳理了客户细分理论和数据挖掘技术基础上，构建了一套基于 RFM 的客户细分指标体系，利用基于属性加权的 K-Means 算法建立了客户细分模型。并采用某连锁超市的 3 万多条会员数据和近 38 万条销售数据进行实证研究，运用以上指标体系和客户细分模型将客户细分为四大类，并针对不同客户群体给出相应的营销策略。实证研究结果表明：该方法在客户聚类紧凑性及行为特征区分能力方面均优于传统基于 RFM 的细分方法，能够更好地解决零售业客户细分问题，提升客户关系管理和营销决策质量^[4]。徐天（2015）的研究认为，CRM 系统可以对客户的数据进行管理和划分，从中细分出客户类型，客户满意度、客户忠诚度、客户有效需求和潜在需求等有效资源，形成以培养以客户为中心的经营理念、业务流程和销售模式^[5]。

2.2.3 大型商场与客户关系研究

邵莉娟（2015）从企业文化、信息管理、员工培训以及评估机制等方面，对家乐福超市的客户关系管理存在的问题入手进行分析，针对这些问题对家乐福超市客户关系管理提出了将客户关系管理和企业文化、员工培训相融合，通过对客户抱怨管理，提高客户满意度进而提高客户忠诚度，建立企业与客户相互依存关系等改进的对策。^[6]王俊峰（2013）研究了郑州丹尼斯百货客户管理系统，分析了丹尼斯百货客户管理系统目前存在的问题，以客户管理系统作为突破点，结合郑州丹尼斯百货客户管理系统的现状，提出建立和完善差异化的客户管理系统^[7]。阳茂春（2012）专门研究了 XD 购物中心客户满意度。通过对 XD 购物中心经营现状及零售业顾客满意度相关指标体系的分析，运用主观赋权法对相应指标的权重赋值，构建顾客满意度模型。通过对 XD 购物中心客户服务工作的初步调查量化相应的指标，设计顾客满意度的调查问卷，并通过约五千份问卷调查的统计结果，较准确地掌握了该购物中心顾客满意度的实情、挖掘 XD 购物中心客户关系管理中存在问题的原因。采用因果分析法和对比分析法探讨改进顾客满意度的具体改进措施^[8]。王琮凯（2014）结合联通全资子公司联通华

盛的新一代销售管理系统，探讨了关联规则在客户相关消费规律中的应用。即借助 Matlab 工具仿真分析比较了各类相关算法的性能。运用 Weka3.7.4 专业数据挖掘建模软件对预处理完毕后的企业销售交易数据进行分析建模，选取合适的关联规则算法分析预处理后的销售数据，最终得到销售数据的关联规则。有益于优化公司的资源配置，帮助联通华盛公司做出合适的营销策略^[9]。

2.2.4 客户分析评述

客户关系管理追求其有效性，要达到客户关系有效管理的目的，需要做深入客户分析。客户分析主要是客户个性化需求、客户行为和客户价值信息等三个维度的分析。**客户个性化需求分析**，须细分客户，了解客户需求特征与商务效益的关系，了解客户需求变化，发现潜在客户，引导客户的新需求，从而进一步扩大商业规模，使企业得到快速的发展。实施 CRM 的一个重要目标就是能够分析出客户的个性化需求特征，并对这种需求采取相应措施，同时分析不同客户对企业效益的不同影响，以便做出正确的决策。这些都使得客户分析成为企业实施 CRM 时不可缺少的组成部分；**客户行为分析**，须了解客户消费偏好、消费习惯乃至消费理念，企业可以利用收集到的客户信息，跟踪并分析每一个客户的信息，以提供“以客户为中心”服务。在客户行为分析中还可以了解客户行为对企业收益的影响，使企业与客户的关系及企业利润得到最优化；**客户价值分析**，是在合法、有效和科学收集客户信息中，挖掘价值信息和价值客户。如挖掘产品的缺陷、产品的受欢迎度、产品改进领域、售后服务问题等；挖掘最能给企业带来效益的客户，帮助企业积累价值客户，发现价值客户，充分利用价值客户关系资源，进而将客户价值转化为企业价值，实现企业价值的转型升级。

目前我国企业对客户的分析还很欠缺，分析手段较为简单，方法也不够系统和完善。而简单的统计方法虽然可以在一定程度上得出分析结果，但因为不同企业发展中存在一定的不平衡性，利用简单的统计模式得出的结论容易有较大的误差，难以满足企业的特殊需求。因而企业需要有更加完善、合理的客户分析方案，进一步提高客户分析的合理性、一致性，并能在对潜在客户的培养和发现中提供更多的决策支持。

2.3 当前客户分析理论研究面临的挑战

在过去，零售业没有专门的业务数据采集手段，业务数据的积累，主要依靠通过零售收银终端所产生的销售流水数据，从而大大限制了可以用于分析决策的数据源。

随着零售企业会员业务和会员系统的建立，零售企业可以逐步获得成为会员的这部分消费者的数据信息，并且在简单的身份识别基础上不断增加丰富的身份信息内容。随着智能手机的普及、室内定位技术的成熟，零售企业又可以进一步对消费者行为数据进行采集，可以通过对智能手机在零售企业门店内的定位位置定时抓取，获得消费者的行为轨迹、购物过程。

另一方面，在消费者主权时代，市场竞争格局最终取决于消费者选择。当价格已经不是消费者最为关注的指标时，如何更好地满足消费者需求，贴合消费者喜好，就成为零售商最需关注的新问题。应对消费升级，零售企业需要从以销售额、利润为核心的经营模式，转变为适合市场、适应消费者的新思路。

由此可见，在数据采集获得了极大丰富的前提下，在以消费者为核心的经营理念和思路下，零售客户分析由单一的自身销售数据，又添加了消费者身份数据、消费者行为数据两大板块，这对零售业企业开展多维度数据的交叉分析、更深入的分析 and 研究客户行为，提出了新的挑战。

3 零售业客户分析方法

从总体上看，客户关系管理(CRM)，是一整套具有先进理念和管理方法的销售方案。它是一项综合的 IT 技术，也是一种新的商业及管理的运作模式，它源于“以客户为中心”的新型商业模式，是一种旨在改善企业与客户关系的新型管理机制。从微观上看，客户关系管理(CRM)，是企业通过富有成效的沟通交流，理解、影响并引导客户消费行为，最终实现客户留存，客户忠诚，客户满意、客户获利和客户创利的目的。从技术上看，客户关系管理(CRM)，是利用软硬件技术和网络技术，为企业建立一个客户信息收集、分析、挖掘、管理、利用的信息处理系统。

3.1 传统分析方法

传统客户关系管理的目标是缩减销售周期和销售成本、增加收入、寻找扩展业务所需的新的市场和渠道以及提高客户的价值、满意度、赢利性和忠实度。客户关系管理要求企业坚持以客户为中心的商业哲学和企业文化来支持有效的市场营销、销售与服务流程，进而形成了以描绘客户为核心的经典客户分析理论。

