**食之韵味，预制相随**

**——以成都市为例的预制菜消费现状发展以及策略研究**

目录

[一、 绪论 1](#_Toc160377981)

[（一） 预制菜的概念 1](#_Toc160377982)

[（二） 调查背景 1](#_Toc160377983)

[（三） 调查意义 3](#_Toc160377984)

[（四） 文献综述 3](#_Toc160377985)

[（五） 研究思路 5](#_Toc160377986)

[（六） 特色与创新 6](#_Toc160377987)

[二、 调查方案设计 6](#_Toc160377988)

[（一） 调查设计架构 6](#_Toc160377989)

[（二） 调查目的与调查任务 6](#_Toc160377990)

[（三） 调查范围及调查对象 7](#_Toc160377991)

[（四） 调查方式和方法 8](#_Toc160377992)

[（五） 抽样方案设计 9](#_Toc160377993)

[（六） 调查项目设计 12](#_Toc160377994)

[三、 调查实施 12](#_Toc160377995)

[（一） 调查实施过程 12](#_Toc160377996)

[四、 数据处理与检验 14](#_Toc160377997)

[（一） 预调查数据检验 14](#_Toc160377998)

[（二） 正式调查数据检验 18](#_Toc160377999)

[五、 基于电商平台评论数据的文本挖掘 21](#_Toc160378000)

[（一） 预制菜评价数据爬取 21](#_Toc160378001)

[（二） 数据分析 22](#_Toc160378002)

[（三） 关键词数据分析 24](#_Toc160378003)

[（四） 情感倾向分析 24](#_Toc160378004)

[（五） 主题分析 25](#_Toc160378005)

[六、 消费者特征描述性统计 27](#_Toc160378006)

[（一） 预制菜消费者特征分析 27](#_Toc160378007)

[（二） 消费者对预制菜的偏好分析 32](#_Toc160378008)

[（三） 基于 K-means 聚类分析的消费群体细分 36](#_Toc160378009)

[（四） 非预制菜消费者分析 38](#_Toc160378010)

[七、 预制菜消费者潜在客户的挖掘 41](#_Toc160378011)

[（一） 数学建模思路 41](#_Toc160378012)

[（二） 模型介绍 42](#_Toc160378013)

[（三） 变量的检验与整合 43](#_Toc160378014)

[（四） 模型建立与修正 47](#_Toc160378015)

[八、 对预制菜行业的结论与建议 55](#_Toc160378016)

[（一） 主要结论 55](#_Toc160378017)

[（二） 提出建议 57](#_Toc160378018)

[**参考文献** 61](#_Toc160378019)

图目录

[图 1 调查设计架构 6](#_Toc160377520)

[图 2 项目开展甘特图 13](#_Toc160377521)

[图 3 人员调查分工 14](#_Toc160377522)

[图 4 总方差解释图 19](#_Toc160377523)

[图 5 因子旋转后的成分矩阵表结果图 20](#_Toc160377524)

[图 6 京东商城评价示意图 21](#_Toc160377525)

[图 7 词云图 24](#_Toc160377526)

[图 8 情感倾向分析图 25](#_Toc160377527)

[图 9 LDA模型图 26](#_Toc160377528)

[图 10 消费者预制菜购买与未购买占比 28](#_Toc160377529)

[图 11 消费者性别分布情况 28](#_Toc160377530)

[图 12 消费者年龄分布情况 29](#_Toc160377531)

[图 13 消费者婚姻状况分析 29](#_Toc160377532)

[图 14 消费者月收入分布 30](#_Toc160377533)

[图 15 消费者职业分析 31](#_Toc160377534)

[图 16 消费者日常饮食在收入中的占比 31](#_Toc160377535)

[图 17 消费者食品需求多元化情况 32](#_Toc160377536)

[图 18 预制菜购买频率 33](#_Toc160377537)

[图 19 消费者对各因素的重视程度 33](#_Toc160377538)

[图 20 消费者购买的预制菜类型占比 34](#_Toc160377539)

[图 21 消费者购买预制菜的主要原因 34](#_Toc160377540)

[图 22 消费者能接受的预制菜价格分布 36](#_Toc160377541)

[图 23 非预制菜消费者性别分布 39](#_Toc160377542)

[图 24 非预制菜消费者月收入分布 39](#_Toc160377543)

[图 25 非预制菜消费者收入分布 40](#_Toc160377544)

[图 26 限制购买预制菜的因素分析 40](#_Toc160377545)

[图 27 吸引非预制菜消费者消费的方式 41](#_Toc160377546)

[图 28 碎石图（提取方法：主成分分析法） 45](#_Toc160377547)

[图 29 10折交叉验证图 50](#_Toc160377548)

[图 30 混淆矩阵 51](#_Toc160377549)

[图 31 各个特征的重要性 53](#_Toc160377550)

表目录

[表 1 预制菜分类 1](#_Toc160377551)

[表 2 2022年至2023年2月期间各地重要政策、标准、规划 1](#_Toc160377552)

[表 3 初级抽样单元 10](#_Toc160377553)

[表 4 整体抽样框 11](#_Toc160377554)

[表 5 时间安排表 13](#_Toc160377555)

[表 6 第8题信度结果 15](#_Toc160377556)

[表 7 第14题信度结果 15](#_Toc160377557)

[表 8 KMO 和巴特利特检验 16](#_Toc160377558)

[表 9 量表8信度结果 18](#_Toc160377559)

[表 10 量表14信度结果 18](#_Toc160377560)

[表 11 正式调查 KMO 和 Bartlett 19](#_Toc160377561)

[表 12京东部分用户评论爬取结果 22](#_Toc160377562)

[表 13 积极情感词汇 26](#_Toc160377563)

[表 14 中性情感词 27](#_Toc160377564)

[表 15 问题整合情况 37](#_Toc160377565)

[表 16 消费者购买意愿聚类情况 37](#_Toc160377566)

[表 17 相关性表 44](#_Toc160377567)

[表 18 因子分析总方差解释表 44](#_Toc160377568)

[表 19 成分得分系数矩阵 45](#_Toc160377569)

[表 20 皮尔逊相关分析 49](#_Toc160377570)

[表 21 模型关键参数 51](#_Toc160377571)

[表 22 准确率、精确率、召回率、AUC值 53](#_Toc160377572)

# 绪论

## 预制菜的概念

2022年6月2日，中国烹饪协会发布标准标号为 T/CCA 024-2022《预制菜》团体标准，正式明确了预制菜的定义。预制菜是指以一种或多种农产品为主要原料，经过标准化、规模化以及流水化作业，预加工（如分切、搅拌、腌制、滚揉、成型、调味等）和预烹调（如炒、炸、烤、煮、蒸等）制成，并进行预包装的半成品或者成品菜肴。

预制菜根据对复热或烹饪的不同要求，可分为即食食品、即热食品、即烹食品、即配食品四大类。详见表1

表 1 预制菜分类

|  |  |
| --- | --- |
| 即食食品 | 开袋即食，无需加工或复热，如豆腐干、牛肉干、罐头等 |
| 即热食品 | 需要加热才能食用的产品，如速冻点心、水饺、火锅料（肉丸等）、方便米饭、方便面等 |
| 即烹食品 | 需要经过简单的烹饪即可出品的菜式，比如腌制牛排、鱼香肉丝等，将调味包与主料混合，简单烹饪，主要是根据自己的喜好控制加入调味包的品种和数量 |
| 即配食品 | 按份分装的净菜、净肉及其辅料，如煲汤组合包、小炒菜组合包等 |

## 调查背景

#### 1．政策背景**：**各地相继出台政策指导文件，持续加码支持预制菜产业发展

2023年中央一号文件《关于做好2023年全面推进乡村振兴重点工作的意见》提出 “提升净菜、中央厨房等产业标准化和规范化水平，培育发展预制菜产业”，这是预制菜首次被写入中央一号文件。预制菜产业一头连着田间地头，一头连着消费者餐桌，其快速发展既顺应了需求侧的变化，满足了消费者对美食的多元需求，也推动了农村一二三产业的融合发展，为助力乡村产业振兴、实现共同富裕发挥重要作用。2022年以来，广东、山东、四川、重庆、福建、浙江、江西、上海、辽宁等多个省市相继出台预制菜相关产业的政策指导文件。

表 2 2022年至2023年2月期间各地重要政策、标准、规划

|  |  |  |
| --- | --- | --- |
| 时间 | 文件标题 | 省份 |
| 2022.3.25 | 《加快推进广东预制菜产业高质量发展十条措施》 | 广东省 |
| 2022.9.30 | 《云南省预制菜产业发展实施意见》 | 云南省 |
| 2022.10.26 | 河南省人民政府办公厅关于印发河南省加快预制菜产业发展行动法案（2022-2025年）的通知 | 河南省 |
| 2022.11.10 | 《加快推进预制菜产业高质量发展的措施》 | 福建省 |
| 2022.11.20 | 《关于推进全省预制菜产业高质量发展的意见》 | 山东省 |
| 2023.1.11 | 关于印发《支持预制菜产业高质量发展的若干措施》 | 四川省 |
| 2023.2.13 | 中共中央《国务院关于做好2023年全面推进乡村振兴重点工作的意见》 | 国务院 |
| 2023.2.16 | 《预制川菜分类和术语》、《预制川菜规范与编制指南》、《预制川菜质量追溯规范要求》等 | 四川省 |

#### **2．**行业背景：市场规模稳步增长，企业数量不断增加，产业链趋于完善

近三年来中国预制菜市场规模稳步增长，市场发展潜力大。数据显示，2022年中国预制菜市场规模达4196亿元，同比增长21.3%。预计未来3-5年，中国预制菜市场规模有望以20%左右的高增长率逐年上升，在2023年达到5165亿元，在2026年达10720亿，我国预制菜产业有望发展成下一个万亿级市场。

在政策红利和市场需求的驱动下，我国预制菜市场扩容趋势明显，预制菜相关企业数量也在快速攀升。数据显示，截至2022年12月29日，我国现有预制菜相关企业7.59万家。2011-2020年，预制菜企业注册数量逐年递增。众多企业纷纷加入预制菜市场，预制菜赛道火热，行业融资数量也呈爆发增长之势。

在产业链方面，已形成相对完整的产业链。上游为原材料供应型行业，主要包括农、牧、渔等种养殖业，米面粮油等农副产品初加工业以及食品包装等包装供应业。中游为预制菜生产和加工业，主要负责净菜、半成品菜的生产加工，包括专业预制菜加工企业、传统速冻食品企业和餐饮企业等。产业链下游是预制菜的消费市场，基本分为B端和C端两大销售渠道。B端销售主要面向连锁餐厅、菜品批发市场、乡厨、酒店等，C端销售渠道是指食品企业或者餐饮企业在工厂完成预制菜加工后，通过餐饮门店、生鲜电商平台、商超等渠道销售给顾客。

#### **3.**市场背景：社会快节奏发展的必然产物，C端市场快速增长

预制菜并非是一种全新的食品品类，而是工业化程度较高，可“即食、即热、即烹和即配”的菜肴。在当今生活节奏不断加快的信息时代，“宅家文化”、“懒人经济”、“烹饪小白”、“单身群体”等理念盛行，预制菜除去了烹饪过程的繁琐、省时省力的优点使之得到年轻消费者的接受，尤其是疫情期间，预制菜满足了消费者在特殊时期对食品的方便、美味、营养的诉求。同时，在消费升级背景下，长时间单一品类、单一菜品已经无法满足市场多样化的需求，多元化的预制菜则能满足消费者多样化的需求。

另一方面，目前我国预制菜市场仍以B端为主，数据显示，我国预制菜行业最大的需求来自餐饮行业，有85%以上的预制菜产品销售至B端，B端与C端占比约为8:2。2020年以来，随着餐饮外卖行业的蓬勃发展和因疫情催生的家庭端消费需求变化，预制菜C端迎来消费加速期，销量大幅提升，预制菜从餐饮后厨逐步走入家庭餐桌。随着时间的推移，后续C端市场有望加快增长，但预计相当长时间内我国预制菜消费市场仍会以B端为主。

## 调查意义

本报告通过调查了解消费者关于预制菜的消费观念，分析消费者消费行为影响因素，并对收集到的数据进行分析，可以帮助企业直观了解预制菜行业痛点，从而探寻“懒宅经济”和后疫情时代背景下预制菜发展道路，具有现实意义。

当前，我国预制菜行业渗透率和市场接受度仍远低于日本和欧美等国家，中国预制菜市场发展潜力巨大。报告通过文献调查、问卷调查、深度访谈等方法，从消费者角度，探寻消费观念、行为影响因素、预制菜了解程度，最后通过数据分析，数据挖掘，定性分析与定量分析相结合等方式，对消费现状、预制菜行业痛点等进行了研究，并给出对策建议，对丰富预制菜理论研究具有重要意义。

## 文献综述

#### **1.**国外研究现状

在20世纪40年代，预制菜诞生于美国。80年代后，传入日本、欧洲和中国台湾地区，并在这些地区日渐兴起和流行。目前，欧美及日本在预制菜工艺、标准和市场等方面已形成较为完善和成熟的体系。预制菜早在1976年就出现在美国俄克拉荷马州立大学Tapp的研究论文中[3], 之后美国学者将其阐述为“烹制菜肴被预制为可即食的产品”[4]。

欧美国家早期对预制菜的研究主要针对原料及其处理对产品感官品质的影响、工业化配方和工艺与传统菜肴营养的差异及其评估等。之后，细化到预制菜污染防控、调味添加料的检测、营养价值评估、预制菜加工灭菌技术、以及工业化预制菜与传统菜对大众消费的影响等。近年来，随着欧美国家预制菜市场渗透率提高、技术和设备水平不断完善，其研究也更侧重于食品安全体系建设完善以及新技术在提升预制菜产品品质和安全性的应用等[5]。

