

楼台千里，烟火万家

地摊经济发展现状及客户特征分析



## 摘 要

在严峻的疫情形势下，失业率上升和就业不充分的现象不容忽视。于是从 2020 年两会前夕，由上海、重庆两市开始，逐步揭开了有序开放马路摊点的序幕，到 2020 年 6 月，全国各地掀起了“地摊经济”的热潮：从政策扶持，到企业援助，马路摊点或成为灵活就业的一个出口，成为缺乏资金技术的普通人能够解决生存问题的重要出路。

在此背景下发展“地摊经济”就具有了更重要的价值，不仅是能应对当下的疫情冲击，它是否还能长远地促进小微企业的发展和个体经营户的就业，是需要考虑的问题。因此我们需要对“地摊经济”的市场进行调查，并且提出以下问题：一、大众、新闻媒体对于“地摊经济”，最为关注的是哪些方面？二、客户的总体特征是怎样的？这些特征对他们的购买行为是否有影响？三、客户对于目前的地摊商品的满意度如何？对哪些因素最为重视？四、除现有客户外，还有哪些客户值得进一步挖掘？

针对第一个问题，我们采用词云、情感分析及 LDA 主题分析的方法对网络报导的文本进行数据挖掘，从中分析出大众对“地摊经济”的舆论态度主要是积极的，对政策保持着关心和支持的倾向，提出了对“地摊经济”多元化发展趋势的期待。此外，也反映出了一些人们心系的问题：“地摊经济”的发展是否会对市容、交通造成影响，并且呼吁着更加规范的市场管制。

接下来，我们以全国人口为调查对象，以每一位常住人口为调查单位，在居住地开展走访调查。我们结合了文献分析的方法，对问卷研究的内容进行了设计，并在特殊时期采用了网络调查与走访调查相结合的方式，在“问卷星”网站上进行问卷发布与收集，通过预调查 100 份样本的信度与效度检验后，收集了共 814 份正式调查的问卷结果。

为了解决第二个问题，我们采用描述性统计方法，对调查指标进行简要分析。调查结果表明，女性客户的占比大于男性客户。受调查者中，88.7% 的人都了解过有关“地摊经济”的扶持政策，他们最主要的信息来源为新闻媒体。在一系列的地摊产品中，最受客户欢迎的为宵夜烧烤、小吃以及生鲜水果，这表明，比起生活用品及服务方面，客户更乐于购买小吃类商品。

接着，为了进一步挖掘客户个人特征与购买行为之间的联系，我们使用 lasso-logistic 模型，针对性别、年龄、学历、家庭成员数目等指标，对客户的基本特征进行建模，分析客户购买行为的影响因素。

对于第三个问题，我们结合熵权法和模糊综合评价方法找到客户在地摊购物时的关注点，从满意度和重视程度两个角度分析影响客户购物行为的影响因素。对于第四个问题，我们利用 k-means 方法对潜在的客户进行聚类，以发掘潜在消费群体，提出合理的用户分析建议。

通过对调查数据的分析可以得到以下的结论：第一，根据文本分析结果，目前而言，民众、政府对地摊经济发展的态度，规范措施的落实以及地摊经济对就业的影响为媒体最关注的问题。第二，女性客户愿意购买地摊商品的比率更高，男性的购买意愿为女性的 77.1%，50 岁以上的年长者更受地摊商品的吸引。并且，政策也是影响因素质疑。政策宣传能够很大程度地促进地摊经济市场的发展，了解帮扶政策的人比没有了解过的购买几率上升了 166.32%。第三，客户前往地摊消费的主要原因因为地摊购物方便快捷、热闹、价格低廉，整体来说，客户对便利性、商品质量最为重视。第四，目标客户与潜在客户并存，最主要的目标人群为收入位于 10-20 万之间的人群，最主要的发展客户为年收入小于 10 万的人群。

根据上述结论，我们提出以下建议：对商贩来说，第一，以女性、50 岁以上和收入位于 10-20 万之间的客户作为主要的目标人群，以年收入小于 10 万的 18-40 岁之间的青年人作为主要的潜在客户人群，进行进一步的客户拓展；第二，保证购物流程的便利性和商品的质量；第三，对客户进行分类处理，开发具有一定针对性的商品。对政府来说，第一，大力发展“地摊经济”，持续刺激消费市场；第二，规范落实措施，保护城市面貌；第三，加大宣传力度，普及经营知识。

**关键词：** 地摊经济    文本挖掘    聚类分析    模糊综合评价

# 目录

1 引言 .....	5
1.1 问题的提出 .....	5
1.2 文献综述 .....	7
1.3 研究思路 .....	8
2 调查方案及数据检验 .....	10
2.1 调查方案设计 .....	10
2.2 调查实施过程 .....	14
2.3 预调查 .....	14
2.4 正式调查过程 .....	16
3 基于文本分析的“地摊经济”关注热点分析 .....	20
3.1 热词的提取 .....	20
3.2 热词的情感表达 .....	21
3.3 新闻文本主题提取 .....	23
4 基于调查样本的地摊购物满意程度分析 .....	25
4.1 消费者个人特征分析 .....	25
4.2 消费者对地摊经济的认知分析 .....	26
4.3 消费者地摊购物情况分析 .....	27
4.4 相关指标重视程度分析 .....	29
4.5 消费者满意度分析 .....	30
5 客户购买影响因素及特征分析 .....	31
5.1 基于 Lasso-Logistic 模型的购买行为影响因素分析 .....	31
5.2 基于模糊综合评价的客户满意与重视程度分析 .....	34
5.3 基于用户聚类的潜在用户挖掘 .....	36
6 结论与建议 .....	40
6.1 结论 .....	40
6.2 建议 .....	41
A 量表的描述性统计分析 .....	44
A.1 相关指标的重视程度 .....	44
A.2 满意度 .....	45

---

B 调查问卷 .....	46
--------------	----

# 插图目录

1.1	百度搜索指数趋势 . . . . .	5
1.2	百度搜索人群画像——地区分析 . . . . .	6
1.3	百度搜索人群画像——年龄分析 . . . . .	6
1.4	研究思路图 . . . . .	9
3.1	新闻文本词云 . . . . .	21
3.2	新闻文本情感得分直方图 . . . . .	22
3.3	LDA 模型参数估计过程 . . . . .	24
4.1	性别分布条形图 . . . . .	25
4.2	年龄分布条形图 . . . . .	25
4.3	学历分布柱形图 . . . . .	26
4.4	是否居住在商业区分布图 . . . . .	26
4.5	认知分析及了解渠道 . . . . .	27
4.6	前往地摊购物频率条形图 . . . . .	27
4.7	前往地摊购物原因分布条形图 . . . . .	28
4.8	购物商品偏好分布条形图 . . . . .	28
4.9	不前往地摊购物原因分布条形图 . . . . .	29
4.10	相关指标重视程度条形图 . . . . .	30
4.11	满意度条形图 . . . . .	30
5.1	$\log(\lambda)$ 变化图 . . . . .	32
5.2	Lasso 解析图 . . . . .	33
5.3	分客户类型重视度打分情况 . . . . .	38

## 表格目录

2.1	经费预算表 . . . . .	12
2.2	调查项目表 . . . . .	13
2.3	重要程度信度检验表 . . . . .	14
2.4	满意度信度检验表 . . . . .	15
2.5	重要性内容效度检验表 . . . . .	15
2.6	满意度内容效度检验表 . . . . .	15
2.7	预调查 KMO 和 Bartlett 球形度检验 . . . . .	16
2.8	预调查 KMO 和 Bartlett 球形度检验 . . . . .	16
2.9	性别与地摊购买状况四格表 . . . . .	17
2.10	居住商业区与地摊购买状况四格表 . . . . .	17
2.11	政策了解与地摊购买状况四格表 . . . . .	17
2.12	四格表独立性检验表 . . . . .	18
2.13	四格表相合性检验表 . . . . .	18
2.14	年龄与购买情况列联表 . . . . .	18
2.15	受教育程度与购买情况列联表 . . . . .	18
2.16	受教育程度与购买情况列联表 . . . . .	19
2.17	kendall 系数与 Somer's D 系数计算及检验结果表 . . . . .	19
3.1	主题提取结果 . . . . .	24
5.1	变量设置 . . . . .	31
5.2	变量设置 . . . . .	33
5.3	熵值法权重计算表 . . . . .	35
5.4	满意度模糊综合评价结果 . . . . .	35
5.5	重视程度模糊评价综合评价表 . . . . .	36
5.6	聚类结果表 . . . . .	37
5.7	客户个人特征分类 . . . . .	37

# 1 引言

## 1.1 问题的提出

### 1.1.1 问题背景

地摊经济，比起其他的经济体制，是一种较为广泛，较难被定义的概念。它主要指一种依靠“摆地摊”获取收入来源的经济形式，是一种合法性仍存在争论的边缘经济体。

今年 6 月 1 日，国务院总理李克强在考察山东烟台时，重点表示：“地摊经济、小店经济是就业岗位的中叶来源，是人间的烟火，和‘高大上’一样，是中国的生机。”一时间，“人间烟火”地摊经济一夜爆红。多地政府皆出台政策，大力扶持地摊经济，互联网巨头也相继发布“地毯经济”帮扶计划。清晨的阳光中，深夜的霓虹里，无处没有“星星之火”们的身影。“地摊经济”也一夜之间成为“热词”。根据百度搜索大数据报告，近 7 天来，“摆摊技巧”词热度创 10 年新高，同比增涨 655%。我们根据“地摊经济”这一关键词，使用百度指数进行了进一步的分析。

地摊经济的热度，自提出以来，一直处于上升状态。根据百度指数显示，自从“地摊经济”关键词于两会上提出之后，其受到的关注越来越多，并在 6 月 5 日达到峰值，自此之后，“地摊经济”热度一直居高不下，并再次在 6 月 9 日达到峰值。30 天以来，“地摊经济”的搜索日均指数达到了 19304，成为了全国人民热切讨论的话题。<sup>1</sup>



图 1.1: 百度搜索指数趋势

对地摊经济最为关心的人来自广东、浙江、江苏。对人群画像进行分析，可以发现，全国

<sup>1</sup>数据来源：<http://index.baidu.com/v2/main/index.html#/trend/>



“地摊经济”搜索指数最高的省份为广东，其次为浙江、江苏等。这说明，这些地区的青年人最关心“地摊经济”的走向，也是“地摊经济”发展最有潜力的几个省份。<sup>2</sup>

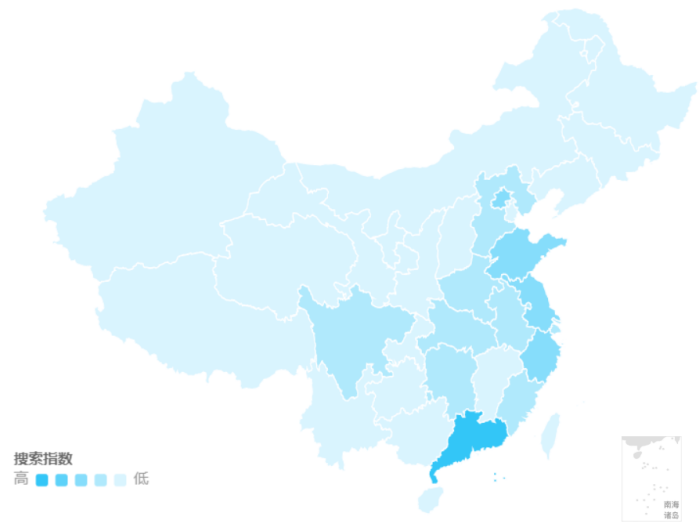


图 1.2: 百度搜索人群画像——地区分析

30-40 岁的青年人，最受地摊经济的吸引。对人群年龄进行分析，可以看到，对“地摊经济”最感兴趣的为 30 至 39 岁的青年人。虽然简简单单的“地摊经济”对如今的后浪们没有足够的吸引力，但其却能够调动一批“中浪”的胃口，这说明，地摊经济能够为大众提供巨大的就业机会，带动人群实现全面小康。<sup>3</sup>

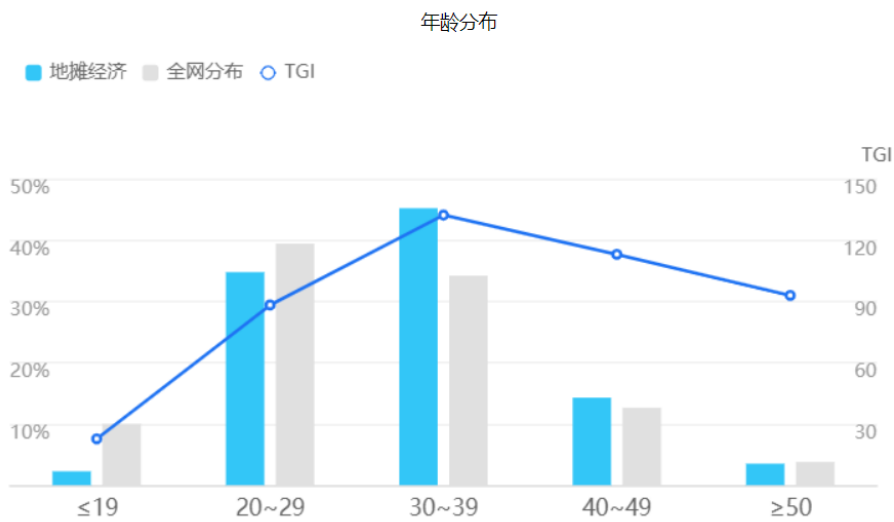


图 1.3: 百度搜索人群画像——年龄分析

<sup>2</sup>数据来源: <http://index.baidu.com/v2/main/index.html#/trend/>

<sup>3</sup>数据来源: <http://index.baidu.com/v2/main/index.html#/trend/>

正如上文所分析的,流动商贩是生活中不可或缺的存在,尤其是在“后疫情”时代,各个行业发展趋势疲软,重振市场、激发购买力是目前发展经济刻不容缓的目标,而“夜市”、“地摊”就是一个非常好的切入口。对于流动商贩来说,摆地摊就像是一种“微创业”——没有创立公司那么复杂,但也没有那么简单。就像搜索指数上反映出来的一样,商贩们需要精准把握需求,实时跟进客户偏好,合适地定价,合理地定位,才能真正地在这一片城市烟火中立足。本调查通过调查地摊商贩的目标用户群体的购买意愿、购买经历,对数据进行模型分析拟合,研究了可能影响用户购买与否的主要变量,并据此为商贩提供了一个可靠的影响因素分析结果,为政府提供了详细的指导政策。