面对多维度的大量数据，以关联性为目标的数据分析，最基本的分析策略是聚类思想。聚类是一种多变量的统计思维，聚类分析也称群分析、点群分析，是研究分类的一种多元统计方法。聚类分析法主要有分层聚类法和迭代聚类法。

根据聚类方式的不同又可以分为：直接聚类法和间接聚类法。

3.1.1 直接聚类法

根据距离最小的原则，先把各个分类对象单独视为一类，然后依次选出一对分类对象，并成新类。如果一对分类对象正好属于已归的两类，则把这两类并为一类；如果其中一个分类对象已归于一类，则把另一个同类也归入该类。每一次归并，都划去该对象所在的列与列序相同的行。经过 $m-1$ 次就可以把全部分类对象归为一类，根据归并的先后顺序可以作出聚类谱系图。

3.1.2 最短距离聚类法

在原来的 $m \times m$ 距离矩阵的非对角元素中找出 分类对象, 把分类对象 G_p 和 G_q 归并为一新类 G_r , 然后按计算公式 计算原来各类与新类之间的距离, 这样就得到一个新的 $(m-1)$ 阶的距离矩阵; 再从新的距离矩阵中选出最小者 d_{ij} , 把 G_i 和 G_j 归并成新类; 再计算各类与新类的距离, 这样一直下去, 直至各分类对象被归为一类为止。这种聚类被视为最短距离聚类法。

3.1.3 最远距离聚类法

最远距离聚类法与最短距离聚类法的区别在于计算原来的类与新类距离时采用的公式不同。最远距离聚类法所用的是最远距离来衡量样本之间的距离。

按照聚类的策略又可以分为: 时间聚类和客群聚类。

(1) 时间聚类

时间聚类即通过客户到店的时间进行聚类。时间聚类的方式分为按日期聚类和按时段聚类两个类型。日期聚类包括工作日、双休日、节假日、活动日四类。时段聚类分为普通时段、高峰时段、活动时段三类。通过将时间序列按日期和时段两个维度进行聚类, 将大量的原始数据进行排序归类, 为进一步分析客户特征数据、客户销售数据、客户行为数据打下基础。

(2) 客群聚类

客群聚类是基于客户特征, 通过多种消费者标签的定义对客群进行聚类, 便于进一步进行细分人群的分析。客群聚类主要通过客户价值聚类、客户生命周期聚类、客户粘性聚类、挖掘语义聚类四种聚类方式。通过四种聚类方式可以对客户建立细分类别, 进而模型化的研究客户数据。

3.2 大数据分析方法

3.2.1 大数据分析数据源构成

在大数据时代，零售业可以从人（客户/消费者）、店（线下实体店铺）、场（商业地产实体）、圈（辐射商圈）、网（线上）五个数据来源做分类，这样可以全面囊括各种数据来源及业务场景，从而满足零售业以线下实体为经营核心，影响和辐射周边商圈内的消费者。

以人为目标的分析，也就是消费者分析，主要研究的内容是会员的特征与偏好，重点落在单个会员分析与跟踪上。以店为目标的分析，也就是商户分析，主要研究各店铺的表征。以场为目标的分析，也就是购物中心分析，即是从购物中心整体为落脚点，研究购物中心场内整体及液态楼层划分后的综合表征。以圈为目标的分析，也就是商圈分析，通过将会员的地域属性充分挖掘，研究会员覆盖商圈的综合表征。以网为目标的分析，也就是线上电子平台分析，主要研究线上电子平台会员及网络上的综合表征。

3.2.2 大数据分析方法的核⼼数据维度

大数据时代下零售行业的分析维度可以通过对客户接触环节分析产生^[10]。在各种数据源及采集方式下，客户与企业发生关联时都会经历客户识别、客户服务（销售）和客户跟踪三个接触环节^[11]。在这三个接触环节中分别产生了三类核心数据，即客户特征数据、客户购买数据、客户行为数据^[12]。这三类核心数据分别引领了三类分析维度：

(1) 客户特征数据是对消费标签特征的统计数据，是对消费者所有标签进行的详细分析。

(2) 客户购买数据是基于原始交易数据并对进行进一步分析与落实, 包括与时间维度结合后对会员进行的会员价值分析等。

(3) 客户行为数据是基于大会员访问、停留等行为的分析, 包括对访问时间、访问频次、停留时长等指标的分析。

3.2.3 大数据分析方法的核⼼指标

客户关系管理是一个获取、保持和增加可获利客户的方法和过程，认为客

户是企业最重要的资产，通过提高客户忠诚度可最终提高企业的利润率^[13]。随着零售行业的发展，传统的数据采集已经无法满足客户关系管理的在理论和技術上的拓展。零售企业应着眼于全渠道构建的数据分析思维。

在大数据时代下，海量数据、多维度数据的分析和挖掘成为可能，进而零售行业可对客户资料进行深入分析^[14]，根据“一对一”营销原则，满足客户的个性化需求，提高客户忠诚度和保有率，从而实现客户价值持续最大化，提升企业竞争力。^[15]

(1) 会员特征分析方法与分析指标^[16]

对客户特征数据的分析，主要是考察不同消费标签下的客户群体统计数据，以人数、转化率为核心指标，详细分析指标见表 3.1。

表 3.1

会员特征指标表

编号	分析指标	计算方法	数据输入
A01	人数	将用户按照多种聚类方式进行聚类后,对选定客群的人数 ^[17] 、在某时间段内的新增人数以及同比环比进行分析,主要包括大会员客群人数、单店客群人数 ^[18] 、全场客群人数、地铁沿线人数、区域人数、线上客群人数、自定义客群人数等 ^[19] 。	时间区间 分析客群
A02	占比	将用户按照多种聚类方式进行聚类后,对选定客群的人数在该分类下的占比以及同比环比进行分析,主要包括客群人数占比、单店客群人数占比、地铁沿线人数占比、区域人数占比、线上客群人数占比、自定义客群人数占比等 ^[20] 。	时间区间 分析客群
A03	转化率	对用户在大会员体系中的会员身份转化的转化率进行分析。主要包括电子会员卡会员转化为实体会员卡会员转化率、WIFI 登陆者转化为实体会员卡会员转化率、全网会员转化为实体会员转化率六种转化类型 ^[21] 。	转化类型
A04	转化时长	将以上五种转化类型在转化时所需的平均时长进行对比分析 ^[22] 。	无
A05	客群匹配度	通过衡量品牌的风格标签、消费水平标签、品牌标签与消费者风格偏好标签、消费水平标签、品牌偏好标签的匹配程度,对单店客群匹配度、全场客群匹配度进行分析 ^[23] 。	时间区间 分析客群

(2) 客户购买分析方法与分析指标

客户购买分析来源于对原始交易数据的进一步分析与挖掘，包括单纯的数量指标，如销售额、销售笔数等；也包括数量指标在时间、空间上的投影，如坪效、提袋率等。客户购买分析指标详见表 3.2。

