



# 上海商学院

## “云拿杯”数据分析大赛报告

论文题目 商业决策与数据分析——以云拿无人便利  
店为例

学生姓名 何贝尔、蒋承浩、张逸晨、陆畅、朱心宁

指导教师 毛志新

---

## 目录

目录 .....	1
摘要 .....	1
第 1 章 数据来源及数据预处理 .....	1
1.1 数据来源 .....	1
1.2 数据预处理 .....	1
1.3 数据概述 .....	1
1.4 数据背景 .....	4
第 2 章 现状分析 .....	5
2.1 退款订单分析 .....	5
2.1.1 筛选所有实际发生金钱交易的订单数据 .....	5
2.1.2 支付成功且无退款订单分析 .....	6
2.1.3 退款订单分析 .....	10
2.1.4 退款用户分析 .....	12
2.2 用户进店时间、时长与行为分析 .....	15
2.2.1 顾客进店时间分布 .....	15
2.2.2 顾客进店时长分布 .....	16
2.2.3 每日客流量分布 .....	17
2.2.4 每周客流量分布 .....	18
2.3 一级品类的商品占市场份额分析 .....	19
2.3.1 一级品类的商品近 7 日销售数量之和比较分析 .....	19

---

2.3.2 一级品类的商品毛利之和比较分析 .....	20
第 3 章 影响因素分析 .....	21
3.1.1 每日客流量是否影响该日商品的拿起与放下次数 .....	21
3.1.2 米饭品类供应商是否影响近 7 日销售数量 .....	23
3.1.3 米饭品类拿起、放下次数，实收金额与近 7 日销售数量的 相关关系 .....	23
3.1.4 米饭品类拿起、放下次数与近 7 日销售数量的影响因素分 析 .....	24
3.1.5 米饭品类实付金额与近 7 日销售数量的影响因素分析 .....	27
第 4 章 顾客类型分析 .....	29
4.1 用户类型趋势 .....	29
4.2 用户阶段 .....	30
4.3 用户分析 .....	31
第 5 章 总结和建议 .....	32
5.1 订单数据 .....	32
5.2 闸机流水数据 .....	33
第 6 章 数据问题汇总 .....	33
6.1 订单列表 .....	33
6.2 用户标签表 .....	33
6.3 闸机流水表 .....	33

---

## 摘要

本文以云拿便利店上海商学院徐汇校区店 2022 年 9 月 1 日至 2022 年 11 月 30 日的商品销售、用户行为等信息为数据基础，通过订单金额、订单时间等指标的累计分布和退款订单、支付且未退款记录订单的对比分析，分析了便利店高频订单和高频退款的时间段，总结便利店应当做出的对应行为。运用描述性分析、影响因素分析、相关关系分析等，分析得出便利店客流量规律、特定品类的拿起放下次数、供应商的选择、实付金额对商品销量影响情况，以期为云拿便利店未来的商业决策提供建议与参考。

**关键词：**无人便利店；数据分析；商业决策

---

## 第1章 数据来源及数据预处理

### 1.1 数据来源

本次数据由上海云拿智能科技有限公司提供，包含九张表，除用户标签表没有时间字段外，其余八张表都记录了 2022 年 9 月 1 日-2022 年 11 月 30 日的数据。

### 1.2 数据预处理

对于给定的九张表格，为了便于进行分析，我们删除了全局唯一值的字段"商店 id"和"渠道"，同时将相同编码方式的字段名进行统一，将 "会员号"修改成 "用户 id"。

### 1.3 数据概述

以下为预处理后的表格数据：

订单列表，共 54176 行数据，11 列。字段为订单号、会员号、交易单号、订单金额、余额支付金额、待付金额、优惠金额、订单类型、支付时间、订单状态、创建时间。其中订单号、会员号、交易单号、订单类型、订单状态为定性数据，订单金额、余额支付金额、待付金额、优惠金额为定量数据，支付时间、创建时间为定时数据。

订单商品详情，共 89107 行数据，13 列。字段为订单 ID、商品 ID、商品名称、一级品类、二级品类、三级品类、商品数量、商品单价、该商品实付总金额、该商品展示总金额、商品税、商品税率、创建时间。其中订单 ID、商品 ID、商品名称、一级品类、二级品类、三级品类为定性数据，商品数量、商品单价、该商品实付总金额、该商品展示总金额、商品税、商品税率为定量数据，创建时间为定时数据。

---

店铺自助报表进店(天)共有 36748 行数据，13 列。字段为：用户类型、日期、进店人次、进店人数、进店未消费人次、进店未消费人数、进店消费人次、进店消费人数。其中用户类型为定性数据，进店人次、进店人数、进店未消费人次、进店未消费人数、进店消费人次、进店消费人数为定量数据，日期为定时数据。

店铺自助报表服务(天)共有 147 行数据，8 列。字段为：用户类型、日期、线下投诉量、线上退款订单量、线上退款商品数量、线上退款金额、线上退款商品成本、线上退款商品不含税成本。其中用户类型为定性数据，线下投诉量、线上退款订单量、线上退款商品数量、线上退款金额、线上退款商品成本、线上退款商品不含税成本为定量数据，日期为定时数据。

店铺自助报表销售(天)共有 500 行数据，21 列。字段为：用户类型、业务类型、日期、销售原价、优惠金额、实收金额、商品销量、销售商品成本、毛利、订单量、客单价、拿起次数、放下次数、不含税销售原价、销售商品不含税成本、不含税毛利、实收金额毛利、净收入、实收金额客单价、首购金额、首购订单数。其中用户类型、业务类型为定性数据，销售原价、优惠金额、实收金额、商品销量、销售商品成本、毛利、订单量、客单价、拿起次数、放下次数、不含税销售原价、销售商品不含税成本、不含税毛利、实收金额毛利、净收入、实收金额客单价、首购金额、首购订单数为定量数据，日期为定时数据。

栏位操作流水共有 531560 行数据，11 列。字段为：流水号、栏位 id、栏位默认商品、商品名称、商品 id、变化前数量、变化后数量、业务类型、操作人、操作时间、创建时间。其中流水号、栏位 id、栏位默认商品、商品名称、商品 id、业务类型、操作人为定性数据，变化前数量、变化后数量为定量数据，操作时间、创建时间为定时数据。

商品自助报表(全指标按日)共有 148913 行数据，63 列。字段为：日期、商品 ID、商品名称、一级品类、二级品类、三级品类、采购入库数量、采购

---

入库含税总价、采购入库不含税总价、采购退货数量、采购退货含税总价、采购退货不含税总价、损耗数量、损耗含税总价、损耗含税售价、调拨入库数量、调拨入库含税总价、调拨入库不含税总价、结存数量、结存含税成本、结存含税成本、强调入库数量、强调出库数量、强调入库成本、强调入库不含税成本、强调出库成本、强调出库不含税成本、商品拿起次数、商品放下次数、近 7 日销售数量、周转天数、最近供应商、最近供应商名称、最近入库的含税单价、最近入库的不含税单价、销售原价、优惠金额、实收金额、商品销量、销售商品成本、毛利、订单量、客单价、不含税销售原价、销售商品不含税成本、不含税毛利、实收金额毛利、净收入、退款订单量、退款商品数量、退款金额、退款商品成本、退款商品不含税成本、未付款订单量、未付款商品数量、未付款金额、未付款商品成本、未付款商品不含税成本、补付款订单、补付款数量、补付款金额、补付款商品成本、补付款商品不含税成本。其中商品 ID、商品名称、一级品类、二级品类、三级品类、最近供应商、最近供应商名称为定性数据，采购入库数量、采购入库含税总价、采购入库不含税总价、采购退货数量、采购退货含税总价、采购退货不含税总价、损耗数量、损耗含税总价、损耗含税售价、调拨入库数量、调拨入库含税总价、调拨入库不含税总价、结存数量、结存含税成本、结存含税成本、强调入库数量、强调出库数量、强调入库成本、强调入库不含税成本、强调出库成本、强调出库不含税成本、商品拿起次数、商品放下次数、近 7 日销售数量、周转天数、最近入库的含税单价、最近入库的不含税单价、销售原价、优惠金额、实收金额、商品销量、销售商品成本、毛利、订单量、客单价、不含税销售原价、销售商品不含税成本、不含税毛利、实收金额毛利、净收入、退款订单量、退款商品数量、退款金额、退款商品成本、退款商品不含税成本、未付款订单量、未付款商品数量、未付款金额、未付款商品成本、未付款商品不含税成本、补付款订单、补付款数量、补付款金额、补付款商品成本、补付款商品不含税成本为定量数据，日期为定时数据。

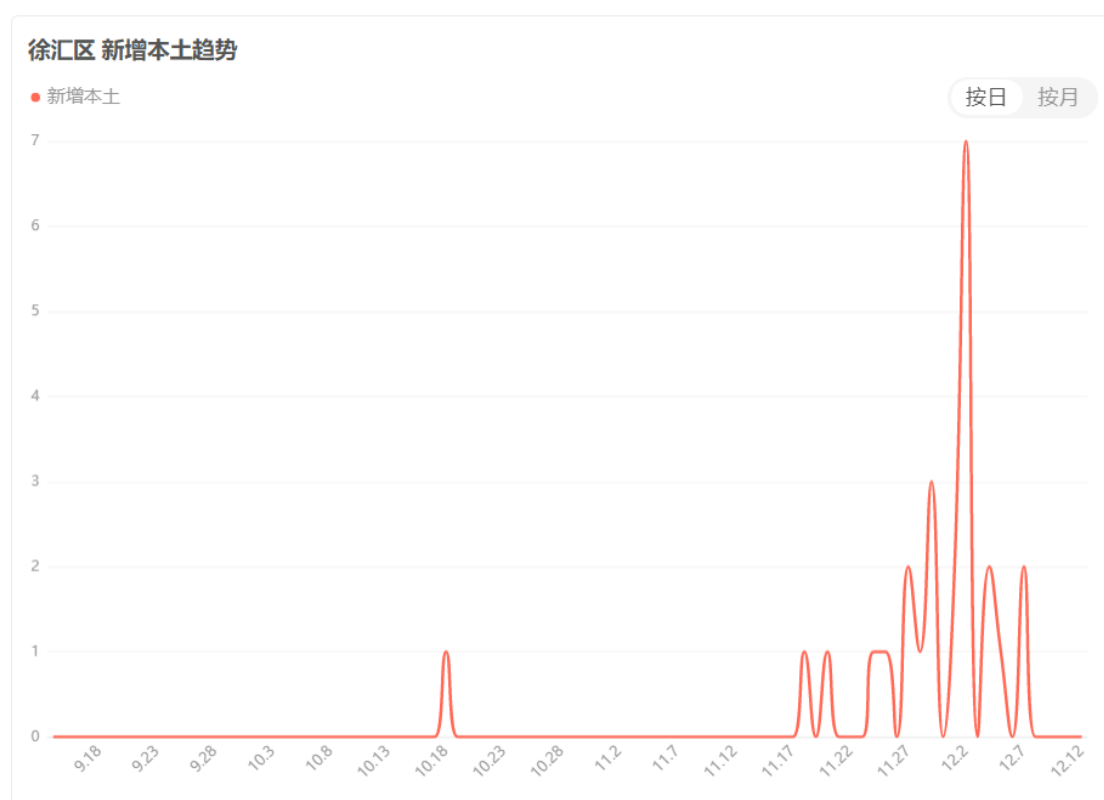
用户标签共有 26331 行数据，3 列。字段为：用户 ID、标签类型、标签值，均为定性数据。

闸机流水共有 163605 行数据，9 列。字段为：流水 id、闸机 id、闸机类型、用户 id、用户类型、顾客 id、动作时间、动作类型、创建时间。其中流水 id、闸机 id、闸机类型、用户 id、用户类型、顾客 id、动作类型为定性数据，动作时间、创建时间为定时数据。

以上六张表经过预处理后，数据记录均合理，无异常值，可以进行数据分析。

## 1.4 数据背景

本数据为云拿上海徐汇商学院店，店址位于校内，因此主要顾客为学生和老师，遵从学校作息规律。根据百度数据区域疫情新增趋势，数据记录时间内徐汇区疫情均控制良好，因此销售数据受疫情干扰不大。