### 1.1.2 研究意义

“地摊经济”并不是一种新兴的经济体,而是一个长久存在于城市中的,非常常见的经济体。然而,由于其社会收益性较低,且对现代城市的市容市貌影响较为严重,所以,很多城市都较大程度上限制了“地摊经济”的发展。但是,在疫情来临之后,大量企业面临着裁员危机,失业人员数量增加,待业人员难以找到工作,就业市场非常疲软。并且,生产厂家的库存堆积,无法有效处理售卖,导致厂家面临着严重的经济损失。在这样的情况下,通过“地摊经济”,来快速提供大量的就业岗位,在为失业人员提供就业岗位的同时,刺激居民消费,盘活市场。

然而,即使现在“地摊经济”在各地掀起了不小的“摆摊热”,其仍然是一个新兴行业,存在着非常多的问题。最大的问题就是经济发展与城市环境之间的博弈,例如,在北京市,很多市民反应个别地点出现混乱的摆摊设点、占道经营等等,对市民出行造成了很大的困扰。可见,“地摊经济”也需要因地制宜。除此以外,对于商贩来说,并不是所有的地摊商品都能招来眼球。只有清楚目标客户以及潜在客户的需求,才能合理安排地摊,规划地摊,真正的盘活经济金融市场。

本文基于上述原因,认为在疫情之后,合理地发展“地摊经济”对我们的社会维稳、经济恢复非常重要,故以“地摊经济”为调查对象进行市场调研与分析。本次调查希望能够深入调查目前“地摊经济”的主要客户群体,分析他们的购买意愿与购买行为,并对此进行建模,发掘“地摊经济”的潜在客户以及这些客户对于地摊购物这一新兴购物方式的建议与意见。调查结果能够为目前的“地摊经济”经营者提供经营建议,减少经营者在进入“地摊经济”市场时会因为不合理的经营政策受到的经济损失,帮助更有效地发展地摊经济。

## 1.2 文献综述

地摊经济由来已久,一直以来由于其合法性备受争议和关注,但国内缺乏对于地摊消费者人群特征、购买意愿及购买意愿影响因素等方面的研究,研究理论仍不完善。魏博通(2013)<sup>[5]</sup>在黄石对当地居民进行实地调研,结果显示,地摊经营的商品类别主要是餐饮小吃、服装

饰品和瓜果蔬菜。

消费需求方面，赵佳宁（2016）<sup>[4]</sup>在安宁区展开实地调研，调查结果显示，经常在地摊消费的居民人数占到 14.5%，偶尔在地摊上消费的居民占比达到 47.0%，而即兴消费的居民占到 31.3%，不去地摊进行消费的居民占比 7.2%。这说明，大量的居民往往是不带有购买的目的性前往地摊消费的，他们常常是在路过的同时，看到地摊商品并对其感兴趣，而决定消费。

发展前景方面，杨满珍（2015）<sup>[2]</sup>研究指出，地摊经济方便群众，活跃市场，有快捷的交易效率，并且能够增加就业岗位，缓解就业压力，丰富城市生活，能够起到带动城区经济发展的作用。该结论说明，在目前疫情导致的市场疲软的情况下，发展“地摊经济”是不容忽视的一方面。

城市管理方面，李坤梁、李娟（2018）<sup>[1]</sup>基于完全竞争视角下地摊经济管理研究指出，地摊的管理和摊主提供商品的需求弹性有密切联系，对地摊经济的管理可以从地摊商品的需求弹性进行思考。王岩（2018）<sup>[3]</sup>基于上海中心城区的实证调查研究指出，地摊经营存在私人生存空间与城市公共空间相互挤压、虚实结合经营模式的灵活性与管控主体多元化、“自由性”发展与商场的“规范性”生存之间的矛盾。

### 1.3 研究思路

本文结合了多方面的调查与研究方法，对数据进行充分挖掘分析。主要的研究思路如下：

第一，对问卷进行设计，并进行预调查检验，获得有效的问卷并开展正式调查。通过正式调查，了解大众参与“地摊经济”的具体情况。我们在问卷中设计了三个部分——用户基础信息、用户选择或不选择地摊购物的原因、以及用户对地摊经济的看法及建议。通过描述性统计分析，我们初步了解了地摊经济在群众中的普及程度，以及大众对地摊经济的认知度。

第二，通过基础的自然语言处理方法对“地摊经济”舆论倾向有一个直观的感受。通过 python 爬虫，收集网络上有关“地摊经济”的报导，对其进行文本分词，词频统计等处理，并使用词云对分词后的结果进行直观展示。然后，使用预训练的情感分析模型，对收集到的新闻标题进行细粒化情感分析，了解网络对于“地摊经济”的情感偏向。最后，使用 LDA 主题模型，提取文本中的隐含主题，了解公众对于“地摊经济”最在意的问题

第三，潜在用户挖掘，了解用户的潜在偏好。运用 Lasso-Logistic 模型，对用户的基础信息进行建模，分析影响用户购买意愿的因素。基于 Logistic 拟合结果，使用 k-means 方法对潜在用户进行分级聚类，发掘用户的潜在特征。

第四，根据评价量表以及模糊综合评价的方法，分析消费者在选择地摊购物时最看重的因素。根据上文所有的分析结果，为政府、消费者、以及地摊商户提出合理的建议，完善研究理论体系。

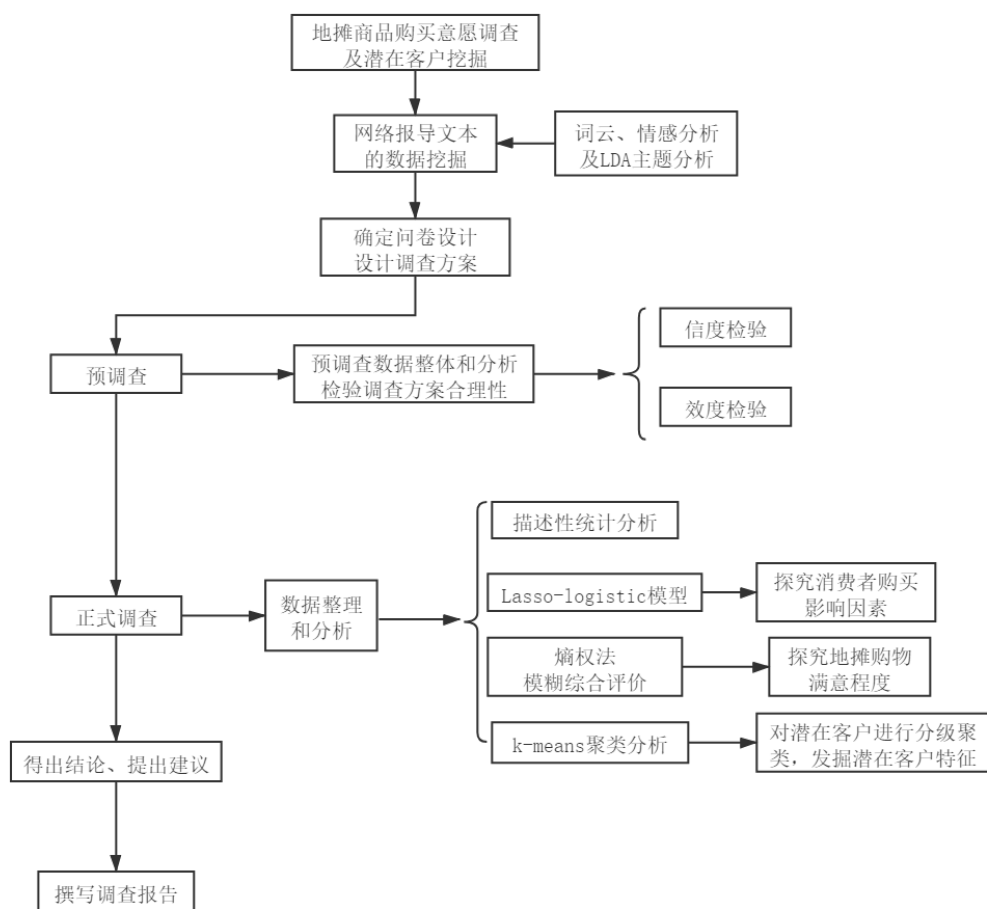


图 1.4: 研究思路图

## 2 调查方案及数据检验

### 2.1 调查方案设计

#### 2.1.1 调查目的

本次调查通过文本挖掘与网络调查，对地摊消费状况进行深入了解，对用户购买意愿进行分析，达到以下几个目的：

第一、爬取网络上关于“地摊经济”的报导，进行文本分词、词频统计，进行情感分析及 LDA 主题分析，了解公众对于“地摊经济”的情感倾向及最在意的问题；

第二、依据性别、年龄、学历、是否居住于商业区、家庭成员数目、家庭年收入水平等基本信息对调研个体进行分类，调查个体是否会前往地摊购物及原因和购物种类深入了解地摊消费基本情况；

第三、运用 Lasso-Logistic 模型，对用户的基本特征进行建模，分析用户购买意愿的影响因素，同时利用熵权法和模糊综合判别分析消费者选择地摊购物时最看重的因素，利用 k-means 方法对潜在用户进行分级聚类，发掘潜在用户特征；

第四、分析调查单位对地摊市场的意见及建议，并根据数据分析的结果，找出影响用户购买意愿的主要因素，消费者选择地摊购物时最看重的因素，及潜在用户特征，提出地摊经济发展的合理建议。

#### 2.1.2 调查对象和调查单位

**调查对象：**全国的全部常住人口

**调查单位：**全国每一个常住人口

#### 2.1.3 调查方式和方法

本次调查主要采用网络调查的方式，采用文献分析法、问卷调查法等方法进行调查。

##### 1. 调查方式

网络调查是通过互联网平台发布问卷，由群体在网络上自行选择填答的调查方法。网络调查是互联网日益普及的背景下经常采用的调查方法，其主要优点是访问者与被访者可以进行互动，即访问者可以即时浏览调查结果。从样本来源角度看，网络调查可以在更为广泛的范围内、对更多的人进行数据收集，资料庞大。

## 2. 调查方法

### (1) 文献分析法

文献分析法用于问卷调查的前期，我们通过查阅研究相关主题的文献资料及新闻报道，了解武汉市居民收入现状和现存问题，并搜集相关文献，分析其成果以及研究材料，为问卷调查提供了大量的背景资料，增强了研究结果的理论性，为调研提供了理论基础。

### (2) 问卷调查法

问卷调查法有标准化和成本低的优点，方便进行数据处理和分析，依据具体性原则、单一性原则、通俗性原则、准确性原则、简明性原则等设计调查问卷，在武汉市发放问卷，对抽出的样本进行问卷调查，进而分析武汉市居民家庭基本情况及收入现状及意见建议等。将调查分为预调查和正式调查两个阶段。在预调查后，根据结果，分析问卷的信度和效度，并对问卷进行改进；得到正式数据后，对问卷进行随机游程检验和卡方独立性检验，从而检验问卷的随机性和问题之间的独立性。

### (3) 任意抽样法

由于疫情的特殊原因，概率抽样难以实施，故我们采用了在居住地附近的地摊市场附近进行街头拦截的抽样方法。这种抽样方法的抽样误差难以控制，故为了适当减少抽样误差，我们对问卷进行了优化，并在样本的收集过程中，实时对样本进行监测，保证样本的真实性与代表性。

## 2.1.4 调查数据整理和分析方法

对预调查数据进行信度检验及效度检验以检验问卷调查结果是否有较好的代表性及可信性。正调查数据借助计算机集中分类整理，并采用定性分析与定量分析相结合的方法，进行列联分析、相关分析、回归分析、及多元分析。

## 2.1.5 调查时间和调查工作期限

1. 调查设计及前期准备：6.1-6.3
2. 预调查：6.4-6.6
3. 正式调查：6.10-6.20
4. 数据处理及分析：6.21-6.30
5. 报告撰写与提交：7.01-7.10

### 2.1.6 调查经费预算

表 2.1: 经费预算表

经费项目	预算金额（元）	备注
文案制作、打印费	100	相关资料搜集准备；问卷、文案的打印
文具	100	中性笔、笔记本等
样本费用	600	受调查者每份样本所需的调查费用
餐饮费	200	调研工作日
总计	400	

### 2.1.7 提交报告的形式

提交纸质报告书，报告地摊市场的发展状况及大众参与“地摊经济”的基本情况，并提出改善建议。

### 2.1.8 调查的组织实施计划

#### 1. 调查的组织管理及调查工作组的设置 调查方案设计：5 人

问卷设计：5 人

调查人员：5 人

数据处理及分析：5 人

报告撰写：5 人

#### 2. 调查人员的选择和培训

选择有经验且专业的调查人员，并在确定了调查项目内容、目的、要求等之后，进一步在专业老师的帮助下，系统地组织调查人员学习并掌握一些必备的调查知识以及注意事项等，全面提升调查人员的调查能力和素质及对本次调查的了解。

#### 3. 调查的质量控制

为了保证获得客观准确、真实有效的数据，提高调研结果的科学性，减少人为误差，对调查过程实施质量控制。

在调查方案及问卷设计之前，首先对调查项目进行深入了解，明确调查目的，随后查阅资料并参照专业人士的意见进行调查问卷的设计。在调查预备及实施阶段，合理安排调查任务，确保抽样员、访问员、复核员各自独立工作，从而实现样本的随机性；同时对小组成员进行调查技巧及注意事项的培训，使其能够应对调查过程中出现的突发状况。