#### 2.国内研究现状

相较于欧美和日本，王卫、张锐、张佳敏等在《预制菜及其研究现状、存在问题和发展展望》（2022 年）中指出我国预制菜行业尚处于发展初期，存在许多亟待解决的问题[6]。国内关于预制菜的研究也主要集中在预制菜行业发展问题及其发展对策、新技术发展及应用等方面。

在预制菜行业发展方面，王娟，高群玉，娄文勇提出，预制菜技术难点主要有食品安全、感官品质、营养品质与保健功能、原材料质量控制及溯源体系、速冻及解冻技术、包装材料与技术、智能装备成套设备等方面，标准化、细分市场和高质量专业技术人才的培养与输送，将成为预制菜行业的发展方向[7]。王卫、张锐、张佳敏等也同样提出存在概念模糊、技术和设备落后、标准不配套、安全隐患突出等问题[6]。

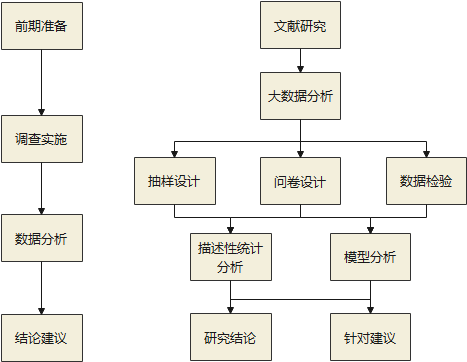
在四川预制菜发展方面，作为在中国八大菜系中占有重要地位的川菜，四川预制菜是中式预制菜的重要发展方向之一。张宇昊、陈海指出，目前川渝预制菜品质形成机制不清，且以高盐、高油、麻辣和重味的川渝预制菜存在一些不可回避的潜在健康风险,因此需要加快川渝菜肴品质形成理论体系构建，进一步研发品质保真技术和健康技术[8]。王绪模,朱克永也给出了类似的发展建议，并且指出近年来90后成为预制菜的主流消费群体，预制菜迎合了城市快节奏生活下无暇下厨的年轻消费者的生活习惯，应尽快加强全省川菜预制菜产业发展战略研究，引导川菜预制菜产业协同整合、强化产业体系建设[9]。

在预制菜技术发展方面，国内早期对预制菜的研究是基于调理肉制品、食品工业化，提出中国传统食品现代化的发展路径。张国宝，王静对我国存在的问题和工业化途径进行了详细的阐述[10]。王卫、赵佳敏等对川菜肉类菜肴工业化及其关键技术进行剖析[11]。近年来，与预制菜一同发展起来的还有保鲜冷藏、冷链物流、无菌包装等技术。税小林、孙钦秀、夏秋瑜等在《预制菜包装技术的发展概况及趋势》中提出了新型的食品包装技术在预制菜领域具有广阔的发展空间，并对预制菜各类包装材料进行了分析[12]。肖欢,曹宏等提出辐照技术在预制菜卫生安全方面的应用[13]。吴晓蒙,饶雷,张洪等详细说明了新型杀菌技术、快速冷却与冷冻技术、冷链物流、智能包装等新型食品加工技术在提升预制菜肴质量与安全的重要应用[14]。

#### 3.文献综述

由此可以看出，相较于欧美日本等国家，我国预制菜产业有渗透率低、理论体系不完善、技术水平有待提高等一系列问题，但同时随着生活节奏的加快，宅家文化、懒人经济、烹饪小白、单身群体等不断扩容，推动了预制菜行业的发展。在B端厨房工业化创新降成本、提效率的要求和C端生活方式变革，以及需求新鲜、健康饮食习惯养成和新零售模式三重共振下，预制菜市场快速增长。未来，我国预制菜行业将不断向多元化、安全化、高质量化方向发展。要善于借鉴外国经验，结合我国实际，推进预制菜行业发展。

## 研究思路



## 特色与创新

#### 1.选题新

预制菜是近年来热门话题，可谓发展迅猛。而近两年，预制菜又从B端火到C 端。特别是疫情期间，消费者对预制菜产品需求增多，各大电商、生鲜超市预制菜销量成倍增长，C 端预制菜产品接受度提升，C端预制菜发展潜力巨大。本报告从消费者视角进行调研分析，选题新颖，结合时代热点以及当前消费趋势。

#### 2.基于数据挖掘

在调查过程中，运用数据挖掘，从而了解消费者对预制菜的评价， 进而分析预制菜的优缺点，使调查研究更真实有效。

# 调查方案设计

## 调查设计架构



图 1 调查设计架构

## 调查目的与调查任务

#### 1、调查目的

（1）了解消费者关于预制菜消费观念的现状：探寻消费者关于预制菜消费观念的基本情况和特点，寻找推进预制菜消费的突破口；

（2）了解消费者预制菜消费行为影响因素：探寻消费者对于预制菜消费行为的影响因素，从而进一步分析消费者的预制菜消费行为；

（3）了解预制菜行业痛点：通过网上爬取数据以及问卷调查结果，调查预制菜的市场反响，精确狙击当前市场上预制食品行业的痛点；

（4）探寻预制菜发展道路：在“懒宅经济”和国家各类政策逐步完善的背景下，探究推动预制食品进一步发展的同时提高预制菜消费水平的创新策略与模式。

#### 2、调查任务

本文通过调查预制菜行业在“懒宅经济”时代以及食品安全问题、消费者权益问题横行背景下的消费现状和影响因素，通过实地调查考察发放问卷和网络问卷的方式向居民了解以下内容：

（1）搜集受访者个人基本信息，如年龄、性别、月收入水平、在家做饭频率等信息，旨在探索受访者基本信息对预制菜认知情况和消费行为的关系；

（2）调查分析受访者在新时代背景下食品消费观念转变的情况，进而分析预制菜应在哪些方面进行改进和完善；

（3）调查分析受访者选择购买预制菜以及限制其购买预制菜的原因，针对性地提出建议，为促进我国预制菜健康发展以及提高预制菜消费水平提供有效参考；

（4）调查分析受访者更偏好哪种宣传方式，从而推动完善当前时代下预制菜的宣传模式。

## 调查范围及调查对象

#### 1、调研地点

调研在四川省省会城市成都市展开。

#### 2、调查范围

本次研究首先在四川省统计局网站上查找相关资料，收集到成都市各主城区人口数量和各城区社区数量，设计抽样调查方式。

调查覆盖范围为成都市14（“12＋2”）个主城区，即锦江区、青羊区、金牛区、武侯区、成华区、成都高新区6个中心城区核心区和新都区、郫都区、温江区、双流区、龙泉驿区、青白羊区、新津区和天府新区8个其他主城区。

#### 3、调查对象

成都市18周岁及以上常住居民。

## 调查方式和方法

#### 1.调查方式

本次调查的调查方式分为文献资料法、访谈法、多阶段抽样调查法、问卷调查法和街头拦截法等。其中，以问卷调查法为主。

#### (1)文献调查法

小组调查在进行开展调查之前，通过线下图书馆以及线上各学科文献网站了解目前针对预制菜的研究现状和国家相关政策，依据文献进行主题和调查方法的梳理，最终确定本次调查主题以及数据处理思路。

#### (2)访谈法

本次调查，小组成员利用课余时间，在成都市不同小区对居民随机进行采访。在初步了解相关背景建立信任后询问其对于预制菜认知程度以及消费行为等相关信息，为问卷的制定以及修改提供信息。

#### (3)多阶段抽样调查法

多阶段抽样是指将抽样过程分阶段进行，每个阶段使用的抽样方法往往不同，将各种抽样方法结合使用，既有便干组织抽样，也有利于提高抽样效率。

本次调查四川省成都市14个主城区作为一级抽样框，分层多阶段进行抽样后确定最终调查样本，对所获样本进行数据的收集处理后，针对本次调查问题提出相关建议。

#### (4)问卷调查法

该方法为本次调查所采用的主要方法，在控制问卷长度的同时，也保证所提问题的有效性。通过线下与线上结合的方式发放以提高整个过程的回收效率。

#### (5)街头拦截法

本次调查前期，为了解居民对于预制菜的消费行为，在商场、固定街头随机拦取行人进行交谈与采访，为后续问卷问题的设计与调查奠定良好的基础。

#### (6)文本挖掘法

利用文本挖掘法，从购物软件、论坛上获取消费者对预制菜评论，看法，认知情况等数据，从数据中挖掘重要信息，分析预制菜优缺点。

#### 2.调查方法

#### (1)定性研究方法

定性研究方法是消费者行为学研究学者最先和比较常用的一种研究方法。它的基本特点就是不要求消费者按照事先安排好的回复类别来回答问题。且答案是文字性质的，不是数量化的，被调查者要用自己的话来陈述答案。因此在本调查过程中凡涉及预制菜消费者行为以上相关方面的问题都运用到该种研究方法。

#### (2)定量研究方法

鉴于定性分析方法的抽象性和一定程度的模糊性，为了更好地了解居民对于预制菜的消费行为、认知情况等信息，本调查将采用全面调查与典型调查相结合的方法，收集一些相关消费者行为资料，并对这些资料运用相关统计软件(如SPSS)进行计量分析，从而对影响预制菜消费者行为的相关自变量和因变量间的关系(相关关系或因果关系)能够进行准确界定，并运用到预制菜企业营销策略的制定过程中。

## 抽样方案设计

本次调查首先将总体分为中心城区核心区（6个）和其他主城区（8个）两层，然后分别在中心城区核心区和其他主城区独立使用三阶段抽样方法。第一阶段，以中心城区核心区和其他主城区的所有行政区作为一级单元的抽样框，通过PPS抽样方法从一级单元的抽样框中各选取3个入样行政区，采用概率比例规模抽样方法，该方法是放回的不等概率抽样，即每个行政区入样概率是不等的，与该区的常住人口数成正比。利用PPS 法进行初级抽样单元的抽取时运用代码法进行实施。即赋予每个行政区与该辖区人口数相同的代码数，将代码数依次进行累加，利用计算机各抽取3个随机数。随机数所属的代码范围对应的行政区入样，则这就构成了初级抽样单元，具体如表3-1所示：

表 3 初级抽样单元

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | 行政区 | 常住人口数 | 累计 | 代码范围 | 随机产生数 | 是否抽中 | 重新编码 |
| 中心城区核心区 | 1 | 锦江区 | 909000 | 909000 | 1-909000 |  | 否 |  |
| 2 | 青羊区 | 961000 | 1870000 | 909001-1870000 | 1417843 | 是 | 1 |
| 3 | 金牛区 | 1280000 | 3150000 | 1870001-3150000 |  | 否 |  |
| 4 | 武侯区 | 1880000 | 5030000 | 3150001-5030000 | 3450445 | 是 | 2 |
| 5 | 成华区 | 1397000 | 6427000 | 5030001-6427000 | 5485466 | 是 | 3 |
| 6 | 成都高新区 | 1303900 | 7730900 | 6427001-7730900 |  | 否 |  |
| 其他主城区 | 7 | 新都区 | 1572000 | 9302900 | 7730901-9302900 |  | 否 |  |
| 8 | 郫都区 | 1683000 | 10985900 | 9302901-10985900 |  | 否 |  |
| 9 | 温江区 | 990000 | 11975900 | 10985901-11975900 | 11052649 | 是 | 4 |
| 10 | 双流区 | 2723000 | 14698900 | 11975901-14698900 |  | 是 | 5 |
| 11 | 龙泉驿区 | 1356000 | 16054900 | 14698901-06054900 |  | 否 |  |
| 12 | 青白江区 | 501000 | 16555900 | 16054901-16555900 |  | 否 |  |
| 13 | 新津区 | 373000 | 16928900 | 16555901-16928900 | 13856741 | 否 |  |
|  | 天府新区 | 866200 | 17795100 | 16928901-17795100 | 17298352 | 是 | 6 |

二级单元的抽样框是第一阶段入样的行政区的所有社区，第二阶段，在各入样的二级抽样单位中采用分层随机抽样方法，每层抽取人数的权数由每个行政区人数比例决定。基于调查实施的便利性以及调查的时间成本与人力成本的考虑，最终决定从中心城区核心区（即编号为1、2、3的行政区）抽6个社区，其他中心城区（即编号为4、5、6的行政区）抽取 3个社区展开调查。

第三阶段的三级单元抽样框为每个入样社区的所有住户，仍然通过比例分配的方式确定每个社区需要发放的问卷数，再使用简单随机抽样的方式选取社区的住户作为入样的三级抽样单位。这样的抽样方式便于组织抽样，简化了抽样框的编制，提高了抽样效率。整体抽样框如表3-2所示：

表 4 整体抽样框

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 一级单元抽样框 | 入样行政区 | 二级单元抽样框 | 入样社区 | 三级单元抽样框 |
| 成  都  所  有  主  城  区 | 青羊区 | 青羊区的所有社区 | 八宝街社区 | 八宝街社区的所有居民 |
| 东坡路社区 | 东坡路社区的所有居民 |
| 武侯区 | 武侯区的所有社区 | 倪家桥社区 | 倪家桥社区的所有居民 |
| 太平社区 | 太平社区的所有居民 |
| 成华区 | 成华区的所有社区 | 桃源社区 | 桃源社区的所有居民 |
| 红桥社区 | 红桥社区的所有居民 |
| 温江区 | 温江区的所有社区 | 合江社区 | 合江社区的所有居民 |
| 双流区 | 双流区的所有社区 | 莲花社区 | 莲花社区的所有居民 |
| 天府新区 | 天府新区的所有社区 | 罗家店社区 | 罗家店社区的所有居民 |