图表 1.1 百度数据 上海市徐汇区 2022.9.18-2022.12.12 疫情新增人数变化趋势



---

## 第2章 现状分析

### 2.1 退款订单分析

#### 2.1.1 筛选所有实际发生金钱交易的订单数据

通过筛选订单列表的订单状态字段，“退款成功”的订单有 267 条，“支付成功”的订单有 53883 条。利用 Access 创建一个新表（如图表 2.1），命名为收退款成功订单列表（如图表 2.2），统计发现共计购买了 94913 件商品，退款了 341 件商品。通过关联订单创建时间、用户 id 和商品名，可以将退款的订单和支付订单关联，标记出退款的商品。经过处理，发现有六条商品退款记录（如图表 2.3）无法关联购买记录，同一用户未在退款订单前购买过这些商品，因此将这六条记录删除，后续退款订单相关分析仅分析可以关联支付记录的退款记录。

```
SELECT 订单列表.订单号, 订单列表.会员号, 订单列表.订单金额, 订单列表.余额支付金额, 订单列表.优惠金额,
订单列表.订单状态, 订单商品详情.商品名称, 订单商品详情.商品数量, 订单商品详情.商品单价, 订单列表.创建时间
FROM 订单列表 INNER JOIN 订单商品详情 ON 订单列表.订单号 = 订单商品详情.订单ID
WHERE (((订单列表.订单状态)="退款成功" Or (订单列表.订单状态)="支付成功"))
ORDER BY 订单列表.会员号, 订单商品详情.商品名称, 订单列表.创建时间;
```

图表 2.1Access 查询 code

订单号	会员号	订单金	余额支	优惠金	订单状	是否退	商品名	商品数	商品单	创建时
220926SHJ	18030614	3.80	0.00	0.38	支付成功	0	可口可乐	1	3.80	2022-09-26 15:46:0
221114SHJ	18083019	6.50	0.00	0.65	支付成功	0	川味肥牛	1	6.50	2022-11-14 13:31:2
221024SHJ	18083019	7.80	0.00	0.78	支付成功	0	乐事自然	1	7.80	2022-10-24 09:34:0
220919SHJ	18083019	5.00	0.00	0.50	支付成功	0	李子园哈	1	5.00	2022-09-19 12:40:3
220922SHJ	18093015	2.00	1.80	0.20	支付成功	0	农夫山泉	1	2.00	2022-09-22 11:38:2
221108SHJ	18093015	4.00	0.00	0.00	支付成功	1	维他豆奶	1	4.00	2022-11-08 21:08:5
221108SHJ	18093015	4.00	0.00	0.00	退款成功	1	维他豆奶	1	4.00	2022-11-08 21:10:1
221108SHJ	18093015	4.00	0.00	0.40	支付成功	1	维他豆奶	1	4.00	2022-11-08 21:40:2
221108SHJ	18093015	3.60	0.00	0.00	退款成功	1	维他豆奶	1	3.60	2022-11-08 21:42:0
221108SHJ	18093015	3.00	0.00	0.00	支付成功	1	鲜肉大包	1	3.00	2022-11-08 21:07:1
221108SHJ	18093015	3.00	0.00	0.00	退款成功	1	鲜肉大包	1	3.00	2022-11-08 21:10:3
221108SHJ	18093015	3.00	0.00	0.30	支付成功	1	鲜肉大包	1	3.00	2022-11-08 21:40:5
221108SHJ	18093015	2.70	0.00	0.00	退款成功	1	鲜肉大包	1	2.70	2022-11-08 21:41:5
220929SHJ	18111917	12.80	0.00	1.28	支付成功	0	川香大盘	1	12.80	2022-09-29 13:41:3
220927SHJ	18111917	20.80	0.00	2.08	支付成功	0	光明咕浓	1	4.00	2022-09-27 21:19:3
220929SHJ	18111917	2.00	0.00	0.20	支付成功	0	糯米烧卖	1	2.00	2022-09-29 13:57:2
220927SHJ	18111917	20.80	0.00	2.08	支付成功	0	甜辣鸭锁	1	16.80	2022-09-27 21:19:3
221022SHJ	18122918	1.80	0.00	0.18	支付成功	0	农夫山泉	1	1.80	2022-10-22 11:10:3
221022SHJ	19020807	4.20	0.00	0.42	支付成功	0	三得利乌	1	4.20	2022-10-22 07:59:2
221030SHJ	19020807	4.20	0.00	0.42	支付成功	0	三得利乌	1	4.20	2022-10-30 13:44:5
220902SHJ	19100515	12.90	0.00	1.29	支付成功	0	脆骨鸡肉	1	3.50	2022-09-02 09:07:1
220915SHJ	19100515	6.00	0.00	0.60	支付成功	0	脆骨鸡肉	1	3.50	2022-09-15 08:56:3
220902SHJ	19100515	12.90	0.00	1.29	支付成功	0	豆褶串	1	2.50	2022-09-02 09:07:1
220915SHJ	19100515	6.00	0.00	0.60	支付成功	0	豆褶串	1	2.50	2022-09-15 08:56:3
220906SHJ	19100515	11.50	0.00	1.15	支付成功	0	光明优+纯	1	6.50	2022-09-06 08:20:1
221027SHJ	19100515	6.00	0.00	0.60	支付成功	0	虎皮卷蛋	1	6.00	2022-10-27 08:25:3
221017SHJ	19100515	8.30	0.00	0.83	支付成功	0	金厨娘32	1	2.80	2022-10-17 12:38:1

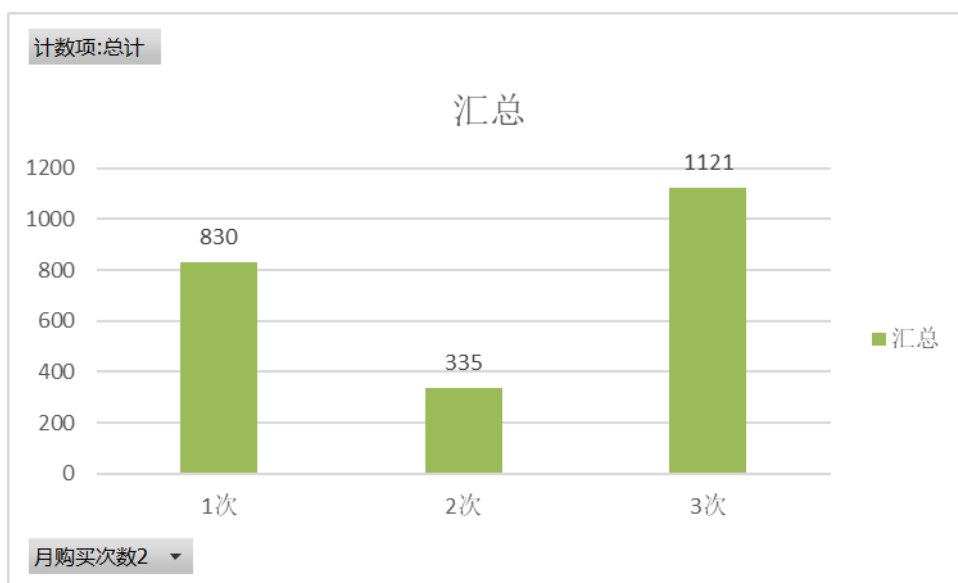
图表 2.2 收退款成功订单列表

订单号	会员号	订单金	余额支	优惠金	订单状	是否退款关联	商品名	商品数	商品单	创建时
220901SHJ	210106220323891UA099404	12.42	0.00	0.00	退款成功		0 奥尔良鸡	1	12.42	2022-09-01 18:39:55.0
220902SHJ	210602124457183UA171732	9.00	0.00	0.00	退款成功		0 BZ快趣果	1	9	2022-09-02 20:02:25.0
220901SHJ	220828100125486UA425737	1.80	0.00	0.00	退款成功		0 怡宝纯净	1	1.8	2022-09-01 17:14:38.0
221128SHJ	220829100704277UA426856	6.30	0.00	0.00	退款成功		0 健达巧克	1	6.30	2022-11-28 18:51:22.0
220905SHJ	220829132655701UA427169	9.00	0.00	0.00	退款成功		0 七度空间	1	9.00	2022-09-05 17:17:49.0
220906SHJ	220829132655701UA427169	18.00	0.00	0.00	退款成功		0 七度空间	2	9.00	2022-09-06 08:15:01.0

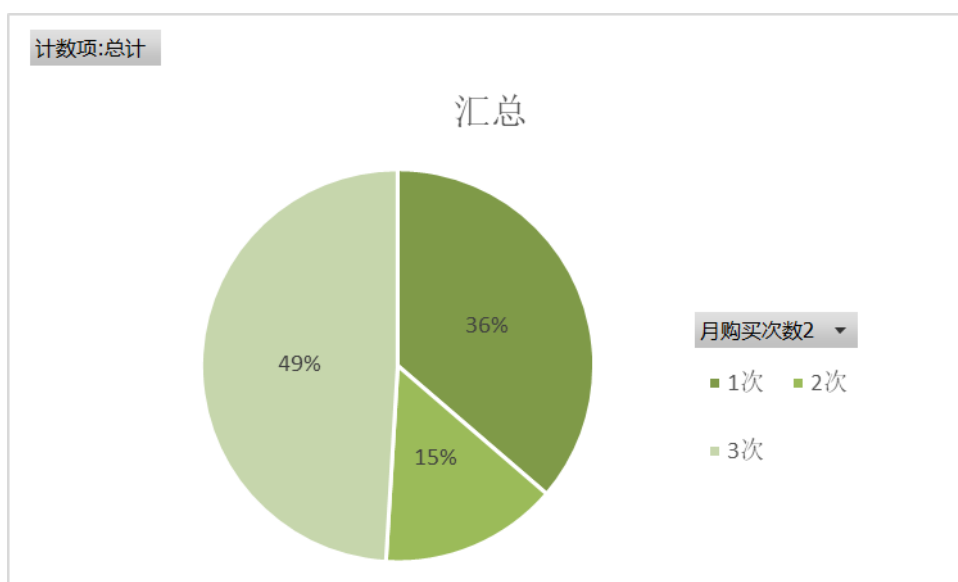
图表 2.3 无法关联购买记录的退款订单

### 2.1.2 支付成功且无退款订单分析

基于支付成功且无退款订单，如图表 2.4 展示，有 2286 人至少有一个月购买过商品，有 1121 人每个月都购买商品。如图表 2.5 展示，两个月及以上购买商品的人数占总购买人数的 64%，每个月都购买商品的人数占总购买人数的 49%。因此可以得出，商店月留存率为 64%，每月会有 1121 位左右的顾客前来购物，顾客粘性大，收益稳定。