在调查过程中，对于回答明显不完整、回答模式明显不符合要求的问卷，将指导调查对象重新填写直至符合要求，若多次指导下仍然无法获得满意问卷，将此问卷作为无效问卷处理。对于以下几种问卷，也将视为无效问卷处理：选择单一选项，李克特量表均选择同一分值；填答问卷时间过短；逻辑错误，前后矛盾；漏答过多。对无效问卷，将进行相应的剔除，并适当增加问卷发放的数量，保证有效问卷的数量足够多，从而能够满足每个最佳样本量的要求。

除此之外，对于一些开放式问题，实现统一开放式问题的处理方法，调查人员之间互相监督，从而降低理解误差，减少调查过程的随机性误差。随时随地检查回收问卷的有效程度，以保证问卷数据的效度和信度。

2.1.9 调查项目

根据前面所述的调查目的将此次的调查项目表分成三个部分：调查单位基本信息；调查单位购物情况及调查单位的购物感受及意见建议。

表 2.2: 调查项目表

类别	项目	编码
调查单位基本信息	性别	1
	年龄（岁）	2
	学历	3
	是否居住于商业区	4
	家庭成员数	5
	家庭年收入水平	6
	对地摊经济政策是否有了解	7
	对地摊经济政策的了解渠道	8
调查单位购物情况	是否去地摊购物	9
	去地摊购物的频率	10
	不去地摊购物的原因	11
	会选择在地摊购物的原因	12
	消费倾向	13
	是否看重商品品牌、是否追求潮流	14
调查单位的购物感受及意见建议	地摊经济对社会发展带来的好处	15
	最担心地摊市场存在的问题	16
	对地摊商品的期望价格区间	17
	对地摊购物相关因素的重视程度	18
	对地摊购物的满意度	19



## 2.2 调查实施过程

## 2.3 预调查

调查研究中的常用工具是问卷，而在其正式投入使用之前必须先进行信度和效度进行分析，只有信度和效度在研究可接受的范围之内的前提下，所收集的数据才有分析价值，其结果才能回答相应的研究问题。因而为了保证问卷的可靠性和有效性，在进行大规模数据收集以前，进行了一次预调研，来分析、纯化问卷的各个测量指标。

预调查中，我们预计收集 100 条调查样本，实际在收集了 100 条调查样本后，停止发放问卷，最终有效样本为 100 份。预调查的样本主要通过“滚雪球”的抽样设计方法进行收集，即由调查人员在“问卷星”网络问卷平台上，设计问卷并发放问卷，然后让已填写问卷的被调查者进一步分享问卷给其他的被调查者，从而扩大问卷的传播。信效度分析的结果如下：

### 2.3.1 信度分析

问卷的信度分析是为了考察问卷测量的可靠性，用来评价测量所得结果内部的一致性。本次使用克朗巴哈（Cronbach） $\alpha$  系数来检测数据信度是否达标，检测各分量表中被测试者对量表中条目回答的一致性。信度系数的取值范围为  $[0, 1]$ 。下面是 Cronbach 系数的计算公式：

$$\alpha = \frac{k}{k-1} \left( 1 - \frac{\sum s_i^2}{s_T^2} \right) \quad (2.1)$$

根据研究普遍观点，测验或量表的信度系数大于 0.9 时，表明该测验和量表的信度非常好，信度系数大于 0.8 是可以普遍被大众接受，信度系数大于 0.7 表示虽然量表价值仍在，但应该进行较大修改；信度系数低于 0.7 表明要重新设计。

信度检验结果如下：

表 2.3: 重要程度信度检验表

项目	Cronbach $\alpha$	项数	信度评价
商品质量	0.746	3	可接受
地摊购物体验	0.769	3	可接受
购物便利性	0.812	3	比较好

综合来看，Cronbach 系数均超过 0.7，信度比较好，问卷有一定可行度，可以进行大样本投放。

表 2.4: 满意度信度检验表

项目	Cronbach $\alpha$	项数	信度评价
商品质量	0.996	3	非常好
地摊购物体验	0.997	3	非常好
购物便利性	0.994	3	非常好

综合来看, Cronbach 系数均超过 0.9, 信度非常好, 问卷有可行度很高, 可以进行大样本投放。

### 2.3.2 效度分析

#### 内容效度

内容效度是指所设计的题项能否代表所要测量的内容, 用各分量表与总量表之间的相关性作为考察量表内容效度的指标, 内容效度指的是量表所能代表的主题的多少。对预调查中量表内容分析。

表 2.5: 重要性内容效度检验表

	商品质量	地摊购物体验	购物便利性
相关系数	0.592	0.616	0.716
Sig (双侧)	0.000	0.000	0.000

由上表可知, 各因子与总量表得分之间的相关性比较显著 ( $p$  值均小于 0.01), 量表内部一致性较好。

客户满意度量表内容效度分析如下:

表 2.6: 满意度内容效度检验表

	商品质量	地摊购物体验	购物便利性
相关系数	0.988	0.991	0.981
Sig (双侧)	0.000	0.000	0.000

由上表可知, 各因子与总量表得分之间的相关性很显著 ( $p$  值均小于 0.01), 量表内部一致性较好。

#### 结构效度

结构效度是指测量结果体现出来的某种结构与测值之间的对应程度。KMO 值用于比较题项间简单相关和偏相关系数, 取值在 0 到 1 之间。Bartlett 球形检验值用以检验题项间相关系数是否显著。

运用 SPSS 计算题目内容重要程度量表的 KMO 的数值是 0.820，超过 0.8，表示问卷结构设计较好。Bartlett（巴特利特）球度检验结果：卡方值为 470.205，数值较大，证明所对应 P 值（为 0.000） $<0.05$ ，因此巴特利特（Bartlett）球度检测具有显著意义。

表 2.7: 预调查 KMO 和 Bartlett 球形度检验

KMO 取样适切性量数	0.820
巴特利特球形度检验	470.205
显著性	.000

运用 SPSS 计算客户满意度量表的 KMO 的数值是 0.955，超过 0.9，表示问卷结构设计较好。Bartlett（巴特利特）球度检验结果：卡方值为 3121.421，数值较大，证明所对应 P 值（为 0.000） $<0.05$ ，因此巴特利特（Bartlett）球度检测具有显著意义。

表 2.8: 预调查 KMO 和 Bartlett 球形度检验

KMO 取样适切性量数	0.955
巴特利特球形度检验	3121.421
显著性	.000

## 2.4 正式调查过程

### 2.4.1 实地调查与抽样

本次调查主要采取了实地调查的方式。调查人员在居住地附近的地摊集中点随机拦截采访，通过“发红包”的方式，保证受调查者填写问卷的准确度，提高问卷的回收效率。在获得了受调查者的同意之后，调查人员将网络问卷链接分享发送给受调查者，进行样本收集。调查人员共进行了 6 天的调查，平均每日人均收集 30 份调查样本，在样本收集数量达到 850 份之后停止问卷发放，故本调查共收集了 850 份调查样本，其中有效样本数量为 814 份，最终提交的数据为 814 份样本的调查数据。

本调查的网络调查问卷在“问卷星”<sup>1</sup>问卷平台上发布，问卷星可以实时对问卷进行描述性分析，根据描述性统计分析的结果，我们可以对目前收集的数据进行监测。在监测过程中，我们剔除了部分填写存在问题，例如，在期望价格方面填写未按要求进行的样本。在收集完所有调查样本之后，我们对数据进行了第二次清洗。由于拦截采访中可能会多次遇到同一个受调查者，我们对受调查者的 IP 地址进行去重，并保留了受调查者第一次填写本问卷时的样本。在两次筛选之后，剩下的有效样本共 814 份。

<sup>1</sup>问卷星：<https://www.wjx.cn/>

### 2.4.2 关于商品购买情况的四格表分析

对收集到的数据进行列联表检验，分析变量之间的依赖情况与独立性。

表 2.9: 性别与地摊购买状况四格表

购买情况 \ 性别	女	男	合计
	女	男	合计
未购买	63	88	151
购买	361	302	663
合计	424	390	814

表 2.10: 居住商业区与地摊购买状况四格表

购买情况 \ 商业区	否	是	合计
	否	是	合计
未购买	65	86	151
购买	302	361	663
合计	367	447	814

表 2.11: 政策了解与地摊购买状况四格表

购买情况 \ 政策了解	否	是	合计
	否	是	合计
未购买	36	115	151
购买	56	607	663
合计	92	722	814

上述表格中给出了性别、是否处于商业区、政策了解情况与消费者地摊购买状况的四格表，通过构造卡方统计量可以实现对其独立性的检验。设置假设检验问题如下：

- $H_01$ : 性别与地摊购买两个变量是相互独立的。
- $H_02$ : 居住在商业区与地摊购买两个变量是相互独立的。
- $H_03$ : 是否了解地摊经济相关政策与地摊购买两个变量是相互独立的。

由于他们都是二维变量，因此构造的检验统计量服从自由度为 1 的卡方分布，具体的独立性检验结果如下表所示。

表 2.12: 四格表独立性检验表

变量	检验统计量值	P 值	结论
性别	7.481	0.006	不独立
商业区	0.219	0.640	独立
政策了解	27.56 0	1.523e-07	不独立

从检验结果来看, 是否居住在商业区与购买地摊商品二者是相互独立, 居住在商业区与不居住在商业区的消费者购买地摊商品的概率是没有显著差异的; 性别、政策了解情况地摊商品的购买状况不是相互独立的, 男消费者和女消费者、了解地摊经济政策与不了解政策的消费者在地摊商品购买选择上是有一定差异的, 但是他们之间到底具有着怎样的联系从目前的检验结果中尚无法得知, 因此我们可以针对这两个变量进行更进一步的相合性检验。

表 2.13: 四格表相合性检验表

变量	检验统计量值	结论
性别	-12742	负相合
政策了解	15412	正相合

根据相合性的检验结果, 性别与地摊商品购买的关系是负相合的, 即女性的消费者更偏向于选择购买地摊商品; 政策了解情况与地摊商品购买状况之间的关系是正相合的, 了解地摊经济相关政策的消费者更倾向于购买商品。

### 2.4.3 关于商品购买情况的列联表分析

表 2.14: 年龄与购买情况列联表

购买情况 \ 年龄	0 17 (1)	18 30 (2)	31 50 (3)	>50 (4)	合计
未购买	2	116	31	2	151
购买	5	477	164	17	663
合计	7	593	195	19	814

表 2.15: 受教育程度与购买情况列联表

购买情况 \ 受教育程度	小学 (1)	初中 (2)	高中 (3)	专科 (4)	本科及以上 (5)	合计
未购买	2	4	16	45	84	151
购买	7	16	65	196	379	663
合计	9	20	81	241	663	814

表 2.16: 受教育程度与购买情况列联表

购买情况 \ 年收入情况						合计
	<10 (1)	10 20 (2)	20 30 (3)	30 40 (4)	>40(5)	
未购买	22	49	57	15	8	151
购买	76	156	305	91	35	663
合计	98	205	362	106	43	814

上述表格中给出了年龄、受教育水平、年收入与购买状况的二维列联表,通过计算 Kendall 相关系数可以度量他们之间的相关程度。但是,考虑到受到列联表中会受到同分对的影响,Kendall 相关系数的仅从数值大小上也许不能很好的反应他们之间的相关程度,我们还可以通过计算 2\*c 列联表的 Somer 's D 系数,从条件概率的角度来对其相关程度进行度量。

表 2.17: kendall 系数与 Somer's D 系数计算及检验结果表

变量	Kendall 系数	P 值	Somer' s D 系数	P 值
年龄	0.055	0.118	0.060	0.119
受教育水平	0.637	0.637	0.017	0.682
年收入	0.073	0.037	0.117	0.002

从 Kendall 系数与 Somer 's D 系数的符号和显著性来看,这三个变量中与地摊商品的联系在统计上是有意义的。年收入水平与地摊商品购买具有正相合性,呈现出年收入增加购买概率增长的趋势,从 kendall 系数值来看,这两个系数之间的相关系数为 0.07,由于 kendall 系数会受到同分对的影响,因此不能准确的度量其相关性,从 somer' s D 系数数值大小来看,这两个变量之间的相关性为 0.117。

## 3 基于文本分析的“地摊经济”关注热点分析

本节通过 python 爬虫，以“地摊经济”为关键词，在新华网新闻站上爬取了 503 条新闻样本。由于 LDA 主题模型主要适用于长文本，对于评论类型的短文本的表现水平很差，所以本文分析中只保留了 116 条长篇新闻。对于这 116 条新闻，本文对其进行了中文分词处理，使用 tf-idf 对其进行文本向量化，根据文本相似度结果，使用主题模型进行主题提取，了解新闻媒体对于“地摊经济”的主要关注点。除此以外，本文也对新闻文本进行了细粒度情感分析，使用预训练语言模型，以了解新闻媒体对于“地摊经济”这一问题的整体情感倾向。

### 3.1 热词的提取

经过对各大新闻网站的 html 格式解析，我们选择了爬虫较为便捷的新华网。新华网的网站结构为 js 框架下的动态网页。通过检查网页的网络结构，可以获取其存储信息的 json 文件。对 json 文件进行逐页分析，便可以爬取该网站上所有已经发布的新闻标题，新闻的发布机构，以及新闻的具体链接。对爬取到的链接进行循环爬取，便可以获取所有的新闻数据。筛选掉不适合用于进行分析的短文本，以及部分重复文本，最后保留了 116 条 300 字以上的长文本。