表 3.2 客户购买分析指标

编号	分析指标	计算方法	数据输入
A06	销售额	对一定日期区间或单日时刻区间内的会员购买额、全场销售额、楼层销售额、业态品类销售额、单店销售额、线上销售额、选定客群销售额进行统计分析和同比环比分析。支持用户选定时间区间分析 ^[24] 。	时间区间 分析客群 分析品类
A07	销售笔数	对一定日期区间内会员购买笔数、全场销售笔数、楼层销售笔数、业态品类销售笔数、单店销售笔数、线上销售笔数、选定客群销售笔数进行统计分析和同比环比分析。支持用户选定时间区间分析 ^[25] 。	时间区间 分析客群
A08	平均客单价	对一定日期区间内会员购买客单价、单店销售客单价、线上销售客单价、选定客群销售客单价进行统计分析和同比环比分析。支持用户选定时间区间分析 ^[26] 。	时间区间 分析客群
A09	SKU 数	销售 SKU 数据占比=商户销售 SKU 数据/库存 SKU 数据统计单店销售 SKU 数、单店库存 SKU 数、品类平均 SKU 数，并对销售 SKU 数占比、库存 SKU 数进行统计分析及同比环比分析 ^[27] 。	分析品类
A10	销售坪效	销售坪效=销售额/面积 对一定日期区间内单店坪效、业态品类坪效进行统计分析和同比环比分析。支持用户选定时间区间分析。	时间区间 分析品类
A11	关联销售次数	若消费者在一次到店期间既在 A 店消费又在 B 店消费，认为 A 和 B 两个商铺产生一次关联。将所有的关联频次排序，得到关联度最大的商铺组合。若关联次数超过一定值，认为两个商铺有关联。支持用户选定时间区间分析 ^[28] 。	时间区间
A12	关联销售频率	对一定日期区间内各品牌的关联销售数量进行统计分析。支持用户选定时间区间分析。	时间区间
A13	销售占比	楼层销售占比=楼层销售额/全场销售额 业态品类销售占比=（业态/品类销售额）/全场销售额 在一定日期区间内，对楼层销售占比、业态品类销售占比进行统计分析和同比环比分析。支持用户选定时间区间分析 ^[29] 。	时间区间 分析品类
A14	提袋率	提袋率=销售笔数/客流量 对一定日期区间或单日时刻区间内全场提袋率、楼层提袋率、业态品类提袋率、单店提袋率进行统计分析和同比环	时间区间 分析品类

编号	分析指标	计算方法	数据输入
		比分析。支持用户选定时间区间分析 ^[30] 。	
A15	价值指数	以本月最后一天为截止日，将所选区域会员购买额、客单价、会员数量作为分析指标。将各指标最小值与最大值作为极值，平均分为 9 个等级，得到各等级的区段划分，将这九个等级区间依次赋以 1~9 的分数，得到销售额得分、客单价得分、会员数量得分 ^[31] 。 价值指数=销售额得分*权重+客单价得分*权重+会员数量得分*权重 在一定日期区间内对地铁沿线价值指数、区域价值指数进行统计分析 & 同比环比分析 ^[32] 。	时间区间 分析范围
A16	重复购买率	重复购买率=销售过 2 次（含）以上的商品/所有产生过销售的商品数量。 对一定日期区间内的商品的重复购买率进行统计分析。支持用户选定时间区间分析 ^[33] 。	时间区间

（3）客户购物行为分析方法与分析指标

客户行为分析主要考察的是客户访问、停留等行为的数字特征，核心指标包括访问频次、停留时长等。客户行为分析指标详见表 3.3。

表 3.3 客户购物行为分析指标

编号	分析指标	计算方法	数据输入
A17	平均访问频次	对一定日期区间内的会员平均访问频次、线上平均访问频次进行统计分析 & 同比环比分析。支持用户选定时间区间分析 ^[34] 。	时间区间 分析客群
A18	停留时长	对一定日期区间内的会员停留时长、单店平均停留时长、线上停留时长进行统计分析 & 同比环比分析。支持用户选定时间区间分析 ^[35] 。	时间区间 分析客群
A19	停留时长分布	对单店停留时长在不同品类中分为 5 个等级，1 为停留时间最短，5 为停留时间最长，分别计算每个等级的人数及占比，得出消费者进店的停留时长分布 ^[36] 。	时间区间 分析品类
A20	客流量	对一定日期区间内的进场客流量、地铁入口客流量、单店进店客流量、单店所在动线客流量进行统计分析和同比环比分析。支持用户选定时间区间分析 ^[37] 。 对在历史客流统计中发生的全场客流量峰值、单店客流量峰值进行独立显示 ^[38] 。	时间区间
A21	动线客流	对一定日期区间内经过动线的客流人次进行统计，分析楼层动线客流、楼层估计动线客流。支持用户选定时间	时间区间

编号	分析指标	计算方法	数据输入
		区间分析 ^[39] 。	
A22	客流热度	将客流点位通过人次数据收集，将客流量以颜色区分热度等级，在楼层地图上对店铺热度进行标记。支持用户选定时间区间统计。支持用户选定时间区间分析 ^[40] 。	时间区间
A23	车流量	对一定日期区间内的进场车流量进行统计分析和同比环比分析。支持用户选定时间区间分析 ^[41] 。	时间区间
A24	车流黏着度指数	以本月最后一天为截止日，将到访汽车的停留总时长、到店频次、最后一次到店时间作为分析指标。将各指标最小值与最大值作为极值，平均分为 9 个等级，得到各等级的区段划分，将这九个等级区间依次赋以 1~9 的分数，得到停留总时长、到店频次、最后一次到店时间得分 ^[42] 。 车流黏着度指数=停留总时长得分*权重+到店频次得分*权重+最后一次到店时间得分*权重 ^[43] 对一定日期区间内车流黏着度指数分布进行统计分析 & 同比环比分析。支持用户选定时间区间分析。 ^[44]	时间区间
A25	客流转化率	店铺所在动线客流转化为进店客流的转化率=进店客流量/店铺所在动线客流量 进店客流转化为消费客流的转化率=交易笔数/进店客流量 ^[45] 进店消费客流转化为会员的转化率=消费者转化为会员人数/总消费人数 进店消费客流再次消费比率=在商铺再次消费人数/总消费人数 对转化率显著的日期、商铺进行标注；对总体转化率进行标注 ^[46]	品牌名称 时间区间
A27	线上平台转化率	下定订单转化率=下定订单访次/网站着陆页访次 注册会员转化率=注册会员数量/访客数量 重度会员转化率=重度会员数量/注册会员数量	时间区间
A28	客流坪效	客流坪效=经过客流人次/面积 对一定日期区间内店铺客流坪效进行统计分析与同比环比分析。支持用户选定时间区间分析 ^[46] 。	时间区间
A29	跳出率	线上平台跳出率=仅阅读了一个页面就离开的用户/该页面拜访次数 对一定日期区间内线上平台跳出率进行统计分析与同比环比分析。支持用户选定时间区间分析 ^[46] 。	时间区间
A30	流量来源	分析线上流量来源分别是什么，例如：直接访问、微信渠道、搜索引擎渠道等。直接访问：直接访问网站。微信渠道：通过微信渠道访问网站。搜索引擎渠道：通过搜索引擎链接访问网站。搜索引擎来源：百度、搜狗、	时间区间