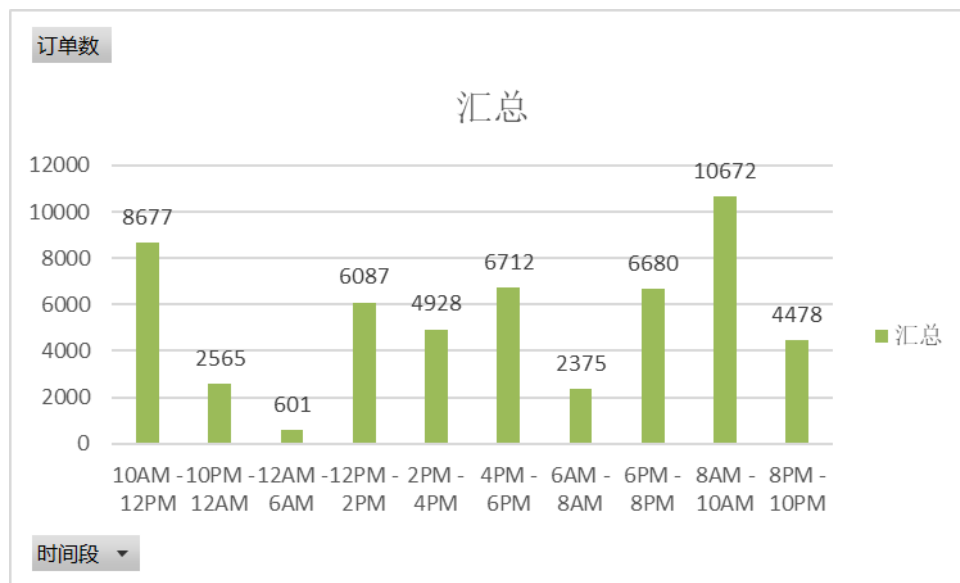
图表 2.4 用户在 9-11 月中有几个月购买过商品统计



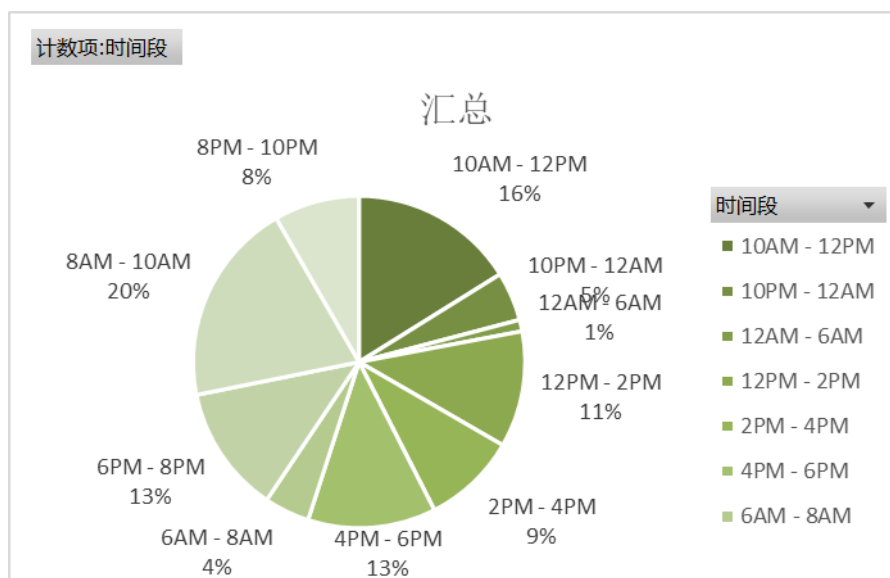
图表 2.5 用户在 9-11 月中有几个月购买过商品的次数占比

将订单创建时间按两个小时分段(0 点-6 点因订单量过少合并为一段), 分时段统计订单数(如图表 2.6)和订单总金额(如图表 2.8)。如图表 2.7 和图表 2.9 展示, 订单在上午 8 点-12 点, 下午 4 点-8 点的数量占全天订单数量的 62%, 总金额占全天订单总金额的 61%。因此可以得出, 在上午 8 点-12 点, 下午 4 点-8 点这八个小时为商店主要盈利时间, 可以适当安排工作人员现场维护秩序, 解决问题, 也需注意在这两个时间段之前做好商品补

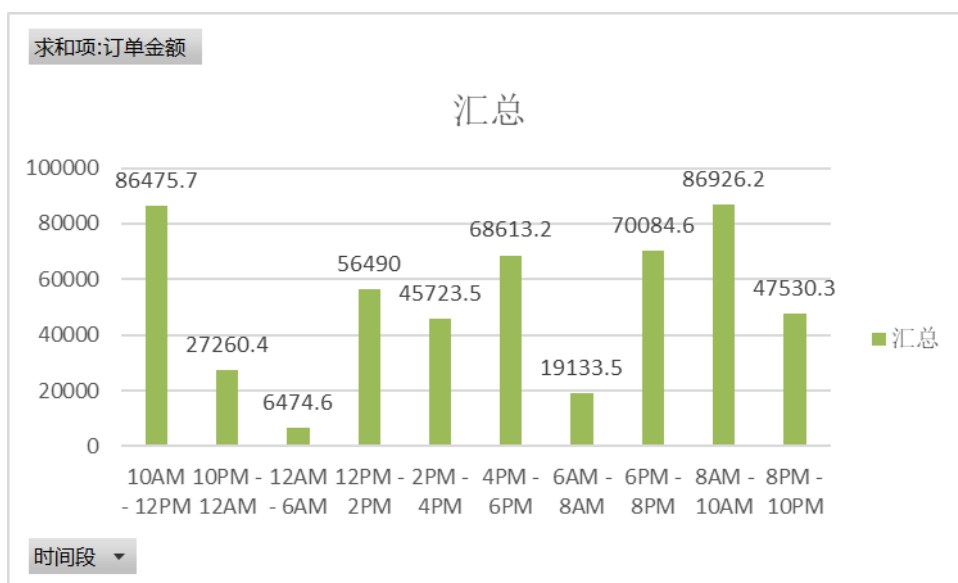
货、货架整理等工作。



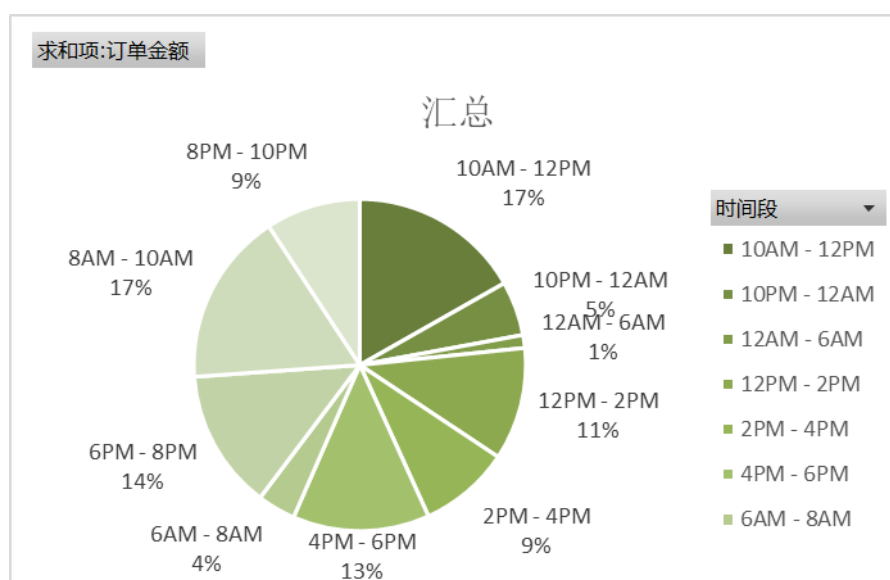
图表 2.6 分时段订单数



图表 2.7 分时段订单数占比



图表 2.8 分时段订单总金额



图表 2.9 分时段订单总金额占比

所有支付且未退款订单中，余额支付总金额在订单总金额的占比为 0.51%，余额支付订单数在总订单数的占比为 0.65%。即几乎全部的用户都使用支付工具支付，会员余额制度的作用不大，可以稍微举办一些使用余额支付的优惠活动，提升使用余额支付的用户比例，稳固用户粘性。

在余额支付的用户中，有 92.05% 用户使用余额全额支付订单。余下 7.95% 的余额支付用户中可能存在用户余额不够才余额部分支付，可以对低余额用户增加消息提醒通知，提高余额系统的存在感。

总订单数

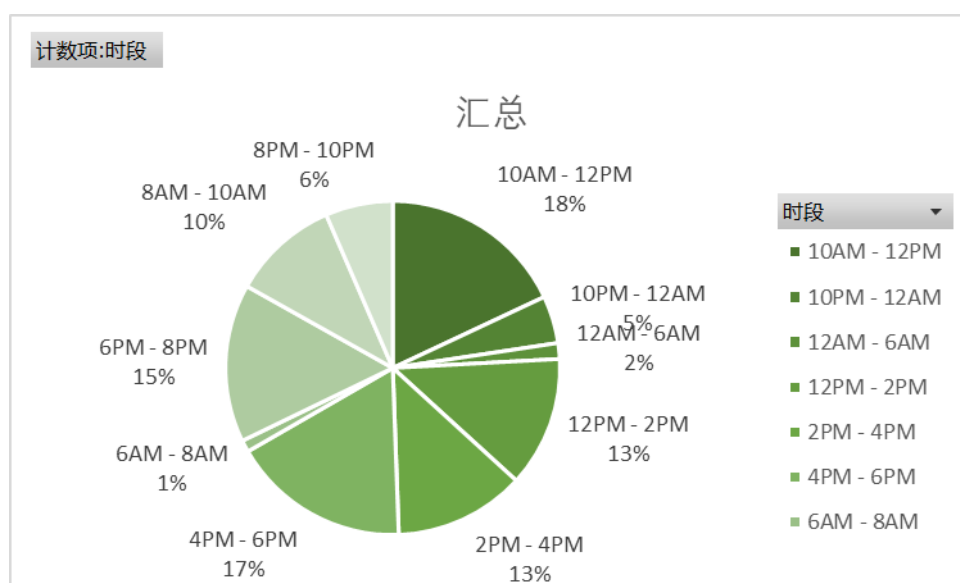
53775

总订单金额	514712
余额支付订单数	352
余额全额支付订单数	324
余额总支付金额	2612.54
余额支付订单数/总订单数	0.65%
余额总支付金额/总订单金额	0.51%
余额总支付金额/余额全额支付订单数	92.05%

表格 1 订单和余额支付订单统计

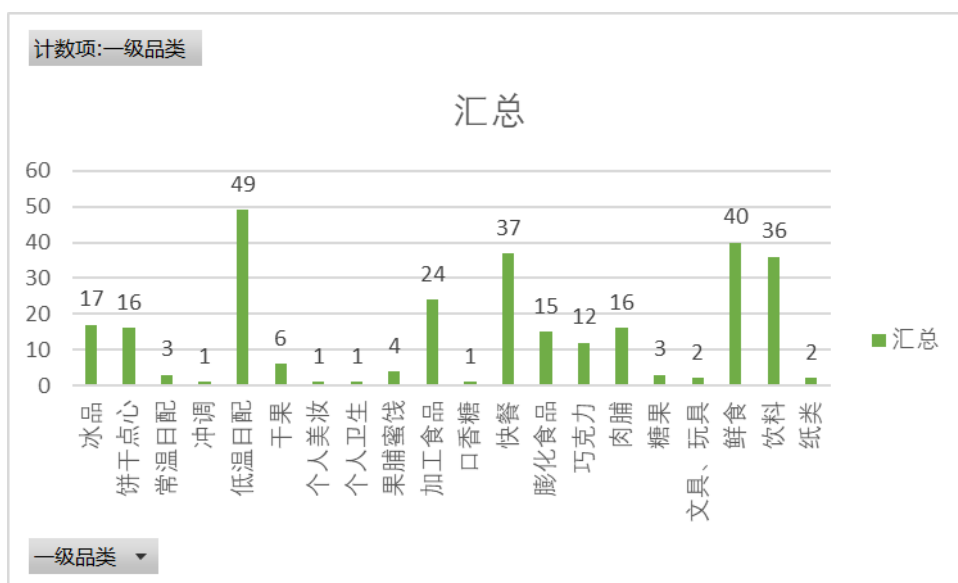
### 2.1.3 退款订单分析

经统计，在 9 月-11 月中，有 203 个用户发起了 261 个退款订单，退款了 198 种，共计 286 件商品。如图表 2.10 展示，在上午 10 点-下午 8 点的投诉订单数占了全天的 76%，而上午 10 点-12 点、下午 4 点-6 点两段饭点退款订单数最多，因此在饭点可以重点关注一下商店的服务质量，减少退款订单，使尽可能多的顾客满意从而提高顾客留存率。