本文使用的分词工具为 python 中的 jieba 分词，停用词表为自定义的停用词表，可以在附件中获得。比起其他文本分析中的分词方法，jieba 往往具有更好的分词能力。这是因为 jieba 分词是一种基于统计的分词方法，根据词频，判断词组属于的组合。因为词是稳定的，所以，在上下文中，相邻的字同时出现的次数越多，就越有可能构成一个词组。对文本中出现的各个字的组合的频率进行统计，就可以计算他们之间的互现信息。互现信息体现了汉字之间的紧密程度，当这种紧密程度高于某个阈值，便可以认为这是一个词组。这就是基于统计的分词方法。

去除停用词，可以更好的提取文本中所包含的主要信息。通过这两步处理，就已经形成了一个比较完整的词袋。这个词袋中应当包含大部分的重要文本信息，且不包含冗余的文本。对这些文本进行词频统计，并且据此做出 100 词的词云如下：



图 3.1: 新闻文本词云

可以看到，较为明显的关键词有“地摊经济”、“就业”、“轿车”、“收获”、“助力”、“互联网”、“时代”、“人间烟火”、“概念股”等等。这说明，地摊经济的发展，并不是一个行业的发展，其也同时带动了汽车行业、互联网等行业的发展。当然，地摊经济也需要其他行业的投资，在相互助力的情况下，才能相互促进，相互扶持。除此以外，“就业”等关键词也说明，大力发展地摊经济能够为目前较为疲软的就业市场提供大量的就业机会，为目前处于失业状态的年轻人提供一个快速的就业平台。地摊的门槛较低，一晚上即可增加十万多个就业岗位。并且，地摊经济是小微企业发展的前身，鼓励地摊经济的发展，可以进一步帮扶小微企业，在“后疫情”时代，为社会经济添加活力，添加“烟火味”。

### 3.2 热词的情感表达

本文使用了 python 中已有的中文自然语言处理模型 snownlp。传统的基于简单正负向情感词典出现词频的分析方法，很难正确地理解目标句的语义，从而错误地对文本进行情感标注。snownlp 使用的是基于贝叶斯推断的情感分析方法，这种方法往往比一般的情感词典准确很多。

贝叶斯的情感分析方法为，对于一个句子，如果不对其进行情感细粒度划分，那么该句子所表达的态度往往是可以分为两类的：积极态度 ( $c_1$ )、消极态度 ( $c_2$ )。我们假设一个句子有特征  $w_1, w_2, \dots, w_n$ ，这  $n$  个特征是相互独立的。这些特征就对应了句子中的单词。



对于一个具有  $(w_1, w_2, \dots, w_n)$  特征（词语）的句子，其属于  $c_i, i = 1 \text{ or } 2$  的概率为：

$$\frac{P(w_1, w_2, \dots, w_n | c_i) \times P(c_i)}{P(w_1, w_2, \dots, w_n)} \quad (3.1)$$

公式表达为：

$$P(c_i | w_1, w_2, \dots, w_n) = \frac{P(w_1, w_2, \dots, w_n) \times P(c_i)}{P(w_1, w_2, \dots, w_n)} \quad (3.2)$$

现在,  $P(w_1, w_2, \dots, w_n)$  仍是未知的。由全概率公式  $P(B) = P(B|A)P(A) + P(B|A')P(A')$ , 可以推出：

$$P(w_1, w_2, \dots, w_n) = P(w_1, w_2, \dots, w_n | c_1) \times P(c_1) + P(w_1, \dots, w_n | c_2) \times P(c_2) \quad (3.3)$$

合并式3.2和3.3，得到：

$$P(c_1 | w_1, w_2, \dots, w_n) = \frac{P(w_1, w_2, \dots, w_n) \times P(c_1)}{P(w_1, w_2, \dots, w_n | c_1) \times P(c_1) + P(w_1, \dots, w_n | c_2) \times P(c_2)} \quad (3.4)$$

这就是 snownlp 推理情感倾向的原理。

使用 snownlp 对每份长文本进行情感分析，对其进行情感极性打分。对情感得分作直方图：

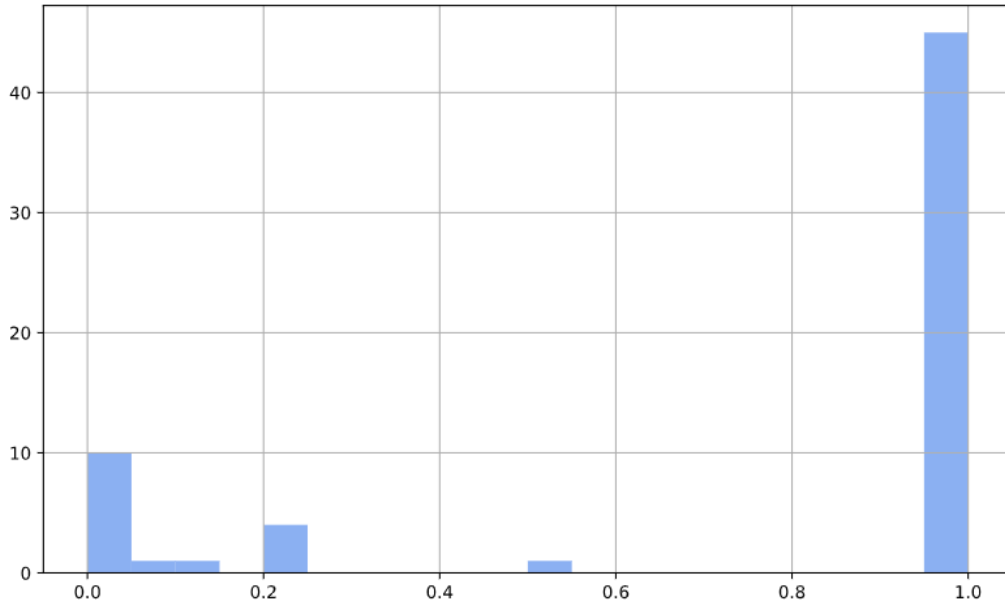


图 3.2: 新闻文本情感得分直方图

可以看到，文本的情感极性分布较为明显，其中，大部分的文本情感都是正向的，接近 0.9。同时，也存在部分情感极性小于 0.5 的文本。综合来看，文本情感得分的均值为 0.75，这说明，文本情感得分总体是偏正面的，然而，正如社会上舆论所表现出来的一样，也存在着一

些负面情绪。这些情绪大多来源于居民对地摊商贩不遵守交通规则，不维护市容市貌的不满。这也说明，“地摊经济”的普及，同样需要因地制宜，需要更精细的规划，更多的政策来约束。

### 3.3 新闻文本主题提取

LDA 主题模型是文本挖掘中经常会使用到的一种分析方法，其基于贝叶斯推断以及隐迪利克雷分布，常常被用于发掘文本、图像等数据中潜在隐变量。LDA 模型实际上是一种双向推断，对于每一篇文档中的每一个词，其都以一定的概率对应某一个主题，而这些主题也会以一定的概率选择某一个词语。LDA 模型的分析结果常常是词汇的线性组合。通过分析这些词汇的系数与内容，便可以对某些词汇代表的主题进行命题，挖掘潜在变量。

假设有一个数据集  $D = w_i^M$ ，并且，有多个主题： $\theta \sim Dir(\alpha)$ ，（ $\alpha$  为迪利克雷函数的参数）。对于每一个词  $w_n$ ，假设：

$$z_n \sim Multinomial(\theta), w_n | z_n, \beta \sim p(w_n | z_n, \beta) \quad (3.5)$$

根据贝叶斯公式，有：

$$P(D | \alpha, \beta) = \prod_{d=1}^M \int p(\theta_d | \alpha) \prod_{n=1}^{N_d} \sum_{z_{dn}} p(w_{dn} | z_{dn}, \beta) d\theta_d \quad (3.6)$$

实际上，LDA 模型提取主题的过程与我们写文章很像。首先从一个迪利克雷分布  $Dir()$  中抽样，对于一篇文章  $document_i$  都有一个主题分布  $\theta_i$ ，且  $\theta_i \sim Multinomial$ 。然后我们从这个多项分布中抽样得到每个文章  $document_i$  中的单词  $word_{ij}$ ，对应主题  $z_{ij}$ 。这样就得到了每个单词的主题。然后，我们在用一个迪利克雷分布  $Dir(\beta)$ ，并从中抽样，来形成每个单词对应的主题分布  $\phi(z_{ij})$ 。最后，从多项分布  $\phi_{z_{ij}}$  中抽样，形成文本， $w_{ij}$ 。通过 Gibbs 抽样的方法，可以对上述参数及逆行估计。下图很好地描述了 LDA 模型参数估计的原理和过程。

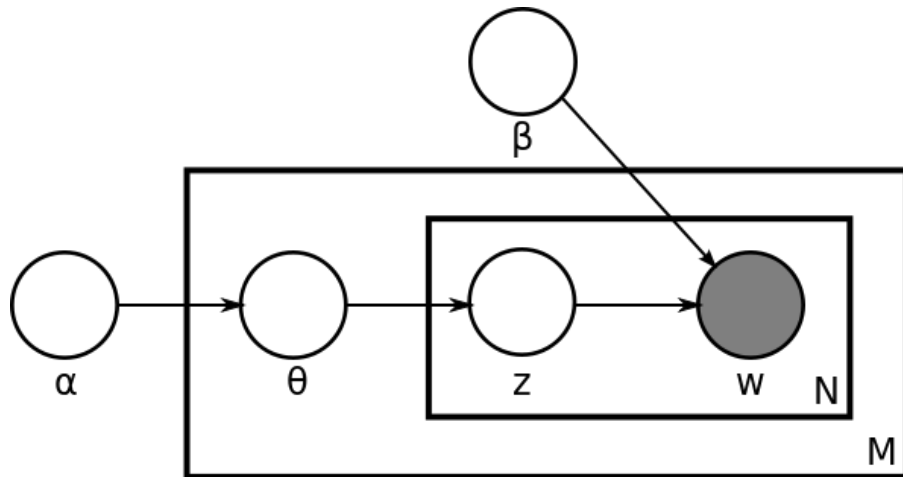


图 3.3: LDA 模型参数估计过程

使用 doc2bow 方法，对文档建立词袋模型 *corpus*。对于 *corpus*，使用 tf-idf 方法计算文本相似度，向量化文本。在已经向量化后的文本的基础上，建立 LDA 主题模型，提取 5 个具体主题，结果如下：

表 3.1: 主题提取结果

Topic1	Topic2	Topic3	Topic4	Topic5
市民	政策	区域	规范	产品
支持	销售	摊位	防控	提供
临时	摆摊	摆摊	临时	公司
摊位	临时	摊主	商贩	支持
提供	提供	摊点	市民	新
小店	商贩	中	提供	摊位

主题 1 中，根据“市民”、“支持”等关键词，我们认为该主题表达的主要为市民的对“地摊经济”的态度。可以看到，大部分的市民对于地摊经济还是持积极态度的。

主题 2 中，根据“政策”等关键词，我们认为该主题表达的主要为政策对“地摊经济”的扶持态度。主题 3 中，根据“区域”等关键词，我们认为该主题表达了媒体对于地摊商贩摆摊的区域的关注，而这也是同时困扰地摊商贩和城市管理者的一点。

主题 4 中，根据“规范”、“防控”等关键词，我们认为该主题反映了媒体对于规范措施的关注。由于地摊经济的一夜成名，城市管理难以跟上节奏，导致城市的市容市貌受到了一定的破坏。加强对地摊商贩的管理、规范，是合理发展地摊经济中非常重要的一步。

主题 5 中，根据“产品”、“提供”、“公司”等关键词，我们认为该主题反映了媒体及大众对地摊商贩所提供产品种类的关注。随着地摊经济逐渐火爆，地摊营业的形式也不仅仅局限于小吃、生活物品、服装服饰这几个方面。如果能够促进地摊经济的多元化发展，地摊经济必将展现出更出彩的一面。

## 4 基于调查样本的地摊购物满意程度分析

### 4.1 消费者个人特征分析

基于调查问卷，我们对特征数据进行了统计，从性别、年龄、学历、是否居住于商业区、家庭成员数目、家庭年收入水平等方面进行分析，并绘制出统计图表，以便对消费者的个人特征有更加直观的了解。

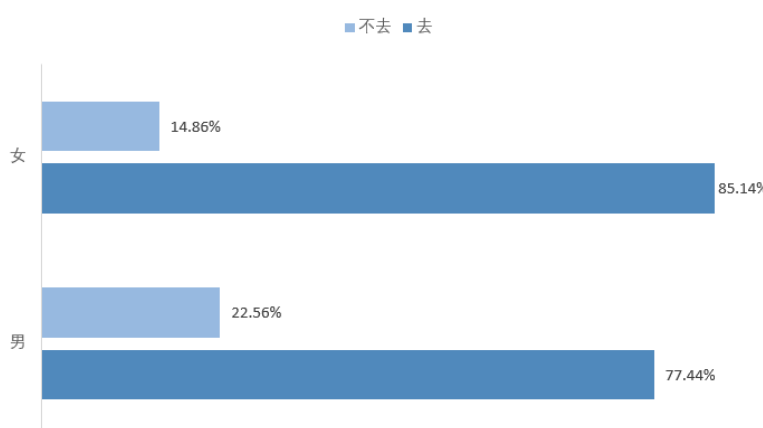


图 4.1: 性别分布条形图

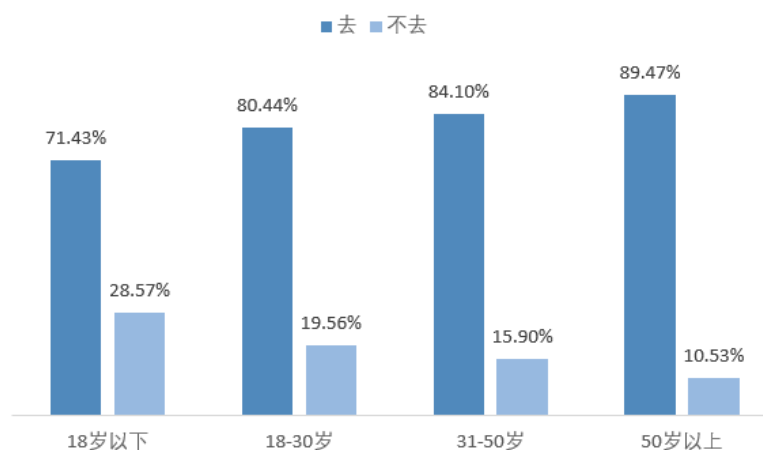


图 4.2: 年龄分布条形图

由图可以看出，相比男性消费者女性的消费者中愿意购买地摊商品的比率更高，年龄越高的消费者中会前往地摊购买商品的比率越高。居住在商业区与不居住在商业区的消费者购买地摊商品的倾向没有太大的差异。