编号	分析指标	计算方法	数据输入
		谷歌等引擎来源分类。 对一定日期区间内线上平台流量来源数量进行统计分析 及同比环比分析。支持用户选定时间区间分析 ^[47] 。	
A31	活跃度指数	以本月最后一天为截止日，将所选区域到访会员数量、 访次总数、平均停留时长作为分析指标。将各指标最小 值与最大值作为极值，平均分为 9 个等级，得到各等级 的区段划分，将这九个等级区间依次赋以 1~9 的分数， 得到到访会员得分、访次得分、平均停留时长得分。 活跃度指数=到访会员量得分*权重+访次得分*权重+平 均停留时长得分*权重 对一定日期区间内地铁沿线活跃度指数、区域活跃度指 数进行统计分析 & 同比环比分析。支持用户选定时间区 间分析 ^[48] 。	时间区间 分析范围

4 零售业大数据客户分析模型设计

目前，国内数据分析在零售领域中的应用，主要集中在客户分析（包括客户行为分析、客户流失分析和客户忠诚度分析等）、产品分析（包括发现管理规则、市场预警和进行描述式数据挖掘等）、促销分析以及改进、企业市场预测机制等方面。在客户分析方面，目前的研究重点是基于数据模型的客户细分方法^[49]。

客户细分的目的就是要找出对企业有重要贡献价值的客户，这些重要客户有哪些特征，重要客户的实际需求是什么。客户响应预测模型就是要估计客户对一个产品或服务的响应概率，是企业采用最为频繁的一种预测模型的方法，旨在发展新的潜在客户^[50]。企业通过某项市场营销活动的响应模型分析，来找到最合适的响应客户，并预测哪些客户能够响应，以便有针对性地开展营销活动。

4.1 客户价值模型（RFM）

客户价值模型是价值聚类的思想，以客户的价值贡献程度为标准，将相同贡献程度的客户划分到同一个价值等级中，进而划分客户结构，实现客户价值等级聚类。同时为了衡量客户价值和交易频次的关系，除了考虑客户在一定时段内的交易总金额外，还要考虑客户购买频次、最近一次交易时间等因素，以反映其消费行为的活跃度^[51]。

RFM 分析是经典客户关系管理理论中重要的分析方法。RFM 分析是对消费者最后购物时间（Recency），到店次数（Frequency）以及购物总金额（Monetary）三个指标的综合评价^[52]。

最后购物时间（Recency）是指以本月最后一天为截止日，将客户最后一次消费时间距截止日的天数作为分析指标。

到店次数（Frequency）是指以本月最后一天为截止日，将客户消费次数作

为分析指标，统计客户总消费次数。

购物总金额（Monetary）是指以本月最后一天为截止日，将客户个人销售总额作为分析指标。

在计算 RFM 分数时，将客户在 R、F、M 三个指标上的表现得分按照一定的权重加和，计算出客户顾客的 RFM 价值得分，实现 RFM 得分等级聚类，用以评价、描述客户的实际价值。

4.2 客户生命周期模型（NESS）

客户生命周期是指从一个客户第一次购买行为开始，直到客户关系完全结束的这段时间。客户的生命周期是企业与客户关系生命周期的演变缩影，描述了客户与企业的关系从建立到发展最终到休眠状态运动的总体特征^[53]。

在实际经营活动当中，最理想的状态是新增客户人数不断增加，现存客户的活跃时间不断延长，睡眠客户能够被激活成为活跃客户，源源不断的为购物中心带来活力^[54]。

本次研究中，初步设计 NESS 模型分析，将客户生命周期划分为四个阶段，便于深度把握处于各生命周期的客户构成，明确他们的消费偏好，通过精准营销实现对客户需求的深度挖掘。

NESS 模型分析是通过界定客户在一定生命周期内的活跃状态来划分客群的方法。NESS 模型将客户分为新增客户（New）、活跃客户（Existing）、沉默客户（Silence）和睡眠客户（Sleeping）四种状态^[55]。

新增客户 New 是指从最后一次消费至本月最后一天为止，不超过 90 天的客户。活跃客户 Existing 是指从最后一次消费至本月最后一天为止达 90 天以上，消费次数达 2 次以上的客户。沉默客户 Silence 是指从最后一次消费至本月最后一天为止达 90 天以上不足 365 天，消费次数仅有 1 次的客户。睡眠客户 Sleeping 是指从最后一次消费至本月最后一天为止达 365 天以上的客户^[56]。

通过对四种状态客户总数在连续月份中的变化，可以直观的观察出各类型客户的积累过程，识别出活动带动或商户经营中的问题节点。企业可以通过观

察 New 客群的新增情况, Existing 客群的忠诚度变化, Silence 客群的沉默和唤醒情况, 以及 Sleeping 客群休眠量, 从而层次、量化的了解企业与客户之间的关系^[57]。

4.3 客户粘性模型 (RFD)

在大数据时代, 客户和企业间的关系不仅限于交易, 各种线上线下的互动都是客户对企业的粘性表现。而随着技术的发展, 店内 wifi 登录、线上关注、互动都可以称为对客户粘性数据采集的方法^[58]。

本次研究中, 初步设计 RFD 模型分析。通过对客户的店内停留总时长 Residence、到店频次 Frequency、最后一次到店时间 Duration 进行统计, 为到店客户进行客户粘性的类型划分。

停留总时长 Residence, 是指拥有该手机号码的客户多次到店中, 在店内停留总时长。在店内停留时间越长的用户, 产生的消费的可能性越大, 产生消费的金额也越高。以本月最后一天为截止日, 将会员各次到店平均停留时间作为分析指标, 统计会员平均停留时间。通过对平均停留时间的分析, 一定程度上反映出该客户利润贡献的价值等级^[59]。

到店频次 Frequency, 是指拥有该手机号码的客户到店登陆 wifi 的总次数。以本月最后一天为截止日, 将会员历史到店频次作为分析指标。一般来说, 到店次数越高的客户的黏着度越高, 也更加容易通过合理引导成为高价值客户^[60]。

最后一次到店时间 Duration, 是指拥有该手机号码的客户在多次到店中, 最后一次到店距本月最末一天的天数。理论上, 最后一次到店时间距本月末最后一天的时间越短的客户, 其粘性越高。通过对最后一次到店时间的把握, 了解客户到店的规律趋势, 分析该客户对零售企业的黏着程度^[61]。

通过将停留总时长 Residence、到店频次 Frequency、最后一次到店时间 Duration 三个指标按照一定权重进行加和, 得到该用户的 RFD 分值。从而反映每一个客户对企业的粘性^[62]。

4.4 客户身份模型（CIM）

客户身份聚类从零售企业会员业务中脱胎，将与零售企业可能发生关联的会员用户分为核心客户、感知客户、全网客户三个大类。核心客户包括线下实体会员、线上平台会员、微生活电子会员。感知客户包括 WIFI 登陆者、微信粉丝、线上平台访客。全网客户是指仅对零售企业现有品牌有感知的会员，以及能够对品牌有感知的潜在会员。客户身份聚类，打破了仅对传统小会员进行管理的瓶颈，实现会员管理的扩展^[63]。