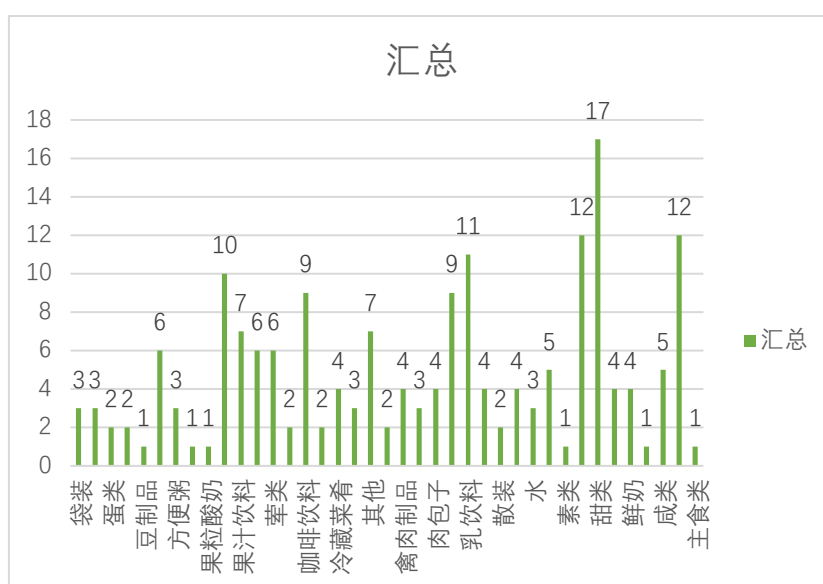
图表 2.10 分时段退款订单数

如图表 2.12 展示，退款商品主要分布在加工食品、低温日配、快餐、鲜食、饮料五个一级分类，共计 186 件商品，占总退款商品的 65.03%。



图表 2.12 退款商品一级分类统计

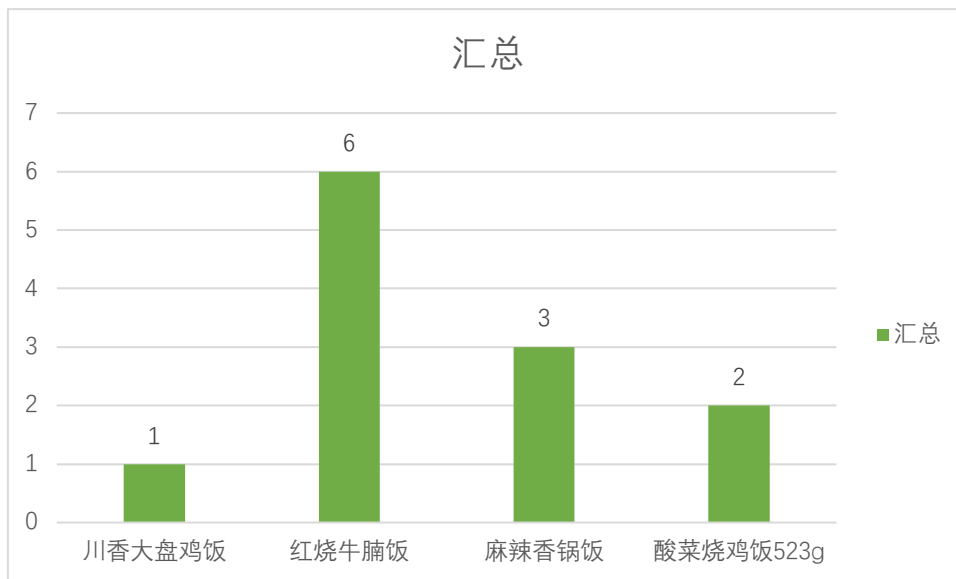
如图表 2.13 展示，在这五类退款商品中，退款商品主要分布在果汁、咖啡饮料、肉类香肠、乳饮料、酸奶、甜类、中式这七个三级分类，占五类退款商品的 43.55%。



图表 2.11 加工食品、低温日配、快餐、鲜食、饮料分类退款商品三级分类统计

无论是总退款商品的主要一级分类退款还是主要一级分类退款下的主要三级分类退款，均为食物饮品。食物饮品是便利店主要售卖品类，也是主流顾客在便利店消费最多的品类，因此更需要谨慎选品，频繁迭代产品。例如图表 2.14 展示，在中式分类下，红烧牛腩饭被退款 6 次，占中式分类退

款的一半，便可以再次考虑是否需要继续进货红烧牛腩饭，或者寻找新品替换红烧牛腩饭。



图表 2.12 中式分类退款商品统计

在尝试将数据和用户标签表联结时发现 203 位退款用户中有一位用户（如图表 2.15）没有用户标签。经过查询闸机流水发现，这位用户是员工身份。因此在退款记录中存在部分记录为员工测试而非真实购物退款行为。

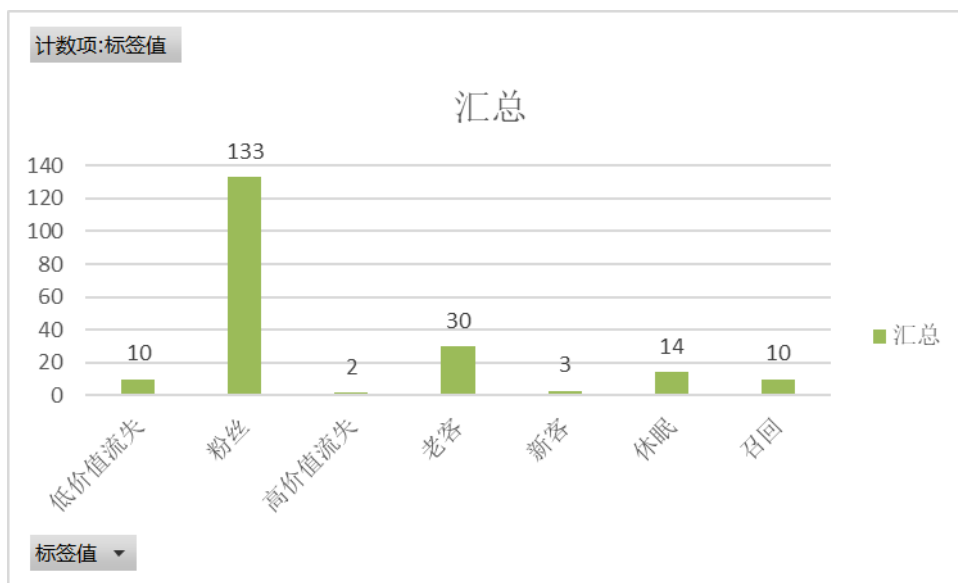
订单号	会员号	订单金	余额支	优惠金	订单状	是否退款关联	商品名	商品数	商品单	创建时间
221108SH	180930153754951UA006078	4.00	0.00	0.00	支付成功		1 维他豆奶(	1	4.00	2022/11/8 21:08
221108SH	180930153754951UA006078	4.00	0.00	0.00	退款成功		1 维他豆奶(	1	4	2022/11/8 21:10
221108SH	180930153754951UA006078	4.00	0.00	0.40	支付成功		1 维他豆奶(	1	4	2022/11/8 21:40
221108SH	180930153754951UA006078	3.60	0.00	0.00	退款成功		1 维他豆奶(	1	3.6	2022/11/8 21:42
221108SH	180930153754951UA006078	3.00	0.00	0.00	支付成功		1 鲜肉大包	1	3	2022/11/8 21:07
221108SH	180930153754951UA006078	3.00	0.00	0.00	退款成功		1 鲜肉大包	1	3	2022/11/8 21:10
221108SH	180930153754951UA006078	3.00	0.00	0.30	支付成功		1 鲜肉大包	1	3	2022/11/8 21:40
221108SH	180930153754951UA006078	2.70	0.00	0.00	退款成功		1 鲜肉大包	1	2.7	2022/11/8 21:41

图表 2.13 员工反复支付退款记录

## 2.1.4 退款用户分析

如图表 2.16 展示，在所有退款用户中，有 133 位用户是粉丝，占所有退款用户的 65.84%。因此可以判断出退款的用户主要由经常购物的用户组成，





图表 2.14 退款用户的用户阶段标签分布

如图表 2.17 展示，202 位有标签的退款用户在即食快餐、酸奶两类标签类型中存在一个用户对应多个标签的情况。

每位用户都有累计购买次数、累计购买金额、用户阶段、用户最近一次消费的商店四个标签。

用户阶段标签为高价值流失、低价值流失的用户都没有客单价、客单价(去除营销商品)、最近一个月消费次数、最近一个月消费金额四个标签。

健康食品、咖啡、牛奶、酸奶、碳酸饮料记录用户购买这六类商品的次数，辣记录用户购买食物的偏好。因为不购买的用户将不会有对应标签，所以记录量稍少。

退款的 202 位用户都没有开通升级卡业务。

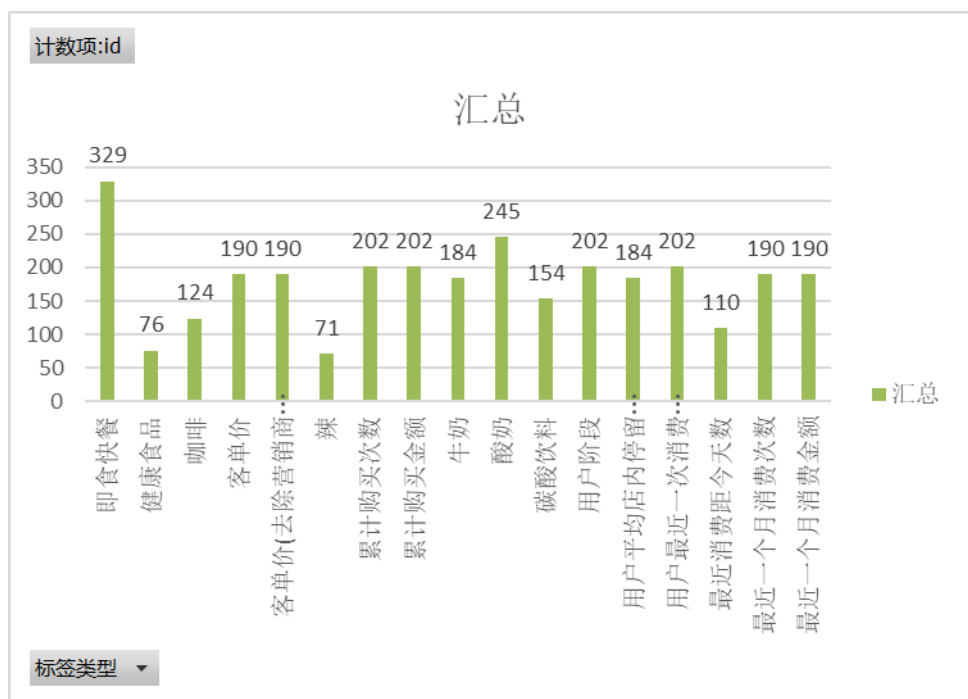


图 2.15 退款用户标签类型数量统计

如表格 2 展示，退款用户平均累计购买 71 次，平均累计购买金额 679 元，最多累计购买次数 456 次，最少累计购买次数 1 次，最多累计购买金额 8215.5 元，最少累计购买金额 3.2 元。

列 1	累计购买次数	列 2	累计购买金额
平均	70.56930693	平均	679.1014851
标准误差	4.110456862	标准误差	54.45761103
中位数	60	中位数	560.25
众数	44	众数	1007.5
标准差	58.42056859	标准差	773.9880765
方差	3412.962834	方差	599057.5425
峰度	13.41755099	峰度	54.85304279
偏度	2.768804634	偏度	6.363195383
区域	455	区域	8212.3
最小值	1	最小值	3.2
最大值	456	最大值	8215.5
求和	14255	求和	137178.5
观测数	202	观测数	202

表格 2 退款用户累计购买次数和累计购买金额的描述性统计

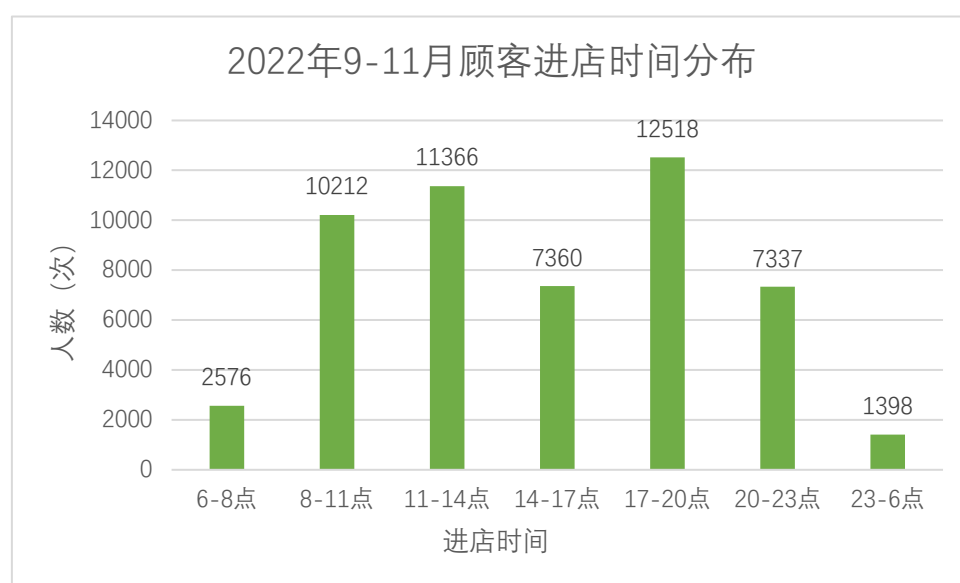
有以上几表可得，退款情况主要来自多次消费的用户，退款没有直接影

响用户不再参与便利店购物，但便利店仍需在退款高频时段提升服务质量和效率，提升用户对便利店的满意度。

## 2.2 用户进店时间、时长与行为分析

### 2.2.1 顾客进店时间分布

由于该云拿无人便利店处在徐汇校区内部，故在分析时将时间段分为：6-8 时（即教师/校工上班以及学生早起的时间），8-11 时（上午）11-14 时（午餐、午休时间）14-17 时（下午），17-20 时（晚餐时间）、20-23 时（晚上）、23-次日 6 时（熄灯时间）。