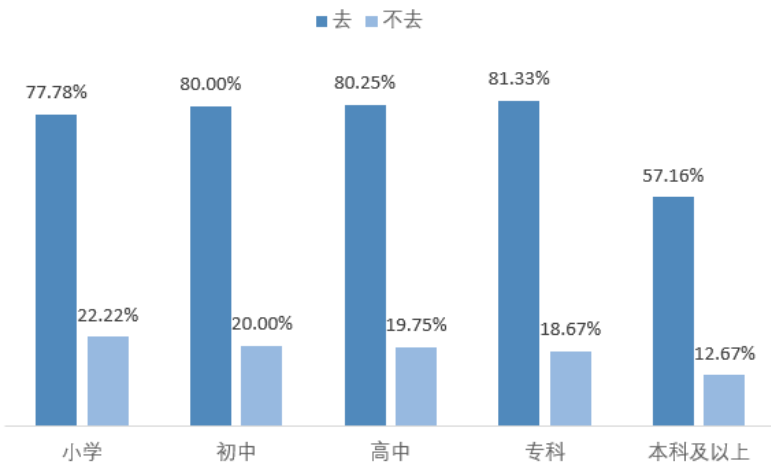


图 4.3: 学历分布柱形图

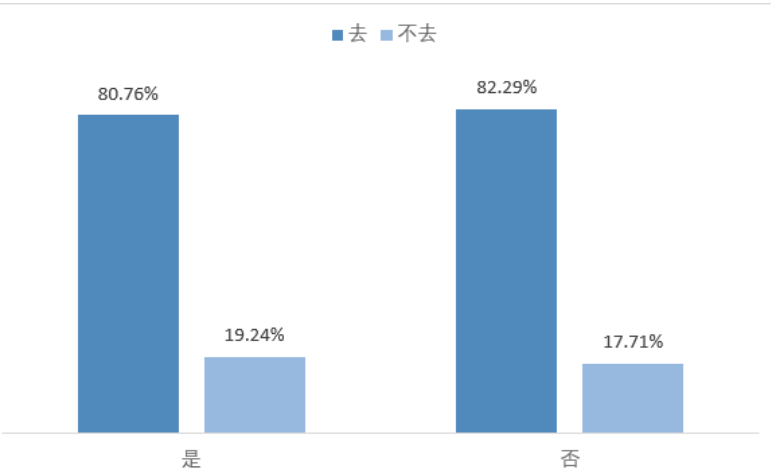


图 4.4: 是否居住在商业区分布图

4.2 消费者对地摊经济的认知分析

在正式调查所得到的样本中，88.7% 对地摊经济相关政策有过了解，11.3% 并不了解相关政策。在对政策有了解的消费者中，主要的了解渠道是新闻媒体，所占比重达到 55.8%，超过了总体样本数的一半，其次是社交媒体，占比为 31%。由此看出，新闻媒体仍然是居民了解国家政策的主要方式，但随着互联网技术的不断发展，社交媒体逐渐也成为居民了解政策导向的主要方式。

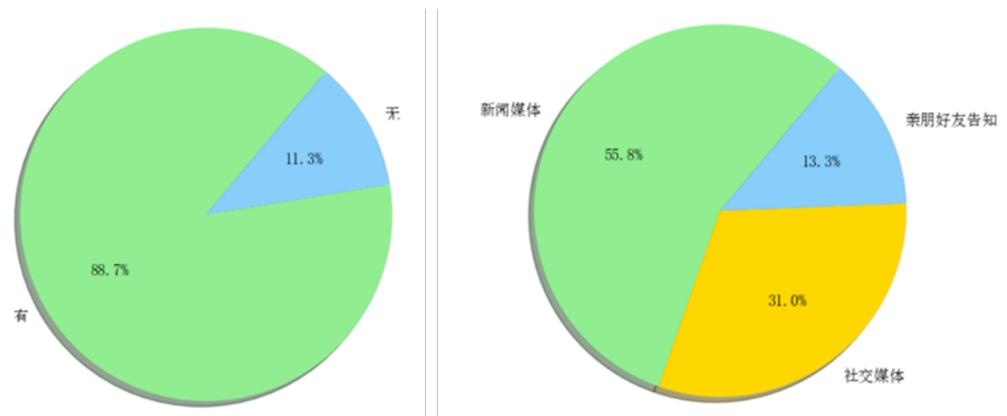


图 4.5: 认知分析及了解渠道

4.3 消费者地摊购物情况分析

正式调查的结果显示，所有的调查单位中 18.05% 不去地摊购物，81.95% 前往地摊购物，前往地摊购买商品的消费者中，15.8% 的消费者非常少去地摊购买商品，47.7% 的消费者经常前往地摊购物，36.5% 的消费者有时会前往地摊购物。

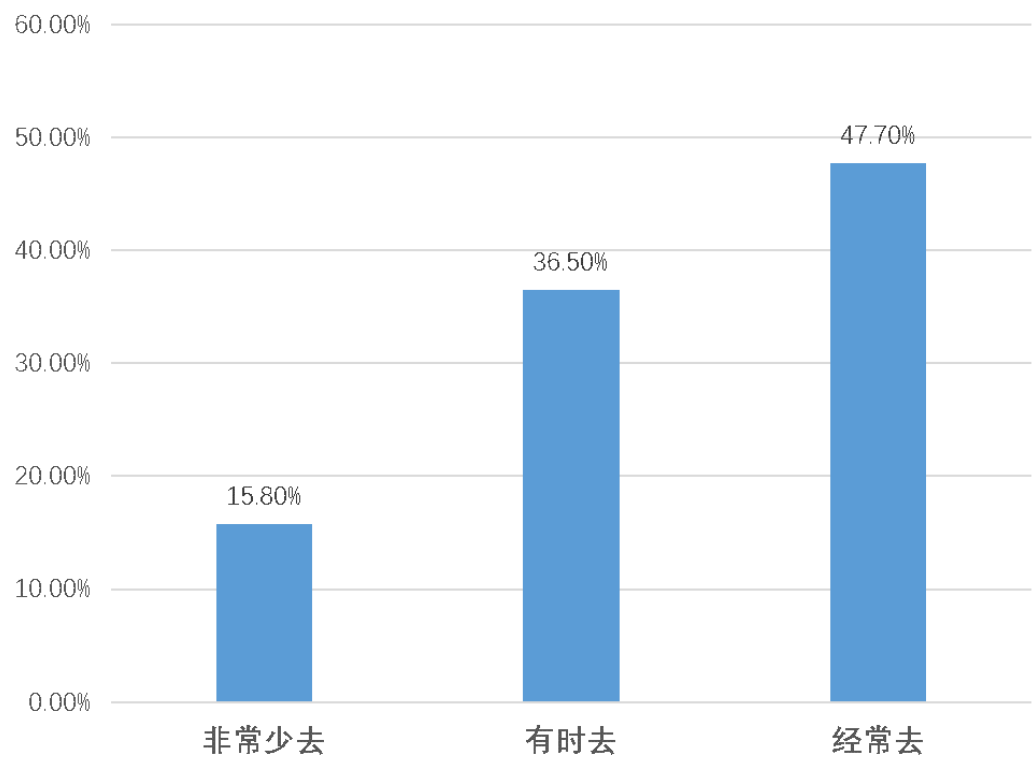


图 4.6: 前往地摊购物频率条形图

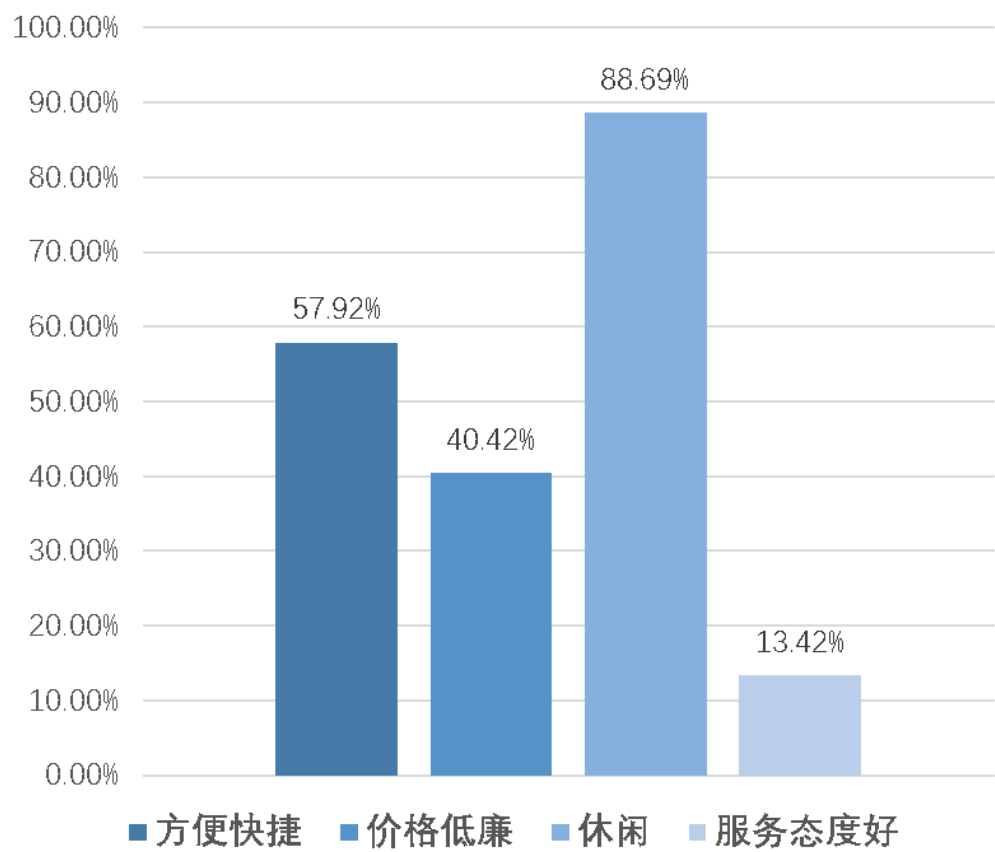


图 4.7: 前往地摊购物原因分布条形图

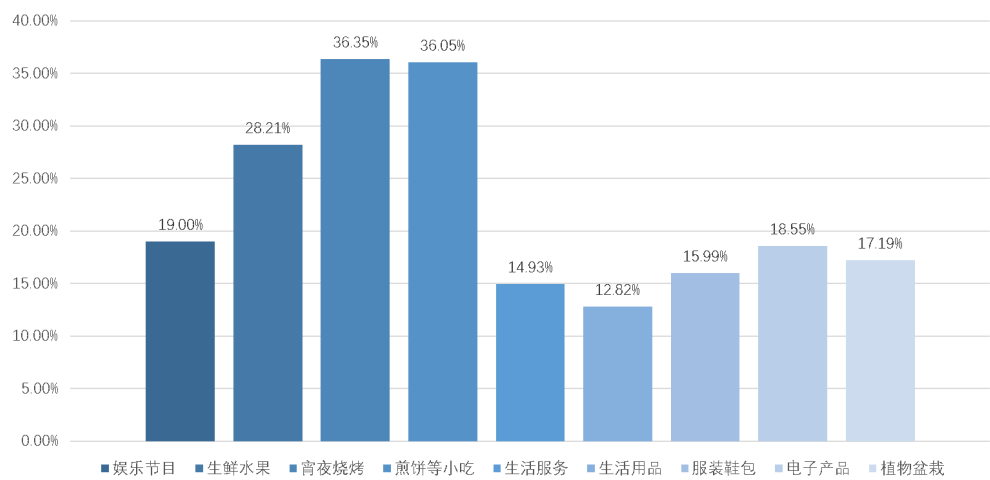


图 4.8: 购物商品偏好分布条形图

消费者前往地摊购物主要是因为在地摊购买商品方便快捷、休闲随便逛逛凑热闹及地摊商品价格低廉，还有少部分是因为摊主服务态度良好。由此看出，地摊经营户应注意服务态度及营销策略以吸引顾客。消费者前往地摊购买的商品种类繁多，其中最受消费者欢迎的是

“宵夜烧烤”、“小吃”及“生鲜水果”，除此以外，“娱乐节目（例如唱歌杂技）”、“电子产品”、“植物盆栽”也较受消费者欢迎。

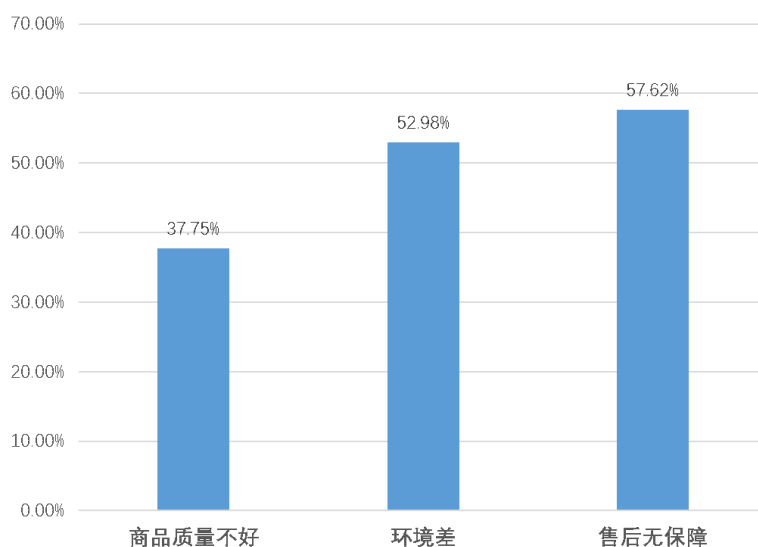


图 4.9: 不前往地摊购物原因分布条形图

而人们不前往地摊购物的主要原因是“商品售后得不到保障”、“地摊环境差”，还有一部分原因是“地摊商品质量不好”。由此可见，地摊商品的质量及售后服务需引起重视，环境脏乱差的地摊也应加以整顿。