核心客户是零售企业的高价值客户，是愿意办理会员卡并暴露部分因私信息获取便利性的会员。这部分会员无论从活跃度还是对零售企业的销售贡献值来讲都是最高的，是企业需要重点经营的会员^[64]。

感知客户是对零售企业的商品、店铺有真实体验和感受的客户，是愿意在多种消费者触点上和企业合作发生联系、除购物外，还愿意参与企业互动的客户。这部分客户通过微信关注、WIFI 登陆的渠道可感知其与零售企业产生的接触。

全网客户是对零售企业现有品牌有感知的客户，是实际线上线下全部和零售企业有联系的客户，以及能够对品牌有感知的潜在客户，包括全网范围内对零售企业品牌有所感知的最大会员范围。

5 北京朝阳大悦城客户分析模型应用的案例分析

5.1 案例背景

朝阳大悦城在国内大数据分析处理领域一直处在领先地位，并已经做了大量的努力和尝试。2015 年，朝阳大悦城实现 POS 机系统和 CRM 系统的打通，电子会员卡与实体会员卡打通，建立了良好的数据准备、打通基础。同时在结合业务的数据分析处理方面也做了大量的创新，真实有效地指导业务开展。

为了进一步提升朝阳大悦城在大数据方面的建设能力，拓展数据来源、综合线上线下数据，强化大数据挖掘和分析能力，实现消费者的精准描述、企业的精准营销，管理机构的精准监管，朝阳大悦城展开了会员大数据体系的建设。

5.2 模型建立及应用

5.2.1 RFM 模型

通过计算消费者 Recency（最后购物时间），Frequency（到店次数）以及 Monetary（购物总金额）得出 RFM 分数，通过 RFM 分数对所有会员进行价值评估和划分。计算公式为：

$$\text{RFM 分数} = \text{Recency 分数} \times \text{权重} + \text{Frequency 分数} \times \text{权重} + \text{Monetary 分数} \times \text{权重}。$$

在实际分析中，RFM 得分设置介于 15~135 之间，保留两位小数。

R 得分：统计每个会员最后一次消费日期距 6 月 18 日的天数。以 0 天为最小值，最长天数为最大值，将最近一次消费天数平均分为 9 个等级，各等级依次赋以 5, 10, 15, 20, 25, 30, 35, 40, 45 的得分。将每个会员的最后一次消费日期距 6 月 18 日的天数归到各个等级中，即可得出各个会员的 R 得分。

F 得分：统计每个会员在 6 月 18 日以前发生过消费的天数。以 1 次为最小值，最多次数为最大值，将发生过消费的天数平均分为 9 个等级，各等级依次赋以 5~45 的得分。将每个会员发生过消费的天数归到各个等级中，即可得出各个会员的 F 得分。

M 得分：统计每个会员在 6 月 18 日以前发生过消费的总金额。以 0 元为最小值，最大金额为最大值，将消费总金额平均分为 9 个等级，各等级依次赋以 5~45 的得分。将每个会员发生过消费的总金额归到各个等级中，即可得出各个会员的 M 得分。

RFM 的分布及变化见图 5.1。



图 5.1 RFM 模型数据分布图

从图 5.1 中可以看到，朝阳大悦城将 15-135 分平均划分为 9 个 RFM 区间，最低分区间为 1，最高分区间为 9，实现 RFM 等级 1-9 划分。通过 RFM 分数及区间划分，朝阳大悦城更清晰的认识了客户价值特征、分布及差异，从而可以集中资源维护最有价值的客户。

5.2.2 NESS 模型

NESS 分析是通过界定会员在项目内的生命周期和活跃状态来划分客群的方法。NESS 将会员分为新增（New）、现存（Existing）和睡眠（Sleeping）三种状态，对应不同的营销阶段目标。其中各状态定义如下：

New: 截止到当月底，第一次购买时间小于等于 90 天的会员总人数，是指项目在三个月以内新吸引到的消费者群体，连续数月 New 值的持续增加直接表明了项目的成长性，相反 New 值的明显萎缩需要立刻引起重视。

Existing: 截止到当月底，第一次购买时间大于 90 天，最后一次购买时间小于等于 365 天的会员总人数；Existing 是指一年之内到店消费的非新增群体，是支持项目会员消费的主体，Existing 的平稳和持续增加是保证项目销售平稳成长的基石。

Sleeping: 截止到当月底，第一次购买时间大于 90 天，最后一次购买时间大于 365 天的会员总人数，Sleeping 是超过一年未到店消费的已经流失的客群，

该数值增速的显著增加是极其危险的信号。

NESS 的分布变化见图 5.2。

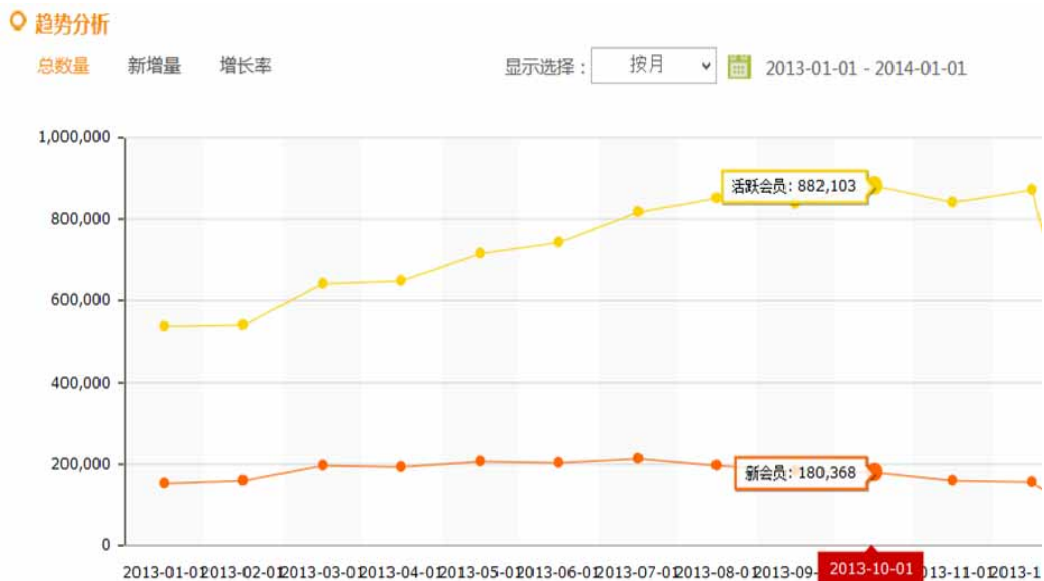


图 5.2 NESS 模型数据分布图

从图 5.2 中可以看到，四种状态会员总数在连续月份里的变化，可以直接观察出项目会员的积累过程，识别出活动带动或者问题节点；描绘会员的活跃度变化，观察 New 客群对项目的理解和接受程度，以及 Existing 客群的忠诚度变化，在这一点上交易笔数也可以作为补充数据。