图表 2.16 2022 年 9-11 月顾客进店时间分布

如图表 2.18 展示，用户进入高峰期为早 8-11 时、11-14 时、17-20 时，即上午、午餐、晚餐时间，这与前文订单高峰期相同。

我们推测主要主要原因为以下 3 点：

- （1）、还没有吃早、午、晚餐的同学进入便利店购买快餐或主食类产品
- （2）、吃完早、午、晚餐的同学进入便利店购买饮料、饭后点心、零食等

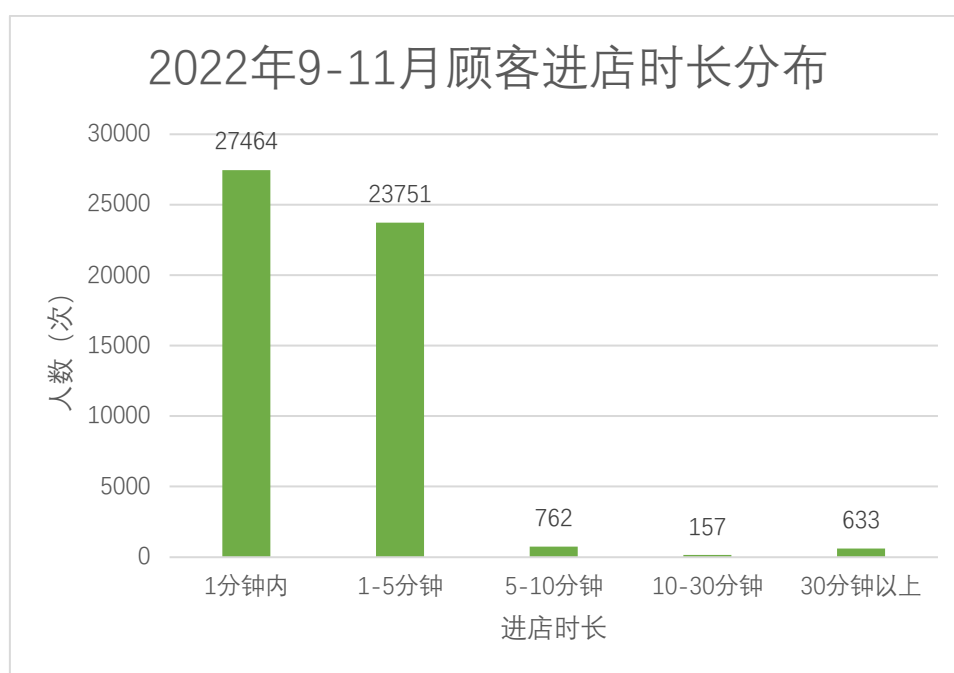
(3) 吃完饭顺路进入便利店采购日常生活所需物品。

### 2.2.2 顾客进店时长分布

对于消费者进店时长，我们将其简要分为 5 类：

(1) 1 分钟以内（对于想要购买的商品有明确目标，且对店内物品摆放位置十分熟悉），(2) 1-5 分钟（已知想要购买商品的品类但却需要对价格、商品口味等作考虑的用户，陪伴其他消费者将进店随便逛逛的用户），

(3) 5-10 分钟（在便利店用餐的消费者，想购买某一类商品的同时还在考虑是否要购买其他商品的消费者），10-30 分钟（等待其他人的消费者，想在便利店短暂休息的消费者），30 分钟以上（消磨时间的消费者）



图表 2.17 2022 年 9-11 月顾客进店时长分布

如图表 2.19 展示，可以发现 97%以上的消费者进店时长都在 5 分钟以内，且其中 55%的消费者均会在 1 分钟内完成消费行为，由此可得大部分的消费者在进店前已经明确了想要购买的商品并且能够迅速找到离去，也证明了云拿无人超市便利程度远超普通超市，能够通过无人化购物加速购物流程，减少耽误的时间。

### 2.2.3 每日客流量分布

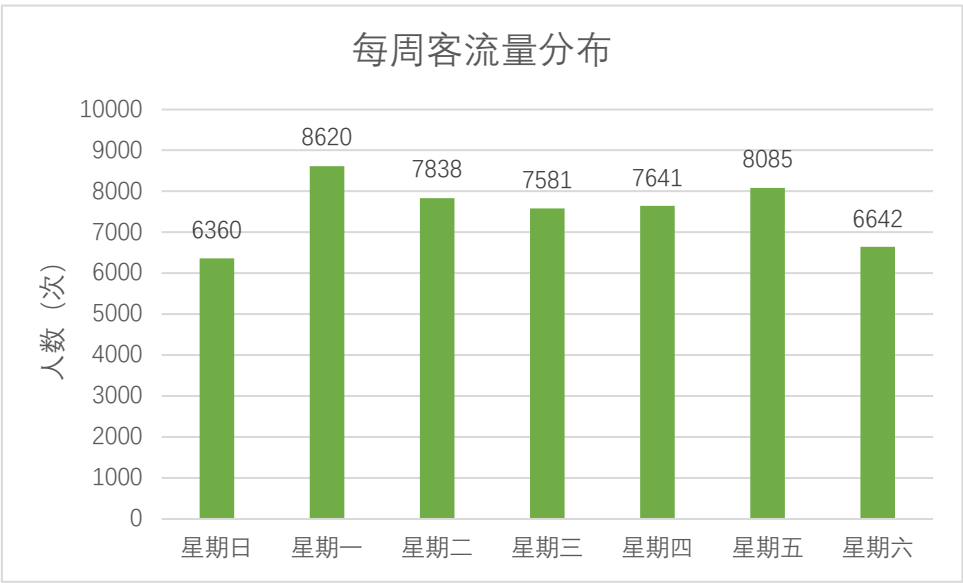


图表 2.18 2022 年 9-11 月每日客流量

由图表 2.20 展示，云拿无人超市在 9-11 月平均每日客流量约为 580 人。其中 10 月 1 日至 10 月 7 日为国庆假期，大量学生外出游玩，10 月 8 日至 10 月 14 日日的 3+4 校园封控，10 月 14 日至 10 月 23 日的线上教学暂缓返校，故而在 10 月 1 日至 10 月 23 日期间平均每日客流量约为 371 人。

排除这段时间的影响，每日客流量约为 650 人。每月的消费最高的 2-3 天均在该月的最后一周的周一，推测是由于学生的生活费在最后一周余裕量较大，选择犒劳自己便有此结果。

2.2.4 每周客流量分布



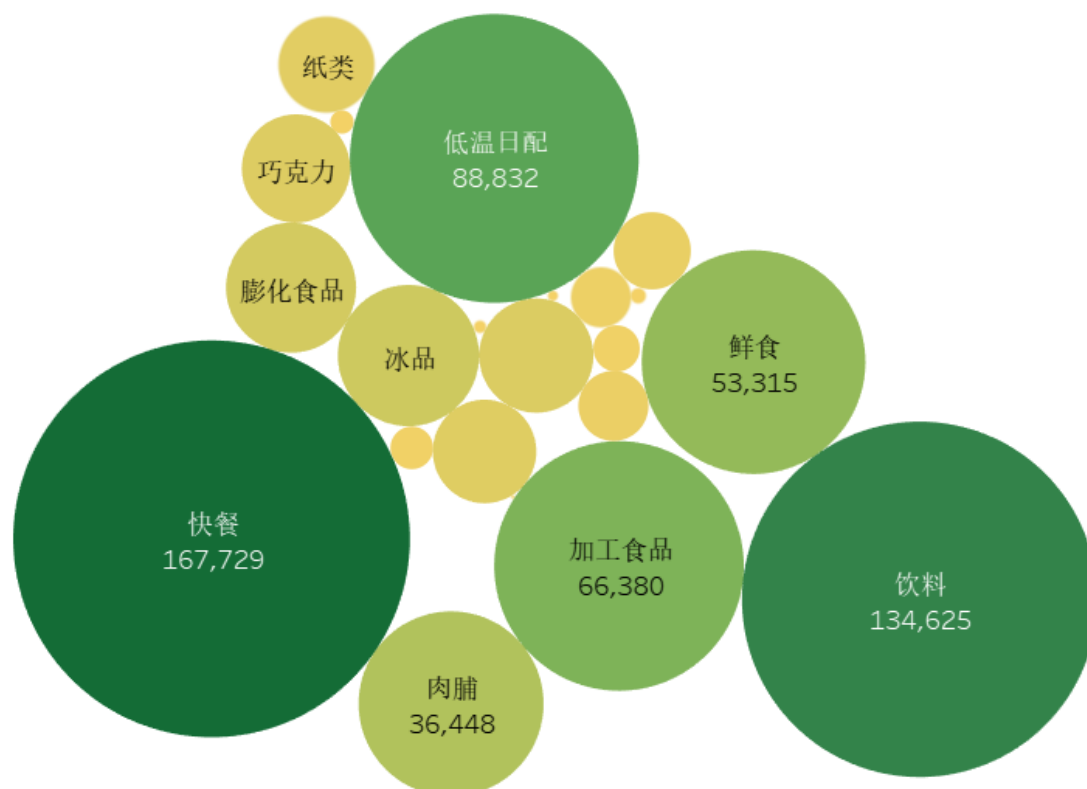
图表 2.19 每周客流量分布

如图表 2.21 展示，云拿无人超市在一周七天中，在周一客流量最大，周五其次，周日最小在周末客流量较小。

推测是由于在周一，教师和学生都倾向于采购该周生活所需用品，而周五有上海学生和教职工想要购买家中需要的物品，在校生想要购买周末需要的物品客流量排在其次，而周日由于周末学生较少，且从家中返校的学生会在家中带当天所需物资前来学校故客流量最少。

## 2.3 一级品类的商品占市场份额分析

### 2.3.1 一级品类的商品近 7 日销售数量之和比较分析



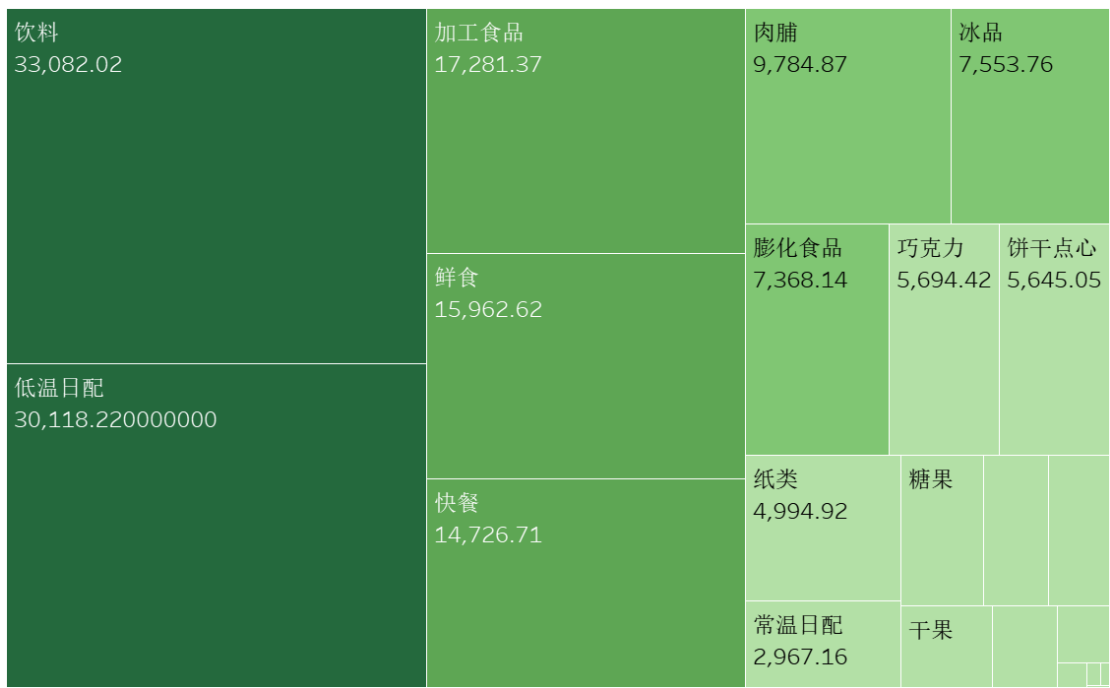
图表 2.22 一级品类的商品近 7 日销售数量之和气泡图

为了探究门店畅销商品，本文先探究了一级品类商品的近 7 日销售数量数据。由一级品类商品的近 7 日销售数量之和气泡图可以看出，位居销量前 6 位且占近 7 日总商品销售额较大的商品品类分别为快餐、饮料、低温日配、加工食品、鲜食、肉脯，说明上述商品品类在近 7 日的销售数量占云拿便利店的总市场份额较大，市场需求高，整体产品结构偏食品类，云拿公司可考虑多进货、补货以及设置更多商品栏位来售卖以上商品，通过增加商品的数量及品类，进一步探寻消费者需求，从而获得更大利润。