## 4.4 相关指标重视程度分析

将消费者购物时所考虑的因素分为“商品”、“购物体验”，“便利性”三个部分，包括商品价格、商品种类丰富程度、商品质量、商贩服务态、等待时长、购物环境、是否便于停车、是否便于支付、购物整体便利性等 9 个指标。并将消费者的重视程度划分为五个等级，即“非常不重要”、“不重要”、“一般”、“重要”，和“非常重要”。计算消费者对于 9 个相关指标的重视程度均值并绘制条形图，可以看出，消费者对商品质量及购物整体便利性最为重视，其次，对等待时长、商品价格、商贩服务态度也比较重视。由此可见，应重视地摊商品的质量保障及商品价格的合理性，也应注重顾客的购物便利性。



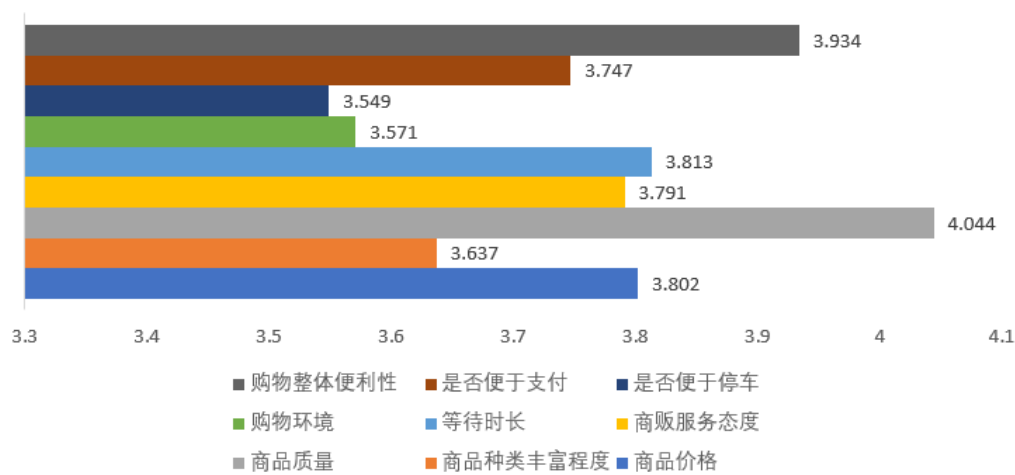


图 4.10: 相关指标重视程度条形图

4.5 消费者满意度分析

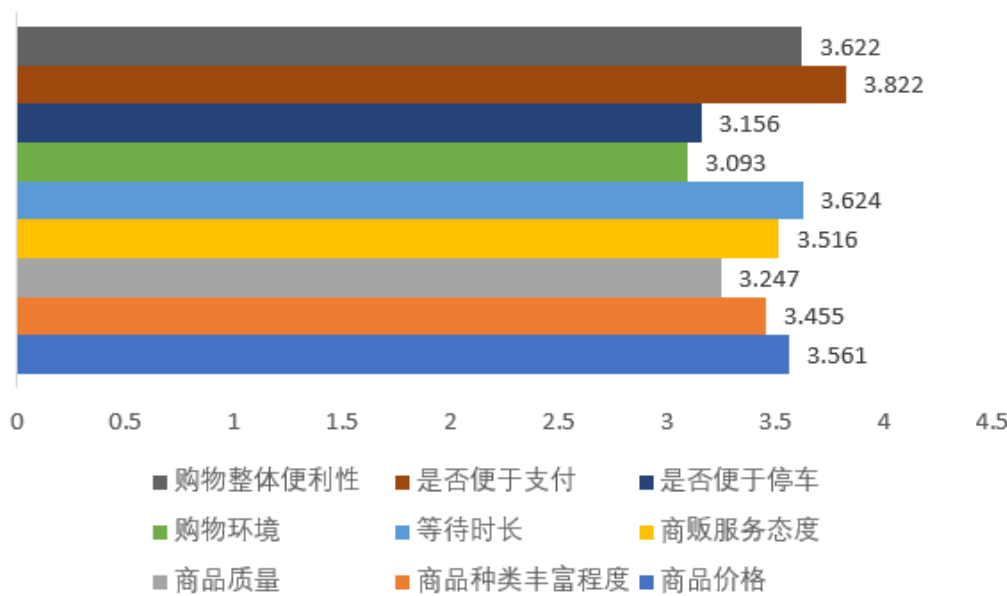


图 4.11: 满意度条形图

将消费者对的满意度划分为五个等级，即“非常不满意”、“不满意”、“一般”、“满意”，和“非常满意”。计算消费者对于 9 个相关指标的满意度均值并绘制条形图，可以看出，消费者整体满意度较好，但对购物环境、是否便于停车、及商品质量的满意度一般，地摊环境还有待改善，地摊商品的质量也应允以重视。

## 5 客户购买影响因素及特征分析

### 5.1 基于 Lasso-Logistic 模型的购买行为影响因素分析

#### 5.1.1 变量处理

将个人基础信息视作自变量,用以观察地摊经济潜在的消费者群体特征,如下表所示,共有七个自变量,分别对应问卷中的一到七题,对定性变量作哑变量的设置处理。

表 5.1: 变量设置

问题	变量设置	变量性质	变量处理
性别	Sex	定性	Sex=1, 男; 0, 女
年龄	Age	定性	Age_18=1, <18; 0, others Age18_30=1, 18 30; 0, others Age31_50=1, 31 50; 0, others NULL: >51
学历	Edu	定性	Edu_1=1, 小学; 0, others Edu_2=1, 初中; 0, others Edu_3=1, 高中; 0, others Edu_4=1, 本科及以上; 0, others NULL: 专科
是否住在商业区	Com	定性	Com=1, 是; 0, 否
家庭成员数目	Mem	定量	Mem=1,2,3,4
家庭年收入水平	Inc	定性	Inc_10=1, <10; 0, others Inc10_20=1, 10 20; 0, others Inc20_30=1, 20 30; 0, others Inc30_40=1, 30 40; 0, others NULL: >40
是否了解政策	Pol	定性	Pol=1, 是; 0, 否

将“是否会在地摊购物”的问题回答作为因变量,处理成 0-1 变量,变量名“Class”。

$$y = Class = \begin{cases} 1 \\ 0 \end{cases}$$

### 5.1.2 模型建立

Logistic 回归是一种广义的线性回归分析模型，常用于数据挖掘，疾病自动诊断，经济预测等领域，这里引入 Lasso 罚函数机制，通过构造一个惩罚函数得到一个较为精炼的模型，即筛选更为显著的自变量，并且避免过拟合的现象产生。

建立如下的 lasso-logistic 模型：

$$\begin{aligned}
 y_i = & \beta_0 + \beta_1 male + \beta_2 age_{18} + \beta_3 age_{18 \sim 30} + \beta_4 age_{31 \sim 50} \\
 & + \beta_5 edu_1 + \beta_6 edu_2 + \beta_7 edu_3 + \beta_8 edu_4 + \beta_9 com + \beta_{10} mem \\
 & + \beta_{11} inc_{10} + \beta_{12} inc_{10 \sim 20} + \beta_{13} inc_{20 \sim 30} \\
 & + \beta_{14} inc_{30 \sim 40} + \beta_{15} pol + \mu_i
 \end{aligned} \tag{5.1}$$

### 5.1.3 模型结果与分析

#### 模型结果

利用 R 语言的 glmnet 包进行结果的分析与展示，用交叉验证的方法求得最佳的  $\lambda$  值。下图中上横轴表示变量个数，下横轴表示  $\log(\lambda)$  的取值。而虚线所对应的红点是拟合度最佳的点，这里  $lambda.min = 0.01510788$ ，最佳变量个数约是 6 个。

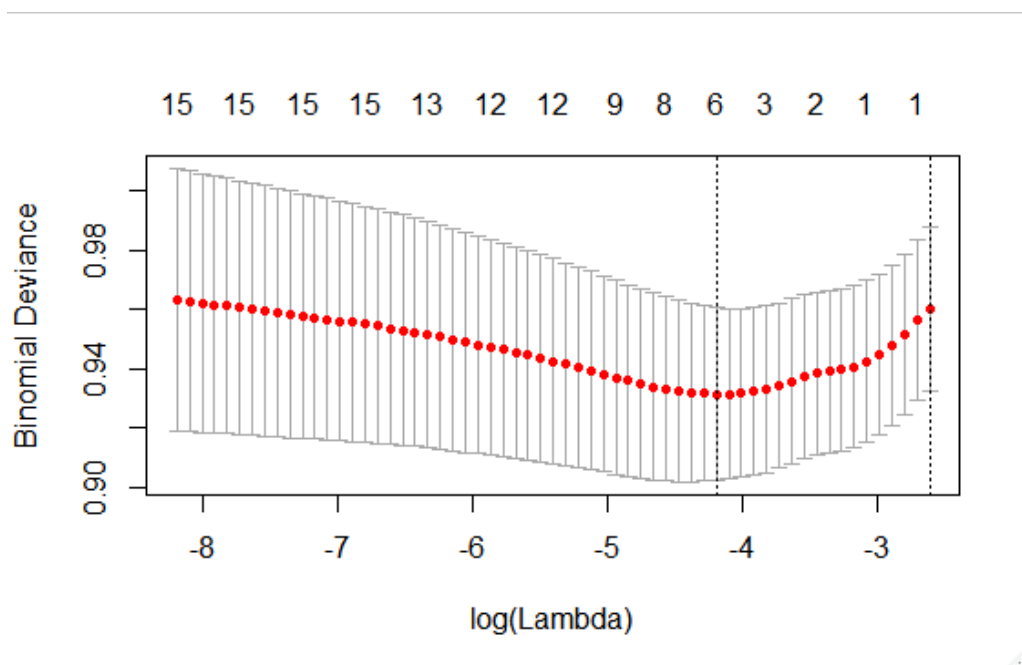


图 5.1:  $\log(\lambda)$  变化图

用  $lambda.min$  作为参数代入建构的 lasso-logistic 模型，所求得的 lasso 解系图如下??所示。图中也展现了各参数估计值随筛选变量个数不同而带来的变化。

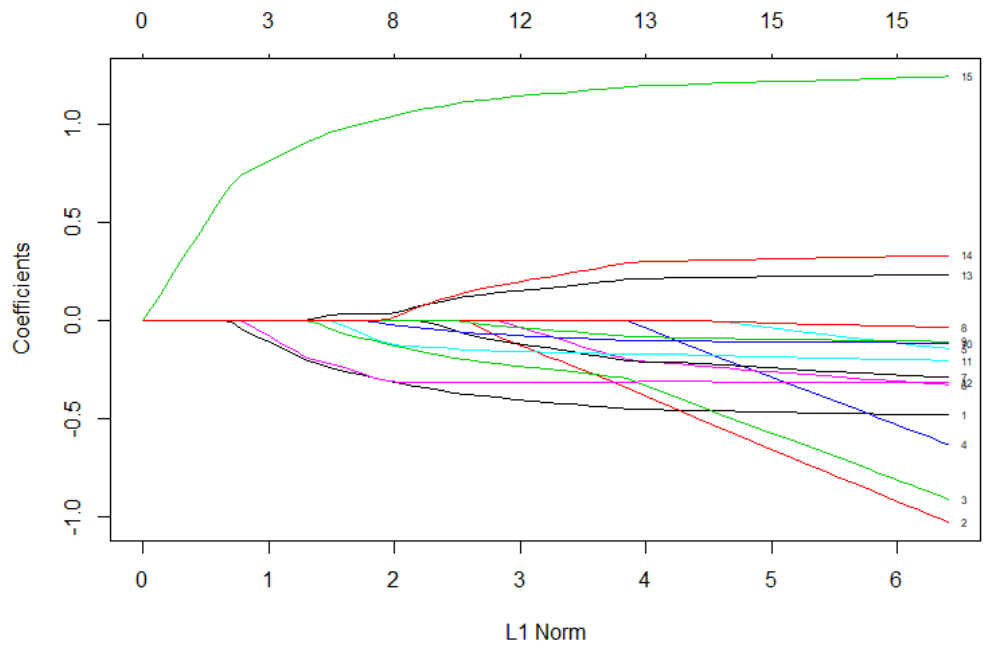


图 5.2: Lasso 解析图

lasso 最终选择的变量有 *sex*、*age*<sub>18~30</sub>、*inc*<sub>1</sub>、*inc*<sub>2</sub>、*inc*<sub>3</sub>、*Pol*，对以上变量进行逻辑回归，得到估计值如下：

表 5.2: 变量设置

变量名称	估计值	OR
(Intercept)	0.8733	2.394801
Sex	-0.25994	0.7711
Age18_30	-0.06566	0.936446
Inc_10	-0.02677	0.973581
Inc10_20	-0.24277	0.784451
Inc20_30	0.032922	1.03347
Pol	0.979528	2.663199