5.2.3 客流动线规划模型

客流动线是指顾客在穿越购物中心时最愿意选择的路线。实际销售活动中，店铺成交量与经过店铺的客流量成正比。因此，全面把握客流动线的状况，进行合理规划与调整，对激活盲区资源和商铺租金的定位与调整都起着决定性的作用。这部分主要分为区域客流及动线规划两个方向：

（1）区域客流分析

外围总体客流量分析：在商场与外界连通的所有通道部署客流统计点，可以获取购物中心整体的客流量数据。以此可以掌握购物中心整体客流量的变化趋势及规律，并可掌握各通道的客流量分布情况。

通过对营销活动前后整体客流量的变化的跟踪，可以对活动效果作出整体的判断和评估并且为优化总体资源的利用提供数据依据。

各区域、各楼层客流量分析：在商场各区域、各楼层通道间部署客流统计点，可以获取购物中心各个区域的客流量数据。以此可以掌握购物中心各个区域在不同时段的客流量分布规律，各个区域的客流贡献(集客力)，以及顾客在

各个区域间的流动规律及相互影响的关系等。

通过这些数据的分析与对比，购物中心对客流的把握更加精准，可以进一步优化购物中心的客动线，以使顾客在购物中心内部更加有效地流动，带来更多的销售机会，对各区域的运营资源进行更加深入的优化。

品牌店铺客流量分析：为商场各品牌店铺部署客流统计点，掌握更为精细的客流动态。在这一层面上，通过客流量数据的分析应用，购物中心可以对每一租户对购物中心客流贡献(集客力)有极其细微的掌控，进一步可辅助对品牌组合优化的决策。

通过对品牌店铺客流量的分析，可以找出购物中心动线的盲区与死角，通过合理的规划与引导，最大化的利用购物中心的场地资源，让购物中心的每一寸面积都可以发挥最大的效力。

(2) 客流动线规划分析

由于监测点的局限性，客流仅在首层各入口，各层直梯、扶梯等主要位置进行统计，是各层客流的唯一既有数据。因此，根据以上既有数据进行合理假设，将每层各入口客流量进行分配，按比例从不同出口离开该楼层，并对每个入口进行逐一计算，通过加总该层所有监测点流动于动线上的客流分布，获得各动线上分布的客流量，以及各店铺的客流量。即，通过既有的客流监测点数据——利用模型进行客流计算——各条动线的客流等级——各店铺的客流等级：

监测点位 A 客流量总入数 (IA) = 该层 A 口客流量出数 (OA) + B 口客流量出数 (OB) + C 口客流量出数 (OC) + ... n 口客流量出数 (On)

a. 动线客流 M 的推论

经过一条动线 M 的客流量 = A 口客流总人数 * 沿途各出口出数占比 0% * 至各出口经过动线 M 的可能性 p + B 口客流总人数 * 沿途各出口出数占比 0% * 至各出口经过动线 M 的可能性 p + C 口客流总人数 * 沿途各出口出数占比 0% * 至各出口经过动线 M 的可能性 p + + n 口客流总人数 * 沿途各出口出数占比 0% * 至各出口经过动线 M 的可能性 p。

b. 店铺 S 客流量推论过程

店铺 S 的客流量可根据经过该店铺的动线的客流量进行标示，若店铺处于岔路口或处于两条动线之间时，店铺 S 的客流量取几条动线的最大值。

c. 主要成果及表现形式

根据以上模型对客流进行了反复计算，得出了全楼所有店铺客流等级的纵向比较，以及每层各店铺客流等级的横向比较两部分主要成果，并以店铺落位图的形式进行了形象表达。纵向比较突出楼内各层间的客流热度，横向比较突出层内各区域间的客流热度。

全楼所有店铺客流等级分类方法依店铺数量均分为 3 个部分，得到整楼客流量前 1/3、中 1/3 和后 1/3 的店铺分类，建立客流量与店铺热度的相关关系，以客流量前 1/3 的店铺为热度最高，客流量中 1/3 的店铺热度次之，客流量后 1/3 的店铺为热度最低，并以颜色进行标示：红色——热度高；橙色——热度居中；蓝色——热度低。（如图 5.1）。

每层各店铺客流等级分类方法与前者类似，但将各店铺所得百分制分值均为 5 档，0-20、20-40、40-60、60-80、80-100，更加细化了层内各区域的热度，并同样以颜色进行标示：红色——热度最高（80-100）、黄色——热度较高（60-80）、绿色——热度居中（40-60）、蓝色——热度较低（20-40）、灰色——热度最低（0-20）（如图 5.2）。

图 5.3 是根据模型计算在某一层中客流的最终分布情况。

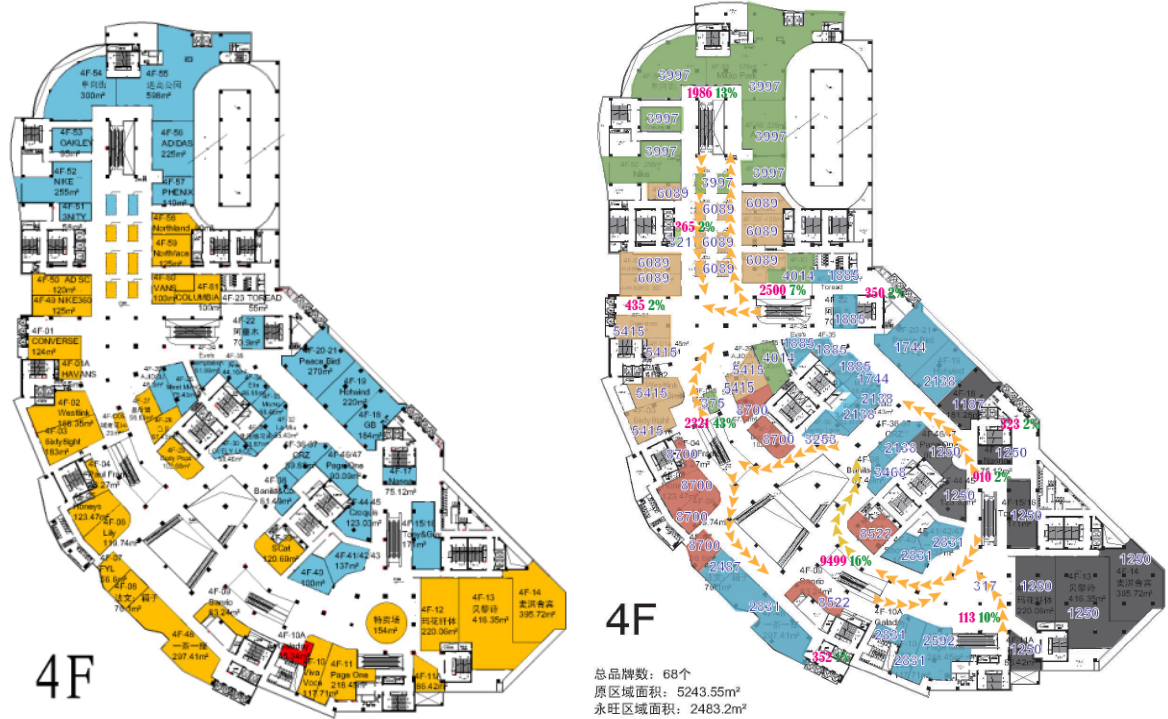


图 5.3 店铺客流等级分布图

5.3 应用总结

朝阳大悦城在上述分析模型的理论研究基础上，经过一段时间的模型因子调整，实现了对每一个加入会员的客户进行模型评价。即每一个会员都有客户价值评分（RFM 模型）、客户生命周期评分（NESS 模型）等，从而在传统的以销售额为唯一评价基础上，提供了更加立体的客户分析与评价方式。