以下商品在上述品类中近 7 日销售数量之和位居前十，以供商家参考：  
怡宝饮用纯净水 555ml

农夫山泉饮用天然水 550ml  
君乐宝慢醇炭烤酸奶袋(150+10)g  
双汇鸡肉肠 70g  
君乐宝简醇酸奶(0 糖)袋(150+10)g  
双汇台式烧烤香肠原味 45g  
双汇玉米肠 40g  
康师傅喝开水 550ml  
三得利乌龙茶无糖 500ML  
水煮鸡蛋（带壳卤制蛋）

2.3.2 一级品类的商品毛利之和比较分析



图表 2.23 一级品类的商品毛利之和树状图

为了探究门店畅销商品的盈利能力，本文探究比较了一级品类的商品毛利之和。由一级品类商品的毛利之和树状图可以看出，位居销量前 8 位的商品品类为饮料、低温日配、加工食品、鲜食、快餐、肉脯、冰品、膨化食品，说明上述商品品类的毛利之和占云拿便利店的利润份额较大，云拿公司可考虑多设计捆绑售卖商品组合，以增加高毛利商品的销售量，最终获得更大利润。



以下商品在上述品类中毛和之和位居前十，以供商家参考：

- 味全当红杨梅 300ml
- 照烧脆骨丸串
- 唐扬鸡块
- 康师傅喝开水 550ml
- 悦鲜活牛乳 450ml\*15
- 三得利乌龙茶无糖 500ML
- 桃李布朗尼蛋糕
- 原味豆浆(热)
- 培根鸡蛋手抓饼
- 奥尔良鸡腿肉包

除此之外，云拿公司还可结合上述商品近几年的总体销售额与行业标准相比较，从而分析企业的业绩变化状况。

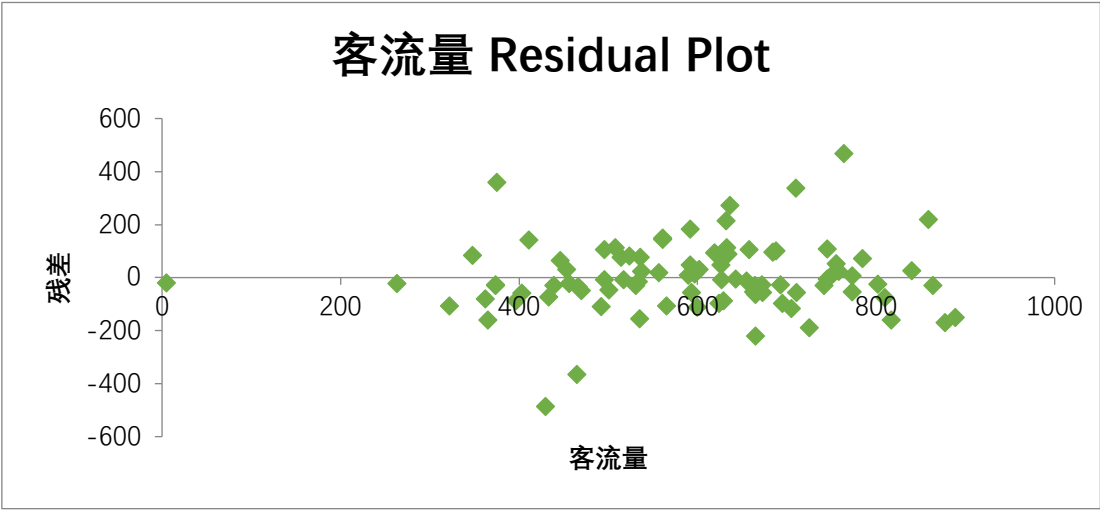
### 第3章 影响因素分析

#### 3.1 每日客流量是否影响该日商品的拿起与放下次数

以每日客流量为自变量，该日商品的拿起与放下次数为因变量构成线性回归。

模型	非标准化系数		标准系数		显著性
	B	标准 误差	Beta	t	
(常量)	31.824	48.229		.660	.511
客流量	1.820	.080	.924	22.877	.000
R <sup>2</sup>				0.853	
F				523.376	
P				<0.001	
因变量：拿起次数					

表格 3 每日客流量和每日商品拿起次数的线性回归

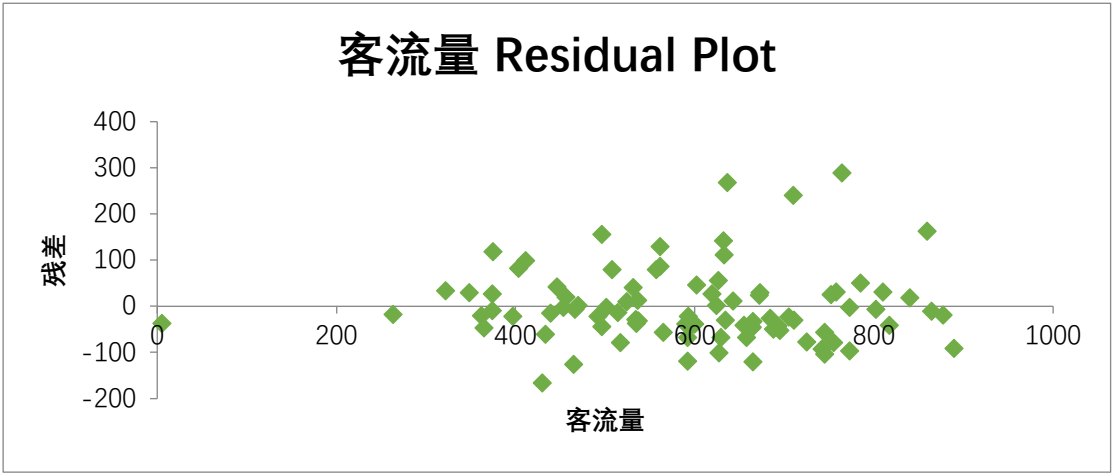


图表 3.1 每日客流量和每日商品拿起次数的线性回归

模型	非标准化系数		标准系数		显著性
	B	标准误差	Beta	t	
(常量)	34.654	29.110		1.190	0.237
客流量	0.475	0.048	0.724	9.896	0.000
	R <sup>2</sup>			0.519	
	F			97.940	
	P			<0.001	

因变量：放下次数

表格 4 每日客流量和每日商品放下次数的线性回归



图表 3.2 每日客流量和每日商品放下次数的线性回归

残差图证明本次回归有效，其中以拿起次数为因变量的回归中  $R^2$  为 0.853，以放下次数为因变量的回归中， $R^2$  值为 0.519，拟合优度均良好，P 值均小于 0.01，且均为正相关。

拿起次数的方程为： $y=1.820x+31.824+\mu$ （ $\mu$  为误差）

放下次数的方程为： $y=0.475x+34.654+\mu$ （ $\mu$  为误差）

即客流量越大，商品被拿起或放下的次数越多且商品被拿起的增长幅度多余商品被放下的幅度。由于前文已经证明商品拿起的次数越多，销售额越高，故我们认为增加客流量能显著影响商品销售额。

### 3.2 米饭品类供应商是否影响近 7 日销售数量

根据上述一级品类商品销售额及盈利能力分析结合订单商品详情表，由于米饭品类的商品销售额多、盈利能力强，且可控因素较明显，所以本文选择探究如何进一步提升米饭品类的商品销售额。

为了检验米饭供应商不同在米饭销量上是否存在显著差异，本文使用独立样本 T 检验进行分析。

表格 5 方差分析结果

	清美	桃李	T	P
米饭销量	12.27±16.113	28.15±10.153	-6.911	0.000

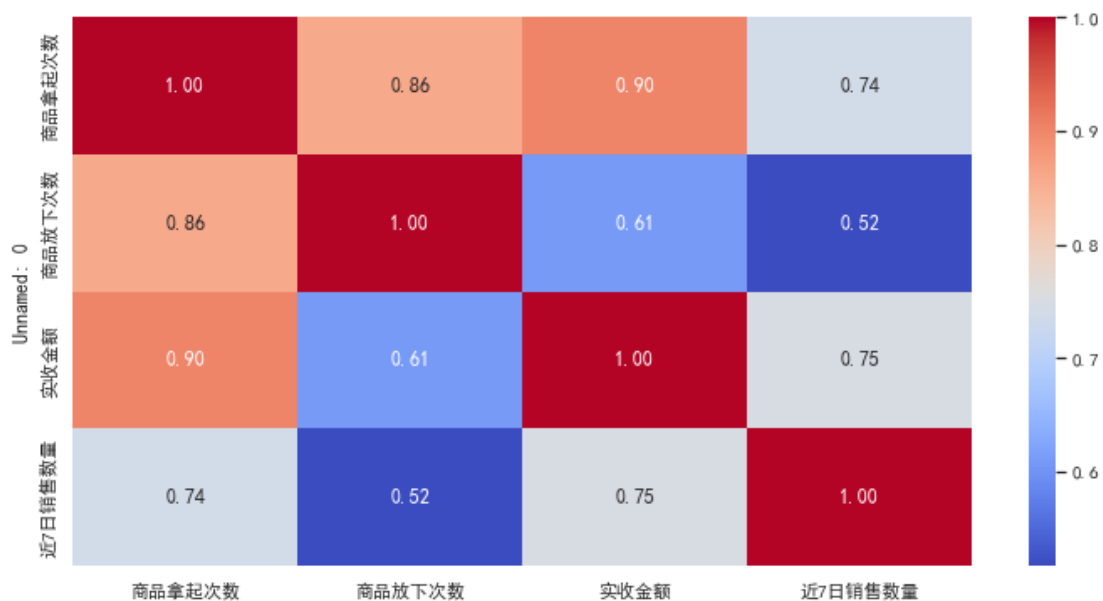
由表 5 显示，米饭销量的双侧 P 值小于显著性 0.05，说明供应商不同在米饭销量上存在显著差异。清美和桃李两家米饭供应商相较之下，桃李的销量更好，故建议商家多进桃李的米饭品类的商品，以提高商品的销量。

### 3.3 米饭品类拿起、放下次数，实收金额与近 7 日销售数量的相关关系

表格 6 相关性分析结果

	商品拿起次数	商品放下次数	实收金额	近 7 日销售数量
商品拿起次数	1			
商品放下次数	0.859**	1		
实收金额	0.903**	0.610**	1	
近 7 日销售数量	0.740**	0.518**	0.748**	1