## 调查项目设计

#### 1.问卷调查项目设计

结合此次调查的目的，我们设计了相应的调查问卷。问卷共包括四个部分：受访者基本信息、食品消费观念和食品消费行为影响因素、对预制菜的认知程度和预制菜消费市场现状。（具体问卷见附录1），本次问卷设计共21题。

问卷结构：说明部分、主体部分、甄别部分和结束语部分。

问卷题型：单选题、多选题、量表题、开放题

甄别部分：设置1道题。为注意力测试题，题目要求选择“否”，未按要求填写问卷当做废卷处理。（看还有没有其他甄别选项）

# 调查实施

## 调查实施过程

#### 1.调查过程

抽样调查分为预调查和正式调查两个阶段进行，在预调查阶段我们在入框样本中共发放了100份问卷，用于检测问卷的合理性、科学性和有效性，为正式调查打下坚实基础。在预调查中，我们发现问卷中部分问题表达方式含糊不清、问题顺序存在不合理之处、问卷中的逻辑方式存在重复之处、量表信效度不达标等问题，经过小组讨论和指导老师的帮助，针对以上问题，我们对部分题目表达方式做了修正，修改部分量表等，最终显示问卷的信度和效度较好。

在正式调查阶段，我们采用线上线下相结合的方式发放问卷，共发放898份问卷，其中有效问卷为738份。同时，我们也通过访谈法，调查消费者在预制菜消费过程中的感受以及对预制菜的发展期望，为问卷调查做补充说明。

图 2 项目开展甘特图

#### 2.调查时间安排

针对本次调查，我们制定了详细的时间安排计划。

表 5 时间安排表

|  |  |
| --- | --- |
| 2023.11.10-2023.11.25 | 与指导老师讨论确定调研主题 |
| 2023.11.26-2023.12.10 | 文献资料查阅，并进行归纳整理 |
| 2023.12.11-2023.12.31 | 撰写调查计划，设计预调查问卷 |
| 2024.1.1-2024.1.14 | 预调查实施，确定最终问卷 |
| 2024.1.15-2024.2.15 | 各小区发放问卷并进行深度访谈 |
| 2024.2.16-2024.3.1 | 撰写调查报告 |

#### 3.调查分工安排



图 3 人员调查分工

# 数据处理与检验

## 预调查数据检验

在正式发放问卷前，为了对问卷进行测试，检验问卷的可行性和有效性，发现潜在问题并进行改进，确保调查的科学性、有效性和可信度，我们进行了预调查。预调查开展地点选取在成都市各主城区，以随机原则实地发放问卷共300份，回收问卷300份，有效问卷255份，有效回收率为85，根据预调查数据对问卷中的量表进行信度和效度检验，从而对问卷问题进行调整，提高问卷的有效性和可靠性。

#### 1．信度检验

信度分析是指问卷调查结果的一致性、稳定性、可靠性，反映客观事物被测 特征真实程度的度量指标，即用于测量问卷中样本回答结果是否可靠，有没有真实作答。检验信度越高，就是表示结果越可信。信度是任何一种测量和必要条件，以确保问卷调查结果是否真实可靠。本小组采用系数信度法对问卷的信度进行分析，计算Cronbach Alpha系数对问卷项目内部一致性进行信度检验，计算公式为：

 公式（1）

其中k为量表中题项的总数，Si方为第i题得分的题内方差，为全部题项总得分的方差。本文选取了消费者购买预制菜时各个方面重视程度评价量表和对预制菜了解程度量表进行分析，即问卷的第8题和第14题进行了信度检验，结果表所示：

表 6 第8题信度结果

|  |  |  |  |
| --- | --- | --- | --- |
| 量表 | Crobach a | 项数 | 信度评价 |
| 产品吸引 | 0.813 | 3 | 好 |
| 情感享受 | 0.886 | 2 | 好 |
| 食品健康与  安全 | 0.718 | 2 | 适合 |
| 产品选择 | 0.825 | 2 | 好 |
| 便利程度 | 0.883 | 1 | 好 |

表 7 第14题信度结果

|  |  |  |  |
| --- | --- | --- | --- |
| 量表 | Crobach a | 项数 | 信度评价 |
| 餐饮行业影响 | 0.787 | 3 | 适合 |
| 预制菜定义 | 0.832 | 2 | 好 |
| 预制菜对健康的影响 | 0.859 | 2 | 好 |
| 预制菜的改进 | 0.918 | 1 | 非常好 |
| 满意度总量表 | 0.923 | 8 | 非常好 |

由上表可知，预调查问卷数据的五个量表的Crobach α系数均在0.7以上，其中满意度总量表的Crobach α系数高达0.923，因此可认为本次调研中量表数据的信度非常好，这说明量表的题目设计具有较高的科学性及合理性，可靠性很高。

#### 2.效度检验

信度检验通过后，还需要考虑设计出来的问卷能否精确测出所需要测量的事物，即问卷效度分析。效度即有效性，它是指所测量到的结果反映所想要考察内容的程度，测量结果与要考察的内容越吻合，则效度越高；反之，则效度越低。

问卷的效度一般用KMO值和Bartlett球型检验检测。KMO值越大，表示所设变量彼此间的共同因素越多，问卷的效度越好；Bartlett球度检验值是检验变量彼此独立程度的指标，该值越大，表明变量之间的独立性越高，当Bartlett球形度检验的伴随概率小于0.05时，说明问卷效度好。

#### （1）KMO 和 Bartlett 球度检验

表 8 KMO 和巴特利特检验

|  |  |  |
| --- | --- | --- |
| KMO 取样适切性量数。 | | .834 |
| 巴特利特球形度检验 | 近似卡方 | 771.401 |
| 自由度 | 45 |
| 显著性 | .000 |

（2）总结

本次预调查问卷量表的 KMO 和 Bartlett 球度检验结果如表所示，可以看 到 KMO 值为 0.834，大于 0.8，表示问卷的相关性很强，适合进行因子分析， 显著性水平为 0.000，通过了显著性水平为 0.05 的 Bartlett 检验，表明问卷具有良好的结构效度。综上，本次预调查问卷量表数据可信，可以用于进行下一 步的数据分析。

#### 3.数据质量控制

数据质量控制的目的是确保收集到的数据准确、完整且具有可信度。这样的流程有助于提高实地调查的数据质量。我们从以下三方面进行了数据质量评估。

#### **（1）调查准备阶段**

调查前先对团队成员进行培训，使每个成员都能充分理解调查目的以及题目设置，使其能够在受访者对问卷提出质疑时进行解答。团队成员遵照明确采用的抽样设计方案，确保被调查样本的代表性。在线下发放问卷时，调查人员在受访者问卷填写中保持沉默，仅在必要时对被访者进行解释和说明，避免诱导性语言，保证了问卷填写的真实有效性。

在正式调查中，通过对八宝街社区、东坡路社区、倪家桥社区、太平社区、桃源社区、红桥社区、合江社区、莲花社区、罗家店社区这 9 个社区路人的拦截式面访调查以及社区周边 路人的拦截式调查进行问卷的发放。调查期间，小组及时对调查问卷的完整性、 正确性、一致性进行检查，对于无回答或者答案不清楚的题目，采取追问的措施以保证回收问卷的质量。

#### **（2）数据整理阶段**

在线下访谈过程中，调查人员要对整个访谈过程进行录音以及文字记录下保证信息的准确。对于线上回收的问卷，填写内容中具有以下问题时，将其视为无效问卷进行剔除，具体的问题如下所示：

1）问卷填写时间小于100s；

2）甄别问题选择“是”

3）量表题或选择题的答案基本没有变化或者持续出现极端结果；

4）问卷问题回答中出现明显的前后矛盾结果；

小组成员在每天结束线下发放问卷后要对自己回收的问卷数据进行甄别、审查，不符合要求的加以剔除，在完成此项工作后，将问卷统一上交。小组成员组建数据复核小组，在固定时间进行轮换，对问卷内容进行二次审核，对有问题数据加以剔除，在保证问卷准确率的基础上最大限度提高问卷的有效率。

#### **（3）数据录入阶段**

1）在数据录入时，采用双重录入的方法，即由两个独立的数据录入员分别录入相同的数据，然后对比两份录入结果，发现并纠正差异。

2）一致性检查：一致性检查是根据每个变量的合理取值范围和内部关系，检查数据是否存在超出正常范围、逻辑上不合理或者存在相互矛盾等问题。运用 SPSS 进行数据的 一致性检查，删去有异常的数据。

3）无效值和缺失值的处理：由于调查、编码和录入等过程，数据中存在一些无效值和缺失值，在数据清洗时需要对此进行处理，将未填写完整的数据进行删除清理。

4）异常值检测：识别和处理异常值，利用SPSS软件设定阈值或规则，检测录入数据中的异常值，并进行适当的处理，如修正或排除。

5）记录日志：记录数据录入过程中的操作和错误，包括录入员的操作记录、错误信息等，便于后续的审查和纠正。

## 正式调查数据检验

#### 1.信度分析

本文选取了消费者购买预制菜的各个环节满意度评价量表进行分析，即问卷 的第 8和14 题进行了信度检验，结果表所示：

表 9 量表8信度结果

|  |  |  |  |
| --- | --- | --- | --- |
| 量表 | Crobach a | 项数 | 信度评价 |
| 产品吸引 | 0.913 | 3 | 非常好 |
| 情感享受 | 0.898 | 2 | 好 |
| 食品健康与安全 | 0.924 | 2 | 非常好 |
| 产品选择 | 0.901 | 2 | 非常好 |
| 便利程度 | 0.893 | 1 | 好 |
| 满意度总量表 | 0.953 | 10 | 非常好 |

表 10 量表14信度结果

|  |  |  |  |
| --- | --- | --- | --- |
| 量表 | Crobach a | 项数 | 信度评价 |
| 餐饮行业影响 | 0.891 | 3 | 适合 |
| 预制菜定义 | 0.941 | 2 | 好 |
| 预制菜对健康的影响 | 0.967 | 2 | 好 |
| 预制菜的改进 | 0.922 | 1 | 非常好 |
| 满意度总量表 | 0.983 | 8 | 非常好 |

由以上两表可知，预调查问卷数据的9个量表的 Crobachα系数均在 0.85 以上, 其中满意度总量表的 Crobachα系数高达 0.953和0.983，因此认为本次调研中量表数据的信度非常好，说明量表的题目设计具有一定的科学性及合理性，可靠性很高。

#### 2.效度分析

为了判断正式问卷是否可以有效地测量调查需要测量的变量，接下来对问卷 进行效度分析。

#### （1）KMO 和 Bartlett 球度检验

表 11 正式调查 KMO 和 Bartlett

|  |  |  |
| --- | --- | --- |
| KMO 和巴特利特检验 | | |
| KMO 取样适切性量数。 | | .931 |
| 巴特利特球形度检验 | 近似卡方 | 1920.988 |
| 自由度 | 45 |
| 显著性 | .000 |

#### （2）因子分析

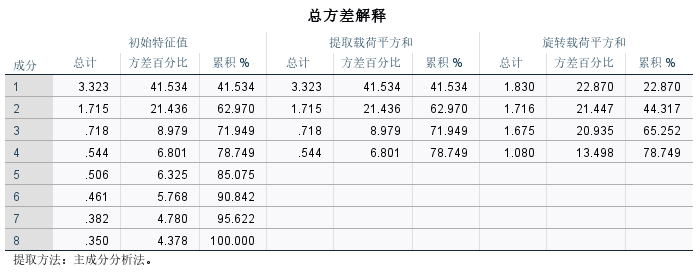
选择探索性因子分析，对问卷量表进行检验，观察检验结果是否符合预期假 设。经过因子分析后，得到下表。

图 4 总方差解释图

通过 SPSS 软件，对自变量和因变量进行降维，利用主成分分析法并使用 最大化方差旋转，对总量表做因子分析，经过总方差解释，第一个量表前 5 个因子的累计贡献率达到 77.123%，第二个量表前4个因子的累计贡献率达到78.749%，都能够包含足够多的信息。

#### （3）旋转后的成分矩

因子旋转是通过改变坐标轴位置，重新分配各个因子所解释的方差比例，使 其载荷数更接近于 1 或 0，能更好地解释和命题。

图 5 因子旋转后的成分矩阵表结果图

如上表所展示旋转后地主成分矩阵，表中数值为各题项的因素负荷量，由结果显示，结果符合原假设量表1的10个题目可由5个主成分进行表示。量表2的8个题目可由4个主成分进行表示。

综上，问卷信度和效度优良，在内容、结构上可信度高，根据此问卷获得的调查数据可以进行接下来的数据分析。

# 基于电商平台评论数据的文本挖掘

在本章中，我们将以谷言品牌的预制菜为研究对象，以京东电商平台作为数据来源，运用网络爬虫技术和文本挖掘方法，对消费者的文本评论数据进行深入的信息挖掘。目标是通过这些技术手段，精准地识别出消费者对预制菜的好评度、关注焦点以及抱怨点，从而初步评价预制菜产品的优缺点。通过这种方式，我们可以更加有针对性地优化调查问卷，完善实地调查方案，以便更加精准地了解消费者的需求和期望，为预制菜产品的改进和市场策略的调整提供有力的数据支持。

## 预制菜评价数据爬取



图 6 京东商城评价示意图

京东商场评论主要包括，用户名头像评论时间，商品名称，评论内容，图片/视频, 评价星级，本文使用 Python 网络爬虫来完成数据采集工作，只保留用户名，评论时间，评论内容，评价星级，共爬取5984条数据，爬取结果示意图如下表所示:

表 12京东部分用户评论爬取结果

|  |  |  |  |
| --- | --- | --- | --- |
| 用户名 | 评价星级 | 评论内容 | 评论时间 |
| 创\*\*\*阳 | star5 | 京东百亿补贴下单，很实惠，发货快，包装安全，发过来冰袋还是硬的，生产日期新鲜，味道可以的，食用方便，蒸上10来分钟。懒人值得拥有。 | 2024-01-03 19:32 |
| x\*\*\*9 | star5 | 谷言方便小碗菜是一款非常实用的产品。它非常方便，无需复杂的操作，随时可以享受美食。每个小碗菜的分量适中，既不会过量也不会不足，非常适合单独食用或与他人分享。而且价格实惠，非常经济实惠。总的来说，谷言方便小碗菜是一款非常值得购买的产品，特别是对于忙碌的现代人来说。 | 2023-11-08 15:54 |
| s\*\*\*0 | star5 | 太满意了，首先用泡沫箱包装，里面好几块生物冰袋还没有融化，菜还是冰块的样子，热菜很快，吃得很爽，方便就一句话，快递也很快，非常满意！ | 2023-10-29 22:19 |
| u\*\*\*k | star3 | 很一般，拌在饭里吃吃还行 | 2023-06-02 05:41 |

## 数据分析

在数据挖掘过程中，原始的庞大数据集往往掺杂着缺失值、重复值以及异常数据，这些因素不仅拖慢了模型的运行效率，更有可能导致挖掘结果的偏离。对于电商平台上的评论数据而言，由于存在商家雇佣水军进行刷单以提高销量的现象，对原始数据进行清洗变得尤为关键。

本次处理的核心任务是剔除由水军发表的重复性评论以及消费者随意发布的无实质性内容。为实现这一目标，我们制定了以下处理规则：

#### 1.评论文本内容与用户名均相同的数据视为水军发表的，予以删除。例如：

|  |  |  |  |
| --- | --- | --- | --- |
| 用户名 | 评价星级 | 评论内容 | 评论时间 |
| 游\*\*\*儿 | star5 | 超方便的预制菜。招即使直接冰箱拿。会回购 | 2023-03-25 10:10 |
| 游\*\*\*儿 | star5 | 超方便的预制菜。招即使直接冰箱拿。会回购 | 2023-03-25 00:05 |

#### 2.评论内容完全为英文字母、数字和标点符号的，视为随意发表的评论，予以删除。例如:

|  |  |  |  |
| --- | --- | --- | --- |
| 用户名 | 评价星级 | 评论内容 | 评论时间 |
| 环\*\*\*天 | star5 | 。。。。。 | 2024-01-01 09:25 |
| 晴\*\*\*郎 | star5 | //// | 2023-11-20 08:05 |

#### 3京东平台系统自动评论，这些数据没有任何分析价值，予以删除。例如：

|  |  |  |  |
| --- | --- | --- | --- |
| 用户名 | 评价星级 | 评论内容 | 评论时间 |
| a\*\*\*e | star5 | 此用户未填写评价内容 | 2022-12-13 16:35 |
| 洪\*\*\*七 | star5 | 此用户未填写评价内容 | 2023-06-26 17:36 |

经过预处理共得到1820条有效数据，无效的数据达4164条，文本有效率为30.4%，我们对经预处理后的有效评论数据进行文本数据分析。

图 7 词云图

## 关键词数据分析

利用Python的jieba库对中文文本进行分词处理（评价内容），并结合WordCloud库绘制词云图，以可视化方式展示分词结果。通过调整词云图的相关参数，可以获得如图所示的视觉效果：

## 情感倾向分析

在情感分析任务中，常见的方法包括基于规则或情感词典的方法，以及基于机器学习的方法。前者依赖于预定义的情感词典或规则集，而后者则依赖于从大量标注数据中学习到的模型。本文所选用的情感分析方法，是基于情感词典的。该方法的主要步骤如下：

1.文本分词：使用jieba分词工具对输入的文本进行分词，将连续的文本拆分为单个的词汇单元。

2.加载词典：加载情感词典和程度副词词典。情感词典包含了词汇的情感倾向（正面、负面或中性），而程度副词词典则用于调整词汇的情感强度。

3.计算情感得分：遍历分词后的文本，对每个词汇根据其在情感词典中的情感倾向进行计分。

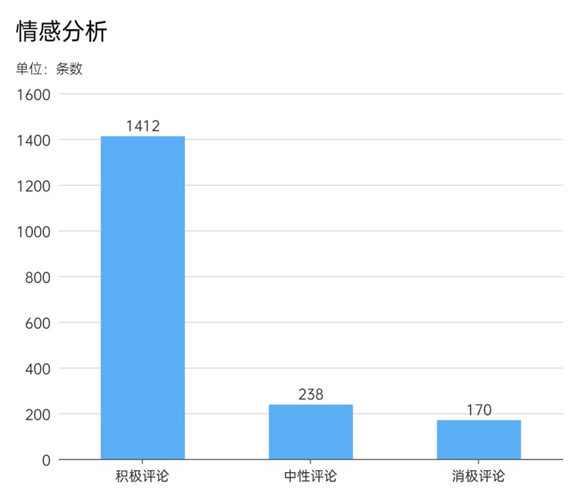
4.情感汇总：将单个词汇的情感得分汇总，得到整个文本的情感倾向得分。这一步骤还可能涉及对分支或子句的情感倾向进行统计和整合。结果如下：

图 8 情感倾向分析图

根据上述结果可知，积极评论1412条，约占77.58%，中性评论238条，约占13.08%，消极评论170条 ，约占9.34 %.这说明大部分消费者对于谷言预制菜持有积极态度，对于谷言品牌预制菜较为满意。

## 主题分析

在主体分析将采取隐含狄利克雷分布（LDA）；LDA模型是一种文档主题生成模型，也被称为三层贝叶斯概率模型，因为它包含了词、主题和文档这三层结构。该模型基于统计方法，可以从语料库中生成主题模型，并据此预测文章所属的主题。LDA模型的核心假设是，文档的生成首先是随机地决定一些主题，然后从每个主题中随机地选择一些词语，这些词语构成了整篇文档。

在情感分析的基础上，为了更好反映不同情感倾向的关注点，本节对不同情感倾向下的主题进行挖掘与分析。

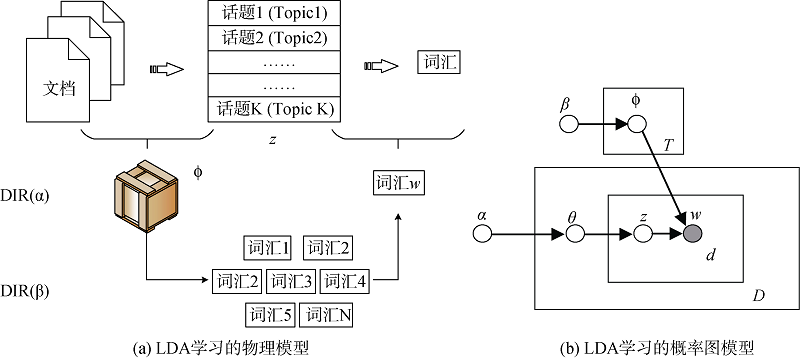


图 9 LDA模型图

LDA分析后，消极和中性情感词均含有三个大主题，每个主题下生成8个最有可能生成的词汇：

表 13 积极情感词汇

|  |  |  |
| --- | --- | --- |
| 物流快 | 旅游 | 可口 |
| 冷链好 | 居家 | 方便 |
| 回购 | 外出 | 新鲜 |
| 无破损 | 民宿 | 入手 |
| 品质保障 | 学生 | 便宜 |
| 售后好 | 老人 | 精美 |
| 包装完整 | 上班族 | 美味 |
| 专营店 | 父母 | 干净 |

#### 1．对谷言预制菜积极评价的3个潜在主题的特征值提取：

主题1的高频词反映出预制菜快递和官方预制菜直营店的发展，物流好，品质保障是预制菜快递行业发展的重要原因，运输过程中的保存完整也为预制菜打好了口碑；不难看出随着物流行业的进一步发展，预制菜行业也会随之不断发展。

主题2的高频次反映出预制菜的使用对象和应用场景，预制菜出现在人们生活方方面面中，各个群体都在广泛的使用，预制菜得到越来越广泛的消费者的认可。同时预制菜厂家应依据消费者外出和居家不同需求设计适合的产品

主题3的高频词反映出消费者对预制菜的满意程度，通过美味，新鲜，可口，我们可以看出消费者对品质的要求，以及其对预制菜的认可。

表 14 中性情感词

|  |  |  |
| --- | --- | --- |
| 营养素 | 保质期 | 小酥肉 |
| 蛋白质 | 生产日期 | 鱼香肉丝 |
| 维生素 | 添加剂 | 小龙虾 |
| 低脂肪 | 食品卫生 | 花椒鸡 |
| 碳水化合物 | 安全监督 | 酸菜鱼 |
| 矿物质 | 风险交流 | 速冻水饺 |
| 纤维素 | 责任承担 | 汉堡包 |
| 必需元素 | 食品检验 | 自热火锅 |

#### 2.对谷言预制菜中性评价的3个潜在主题的特征值提取：

主题1反映出预制菜的营养成分和价值，从特征值提取出的营养素，必需元素等可以看出，人们对于食物营养价值的考量愈发重要，预制菜营销店应关注必要的营养搭配和避免营养流失。

主题2反映出预制菜食品安全相关问题，群众对预制菜食品的食品安全提出多方面要求，侧面反映部分消费者对预制菜食品安全的怀疑。

主题3反映出预制菜食品的多样化，各式各样的预制菜让其具有了排它性，保持了预制菜对消费者持续的吸引力，各大厂家在保障食品安全，营养价值，自身基本利益的基础上应尽可能的实现食品的多样化。

# 消费者特征描述性统计

## 预制菜消费者特征分析

#### 1.基本信息分析

基于调查问卷，我们对客户特征数据进行了统计，绘制出统计图表，从年龄、职业、收入和宣传方式等方面进行分析，以便对消费者的个人特征有直观的了解。 在本次调查中，总计 738份问卷回收。如图所示，根据问卷第15题，购买过预制菜所占比例为 67.07%，共有 495 人，而未购买过预制菜的所占比例为 32.93%，共有 243 人。

图 10 消费者预制菜购买与未购买占比

#### 2.性别分布

在被调查的预制菜消费者中，我们可以根据图中看出女性所占比例为 67% ，明显高于男性。由此可以得出，女性在日常生活中对预制菜的

购买需求要远远大于男性，这与女性总体做菜频率高于男性有关。

图 11 消费者性别分布情况

#### 3.年龄分布

在预制菜的消费者中，18岁到40岁的消费者占大多数，该年龄大多是在校大学生和已经工作的青壮年，其他年龄段的消费者人数少说明这些年龄段对预制菜的接受程度不高。

图 12 消费者年龄分布情况

#### 4.婚姻状况

在预制菜的消费者中，已婚的消费者占了63.28%，未婚的消费者占比36.72%，已婚人群在家做菜的总体频率高于未婚人群，这说明预制菜的便捷方式被经常在家做饭的已婚人群接受，对于未婚人群预制菜市场，可以朝着即食方向发展。

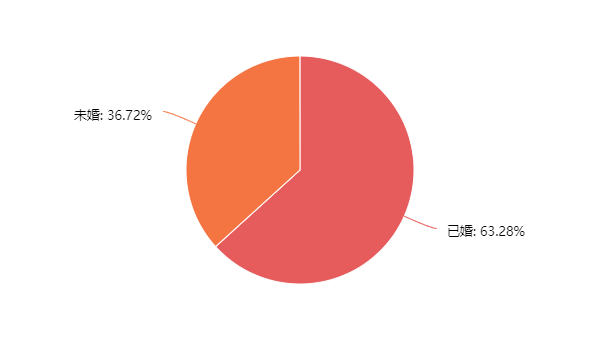


图 13 消费者婚姻状况分析

#### 5.月收入分布

根据预制菜消费人群的月收入分布，我们可以看到1000到4000的收入人群占总消费者的72%，说明预制菜的价格比较亲民，可以满足大部分消费者，而高收入人群占比较少，可能与他们的消费习惯有关系，这部分市场还可以进行挖掘。

图 14 消费者月收入分布

#### 6.职业分布

根据预制菜消费者的职业分布柱状图我们可以看到在校学生和普通职员为预制菜的主要消费者，这表明预制菜的便利性很好的符合学生和普通职员的需求有关。其他职业的预制菜消费者虽然占比不如前两者，但也有意愿购买预制菜，伴随着社会节奏的加快，预制菜还有很好的发展市场和空间。

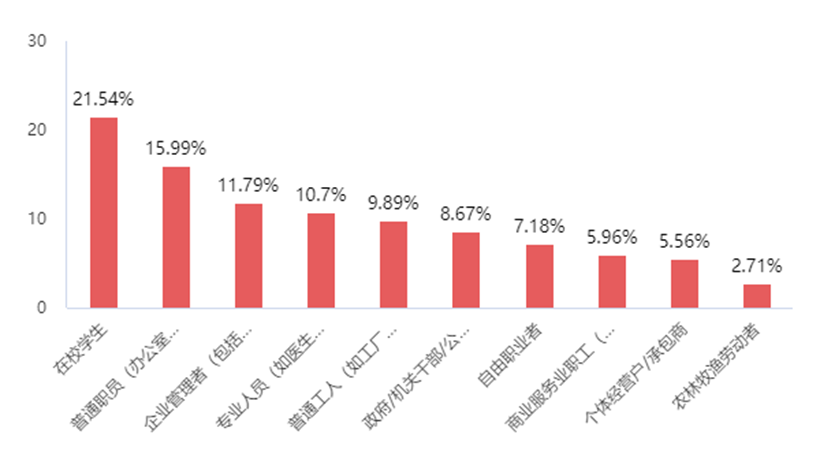


图 15 消费者职业分析

#### 7．饮食在收入中占比

根据下图我们可以看出预制菜消费者的日常饮食在收入中的占比基本上在40%以上，说明预制菜更能满足于中低收入人群，或者说价格高的预制菜不被高收入人群所认可，预制菜未打开高端市场。