模型结果分析

- **性别：**参数估计值的符号为负，且变量 *Sex* 通过显著性检验，说明男性与女性间的购买倾向有显著差异，且男性在地摊购买物品的可能性是女性的 77.1% 左右。可以考虑进一步挖掘男性的潜在客户，并且继续发展女性客户市场。
- **年龄：**只有代表年龄段 18-30 的变量 *Age*<sub>18\_30</sub> 被 Lasso 模型选择，说明年龄大小与购买意愿之间不成一致的倾向性。仅代表 18-30 岁的客户购买意愿要稍微弱于 50 岁以上的客户。

- **教育水平**：没有变量通过显著性检验，说明不同教育水平的人没有明显的购物意愿差别。
- **收入**：因为以年收入大于 40 万的人群作为基准，Lasso 模型中变量  $Inc_{10}$  的参数估计值符号为负，说明年收入低于 20 万的人群对地摊商品的购买可能性会低于高收入人群；而变量  $Inc_{20-30}$  的符号为正，说明年收入 20 万到 30 万的人群购买意愿较高。因此，收入的增减并没有与购买倾向保持一致性，而是因具体水平而带来不同变化。这样的结果并不是说收入越高，购买地摊商品的可能性就越大。比较可能的解释为，当年收入很高或很低时，都不大可能成为地摊经济的目标客户。而收入位于中等偏低或中等水平时，购买意愿较大。这种现象需要综合年龄等因素来说明。总体来说，可以认为中等的人群为更可能的购买对象。
- **政策**：在所有的变量当中， $pol$  最为显著，而从参数估计值看，听说过帮扶“地摊经济”相关政策的比没有听过政策的人购买几率上升了 166.32% 左右，是很显著的提升，足以说明落实政策、宣传政策对进一步发展“地摊经济”，促进商贩就业是很有效的，也可以说明大众对帮扶“地摊经济”的政策是保持关心并且支持的态度。

## 5.2 基于模糊综合评价的客户满意与重视程度分析

为了找到消费者在地摊购物时的关注点，我们运用模糊综合判别分析量表从满意度和重视程度两个角度进行分析，找出影响消费者地摊购物行为的因素，从而为商家提供建议。

### 5.2.1 量表数据说明

我们采用李克特量表的形式，将消费者在购物时所考虑的因素分为“商品”、“购物体验”、“便利性”三个部分，共 9 个问题。将消费者的满意程度或重视程度划分为五个等级，即“非常不重要”、“不重要”、“一般”、“重要”和“非常重要”。编码时，分别对应“1”，“2”，“3”，“4”，“5”。对量表的描述性统计分析如附录所示，可以看出，消费者对于支付的便利性、购物整体便利性、等待时长非常满意，对商贩服务态度，商品价格比较满意。

### 5.2.2 地摊购物满意程度分析

#### 熵权法

熵值（Entropy）是一种物理计量单位；熵越大说明数据越混乱，携带的信息越少，效用值越小，因而权重也越小。熵值法则是结合熵值提供的信息值来确定权重的一种研究方法。

表 5.3: 熵值法权重计算表

指标	权重	指标	权重	指标	权重
商品价格	0.34	商贩服务态度	0.38	是否便于停车	0.44
商品种类丰富程度	0.41	等待时长	0.24	是否便于支付	0.34
商品质量	0.26	购物环境	0.38	购物整体便利性	0.22

使用熵值法对商品等总共 3 项进行权重计算, 从上表可以看出: 商品价格, 商品种类丰富程度, 商品质量总共 3 项, 它们的权重值分别是 0.336, 0.405, 0.259。并且各项间的权重相对较为均匀, 均在 0.333 附近。

### 模糊综合评价

为了进一步研究调查者对这三个方面的满意程度, 我们采取模糊综合评价的方法, 以我们考察的每个方面对应的比例矩阵作为模糊判断矩阵  $R$ , 以归一化的熵权法向量作为权重模糊子集  $A$ , 从而得出判断结果  $B$ 。

例如:

$$\begin{aligned}
 R_1 &= \begin{bmatrix} 0.01 & 0.02 & 0.42 & 0.51 & 0.04 \\ 0.01 & 0.01 & 0.56 & 0.35 & 0.07 \\ 0.02 & 0.05 & 0.62 & 0.29 & 0.02 \end{bmatrix} \\
 A_1 &= [0.34 \quad 0.41 \quad 0.26] \\
 B_1 &= A_1 \times R_1 \\
 &= [0.01 \quad 0.02 \quad 0.53 \quad 0.39 \quad 0.05]
 \end{aligned} \tag{5.2}$$

表 5.4: 满意度模糊综合评价结果

	非常不满意	不满意	一般	满意	非常满意
商品	0.013	0.023	0.528	0.389	0.048
购物体验	0.024	0.055	0.472	0.415	0.035
便利性	0.010	0.044	0.452	0.439	0.055

从表中可以看出, 消费者对地摊购物的便利性最满意, 从便利性方面包括的具体细节来看, 主要体现在支付的便利性和购物整体的便利性, 从停车的便利性所反映出来的通行便利性来看, 满意度还相对不够。消费者对购物体验的满意度紧随其后。商品排在第三位, 说明从总体上来看, 消费者在地摊购物的过程中, 商品价格、商品质量、商品种类的满意程度不及购物体验和便利性两个方面, 这说明为了吸引更多的消费者产生地摊消费的行为, 可以从商贩提供的商品上做出一定的调整和改善。

### 地摊购物满意程度分析

采用与满意度分析类似的分析方法，我们可以构建权重向量和模糊判断矩阵，对消费者在地摊购物过程中所考量的因素进行分析。

表 5.5: 重视程度模糊评价综合评价表

	非常不重视	不重视	一般	重视	非常重视
商品	0.0184	0.0547	0.2679	0.399	0.26
购物体验	0.0147	0.0623	0.2962	0.4359	0.1909
便利性	0.0259	0.0662	0.2534	0.4498	0.2048

从表中可以看出，消费者对于地摊购物过程中的对商品、购物体验、便利性三个方面表现的特征与满意度分析呈现的特征一致，消费者对于便利性的重视程度最高，根据这一指标所蕴含的具体细节，可以从交通便利方面做出改善，因为这一指标上满意度有所欠缺，但却是消费者购物过程中重视的因素。其次是购物体验。商品排在第三位，说明消费者在地摊购物时，对商品的考虑并不摆在首要位置。

## 5.3 基于用户聚类的潜在用户挖掘

为了进一步地对用户进行分析，为地摊经济市场提供合理的用户分析，我们对受访者的基础信息进行聚类，对所有的受调查者进行进一步的细分，并提供对应的营销策略。

### 5.3.1 K-means 聚类原理

$K - Means$  算法是一种经典的聚类分析算法，通过不断迭代分析种子点最近均值，对数据实行无监督学习。 $K - means$  首先会在数据中随机选取  $K$  个初始点，然后计算所有的点到该点的距离。设一点  $P_i$  目前距离种子点  $S_i$  最近，那么就将  $P_i$  算入到  $S_i$  的中。这样重复迭代下去，直到所有的点都被归为类中为止。 $K - means$  的原理非常简单，但这也使得其结果不够稳定—— $K - means$  的结果常常与选择的初始点有关。这就使得初始点以及目标类数  $K$  的选择非常重要。

### 5.3.2 聚类分析

#### 聚类因子的选择

根据前文中模型的分析结果，我们选择了性别、年龄、收入以及对政策的了解为聚类因子。另外，我们对受访者进行了期望的地毯商品价格区间调查。该变量将作为主要的分析变量，在聚类结果中分析。

对这些变量进行哑变量设置，这样对其进行聚类结果分析时，可以直接计算某一类的占比。

聚类结果

对数据进行聚类，并对定性指标计算占比，定量指标计算均值，结果如下：

表 5.6: 聚类结果表

label	gender	lower	higher	pol	class	age1	age2	age3	wage1	wage2	wage3	wage4
0	0.845	3.375	83.438	0.938	0.500	0.063	0.875	0.000	0.920	0.000	0.07	0.00
1	0.524	2.405	67.000	0.952	0.286	0.048	0.014	0.843	0.003	0.870	0.100	0.00
2	0.350	9.538	122.308	0.846	0.615	0.000	0.615	0.231	0.920	0.079	0.00	0.000
3	0.724	3.400	127.000	0.800	0.700	0.000	0.047	0.953	0.110	0.290	0.10	0.20
4	0.538	2.923	112.385	1.000	0.462	0.000	1.000	0.000	0.000	0.000	0.16	0.61

根据聚类所得到的结果，我们做出如下分析。

5.3.3 聚类结果分析

根据 *class* 变量中计算的，每一类选择在地摊购物的比例，我们将聚类得到的 5 组用户分为 5 类，并结合量表 1 选项进行分析。分类结果如下：

表 5.7: 客户个人特征分类

类型	A 类客户	B 类客户	C 类客户	D 类客户	E 类客户
性别	大多为男性	大多为女性	大多为男性	平均	平均
期望价格平均区间	3.4 ~ 127	9.54 ~ 122.31	3.38 ~ 83.44	2.92 ~ 112.38	2.40 ~ 67
年龄	30-50	18-30/30-50	大多为 18-30	18-30	30-50
年收入	10-20 万	小于 10 万	小于 10 万	大于 30 万	大于 30 万
价格重视程度	3.90	4.08	3.75	3.46	3.76
种类重视程度	3.30	4.08	3.69	3.15	3.76
商品质量重视程度	4.10	4.08	3.63	4.00	4.10
服务重视程度	3.50	3.92	3.56	3.46	4.10
等待时间重视程度	4.10	3.92	3.56	3.85	4.00
环境重视程度	3.20	3.85	3.50	3.31	3.62
停车重视程度	4.00	3.62	3.25	3.62	3.52
付款重视程度	3.30	3.92	3.38	3.54	4.00
总体便利重视程度	3.60	4.00	3.69	4.00	4.14

其中，五类用户的购买比例分别为：70%，61%，50%，46%，29%。根据这一结果，我们将这五类用户分别看作为重要潜在客户、重要发展客户、次要潜在客户、一般潜在客户以及低价值客户。根据重视度量表，对客户重视程度进行平均值分析，如下图所示：



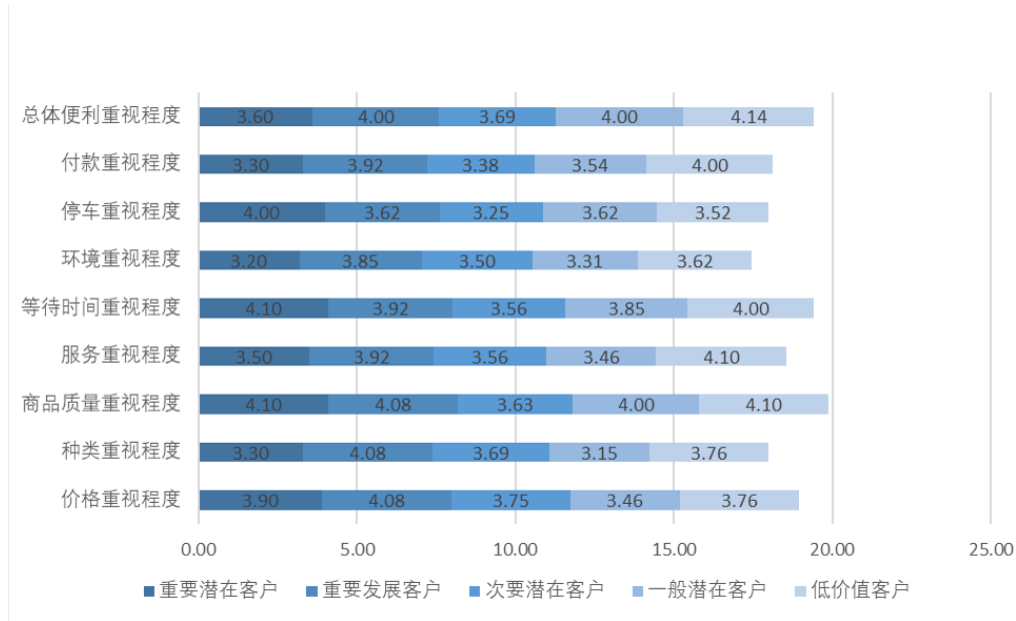


图 5.3: 分客户类型重视度打分情况

**重要潜在客户：**该类消费者为 A 类客户。此类消费者大多为男性，年龄在 30 ~ 50 岁之间，家庭年收入为 10 ~ 20 万。他们对商品的质量、购买时的等待时间以及停车的方便程度最为重视。对购物环境、付款方式的便利程度要求较低。他们预期的地毯商品价格区间为 3 元至 130 元之间。该类消费者为地摊商贩应当重点关注的对象，为了迎合该部分消费者的爱好，商贩应当确定好地摊的选址，尽量在便于停车的地区，例如综合性的步行街、广场等，有统一的停车点。除此以外，地摊商贩应当选择价格位于 130 元以下的商品，且保证商品的质量与其价格相符。

**重要发展客户：**该类消费者为 B 类客户。此类消费者大多为女性，他们的期望价格区间为 9.54 ~ 122.3。从年龄上来看，B 类用户以青年为主，且年收入小于 10 万。与 A 类客户相比，B 类客户对于地摊的要求都较高，对于大部分的因素都持较为看重的态度。其中，最看重的为物品价格、种类以及质量。最不看中的为停车的方便程度。为了迎合该部分消费者的喜好，商贩应当尽量保证商品的价格小于 122 元，并且保证商品的种类齐全，商品的性价比比较高。

**次要潜在客户：**该类消费者为 C 类客户。他们大多为男性，年龄多位于 18 至 30 岁之间，年收入小于 10 万元。他们期望的价格区间为 3.38 ~ 83.44 元。与前面的 A、B 类客户相比，C 类客户的期望价格水平明显下降，并且，他们对于地摊的要求也较低。在量表选项中，该类消费者最看重的是商品的价格、种类丰富程度以及总体便利程度，最不看中的是停车的方便程度以及付款的便利程度。这类群体的要求比较容易满足，地摊商贩可以主要发展该类客户。针对这类群体，地摊商贩们应当尽量保证总体的便利程度以及商品的性价比。