\*\* 在 0.01 级别（双尾），相关性显著。



图表 3.3 相关性分析热力图

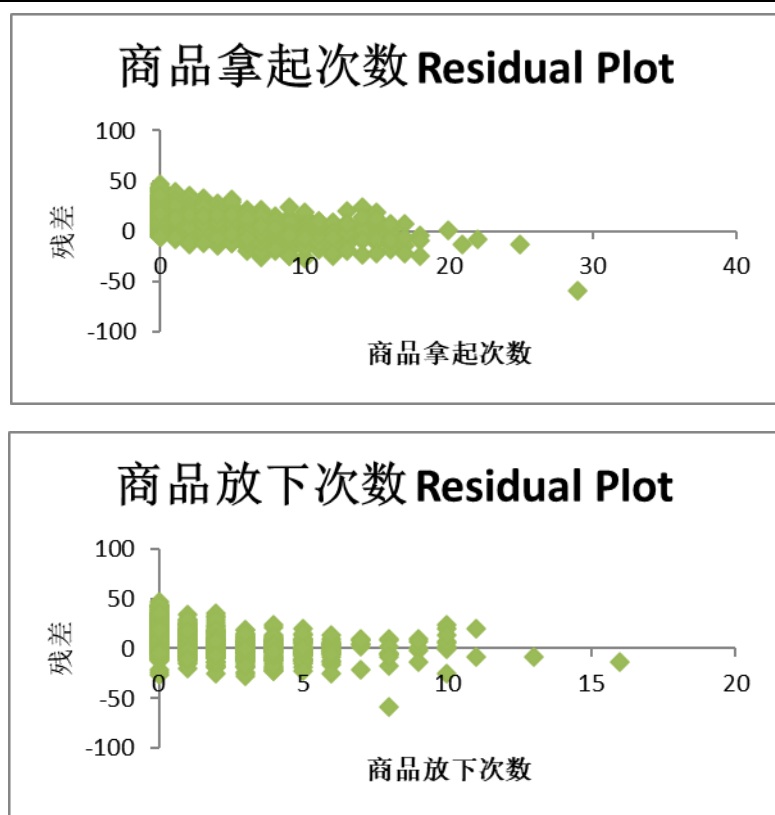
为了初步探究米饭品类的商品拿起次数、商品放下次数、实收金额与近 7 日销售数量之间有无线性相关关系以及线性相关关系的密切程度、方向等，本文利用统计学软件 SPSS26.0 计算了各变量间的 Pearson 简单相关系数（即  $r$ ），并绘制热力图。

得出以下结果：米饭品类的商品拿起次数、商品放下次数，实收金额与近 7 日销售数量均呈现出显著的正向相关关系（见表 6、图 3.3）。

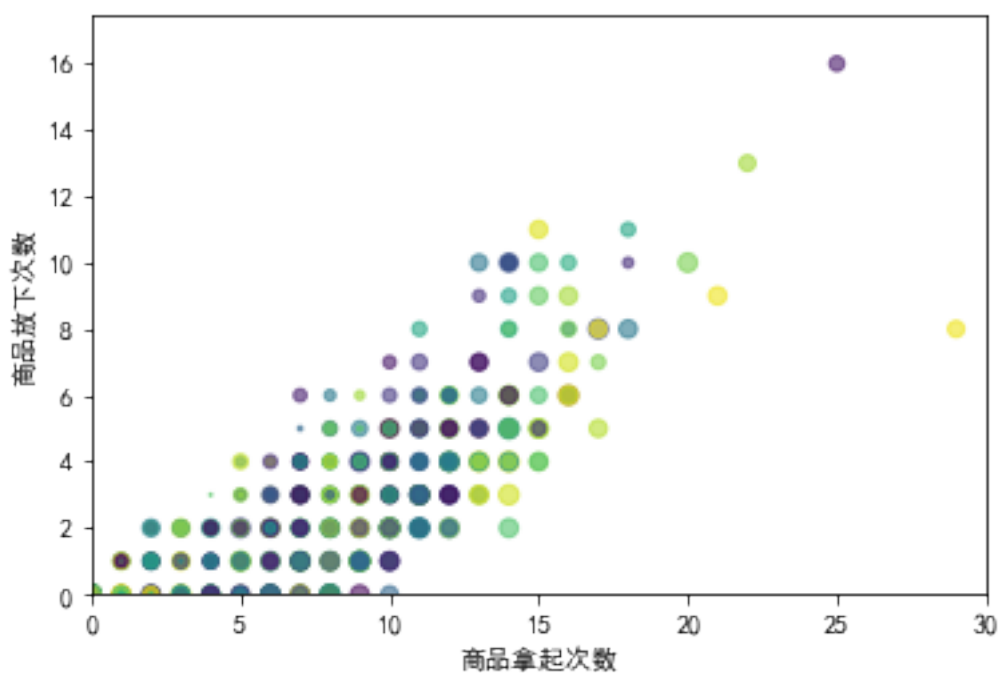
### 3.4 米饭品类拿起、放下次数与近 7 日销售数量的影响因素分析

表格 7 线性回归分析结果

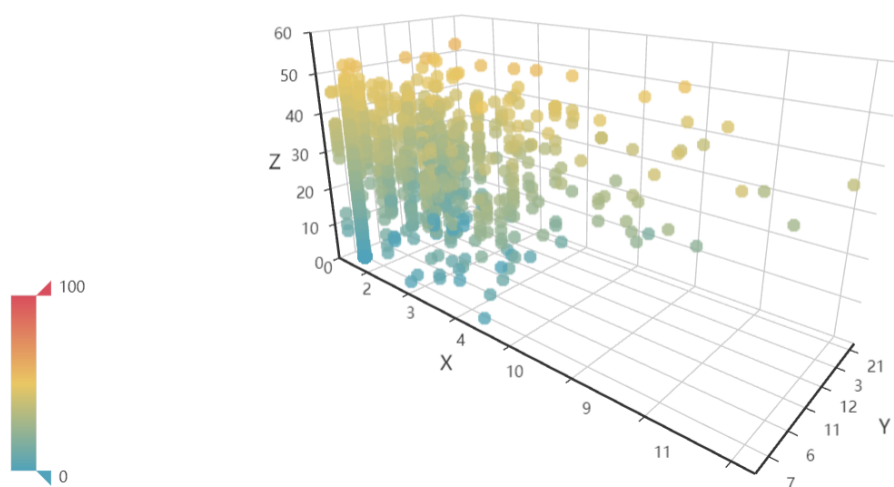
模型	未标准化系数		标准化系数		t	显著性	VIF
	B	标准错误	Beta				
(常量)	4.840	0.261			18.554	0.000	
商品拿起次数	4.311	0.103	1.127		41.725	0.000	3.826
商品放下次数	-4.253	0.255	-0.450		-16.664	0.000	3.826
		R <sup>2</sup>				0.601	
		F				1575.541	
		P				<0.001	
因变量：近 7 日销售数量							



图表 3.4 商品拿起、放下次数残差图



图表 3.5 商品拿起、放下次数与近 7 日商品销量的气泡图



图表 3.6 商品拿起、放下次数与近 7 日商品销量的 3D 散点图

为了进一步探究米饭品类的商品拿起次数、商品放下次数与近 7 日销售数量之间存在的线性相关关系，本文利用统计学软件 SPSS26.0 以米饭品类的商品拿起次数、商品放下次数为自变量，近 7 日销售数量为因变量进行线性回归分析。商品拿起次数、商品放下次数的残差图证明回归有效，并绘制气泡图、3D 散点图可视化商品拿起次数、商品放下次数与近 7 日销售数量

之间的线性相关关系。

本次线性回归模型的拟合度良好， $R^2=0.601$ ，意味着本次线性回归的运算结果可以比较真实可靠地反映出米饭品类的商品拿起次数、商品放下次数对近 7 日销售数量的影响情况。两个自变量之间不存在多重共线性，VIF 全部小于 5。回归方程显著， $F=1575.541$ ， $P<0.001$ ，意味着两个自变量商品拿起次数、商品放下次数中至少有一个可以显著影响因变量近 7 日销售数量。

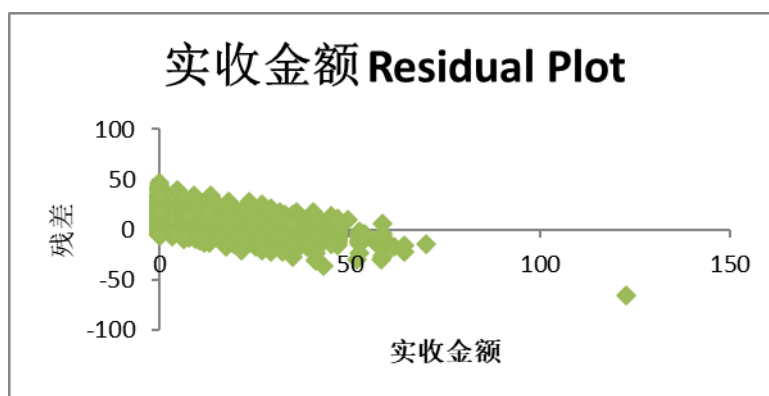
由表 7 可得，商品拿起次数可以显著正向影响近 7 日销售数量（ $\beta=4.311>0$ ， $P<0.05$ ）；商品放下次数可以显著负向影响近 7 日销售数量（ $\beta=-4.523<0$ ， $P<0.05$ ）。最后，变量之间得出如下的回归方程：**近 7 日销售数量=4.84+4.311\*商品拿起次数-4.253\*商品放下次数。**

根据以上分析结论，本文猜测：消费者在拿起米饭品类的商品时一般就决定会选择购买该商品，所以建议商家在售卖时可以适当美化包装，增加对消费者拿起商品的吸引力，从而增加米饭品类商品的销售量。

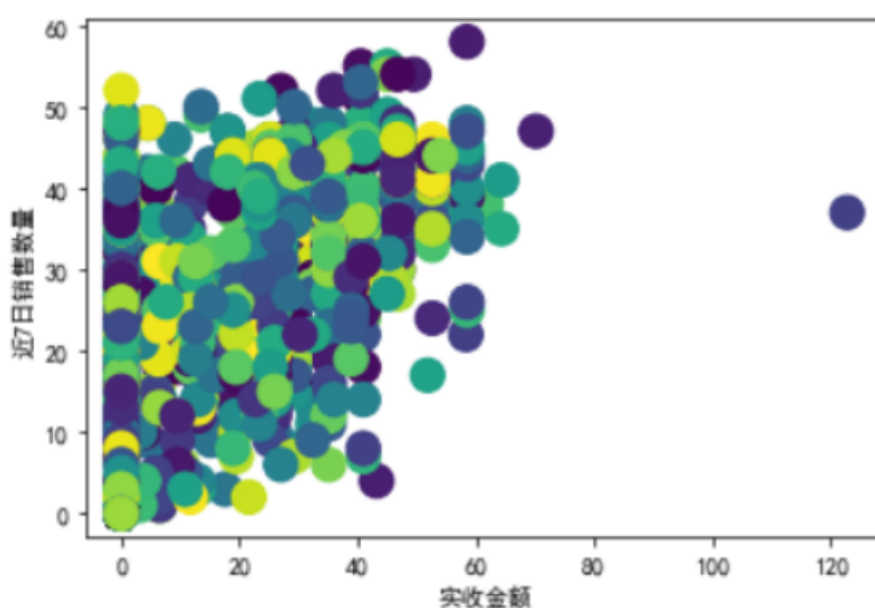
### 3.5 米饭品类实付金额与近 7 日销售数量的影响因素分析

表格 8 线性回归分析结果

模型	未标准化系数		标准化系数	t	显著性	VIF
	B	标准错误	Beta			
(常量)	5.236	0.272		19.246	0.000	
实付金额	0.794	0.015	0.748	51.656	0.000	1.000
		$R^2$			0.560	
		F			2668.310	
		P			<0.001	
因变量：近 7 日销售数量						



图表 3.7 实收金额残差图



图表 3.8 实收金额与近 7 日商品销量的散点图

为了进一步探究米饭品类的商品实付金额与近 7 日销售数量之间存在的线性相关关系，本文利用统计学软件 SPSS26.0 以米饭实付金额为自变量，近 7 日销售数量为因变量进行线性回归分析。实付金额的残差图证明回归有效，并绘制散点图可视化实收金额与近 7 日销售数量之间的线性相关关系。

本次线性回归模型的拟合度良好， $R^2=0.560$ ，意味着本次线性回归的运算结果可以比较真实可靠地反映出米饭品类的商品实付金额对近 7 日销售数量的影响情况。该自变量不存在多重共线性，VIF 小于 5。回归方程显著，