图 16 消费者日常饮食在收入中的占比

#### 8.对食品需求的多元化

根据下面图表我们可以得知预制菜消费者对于食品需求更加多元化，要想加快预制菜的发展，我们也应该跟随潮流，让预制菜发展更加多元化，增加预制菜菜品和种类。

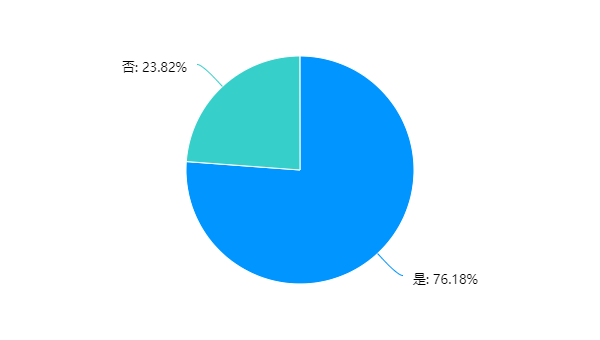


图 17 消费者食品需求多元化情况

## 消费者对预制菜的偏好分析

#### 1.预制菜购买频率

根据预制菜消费者的购买频率扇形图我们可以看到消费者预制菜购买最多的频率是一周1-3次，而几乎不买的消费者占了7%，我们可以通过一些宣传促销活动吸引他们继续购买预制菜，发现预制菜的优点，从而留住这部分消费者。

图 18 预制菜购买频率

#### 2.消费者购买预制菜时重视程度

根据下图我们可以看到，预制菜消费者更加看重食品安全、营养健康价格、口碑、便利程度这几个方面，不怎么看重情感享受、品牌、包装和营销这几个方面，这提示我们在进行预制菜宣传时可以着重宣传营养安全健康等方面。

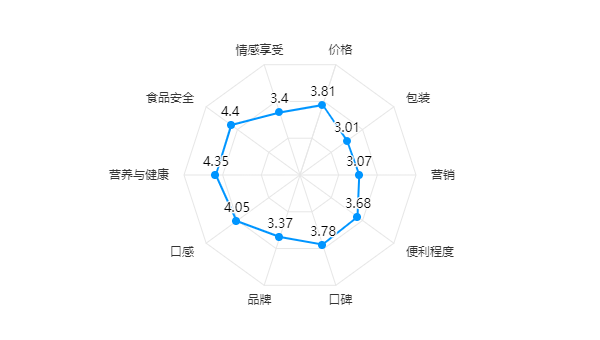


图 19 消费者对各因素的重视程度

#### 3预制菜消费类型

如下图所示，即食食品和即热食品是预制菜主要消费类型，说明这两个产品得到消费者的认可，而即烹食品和即配食品消费者消费较少，可以加大这两类产品的市场宣传与产品改进。

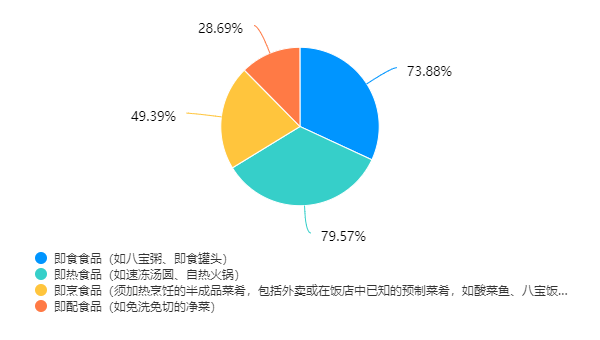


图 20 消费者购买的预制菜类型占比

#### 4.您消费预制菜的主要原因

如下图所示省时省力是消费者消费预制菜的主要原因，其次就是认为预制菜可以帮助想做饭的初学者做出味道更好的食物且预制菜价格较为低廉，其他消费者更多图的是预制菜的便捷性。

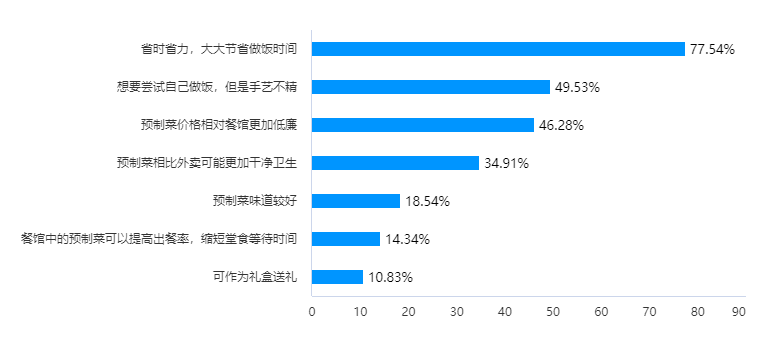


图 21 消费者购买预制菜的主要原因

#### 5.预制菜了解程度

根据下图可知，大多数预制菜消费者认为预制菜多油多盐且不利于健康，且对预制菜的定义不太清楚，从而导致消费者在购买预制菜的时候产生犹豫，担心预制菜的加工添加过多防腐剂等其他因素。要想更好的发展预制菜，我们必须让消费者明确预制菜的概念，改变预制菜就是不健康的观念。

图表 1 消费者对预制菜的了解情况

#### 6.预制菜的价格范围

如下图所示，大多数预制菜消费者认为对于100元一份的小炒菜的预制菜的价格应该在0-75元左右，说明大部分消费者认为预制菜的价格应该低于现炒菜。



图 22 消费者能接受的预制菜价格分布

## 基于 K-means 聚类分析的消费群体细分

#### 1.聚类方法的选择

由于此次问卷调查共收集738份问卷，我们选择适合大样本的K-menas聚类分析方法。

聚类分析是按照个体的特征将分类，目的在于让同一个类别内的个体之间具有较高的相似度，而不同类别之间具有较大的差异性，这样便于根据不同类别的 特征进行有效地分析。

先随机选取 K个对象作为初始的聚类中心，然后计算每个对象与各个种子聚 类中心之间的距离，把每个对象分配给距离它最近的聚类中心。聚类中心以及分 配给它们的对象就代表一个聚类。一旦全部对象都被分配了，每个聚类的聚类中 心会根据聚类中现有的对象被重新计算。这个过程将不断重复直到满足某个终止 条件。 终止条件可以是以下任何一个：

1．没有（或最小数目）对象被重新分配给不同的聚类。

2.没有（或最小数目）聚类中心再发生变化。

3.误差平方和局部最小。

#### 2.聚类因子的选取

进行聚类前，我们从问卷中选取 12个问题，根据问题间的关系进行整合最 终得到购买意愿、购买情况、了解程度、购买偏好，这 4 个变量，采用 K-Means 聚类的方法，它将参与聚类的个案各视为一类，然后根据两个类别之间的距离或 者相似性行逐步合并，直到最后个案合并为一个大类为止，最终得到三种类型的用户。 由于 15、17、21多选题，对这些题目的量化处理方法如下：将选择其中一个选项的赋值“1”，否则赋值为“0”，最后将所有选项赋值相加，最后得到该项的指标值。

表 15 问题整合情况

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 题目 | 指标名称 | 解释情况 |
| 1 | 17 | 购买意愿 | 对预制菜的期望 |
| 2 | 15 | 购买情况 | 预制菜购买频率 |
| 3 | 14、21 | 了解程度 | 预制菜的了解程度，和购买方式 |
| 4 | 8、16 | 购买偏好 | 预制菜的购买种类和对预制菜各方面重视程度 |

#### 3.聚类结果分析

通过对选取的聚类因子进行聚类群分，对每个类别的人群计算其各个变量的均值，最终预制菜消费者被聚成三类。根据聚类结果数据，将其转化 为指标，每个类别的特征如表所示。

表 16 消费者购买意愿聚类情况

|  |  |  |  |
| --- | --- | --- | --- |
| 指标名称 | 类别一 | 类别二 | 类别三 |
| 购买意愿 | 高 | 中 | 低 |
| 购买情况 | 购买频率高 | 购买频率一般 | 购买频率低 |
| 了解程度 | 多样的购买方式，了解程度高 | 了解程度中等，购买方式稍多 | 了解程度低，购买方式单一 |
| 购买偏好 | 关注点多，选择广 | 关注度一般，选择性一般 | 关注点单一，选择性少 |

基于上述特征，我们将消费者分为忠实人群、保守人群、跟风人群三个类别， 具体特征如下：

**忠实人群**：此类人群购买意愿高，且购买频次高，且对预制菜产品非常了解，对预制菜种类选择广，要求低。

**保守人群**：此类人群购买意愿低，购买频率一般，对预制菜了解程度一般，对预制菜产品无明显偏好和选择。

**跟风人群**：此类人群购买意愿中等，购买情况为买过以后不会再买，且对预制菜了解程度少，对预制菜产品关注点单一，选择性少。

其中忠实人群意愿高，对预制菜了解程度高，在预制菜的种类上无明显偏好，选择性广，是预制菜产品的核心消费者。 对于我们优化预制菜的营销策略有极大帮助，可以根据他们的问卷设计有利于预制菜推广的方案。

## 非预制菜消费者分析

基于调查问卷，我们对客户特征数据进行了统计，绘制出统计图表，从年龄、职业、收入和宣传方式等方面进行分析，以便对消费者的个人特征有直观的了解。 在本次调查中，总计 738份问卷回收。根据问卷第15题，购买过预制菜所占比例为 67.07%，共有 495 人，而未购买过预制菜的所占比例为 32.93%，共有 243 人。

#### 1.基本信息分析

#### （1）性别分布

如下图所示，非预制菜消费者中女性也是占60%，要多于男性，可见在对于预制菜调查过程中，女性展示出对预制菜的兴趣要高于男性，说明女性对预制菜的购买意愿高于男性。

图 23 非预制菜消费者性别分布

#### （2）年龄分布

如下图所示，相比于预制菜消费者，非预制菜消费者年龄分布要更加均匀，但18-30岁的群体依然占据大头，而60岁以上的群体占比相比预制菜消费者年龄分布中60岁以上占比有所提高，这些表明预制菜的主要消费群体为青壮年和青少年，预制菜市场应该注重他们的建议。

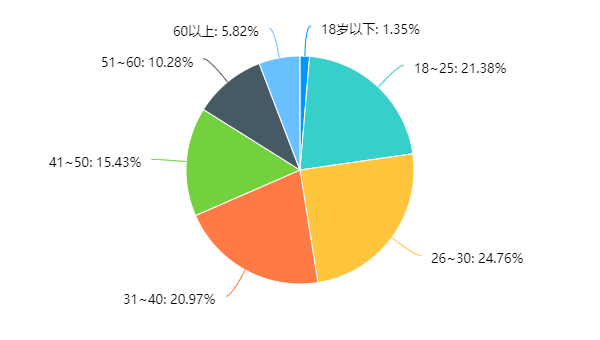


图 24 非预制菜消费者月收入分布

如表所示，可以看到月收入在3000以下的群体占了30%以上，说明在这个群体看来预制菜的性价比不高，不是消费的最佳选择，对于低收入人群预制菜的价格与产品不匹配。收入在10000以上的群体占比很少，说明预制菜可以满足大部分消费者消费情况**。**

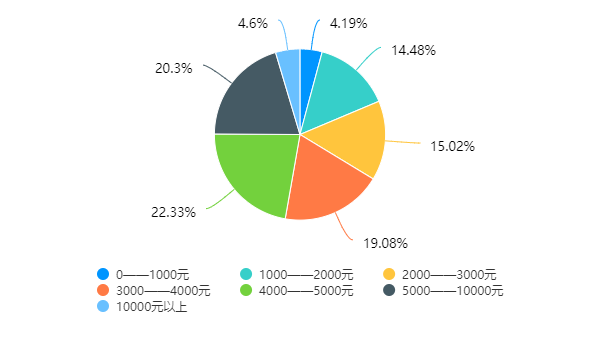


图 25 非预制菜消费者收入分布

#### 2.限制消费预制菜的原因

如下图所示，影响非预制菜消费者消费预制菜的主要原因是他们担心预制菜存在食品安全问题，且预制菜让菜品失去锅气和导致营养流失，所以预制菜商家可以在这三个方面努力改进，同时加大预制菜的宣传，改变这些群体对预制菜的观念。

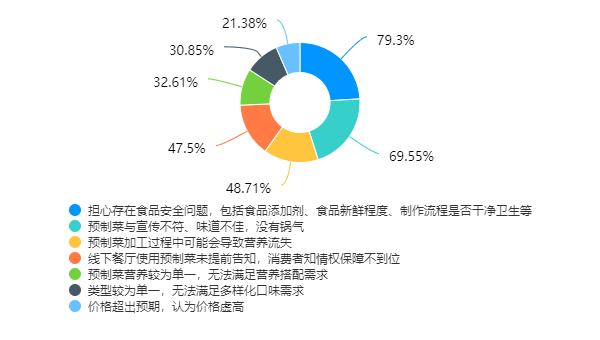


图 26 限制购买预制菜的因素分析

#### 3.吸引非预制菜消费者消费的方式

如图所示，通过朋友或者邻居的推荐是最受欢迎的吸引方法，促销活动和自媒体宣传其次， 说明商家在推广预制菜时可以先做促销活动吸引消费者购买，获得消费者认可后自然能向朋友进行宣传，提高预制菜的知名度。