**一般潜在客户：**该类消费者为 D 类客户。他们性别占比平均，年龄多位于 18-30 岁之间，年收入为 30 万元。他们的期望价格区间为 2.92 ~ 112.38 元。该部分消费者特点突出，评分

量表展现出两级分化的态势。该类消费者最重视购物的总体便利程度和商品的质量，最不关心商品种类的多样化和购物环境。该类消费者关注的重点与第一类消费者较为相似，然而该类消费者大多为高收入的工作人群，其工作强度通常较大，很少有时间也不太愿意前往地摊购物。故该类消费者在短时间内难以发展成为实际客户。

**低价值客户：**该类消费者为 E 类客户，他们中仅有 29% 的人前往过地摊进行购物。他们性别占比平均，年龄多位于 30-50 岁之间，他们的年收入大多大于 30 万。该类客户对地摊购物的要求较高，大多数要求平均分都在 4 分以上。对该类消费者而言，他们最重视的是商品的质量、商贩的服务水平以及总体的便利度，最不重视的为停车的便利度。该类用户大多是高收入的青年人，他们普遍更适应超市购物等传统购物方式，故在这一类人里，地摊购物的市场较小。

## 6 结论与建议

### 6.1 结论

**第一，女性消费者中愿意购买地摊商品的比率更高，年长者更受地摊商品的吸引。**根据描述性统计的结果，在受调查人群中，相对于男性消费者，女性消费者中愿意购买地摊商品的比率更高。根据 Lasso-Logistic 分析的结果，男性购买地摊物品的可能性较女性而言要低 30% 左右。就年龄而言，地摊商品对年长者的吸引力更大，对 18-30 岁的购买者吸引力较弱。收入水平上，对中等收入群体的吸引力更大，对收入很高或者很低的客户吸引力都不高。

**第二，政策宣传能够很大程度地促进地摊经济市场的发展。**根据 Lasso-Logistic 分析的结果，在所有的因素中，政策变量对结果的差异影响最为显著，了解关于帮扶“地摊经济”政策的人群购买意愿是未了解过的 2.7 倍左右。根据描述性统计分析的结果，居民了解“地摊经济”帮扶政策的主要渠道为新闻媒体，但是随着互联网技术的不断发展，社交媒体，例如微博等，也成为了居民了解政策导向的主要方式之一。

**第三，消费者前往地摊消费的主要原因为地摊购物方便快捷、热闹、价格低廉，整体来说，消费者对便利性、商品质量最为重视。**根据描述性统计分析的结果，选择前往地摊购物的消费者主要因为地摊购物方便快捷，价格低廉。最受消费者欢迎的种类为宵夜烧烤、小吃等等。而选择不去地摊购物的消费者，主要担心商品受不到保障、地摊环境差。从消费者对地摊购物几大影响因素的重视程度来看，消费者最重视购物的便利性，其次则是商品的质量和性价比。从消费者对地摊产品的满意度来看，根据模糊评价的结论，消费者对购物体验 and 便利性的满意度较高，对商品本身的满意度较低。

**第四，目前而言，民众、政府对地摊经济发展的态度，规范措施的落实以及地摊经济对就业的影响为媒体最关注的问题。**根据词云聚合分析，“地摊经济”一夜爆红，促进了多个行业之间的交叉发展，“五菱宏光”迅速成为人们讨论的新话题，这说明，地摊经济不仅仅促进了小商品行业的发展，同时还带动了国产汽车行业的发展。除此以外，对新闻长文本的挖掘结果，体现了“地摊经济”在提供就业岗位这一方面的巨大潜力。地摊行业门槛低，摆摊并不需要创立企业那样复杂的流程，对于正处于暂时性失业中的人群来说，是一个很好的过渡期的选择。地摊经济能够在一夜之间，提供 10 万多个就业岗位，在现在就业市场较为疲软的“后疫情”时期，这样的“人间烟火”更能为市场注入活力。然而，对于地摊经济，媒体并不是完全“一边倒”的报道，也存在着一些质疑的声音。地摊经济的不规范运行，导致城市环境遭到破坏，这说明，地摊经济的发展同样需要精细规划，同样需要因地制宜。

## 6.2 建议

本文主要通过挖掘“地摊经济”的潜在消费者，尝试研究消费者的真正需求和偏好，从而扩大地摊商品对这部分潜在消费者的吸引力，扩大市场需求。在如今的“后疫情”时代，就业市场、小商品市场疲软，大量劳动力、商品处于闲置之中。本文通过研究“地摊经济”的潜在消费者，为商贩提供合理精准的消费需求定位，促进闲置劳动力选择通过地摊经济暂时性再就业，一方面，提供了众多的就业岗位，促进小微企业创业；另一方面，通过地摊经济，处理闲置的存货资源，刺激消费，让消费者享受物美价廉的商品，让商贩享受低成本的经营渠道。本文根据前面的分析结果，分别对商贩以及政府提出如下建议：

### 6.2.1 对商贩的建议

**第一，以女性、年长者，收入主要位于 10-20 万之间的消费者，为主要的目标人群。**这类消费者的市场潜在价值较高，对此，可以主要针对这类消费者，选择贩卖的商品，例如，调查中最受欢迎的小吃、水果等。

**第二，保证购物流程的便利性和商品的质量。**在所有的研究变量中，消费者最看重的是购物流程的便利性。消费者前往地摊消费，时常是在娱乐、休闲中产生的，并不具备逛超市时所拥有的强目的性，因此对商品种类等并没有表现的非常重视。这说明，地摊商贩应当尽量保证购物流程的便利，例如选址应当位于方便停车的地方，支付方式应当同时支持现金与电子支付，服务速度要足够快，尽量减少消费者的等待时间。同时，商品也需要合理定价，应注重商品的性价比，提高客户的回头率。

**第三，对客户进行分类处理，开发具有一定针对性的产品。**商贩可以按照本文中给出的五类潜在消费者的条件，对到店客户进行消费者分类，并且针对消费者所属分类提供相应的服务以及商品建议。

### 6.2.2 对政府的建议

**第一，大力发展“地摊经济”，持续刺激消费市场。**词云分析结果显示，目前，地摊经济能够为社会提供大量的就业岗位，在如今的“后疫情”时代，为了促进消费，刺激经济发展，重振就业市场，应当大力宣传“地摊经济”，鼓励年轻人以及部分失业人群走上街头就业，成为“人间烟火”的一部分，维护社会稳定发展。

**第二，规范落实措施，保护城市面貌。**根据 LDA 主题分析以及情感分析的结果，群众对于地摊经济的发展仍然抱有一种质疑的态度。其主要原因是，地摊经济在一夜之间爆红，城市并没有足够的时间来应对这样的现象。在接下来的发展中，政府应当在保护居民生活环境的同时，合理规划地摊商户的分布，例如，设立专门的地摊商业区、地摊广场等等，不能让地摊经济从“升温”变成“发烧”。

**第三，加大宣传力度，普及经营知识。**根据调查结果，政府合理的政策宣传，能够进一步加强消费者的购买意愿。政府应当通过新闻报道、社区宣传等方式，进一步为民众普及相关知识。同时，地摊经济虽然是一个门槛较低的行业，但仍然存在一些风险，一些经营者在未做足充分准备和风险意识薄弱的情况下，仍然选择进入市场，他们很可能面临着血本无归的结局。针对这些商贩，政府部门不能选择只放任不扶持，否则将会导致商贩群体虽大，当个体利润却很小的低效率运行。

## 参考文献

- [1] 李娟李坤梁. 基于完全竞争视角下地摊经济管理研究——以贵州省凤冈县为例. 现代商业, pages 68–69, 2018(27).
- [2] 杨满珍. 议地摊经济对城市发展的影响——以楼底城区为例. 北方经贸, pages 51–53, 2015(9).
- [3] 王岩. 城市精细化管理如何包容地摊经济——基于上海中心城区的实证调查研究. 中国发展观察, pages 47–51, 2018(13).
- [4] 赵佳宁. 安宁区地摊经济发展研究. 青春岁月, pages 458,457, 2016(12).
- [5] 魏博通. 黄石地摊经济发展现状及对策研究. 商场现代化, pages 46–47, 2013(29).

## 附录 A 量表的描述性统计分析

### A.1 相关指标的重视程度

一级指标	二级指标	1(%)	2(%)	3(%)	4(%)	5(%)	均值
商品	商品价格	2.2	3.3	31.87	37.36	25.27	3.802
	商品种类丰富程度	2.2	9.89	29.67	38.46	19.78	3.637
	商品质量	1.1	3.3	18.68	43.96	32.97	4.044
购物体验	商贩服务态度	2.2	6.59	24.18	43.96	23.08	3.791
	等待时长	1.1	1.1	31.87	47.25	18.68	3.813
	购物环境	1.1	10.99	32.97	39.56	15.38	3.571
便利性	是否便于停车	4.4	8.79	31.87	37.36	17.58	3.549
	是否便于支付	2.2	7.69	24.18	45.05	20.88	3.747
	购物整体便利性	1.1	3.3	19.78	52.75	23.08	3.934

注：“1”表示非常不重视，“2”表示不重视，“3”表示一般，“4”表示重视，“5”表示非常重视

## A.2 满意度

一级指标	二级指标	1(%)	2(%)	3(%)	4(%)	5(%)	均值
商品	商品价格	0.74	1.97	42.01	50.98	4.3	3.561
	商品种类丰富程度	1.11	1.11	55.77	35.26	6.76	3.455
	商品质量	2.09	4.55	62.16	28.99	2.21	3.247
购物体验	商贩服务态度	1.97	4.42	38.08	51.11	4.42	3.516
	等待时长	0	2	38	55.56	4.44	3.624
	购物环境	4.3	8.85	62.04	22.85	1.97	3.093
便利性	是否便于停车	2.22	8.89	62.22	24.44	2.22	3.156
	是否便于支付	0	0	26.67	64.44	8.89	3.822
	购物整体便利性	0	2.22	40	51.11	6.67	3.622

注：“1”表示非常不满意，“2”表示不满意，“3”表示一般，“4”表示满意，“5”表示非常满意



## 附录 B 调查问卷

### 地摊经济消费市场及发展前景调查问卷

尊敬的朋友：您好！我们是中南财经政法大学的学生，为了参加“第十届全国大学生市场调查与分析大赛”，希望了解您对于地摊经济的看法与建议，非常感谢您能够接受此次调查活动，本次问卷采用不记名的形式，将对您提供的信息进行严格保密，所有回答仅用于学术研究，没有对错之分，您只需根据自己的实际情况填写本问卷。衷心感谢您的支持与配合！

1. 您的性别是：  
A. 男 B. 女
2. 您的年龄是：  
A. 18 岁以下 B. 18-30 岁 C. 31-50 岁 D. 51-65 岁 E. 65 岁以上
3. 您的学历是：  
A. 小学 B. 初中 C. 高中 D. 专科 E. 本科及以上 F. 其他
4. 您是否居住于商业区：  
A. 是 B. 否
5. 您的家庭成员数目：  
A. 1 B. 2 C. 3 D. 4 E. 其他
6. 您目前的家庭年收入水平为：  
A. 10 万以下 B. 10-20 万 C. 20-30 万 D. 30-40 万 E. 40 万以上
7. 您对地摊经济政策是否有了解：  
A. 有（跳至第 8 题） B. 无（跳至第 9 题）
8. 您的了解渠道：  
A. 新闻媒体 B. 社交媒体 C. 亲朋好友告知 D. 其他
9. 您是否去地摊购物：  
A. 是（请跳至第 10 题） B. 否（请跳至第 11 题）
10. 您去地摊购物的频率（填写完该题，请跳至第 12 题）：  
A. 偶尔 B. 有时去 C. 经常去

11. 您不去地摊购物的原因：[多选题]  
A. 地摊商品质量不好 B. 地摊环境差 C. 商品售后得不到保障 D. 其他
12. 您会选择在地摊购物的原因：[多选题]  
A 方便快捷 B 价格低廉 C 散心时顺便逛逛 D 凑凑热闹 E 喜欢摊主服务态度
13. 您在地摊的消费倾向：[多选题]  
A 生活服务（例如手机贴膜） B 娱乐节目（例如唱歌杂技） C 生鲜水果 D 宵夜烧烤  
E 一般小吃 F 植物盆栽 G 电子产品 H 服装鞋包 J 生活用品
14. 您是否看重商品品牌、是否追求潮流：  
A 不看重 B 一般看重 C 比较看重 D 非常看重
15. 您认为地摊经济对社会发展带来的好处是：[多选题]  
A 缓解就业压力 B 增加收入，减轻家庭生活负担 C 能够拉动促进居民消费  
D 方便居民购物，减少购物成本 E 其他
16. 您最担心地摊市场，马路市场存在哪些问题：[多选题]  
A 商品质量无法保证，消费者权益得不到保障  
B 市场缺乏规范管理，市场秩序混乱  
C 影响市容环境，脏乱差  
D 阻碍交通，影响居民生活  
E 占用消防通道，盲道  
F 其他（注明）
17. 请问您对地摊商品的期望价格为：  
最低（单位：元）：  
最高（单位：元）：
18. 请您对地摊购物时，您对下列因素的重视程度进行打分：

	非常不重视	不重视	一般	重视	非常重视
商品价格					
商品种类丰富程度					
商品质量					
商贩服务态度					
等待时长					
购物环境					
是否便于停车					
是否便于支付					
购物整体便利性					

19. 请您对地摊购物时，您对下列因素的满意程度进行打分（本题仅限去过地摊进行购物的受调查者回答）：

	非常不满意	不满意	一般	满意	非常满意
商品价格					
商品种类丰富程度					
商品质量					
商贩服务态度					
等待时长					
购物环境					
是否便于停车					
是否便于支付					
购物整体便利性					