以客户价值评分为例，RFM 模型提供了每一个客户在任何一個经营时间点的即时价值，从而可以从趋势、分布及数据关联等维度深入研究客户群体。

在趋势研究方面，按季度观察会员价值结构，可以发现自 2012 年 1 月到 2014 年 1 月，朝阳大悦城高价值会员（RFM6-RFM9）在稳步上升。但从整体分布情况看，会员主要集中在 RFM1-RFM4（51%），即朝阳大悦城会员价值还有很大的挖掘空间。

在分布研究方面，朝阳大悦城的客户 RFM 值有明显的聚集特征。最高价值会员的 RFM 评分集中于高位，说明高价值会员但消费能力高，其消费频率也远高于普通会员，而且发现会员级别的激励有助于会员价值挖掘。

朝阳大悦城通过上述分析模型应用，在客户分析与研究方面，得出了相对传统分析方法更深度和广泛的结论，并根据这些分析结论，制定了针对性、差异化的客群营销策略、品牌招商调整策略。在实体零售业受到广泛冲击的今天，朝阳大悦城依旧连续 2 年取得了 30%的业绩增长。

6 结论

6.1 主要结论

传统零售业的数据分析以企业自身为视角，以商品销售为侧重点，忽略和忽视了对消费者的分析和研究。随着零售业以客户为核心的经营理念的转变，随着海量客户数据的沉淀，围绕客户的数据分析成为企业经营的重要抓手。通过对零售业客户管理理论、当前零售业发展和传统客户分析方法的回顾及分析，本文提出了新的客户分析方法和一些新的客户分析模型，并结合实际案例对模型应用效果做了说明。本文的结论如下：

(1) 本文提出了大数据时代下零售业的一系列客户分析模型。其中 RFM 模型、NESS 模型和 RFD 模型分别符合对客户价值、客户生命周期、客户粘性的数据抽象，为上述维度的量化描述提供了方法和度量，其数据结论也符合传统经营经验的判断。本文提出的分析模型，符合客户管理理论和模型设计初衷，在零售企业经营管理中具备一定的实际指导意义。

(2) 在应用实践中，本文提出的分析模型的应用价值得到了进一步的验证。对零售企业的客户资产评估、客户群体细分有直接的帮助，进而对经营决策、营销规划等经营活动提供了直接的量化数据支持。

6.2 研究不足

首先，在零售业实际经营的实践中，由于市场、数据、资源环境的变化和客户消费偏好的变化，大数据分析技术的进步及数据采集方式的变革，还会衍生出新的客户分析主题，由于笔者自身研究水平等方面的原因，目前的客户分析模型还不够全面、不成体系，。

其次，关于各分析模型在实际使用中的因子设置，本文还缺少较为精确的量化比较和分析，这是由于数据可获得性以及相关因素难以量化等方面的原因造成的。以上两点是本文的研究存在的不足，有待深入探索。

参考文献

- [1]王丽芳. 基于客户关系管理的零售企业竞争力提升研究[J]. 河南财政税务高等专科学校学报, 2016(12):10-12.
- [2]田秀政. 商业零售业客户关系管理研究[J]. 商场现代化, 2015(6):7-10.
- [3]王永波. 面向大规模零售业数据的客户分析方法及系统研究[D]:[硕士学位论文]. 深圳: 深圳大学, 深圳大学硕士论文, 2016.
- [4]蔡玖琳. 基于数据挖掘的零售业客户细分方法[D]:[硕士学位论文]. 青岛大学, 青岛大学硕士论文, 2015. 蔡玖琳, 张磊, 张秋三. 一种基于数据挖掘的零售业客户细分方法研究[J], 重庆工商大学学报(自然科学版), 2015-2-25[2].
- [5]徐天. 分析型 CRM 零售业不可缺少的软件工具[J], 电子测试, 2015(5):24-28.
- [6]邵莉娟. 家乐福客户关系管理研究[D]:[硕士学位论文]. 中国知网, 云南大学硕士论文, 2015.
- [7]王俊峰. 郑州丹尼斯百货客户管理系统研究[D]:[硕士学位论文], 中国知网, 郑州大学硕士论文, 2013.
- [8]阳茂春. 购物中心客户满意度研究[J]. 电子科技大学学报, 2012(10):29-33.
- [9]王琮凯. 基于关联规则的联通华盛通信有限公司新一代零售系统客户消费行为分析研究[D]:[硕士学位论文]. 中国知网, 南京邮电大学硕士论文, 2014.
- [10] William G. Zikmund, Raymond McLeod, Jr., Fay W. Gilbert. Customer Relationship Management[M]. Oklahoma State University, 2008. 325-337.
- [11]R Cuthbertson, A Laine. The role of CRM within retail loyalty marketing[J]. Journal of Targeting Measurement & Analysis for Marketing, 2004(6):71-72.
- [12]唐璐瑶. 浅析大数据时代网络营销的客户关系管理[J]. 现代经济信息, 2015(6):33-35.
- [13]武靖恺, 刘哲. 刍议客户关系管理在“大数据”时代下的数据挖掘[J]. 无线互联科技, 2014(9):12-14.
- [14]郭崇. 基于大数据分析的网购客户忠诚度建模仿真[J]. 计算机仿真, 2015(1):61-63.
- [15]Urvashi Makkar. Customer Relationship Management[M]. University of California, 2006. 200-204.
- [16]陆岷峰, 虞鹏飞. 大数据分析在商业银行零售业务中的应用[J]. 金融理论与教学, 2015(6):42-43.
- [17]张晓诺. 利用大数据技术在电子商务中对客户忠诚度分析[J]. 中国科技信息, 2015(12):37-41.