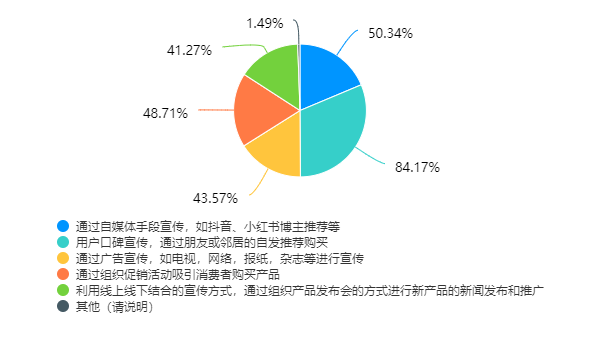


图 27 吸引非预制菜消费者消费的方式

# 预制菜消费者潜在客户的挖掘

预制菜潜在用户是指那些尚未购买过预制菜，但基于他们的生活方式、饮食习惯、消费观念等因素，有可能在未来成为预制菜消费者的群体。这些潜在用户可能具备一些共同的特征，如忙碌的生活方式、对方便快捷食品的需求、对健康饮食的关注等。

在对预制菜消费者和非预制菜消费者特征分析的基础上，本章将着重并深入分析非预制菜消费者，挖掘出潜在的预制菜消费者。

## 数学建模思路

#### 1.样本数据的分析

在对非预制菜消费者特征分析后发现非预制菜消费群体存在较为明显的两类群体：一类是之前从未购买过预制菜，以后也不会购买的群体；一类是此之前未购买过预制菜，但未来很有可能 成为预制菜消费者的群体。这进一步证实了数据适合进行二分类分析。对于这种情况，逻辑回归模型是一个合适的选择，因为它能够处理二分类问题，并且可以提供每个特征对分类结果的权重和系数，从而帮助理解哪些因素在影响消费者的购买决策。

#### 2.变量整合

由于本文数据的自变量过多，需要对变量进行整合：首先对于基本信息中的性别、年龄、职业、月收入等自变量进行相关性检验，我们选取被调查者对预制菜的购买意愿影响因素9个变量，通过因子分析，找出变量之间的公共因子，利用公共因子与变量之间的线性关系对变量进行合并。

#### 3.模型建立

使用训练集数据训练逻辑回归模型。设置合适的损失函数（如对数损失函数）和优化算法（如梯度下降算法）。调整模型参数，如正则化强度（L1或L2正则化），以防止过拟合。

## 模型介绍

#### 1.逻辑回归模型简介

逻辑回归（Logistic Regression）是一种用于解决二分类（0 or 1）问题的机器学习方法。虽然名字中带有“回归”，但逻辑回归实际上是一种分类算法，用于估计某种事物的可能性，而非数学上的概率。逻辑回归的输入通常是线性回归的输出，通过sigmoid函数将整体值映射到（0，1）范围内，并设置一个阈值来进行分类判断。

#### 2.逻辑回归原理

线性回归与Sigmoid函数：逻辑回归的本质是在线性回归的基础上，通过引入Sigmoid函数将线性回归的输出转换为0到1之间的概率值。Sigmoid函数的形式为1 / (1 + e^(-z))，其中z是线性回归的输出。这样做可以将线性回归的连续输出映射到0和1之间，从而表示某个样本属于正例的概率。

Sigmoid函数的表达式为：

 公式（2）

其中，e是自然对数的底数，约等于2.71828。Sigmoid函数是一种非线性函数，它可以将任意实数值映射到0到1之间。当输入值x趋向于正无穷时，函数值f(x)趋向于1；当输入值x趋向于负无穷时，函数值f(x)趋向于0。这种特性使得Sigmoid函数非常适合用于将线性回归的输出转换为概率值，从而实现逻辑回归的二分类功能。



在逻辑回归中，确实经常使用交叉熵作为损失函数。交叉熵损失函数衡量了模型预测的概率分布与真实概率分布之间的差异。

对于包含(n)个样本的数据集，整体的交叉熵损失函数(L)是所有样本损失函数的平均（或求和，取决于具体实现）：

 公式（3）

或者，如果是求和形式：

 公式（4）

在优化过程中，我们的目标是找到一组参数，使得这个损失函数 (L) 最小。这通常通过梯度下降算法或其他优化算法来实现，通过不断迭代更新参数，直到损失函数收敛到一个较小的值。

## 变量的检验与整合

#### 1.基本变量的检验

我们通过对变量进行相关性检验来检验对于影响是否会购买预制菜的基本信息变量间的相关性公式如下：

 公式（5）

SPSS检查结果如下：

表 17 相关性表



一般，相关性系数越接近 1，表明变量之间的相关性越强，反之则越弱。以上变量均不存在较强相关性，在后续建模中以上基本信息变量可以直接保留。

#### 2.用因子分析对变量进行合并

由于问卷中涉及被调查者在消费时关注的多个变量，并且这些变量之间存在较强的相关性，为了简化模型并提高分析的准确性，我们对购买关注变量进行了因子分析。这一步骤的目的是减少输入变量的数量，同时保留原始数据中的关键信息，以便后续建模和分析。通过因子分析，我们可以将多个相关变量综合成少数几个独立的因子，这些因子能够代表原始变量的主要特征，从而实现对模型的优化和简化。因子分析结果如下：

表 18 因子分析总方差解释表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 成分 | 初始特征值 | | | 提取载荷平方和 | | |
| 总计 | 方差百分比 | 累积 % | 总计 | 方差百分比 | 累积 % |
| 1 | 3.102 | 31.016 | 31.016 | 3.102 | 31.016 | 31.016 |
| 2 | 1.713 | 17.125 | 48.141 | 1.713 | 17.125 | 48.141 |
| 3 | 1.215 | 12.146 | 60.287 | 1.215 | 12.146 | 60.287 |
| 4 | .825 | 8.255 | 68.541 | .825 | 8.255 | 68.541 |
| 5 | .688 | 6.884 | 75.425 | .688 | 6.884 | 75.425 |
| 6 | .633 | 6.334 | 81.760 | .633 | 6.334 | 81.760 |
| 7 | .585 | 5.855 | 87.614 |  |  |  |
| 8 | .519 | 5.188 | 92.803 |  |  |  |
| 9 | .409 | 4.088 | 96.891 |  |  |  |
| 10 | .311 | 3.109 | 100.000 |  |  |  |

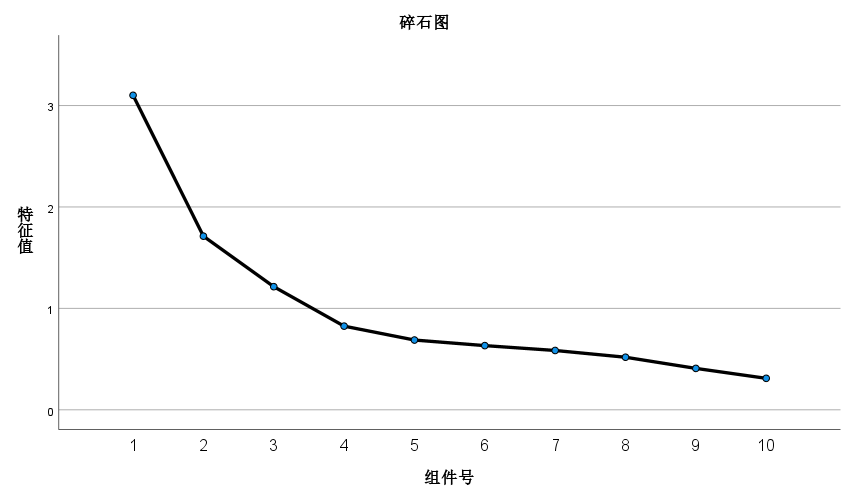


图 28 碎石图（提取方法：主成分分析法）

由表知，我们从表中提取6个公共因子，这6个公共因子包括了原始数据81.760%的信息

表 19 成分得分系数矩阵

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | 1 | 2 | 3 | | 4 | 5 | 6 |
| 8、在您考虑购买预制菜时，您对以下方面的重视程度为—价格 | | -.136 | -.011 | .020 | | .967 | -.094 | -.043 |
| 8、包装 | | -.184 | .029 | .278 | | .220 | .033 | .185 |
| 8、营销 | | .062 | -.042 | -.141 | | -.121 | 1.003 | -.081 |
| 8、便利程度 | | -.009 | -.164 | -.081 | | .046 | -.036 | 1.008 |
| 8、口碑 | | .004 | .619 | .012 | | -.037 | -.144 | -.276 |
| 8、品牌 | | .032 | -.186 | .987 | | -.064 | -.157 | -.152 |
| 8、口感 | | .435 | -.187 | -.094 | | .097 | .018 | .211 |
| 8、营养与健康 | | .373 | -.075 | .020 | | -.129 | .025 | -.030 |
| 8、食品安全 | | .452 | .055 | -.064 | | -.193 | .012 | -.139 |
| 8、情感享受 | | -.109 | .700 | -.264 | | -.030 | .105 | .014 |
| 提取方法：主成分分析法。  旋转方法：凯撒正态化最大方差法。  组件得分。  由该表可知公共因子与原始观察变量的线性关系如下：  F1=-0.136x1-0.184x2+0.062x3-0.009x4+0.004x5+0.032x6+0.435x7+0.373x8+0.452x9-0.109x10  F1=-0.011x1+0.029x2-0.042x3-0.164x4+0.619x5-0.186x6-0.187x7-0.075x8+0.055x9+0.700x10  以此类推，后面几个线性关系也类似。  根据公共因子与原始观测变量的关系强弱，我们将变量归为6大类，并重新命名：  公因子命名表 | | | | | | | | |
| 主成分 | 公因子名称 | | | | 解释的关系变量 | | | | |
| F1 | 食品安全健康口感重视程度 | | | | 食品安全，营养和口感 | | | | |
| F2 | 情感价值重视程度 | | | | 情感享受和口碑 | | | | |
| F3 | 知名度重视程度 | | | | 品牌 | | | | |
| F4 | 营销手段重视程度 | | | | 营销 | | | | |
| F5 | 性价比重视程度 | | | | 价格 | | | | |
| F6 | 食品便捷性重视程度 | | | | 便利程度 | | | | |
| 通过下表可知，公因子间完全独立。  成分得分协方差矩阵   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | 成分 | 1 | 2 | 3 | 4 | 5 | 6 | | 1 | 1.000 | .000 | .000 | .000 | .000 | .000 | | 2 | .000 | 1.000 | .000 | .000 | .000 | .000 | | 3 | .000 | .000 | 1.000 | .000 | .000 | .000 | | 4 | .000 | .000 | .000 | 1.000 | .000 | .000 | | 5 | .000 | .000 | .000 | .000 | 1.000 | .000 | | 6 | .000 | .000 | .000 | .000 | .000 | 1.000 | | 提取方法：主成分分析法。  旋转方法：凯撒正态化最大方差法。  组件得分。 | | | | | | | | | | | | | | | |

#### 3.特征变量的整合

经过因子分析后，我们成功地从原始的基本信息变量中提取出了4个关键因子。结合这6个因子变量与原始的性别、年龄、月收入、职业等4个基本信息变量，我们最终整合得到了10个自变量。这些变量涵盖了购买预制菜时的多个方面，接下来，我们将利用这10个整合后的变量进行建模分析，以深入探究它们对预制菜购买决策的影响。

## 模型建立与修正

由于逻辑回归模型是一个“白盒模型”，它的预测过程十分透明，具有很好的可解释性，可以对非预制菜消费者的各个特征变量进行分析，找到影响非预制菜消费者未来是否购买预制菜最大的几个特征。并且数据集中有很多离散型特征，逻辑回归对此有较好的处理能力，因此选择逻辑回归模型对特征的重要性进行分析。

#### 1.特征变量与非预制菜消费者未来消费预制菜状态特征分析

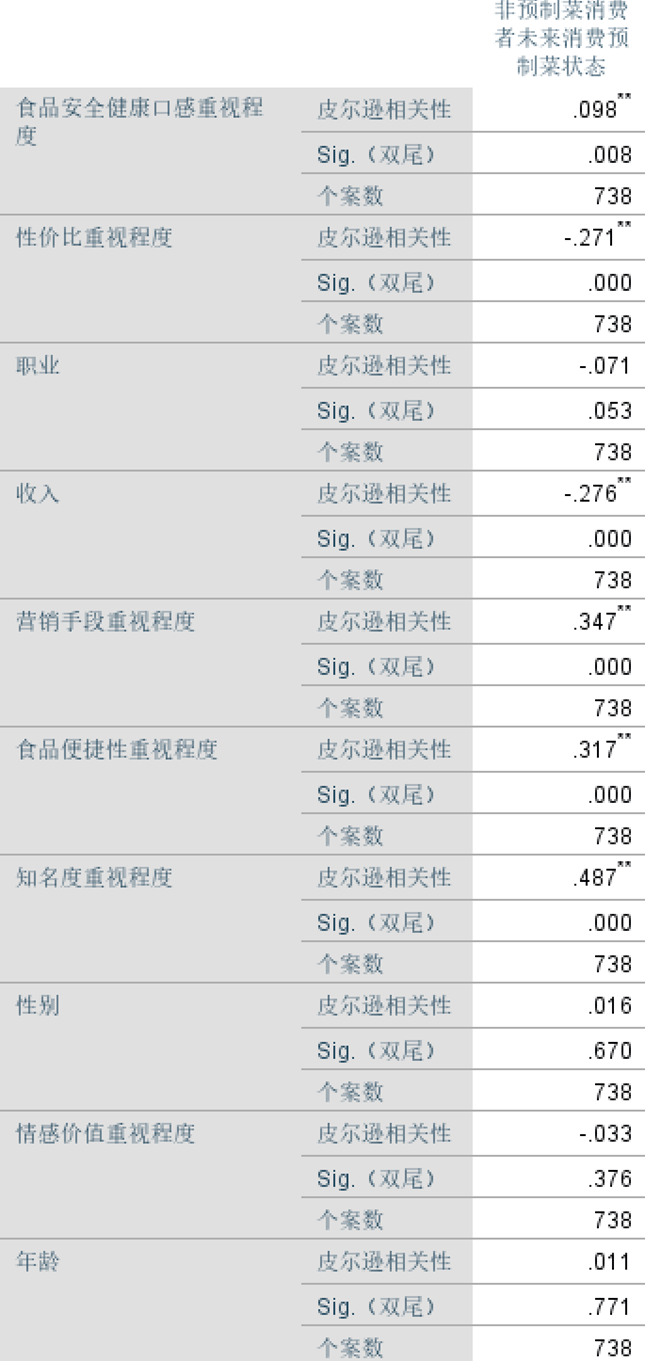
对各个特征与非预制菜消费者未来消费预制菜状态进行单因素相关性分析，可以判断它们之间是否存在联系，如果存在联系便可以选择这一特征进行逻辑回归模型的构建，如果不存在联系则就将这一特征舍弃。

研究两个变量之间相关性的检验方式有很多，选取哪个检验方式取决于变量的类型，看变量的类型是否是连续变量、二分类变量、无序分类变量和有序分类变量等。

#### **（1）连续特征**

当分析非预制菜消费者未来是否会消费预制菜（作为一个二分类因变量）是记作1，否记作0，与另一个连续特征变量之间的相关性时，可以采用点二列相关系数来量化这种关系。虽然点二列相关是皮尔逊相关系数的一个特例，但它们在数据假设上是相似的。为了使用SPSS软件进行分析，我们将二分类变量编码为数值型（例如，0和1），然后运行皮尔逊相关分析：

表 20 皮尔逊相关分析



经过皮尔逊相关分析后衡量特征变量和非预制菜消费者未来消费预制菜状态变量之间的线性关系强度和方向，由相关分析结果可知食品安全健康口感重视程度、性价比重视程度、收入、营销手段重视程度、食品便携式重视程度、知名度重视程度六个特征与非预制菜消费者未来消费预制菜状态在0.05水平上显著线性相关。而职业、性别、情感重视程度、年龄三个特征则于非预制菜消费者未来消费预制菜状态之间不存在线性相关关系。

#### 2.逻辑回归模型的实现

在对特征和离职状态的相关性进行分析后，利用Python软件中的sklearn模块中的LogisticRegression 函数以及 cross\_val\_score 进行逻辑回归模型的构建。

#### **（1）10折交叉验证**

在10折交叉验证中，原始数据集D被划分为10个大小尽可能相似的互斥子集，这些子集分别标记为D1, D2, ..., D10。每个子集都尽可能保持与原始数据集D相同的数据分布，以确保每次训练和测试时数据的代表性。接下来，这10个子集将按顺序（或其他预定的顺序）轮流作为测试集，而其余9个子集则作为训练集。例如，在第一次迭代中，D1作为测试集，D2到D10作为训练集；在第二次迭代中，D2作为测试集，D1、D3到D10作为训练集，以此类推，直到D10作为测试集，D1到D9作为训练集。示意图如下：

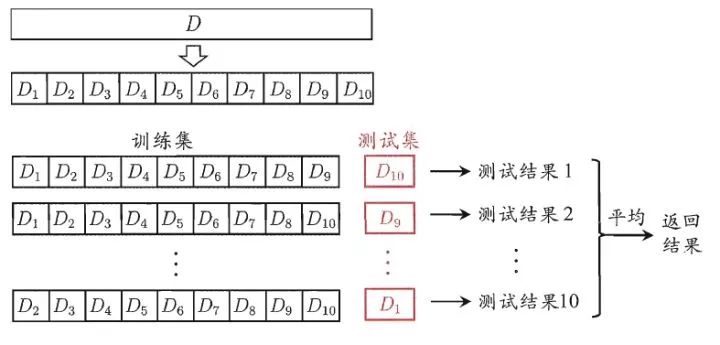


图 29 10折交叉验证图

#### **（2）数学模型的实现与评价**

由于本章研究样本容量不是很大，而在小样本数据场景下，相较于SAGA和SAG求解，L-BFGS求解器由于其拟牛顿法的特性，通常能够提供更稳定和准确的优化结果，下面将常用L-BFGS求解器。

表 21 模型关键参数

|  |  |  |
| --- | --- | --- |
| 函数 | 参数 | 参数设置 |
|  | 惩罚项 | L |
| LogisticRegression | 分类类型权重 | balanced |
|  | 求解器 | L-BFGS |
|  | 迭代次数 | 1000 |
| cross\_val\_score | 交叉验证次数 | 10 |
|  | 函数返回值 | Accuracy;precision等 |

#### **（3）模型评价指标与模型评价**

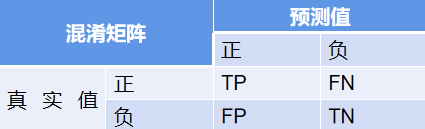
在机器学习中对于分类模型来说常用混淆矩阵来评价模型的效果，如下图所示

图 30 混淆矩阵

逻辑回归模型中的正例表示未来购买预制菜，反例表示未来不购买预制菜，TP表示将未来购买预制菜预测成未来购买预制菜的个数，FN表示将未来购买预制菜预测成未来不购买预制菜的个数，FP表示将未来不购买预制菜预测成未来购买预制菜的个数，TN表示将未来不购买预制菜预测成未来不购买预制菜的个数。在评价模型时常用准确率、精确率、召回率、受试者工作特征曲线（ROC曲线）和和ROC曲线下面积值（AUC值）来评价模型的好坏，他们的定义分别为

1）准确率表示预测正确消费者未来预制菜状态的样本数占总样本数中的比例，定义为：

 公式（6）

2）精确率和召回率。前者表示模型正确预测的未来购买预制菜样本数占模型预测的所有未来预制菜样本数，后者表示模型正确预测的未来购买预制菜样本数占所有实际未来购买预制菜样本数。

 公式（7）

 公式（8）

3）ROC曲线和AUC值。ROC曲线是以真正例率TPR为纵轴，假正例率FPR为横轴，AUC值相较于ROC曲线可以更加直观合理的判断模型的优劣，即取值越大模型的效果越好。它的取值范围在0.5到1之间，当低于或等于0.5时表示模型没有应用价值，当 AUC值在0.8以上时就表示模型效果很好了。TPR和FPP分别定义为

 公式（9）

 公式（10）

构建逻辑回归的模型的目的是为了得到各个特征对消费者未来购买预制菜状态影响的程度，在进行预测时更期望将样本中的未来购买可能购买预制菜消费者都找出来，并有不错的性能。因此模型的召回率和AUC值是重点关注的指标。

4）逻辑回归模型10折交叉验证的准确率、精确率、召回率、AUC值如下表所示。

表 22 准确率、精确率、召回率、AUC值

|  |  |  |  |
| --- | --- | --- | --- |
| 准确率 | 精确率 | 召回率 | AUC值 |
| 0.78 | 0.41 | 0.76 | 0.85 |

逻辑回归模型经过10折交叉验证的准确率是78%，精确率是41%，召回率是76%，AUC值是0.85，表明所构建的模型在召回率和AUC值上都有较好的表现，能够找到样本中73%的未来购买预制菜的消费者，有很好的预测性能。由此得到的各个特征的重要性也有着很高的可信度。

#### **（4）**特征变量重要性

利用逻辑回归分析得到各个特征的重要性，结果如下所示：

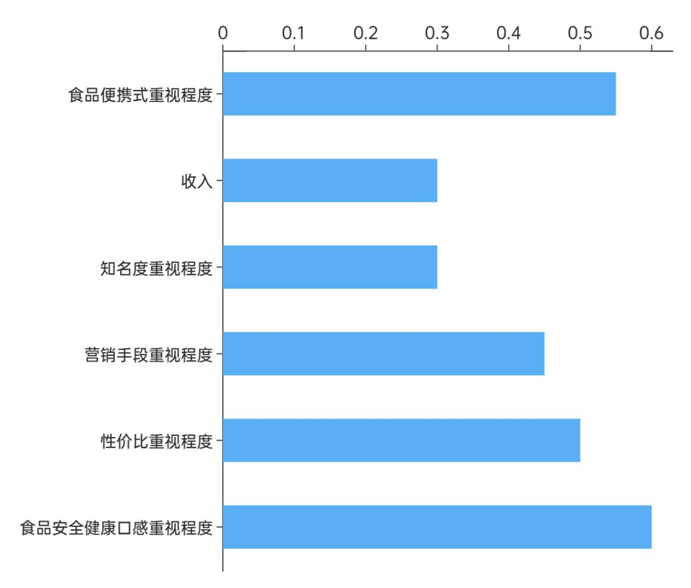


图 31 各个特征的重要性

经过分析各特征变量与非预制菜消费者未来购买预制菜均呈现正相关，从重要度由高到低依次是：食品安全健康口感、食品便携式、性价比、营销手段、知名度、收入。

#### **（5）**模型结果分析及模型现实意义

1）模型结果分析与建议

可以看出食品品安全健康口感、食品便携式、性价比，是影响非预制菜消费者未来购买的重要度最高的三大特征变量。

建议1：随着消费者对食品安全和健康的关注度日益提升，非预制菜消费者在选择食品时，首要考虑的因素便是食品的安全性和健康性。这包括食品原材料的来源、加工过程的卫生标准、添加剂的使用情况等多个方面。同时，口感的好坏也是影响消费者购买决策的重要因素。因此，对于非预制菜企业而言，确保产品的安全健康、提供优质的口感体验是吸引和留住消费者的关键。

建议2：在现代社会，人们的生活节奏越来越快，对于食品的便捷性需求也越来越高。非预制菜消费者在选择食品时，往往会考虑食品的便携性。这包括食品的大小、重量、包装等方面是否方便携带和食用。因此，非预制菜企业可以通过改进产品的包装设计、提高食品的便携性来满足消费者的这一需求，从而增加产品的吸引力。

建议3：性价比是消费者在购买商品时普遍关注的一个因素。非预制菜消费者也不例外。他们在购买食品时，会综合考虑产品的价格和质量，寻求最佳的性价比。这意味着非预制菜企业需要在保证产品质量的前提下，合理控制成本，提供价格合理、物有所值的产品。同时，企业还可以通过促销活动、会员优惠等方式提高产品的性价比，吸引更多消费者购买。

2）模型的现实意义

意义1：通过此模型，我们可以在以前从未尝试过预制菜的消费者中预测出有购买可能的预制菜消费者，进行潜在用户挖掘。

意义2：利用此模型，预制菜企业通过掌握消费者的一些基本信息，就能够精准预测消费者的购买意愿，从而精准定位潜在用户群体。这样的定位策略使得企业无需在广泛的市场中进行盲目推广，而只需针对潜在用户进行有针对性的预制菜推广。这种策略不仅显著降低了企业在人力、物力和财力上的投入，还能极大提高推广的效率和效果，确保每一分投入都能转化为实际的销售增长。通过精准推广，预制菜企业能够迅速扩大市场份额，吸引更多消费者，实现企业的快速发展。

意义3：在当今这个高速运转的社会中，快节奏和碎片化生活成为常态。人们追求高效，尤其是那些长期加班的职员和上班族，他们往往在紧张的工作节奏中忽视了健康饮食的重要性。亚健康状态逐渐成为这一群体的普遍现象。预制菜模型的出现，正是为了回应这一生活挑战。通过精准识别新用户并挖掘潜在用户，预制菜不仅提供了便捷的就餐选择，还确保了营养与口感的双重满足。这样，即使在忙碌的工作间隙，消费者也能享受到健康、美味的餐食，满足高效生活的餐饮需求。

# 对预制菜行业的结论与建议

## 主要结论

#### **1、消费者对于预制菜认知程度不足，抱有偏见**

在调查过程中我们不难看出，部分消费者对预制菜的认知程度不足，先入为主地认为预制菜高油高盐、不利于健康。食品安全与健康是阻碍消费者购买预制菜的重要因素。

事实上，预制菜是一种便捷、安全和经济的选择。首先，预制菜经过严格的质量控制和卫生检查，通常比在家中制作更为安全。其次，由于冷链物流技术的改善和发展，预制菜的加工和保存技术不断提升，使得其保持新鲜度和营养成分的能力逐渐增强。

#### **2、在众多影响消费者购买食品的因素中，食品安全、营养与健康、口感及价格首当其冲**

由统计结果可以看出，在十个影响消费者购买食品的因素中，食品安全、营养与健康排在首位，口感与价格紧随其后。

食品安全是民生工程、民心工程，我国政府高度重视食品安全问题，随着食品安全问题一次次登上舆论的风口浪尖，消费者也越来越重视食品安全问题，这一因素成为消费者在购买食品时最看重的因素；其次，由于人们生活水平提高以及健康意识的增强，人们对食品中各种营养成分以及食物功效更加重视，因此营养与健康排在食品购买考虑因素；第三，人们生活质量的提高也对食品口感提出了更高要求，但是在当前追求极致性价比的时代，人们不仅关注食品安全、营养与健康、口感等因素，也要求更加实惠的价格。因此口感和价格也排在食品购买考虑因素前列。

#### 3、预制菜在便利程度以及价格方面迎合了消费者的需求，但是在食品安全保障方面以及口感味道方面还有较大的提升空间

预制菜在当今社会中成为了许多消费者日常生活中的便捷选择，主要得益于其在便利程度和价格方面的优势。这种方便性不仅满足了现代快节奏生活的需求，同时也为那些缺乏烹饪时间或技能的人提供了一种方便而美味的饮食选择。预制菜的广泛受欢迎，部分原因在于其提供了即时的解决方案，使得人们能够更加轻松地满足饥饿感，减轻了繁忙生活中的压力。