- [18]方献梅,刘亮龙,高晓波.大数据时代数据挖掘在证券公司客户关系管理中的应用研究[J].信息与电脑,2015(3):12-14.
- [19]贺本岚.大数据时代数据挖掘在银行客户关系管理中的应用研究[C].-中国管理学年会,2013.44-47.
- [20] G Blum. Analytisches Customer Relationship Management (CRM) und Big Data[J]. Springer Fachmedien Wiesbaden, 2014.225-248.
- [21]李修林.基于数据挖掘的商业银行客户关系管理经验研究.经济期刊[J],2015(6):23-25.
- [22]张喆.数据挖掘及其在客户关系管理中的应用[M].复旦大学出版社,2007.129-130.
- [23]邓博.基于数据挖掘技术构建电信 4G 客户预测模型的研究[D].兰州大学,2015(2):14-16.
- [24]杨玉强.大数据时代网络购物平台营销战略研究[J].企业文化旬刊,2014(4):75-76.
- [25]汪晓曦.面向商务智能的数据挖掘及其应用研究[D]:[硕士学位论文].中国知网,武汉理工大学,2014.
- [26]王征硕.基于电商平台国际贸易中 CRM 的研究[J].电脑开发与应用,2014(3):22-23.
- [27]何树称.数据挖掘技术在客户关系管理中的应用分析[D]:[硕士学位论文].中国知网,云南大学,2004.
- [28]颜富强.遗传算法在数据挖掘中的应用研究[D]:[硕士学位论文].中国知网,湖南大学,2008.
- [29]冯倩.网购客户流失的实证分析[D]:[硕士学位论文].中国知网,西南财经大学,2013.
- [30]胡晓峰.聚类分析在银行客户细分中的应用研究[D]:[硕士学位论文].中国知网,暨南大学,2015.
- [31]徐龙飞.浅析大数据时代下商业银行的精准营销策略[J].现代营销,2014(11):56-57.
- [32]刘淳猛,何利力,李铨瀚.数据挖掘技术在 CRM 系统中的应用[J].工业控制计算机,2014(4):33-35.
- [33]陈萍.浅谈数据挖掘在个人金融服务客户关系管理中的应用[J].科技信息,2011(11),31-33.
- [34]郭百岭.数据挖掘方法在客户流失管理中的应用[D]:[硕士学位论文].中国知网,华中科技大学,2014.
- [35]李志玲.基于数据挖掘的客户关系管理研究[D]:[硕士学位论文].中国知网,.河北大学,2010.

- [36]柯林. 基于粗糙集的 CRM 数据挖掘理论与方法研究[D]:[硕士学位论文]. 中国知网, . 北京工业大学, 2009.
- [37]谭康荣. 基于数据挖掘的银行客户关系管理的构建与实施[D]:[硕士学位论文]. 中国知网, , 重庆大学, 2007.
- [38]李兵, 游文丽, 薛劲松, 朱云龙, 赖翔宇. 面向客户生命周期的客户关系管理模型[J]. 信息与控制, 2003(3):17-18.
- [39]李纯青, 姬升良, 董铁. 牛动态客户关系管理模型及应用[J], 西安工业学院学报, 2003(6):49-50.
- [40]肖铁. 客户关系价值模型分析与运用[J]. 商业研究, 2004(9):8-10.
- [41]黄敏如. 以客户为中心的客户关系管理模型探讨[J]. 环球市场信息导报, 2011(5):11-13.
- [42]李艳, 何建民, 崔耿. 客户关系分析模型框架及营销策略的设计[J]. 价值工程, 2006(12):33-34.
- [43]刘东升. 面向连锁零售企业的客户关系管理模型(R-CRM)研究[D]:[硕士学位论文]. 中国知网, 浙江工商大学, 2008.
- [44]冯兵, 罗新星, 周永生. 非契约型客户关系演变模型及其运用研究[J]. 预测, 2006(9):7-11.
- [45]吴睿. 基于 Agent 的客户关系管理模型研究[D]:[硕士学位论文]. 中国知网, 合肥工业大学, 2007.
- [46]肖威, 黄健. 对传统客户关系指数模型的改进. 南通大学学报:教育科学版[J], 2004(24):29-30.
- [47]焦娟妮. 020 模式下生产企业与传统渠道客户关系模型的构建[J]. 企业导报, 2014(7):9-12.
- [48]M Han. Customer Segmentation Model Based on Retail Consumer Behavior Analysis[N]. International Symposium on Intelligent Information Technology Application Workshops, 2008-8-13(4).
- [49]罗莹. 基于 RS 的客户关系管理模型设计[J]. Coal Technology, 2009(6):61-65.
- [50]黄敏如. “以客户为中心”的客户关系管理模型[J]. 农电管理, 2011(12):11-15.
- [51]孙树垒, 韩伯棠, 孙健. 全企业与客户关系维度模型及契约关系分析[J]. 北京理工大学学报:社会科学版, 2007(5):12-16.
- [52]肖威, 傅铅生. CRM 中的客户关系模型设计与改进[D]:[硕士学位论文]. 中国知网, 河北理工学院学报:社会科学版, 2004.
- [53]李嵩, 毛健. 基于数据挖掘技术的客户关系进展模型的研究[J]. 科技信息, 2006(22):29-32.
- [54]李纯青, 赵平, 徐寅峰. 动态客户关系管理的内涵及其模型[J]. 管理工程学报, 2005(17):31-33.

- [55]郑玉香, 佟志伟. 客户关系价值的企业市场细分模型与新方法[J]. 经济问题探索, 2006(4):8-12.
- [56]赵民杰, 金萍, 任晓娜. 客户关系管理的价值链模型研究[J]. 现代管理科学, 2005(2):71-75.
- [57]谭元戎, 孙剑平. 聚类模型在客户关系管理中的应用以及对特征提取的探讨[J]. 技术经济, 2007(10):55-59.
- [58]胡新, 殷宝, 陈军. 学基于客户价值评价模型的客户关系管理应用研究[J]. 企业经济, 2010(13):40-44.
- [59]施生旭, 吴声怡. 客户知识与客户价值关系模型及实证研究——基于客户关系生命周期视角[J]. 技术经, 2011(6):38-41.
- [60]赵敏祥, 张大亮. 客户关系价值分析模型与应用[J]. 浙江工业大学学报:社会科学版, 2004(5):4-8.
- [61]周洁如, 张新安, 周朝民. 客户关系维度及其度量模型[J]. 华东交通大学学报, 2007(10):51-55.
- [62]任朝江, 吴祈宗, 卢艳秋. 客户关系全生命周期的状态转移分析模型研究[J]. 生产力研究, 2007(12):21-25.
- [63]李品睿, 许守任, 许晖. 基于 RFM 模型的核心客户识别与关系管理研究——以保险业为例[J]. 现代管理科学, 2015(7):12-16.
- [64]潘书勤. 基于数据仓库技术的客户忠诚度模型研究[J]. 有线电视技术, 2015(1):67-69.

致谢

一转眼间，三年的研究生生活就要结束了，回望过去的学习生活，心中感慨万千，三年的时光使我得到了很大程度的提升。导师无微不至的关怀与教导一直响彻耳际，带给我太多的感动、感悟乃至成长。如今，研究生生活即将结束，我在这里向所有关心、帮助过我的老师、家人以及朋友呈上我最诚挚的感谢。

首先，我要感谢我的导师李春艳教授对我的教诲与关怀，感谢您在我完成论文期间的耐心指导。您言传身教，对学生严格要求，给我们提供了很多实践以及学习的机会，您的耐心包容了所有的无知以及肤浅，让我在学习、生活以及工作中都满载而归。无论是做学问还是为人处事，您都给我们树立了一生学习的榜样。

其次，我要感谢学院的恩师们，感谢您们一直以来的教导，我在您们的身上看到了深厚广博的学术造诣以及严谨的治学精神，它们深深的影响着我，在此我向您们表达我深深的敬意！

最后，我还要感谢我的父母以及朋友，感谢三年来你们一直陪伴我，支持我。