然而，尽管预制菜在便利性和价格上有诸多优势，却仍然面临着一些挑战，其中之一是食品安全保障的问题。从问卷结果和大数据文本挖掘中我们可以看出，很多消费者依然担心预制菜食品安全保障问题。

此外，尽管预制菜在提供便利的同时努力保持口感和味道，但一些消费者仍然认为其在这方面有改进的空间。提升预制菜的口感和味道，不仅可以增加其吸引力，还能够满足更广泛的消费者口味的需求。

#### 4、消费者食品消费观念的转换会影响其预制菜消费行为

在懒宅经济和后疫情时代背景下，我们发现大部分消费者与三年前相比食品消费观念发生了如下转变：更多地在网上购买食品、对食品的需求更加多元化和个性化、更在乎信息透明和消费者权益保护。

上述食品观念的转变影响了预制菜消费行为。例如，预制菜企业由传统的线下售卖转变为线上线下同时销售；消费者购买预制菜有更强的个性化、多元化趋势，例如处于减脂期的消费者更加青睐于低脂低糖高蛋白的预制菜等；消费者更倾向于购买透明度高以及可追溯食材、添加剂的预制菜品。

#### 5、预制菜宣传手段中作用最突出的为用户口碑宣传

在互联网时代，大多数商品的宣传手段由传统的广告宣传转为当前的线上宣传，对于预制菜行业来说，最突出的宣传手段是用户口碑宣传。用户口碑宣传指的是消费者通过口口相传、社交媒体分享、在线评论等方式传播对产品或服务的正面评价和推荐，从而影响其他潜在消费者的购买决策。

用户口碑宣传能够帮助预制菜企业建立信任，具有较高的市场影响力。消费者更倾向于相信其他消费者的真实体验，而当他们看到其他人对预制菜产品的积极评价时，会更加倾向于选择购买。通过用户口碑宣传，预制菜企业可以逐步塑造良好的品牌形象。当消费者不断传播积极的使用体验和评价时，预制菜品牌的形象也会随之提升，进而吸引更多消费者的关注和认可。

因此，针对预制菜企业而言，有效地管理和利用用户口碑宣传是非常重要的，可以帮助企业建立信任、提升市场影响力，同时也有助于塑造良好的品牌形象，促进企业的发展和壮大。

## 提出建议

#### **1、政府层面**

#### **（1）完善预制菜行业监管体系和追溯体系，切实保障食品安全问题**

政府应建立健全预制菜标准体系，完善预制菜食品安全国家标准，推进预制菜推荐性国家标准制定，按照预制菜产品原料、加工工艺等的不同，在各个领域颁布适宜的推荐性国家标准；推动预制菜地方标准制定，例如重庆发布《预制菜产业园区建设指南》《预制菜生产加工行为规范》等2项地方标准和《预制菜生产经营安全监管标准体系》工作指导性文件，有力地支撑重庆预制菜产业集群化升级发展，为推进产业标准化、规范化发展夯实技术基础。

政府应规范预制菜追溯体系建设，推动预制菜追溯标准制订，政府通过制定相关法律法规或政策文件，明确食品追溯标准的必要性和重要性，并提出相应的支持措施，如资金支持、技术指导等，以促进食品追溯标准的制定和实施；在某些地区或企业率先开展预制菜食品追溯标准的试点示范工作，总结经验，不断完善标准，逐步推广应用到整个行业中。

政府应严格预制菜生产经营监管。严格生产经营许可管理，对预制菜生产经营者进行严格审查；加强预制菜生产经营者监督检查，督促预制菜生产企业加强对原辅料采购、生产场所、设备设施、人员制度、贮存运输、追溯记录、标签标识等关键环节的控制，保障预制菜类产品的食品安全；督促企业落实食品安全主体责任。各地市场监管部门督促预制菜生产企业严格落实食品安全主体责任，按照《企业落实食品安全主体责任监督管理规定》要求，细化生产加工各个环节的职责，把责任落实到人。建立预制菜行业组织和协会。

#### **（2）建立行业组织和协会**

政府积极推动预制菜行业建立行业协会和组织，促进行业内部合作和交流，推动行业的健康发展和提升。

建立行业组织和协会，首先可以促进行业内部的交流与合作，推动技术创新、产品研发和市场开拓，有利于行业的健康发展；其次行业组织与协会参与制定行业内的自律性标准和规范及政策制定和法规修订，规范行业内部的行为和操作，监督行业内部的自律执行，维护行业的稳定和健康发展，有利于提升产品质量和服务水平。最后行业组织协会也促进了行业内部企业之间的合作与共享，推动资源整合和互惠互利，提升行业整体竞争力。

#### **（3）预制菜产业发展与乡村振兴战略紧密结合**

预制菜产业发展与乡村振兴战略密切相关，可以成为推动农村经济发展、实现乡村振兴的重要途径之一。政府可以从以下几个方面加以引导：

其一，支持农村产业发展。 预制菜产业需要大量的原材料供应，包括蔬菜、肉类等农产品。因此，发展预制菜产业可以带动农村农产品的生产，推动农村产业结构升级和农业现代化发展。

其二，促进农产品加工与销售。 预制菜产业的发展可以促进农产品的深加工和价值提升，通过加工、包装和销售预制菜，可以创造更多的农村就业机会，增加农民收入，提高农村经济发展水平。

其三，优化农村经济布局： 通过引导预制菜企业向乡村地区扩展投资和生产基地建设，可以优化农村经济布局，促进农村产业升级和转型发展，推动农村产业结构调整和品牌建设，实现城乡经济协调发展。

其四、推动乡村旅游与文化产业发展： 发展预制菜产业可以带动乡村旅游和文化产业的发展，通过打造特色农产品加工基地和农村休闲度假村，吸引更多游客和投资者到农村地区，推动乡村振兴战略的实施和落实。

#### **2、企业层面**

#### **（1）提高预制菜产业透明度**

根据国家发展改革委促消费“20条”中提出的预制菜发展模式——种养殖基地+中央厨房+冷链物流+餐饮门店，预制菜的信息化要注重提高这四大环节的透明度。

种养殖基地要提供“透明食材”，预制菜高质量发展，不仅要在制作工艺上进行创新，还要加强对农业食材的升级优化，把好质量第一关；

中央厨房升级为“透明厨房”，中央厨房作为餐饮和食品产业链的一种特色生产供应模式，上游连接“三农”，下游连接餐桌，实现了对传统食品产业链的全新组织和变革，开启了预制菜产业发展的新纪元。制定中央厨房质量评价规范，树立中央厨房品质标准，确保消费者能够放心购买、安心食用，从而更好满足人民群众日益多元的预制菜消费需求。

冷链物流实现“透明车间”。预制菜品质好坏的关键，在于预制菜到达消费终端时的新鲜度和口味，冷链物流是预制菜从食材、半成品、成品再到消费餐桌全流程品质保证的关键。企业通过建立冷链物流信息平台，让消费者能够查询商品的产地信息、安全信息、储存信息、“移动冷库”数据信息等，保证信息透明化。

餐饮门店提供“透明消费”。通过问卷调查结果以及网络信息爬取不难得知，消费者对于线下餐馆以及外卖中使用预制菜但是未提前告知的情况感到不满。当前消费者权益保障问题屡次登上风口浪尖，确保消费者的知情权和选择权至关重要，因此各餐馆和外卖应共同努力，建构并实施预制菜提前告知制度，打造“透明消费”势在必行。

#### **（2）提升预制菜口感**

企业应在预制菜口感研发方面加大投资力度，保证预制菜口感的稳步提升。预制菜企业应选择新鲜、优质的食材，与供应商建立稳定的合作关系，确保原材料的质量和新鲜度；同时不断创新配方和菜品搭配，结合当地特色和消费者口味，推出新颖美味的预制菜品；预制菜企业还应建立完善的顾客反馈机制，及时收集和处理顾客的意见和建议，不断改进产品口感和品质。

#### **（3）顺应消费者食品消费观念的转变，与时俱进创新预制菜产品**

随着互联网时代的发展，线上购物已经成为消费的主要方式。因此预制菜应采用线上线下双结合的销售方式。

针对线上销售，预制菜企业可以在知名电商平台上开设官方旗舰店或店铺，如淘宝、天猫、京东等，通过线上渠道销售预制菜产品；其次，预制菜企业可以利用社交媒体平台，如微信公众号、微博、抖音等，发布产品信息和促销活动，引导用户通过社交媒体购买；第三，预制菜企业通过搭建自己的官方网站，提供在线购买预制菜的功能，结合搜索引擎优化（SEO）和内容营销，吸引流量并促进销售。

针对线下销售，预制菜企业可以在城市繁华商圈或写字楼区域开设实体门店或专柜，推出即时取餐和品尝体验服务，也可以与超市或便利店合作，将预制菜产品放置在超市或便利店的冷藏柜或特定货架上销售。

其次，针对不同的消费群体做出更具个性化的预制菜产品。预制菜企业应基于消费者的生活习惯、健康需求、口味偏好等因素细分消费者，设计更具个性化的预制菜产品，以迎合不同消费者的口味偏好和饮食习惯。例如，针对上班族，可以设计快速便捷型健康盒饭，如5分钟快速加热即可食用的健康盒饭，注重营养均衡，如低脂高蛋白的鸡胸肉蔬菜盒，适合忙碌的上班族中午快速用餐；针对健身人群推出高蛋白、低脂肪的健身餐，如烤鸡胸肉、蒸鱼等，帮助健身人群补充能量，促进肌肉恢复；针对老年群体推出软食易消化套餐，由于咀嚼能力较弱的特点，提供软食、易消化的预制菜，如蒸鱼、肉末蒸蛋等，注重食材的温和和营养。

#### **（4）加大力度投入用户口碑宣传**

由问卷调查结果可知，对于预制菜行业来说，最有效的宣传手段是用户口碑宣传。因此预制菜企业应加大力度投入用户口碑宣传。例如通过社交媒体做营销，利用社交媒体平台（如小红书、微博、微信公众号等）发布用户分享的照片、评论和视频，展示他们对产品的使用体验和好评，吸引更多用户关注和参与互动；

给分享评论或推荐亲朋好友购买的用户优惠券和折扣：提供给用户优惠券或折扣码，作为鼓励他们购买并分享产品的奖励。这样的促销活动既可以增加销量，也可以激励用户积极参与口碑宣传；与合作伙伴合作推广，合作伙伴可以是一些有影响力的博主、社交媒体上的网红、业内专家或其他相关品牌。通过与合作伙伴合作推广，品牌可以获得更多的曝光率和更多的信任度；提供优质客服服务，建立良好的客户服务团队，及时回复用户的问题和反馈，解决用户的疑虑和困扰，提升用户满意度，促进口碑传播。

## 参考文献

[1] 人民网研究院.预制菜行业发展报告

<http://yjy.people.com.cn/n1/2023/0710/c440911-40031856.html>.2023年7月11日发布.

[2]中共中央.国务院关于做好2023年全面推进乡村振兴重点工作的意见

<http://www.lswz.gov.cn/html/xinwen/2023-02/13/content_273655.shtml>.

2023年2月13日发布.

[3]TAPP S M. Taste panel acceptance of mechanically deboned meat in prepared dishes[D]. Oklahoma: Oklahoma State University,1978:1-4.

[4] RYLEY J. Nutrition and catering:2.The nutritional quality of school meals made from convenience foods compared with traditionally prepared dishes[J].Human Nutrition-Applied Nutrition,1984,38(1):42-49.DOI:10.1016/0306-9192(84)90075-7.

[5] HONG J,JEONG B G,CHUN J, et al. Folate content of Korean vegetable dishes prepared outside the home: comparison between analyzed and calculated values[J].Journal of Food Composition and Analysis,2021,103:104-108.DOI:10.1016/j.jfca.2021.104088.

[6]王卫,张锐,张佳敏等.预制菜及其研究现状、存在问题和发展展望[J].肉类研究,2022,36(09):37-42.

[7]王娟,高群玉,娄文勇.我国预制菜行业的发展现状及趋势[J].现代食品科技,2023,39(02):99-103.DOI:10.13982/j.mfst.1673-9078.2023.2.0388.

[8]张宇昊,陈海.川渝预制菜产业现状与发展路径分析[J].中国食品学报,2022,22(10):9-19.DOI:10.16429/j.1009-7848.2022.10.002.

[9]王绪模,朱克永.顺应大势 形成合力 促进川菜预制菜产业加快发展[J].现代食品,2022,28(05):49-52.DOI:10.16736/j.cnki.cn41-1434/ts.2022.05.013.

[10]孙宝国,王静.中国传统食品现代化[J].中国工程科学,2013,15(04):4-8+2.

[11]王卫,张佳敏,赵志平等.川菜肉类菜肴工业化及其关键技术[J].肉类研究,2020,34(05):98-103.

[12]税小林,孙钦秀,夏秋瑜等.预制菜包装技术的发展概况及趋势[J].包装工程,2023,44(13):132-140.DOI:10.19554/j.cnki.1001-3563.2023.13.016.

[13]肖欢,曹宏,陈士强等.预制菜的卫生安全与辐照技术在其应用进展[J].核农学报,2023,37(07):1428-1434.

[14]吴晓蒙,饶雷,张洪超等.新型食品加工技术提升预制菜肴质量与安全[J].食品科学技术学报,2022,40(05):1-13.

[15]王瑞. 基于逻辑回归模型的员工离职问题研究[D].昆明理工大学,2022.DOI:10.27200/d.cnki.gkmlu.2021.000194.

[16] 国务院办公厅. 国务院办公厅转发国家发展改革委关于恢复和扩大消费措施的通知<https://www.gov.cn/zhengce/content/202307/content_6895599.htm> 2023年7月28日发布.