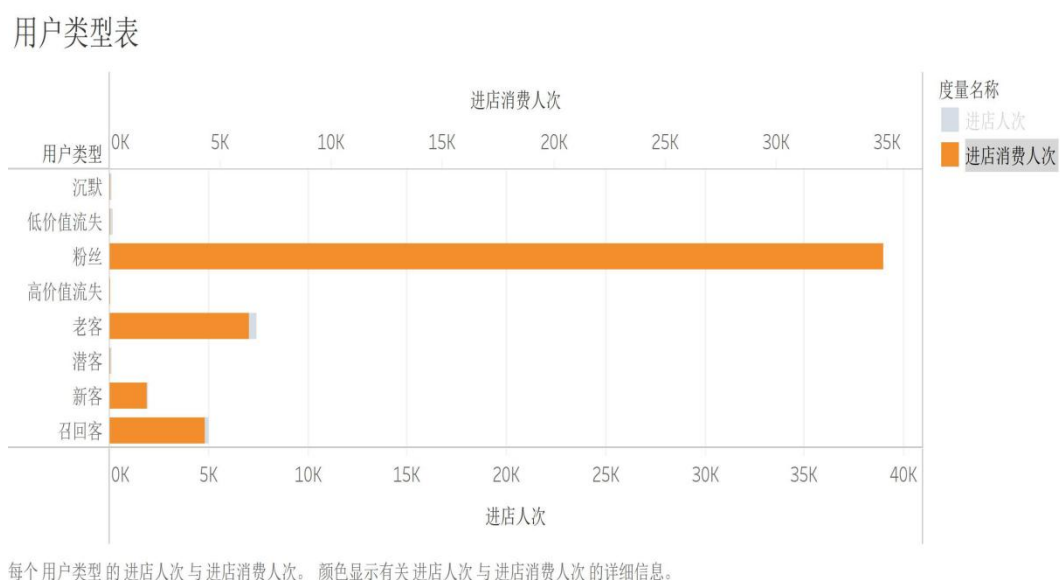


$F=2668.310$ ,  $P<0.001$ , 意味着自变量米饭品类的商品实付金额可以显著影响因变量近 7 日销售数量。

由表 8 可得, 商品实付金额可以显著正向影响近 7 日销售数量 ( $\beta=0.794>0$ ,  $P<0.05$ )。最后, 变量之间得出如下的回归方程: **近 7 日销售数量**  
 **$=5.236+0.794*\text{实收金额}$** 。

## 第4章 顾客类型分析

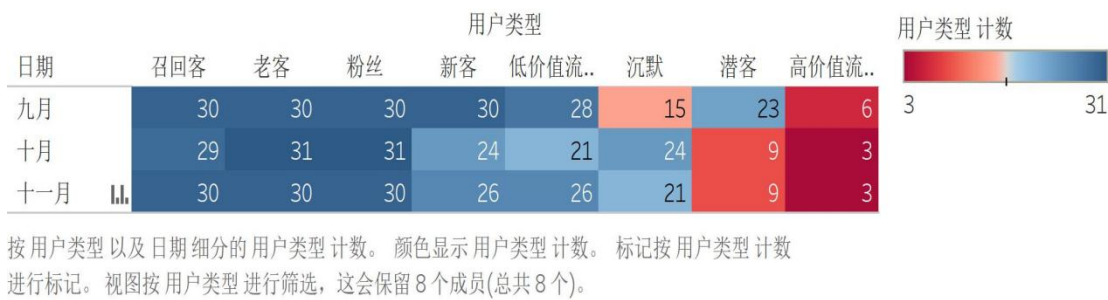
### 4.1 用户类型趋势



图表 4.1 用户类型表

本文关注到云拿将进店的顾客分为沉默、低价值流失、粉丝、高价值流失、老客、潜在客户、新客和召回客八种类型。通过使用 tableau, 根据相关进店消费数据, 本文初步将进店后是否消费作为不同类型的用户的特征之一。其中, 沉默、低价值流失和高价值流失进店后没有任何消费, 且并无对商品产生明显兴趣; 本文猜测潜在客户虽然进店后也并无产生消费, 但有对商品产生兴趣的行为。

不同用户类型三月趋势图



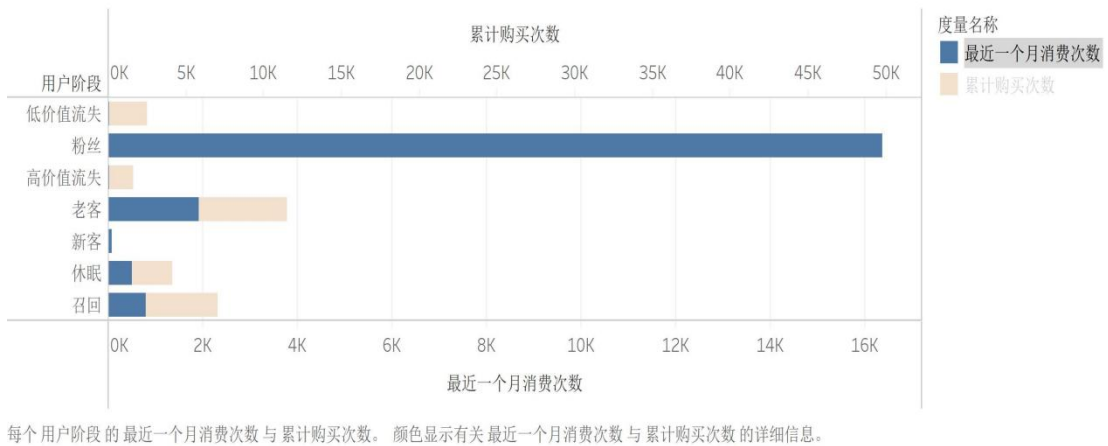
图表 4.2 不同用户类型三月趋势表

本文首先通过 excel 整合数据，将不同日期整合成九月、十月和十一月三月份，探究这三个月不同用户类型数量变化的趋势。

可以看到，云拿的粉丝数量一直保持平稳，值得关注的是潜客的数量在十月和十一月大幅下降，本文猜测可能与疫情相关或是当地市场需求已至饱和，一方面说明云拿对新用户吸引力不足的问题，建议在门店增设装饰或讲解海报来吸引顾客，使顾客了解云拿，另一方面说明该地市场开发完毕，可以寻找下一市场开发新店；每月低价值流失顾客数量也较为平稳，说明云拿的商品对于顾客的吸附能力有待加强，需要增加商品种类或适当调整商品价格来吸引顾客。

4.2 用户阶段

顾客用户阶段分类表



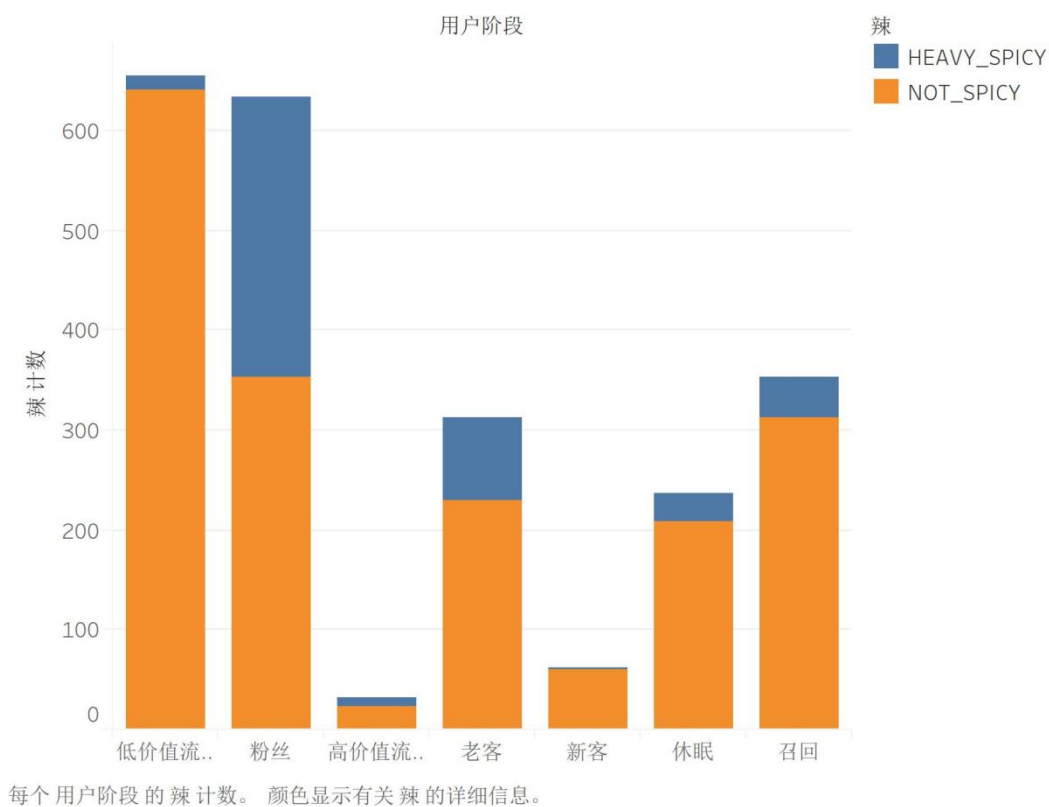
图表 4.3 顾客用户阶段分类表

在用户阶段中，云拿将购买过商品的顾客分为低价值流失、粉丝、高价值流失、老客、新客、休眠和召回 7 种类型。本文初步将最近一个月的消费次数与累计购买次数之比作为不同类型的用户的特征之一。

### 4.3 用户分析

本文接下来将对不同用户行为或喜好进行分析，从而给出有关建议，提高云拿顾客吸引力或销售额。

辣味喜好程度



图表 4.4 顾客辣味喜好程度表

本文先就不同类型用户的口味进行探究，发现除粉丝这一用户类型外，其余用户类型都体现出不爱吃辣的特点，因此建议减少辣的商品品类的数量，在店内保留数量较少的辣的商品品类维持核心顾客；增加不辣的商品品类的数量提高销售额，吸引顾客。

## 用户偏好商品

用户阶段	健康食品	即食快餐	咖啡	牛奶	碳酸饮料
低价值流失	0	516	112	155	156
粉丝	963	12,228	1,291	2,905	3,603
高价值流失	0	314	68	130	134
老客	90	2,974	266	677	901
新客	7	19	4	1	4
休眠	32	936	149	326	309
召回	38	1,909	189	475	398

按用户阶段细分的健康食品, 即食快餐, 咖啡, 牛奶与碳酸饮料。

图表 4.5 用户偏好商品表

云拿作为便利店, 其主要销售来源便来自于即食快餐等食品品类的商品, 这点与数据是相符合的。但发现饮品和食品的数量差异很大, 建议可以研发一些组合商品进行出售, 将食品和饮品捆绑出售, 或是将食品和饮品的栏位放置靠近, 便于用户选取。

## 第5章 总结和建议

### 5.1 订单数据

9-11 月每个月都购买过商品的用户占三个月有购买记录用户的 49%, 便利店每月固定用户在一千人以上, 用户粘性大, 留存率高。

店铺内主要订单、退款发生在上午 8 点-12 点, 下午 4 点-8 点, 可以在这段时间之前整理货架, 补充商品, 并且在这段时间增加工作人员, 保证商店正常运行。

几乎全部用户使用第三方支付工具进行支付, 会员余额系统存在感薄弱, 存在部分用户不及时充值余额的情况。可以在余额较低时增加提醒, 也可以尝试余额支付优惠一类活动, 吸引用户采用余额支付。

退款商品主要集中在饮食品类, 需要定期整理退款商品信息, 若一类或一种商品在退款商品中占比较大, 可以考虑不在进货或寻找新的商品替换掉。

---

## 5.2 闸机流水数据

平均每日客流量约为 580 人，消费高峰为每月最后一个周一附近，消费者主要集中在每周周一、周五进行消费，周日消费者较少。绝大多数的消费者在进入超市时已经具有明确的目标可以做到即拿即走，消费者主要进入时间为上午 8-11 时，中午 11-14 时，以及下午 17-20 时。

基于以上数据结果，我们提出以下建议：

（1）由于消费者倾向于在周一消费，高峰为每月最后一个周一，且大部分消费者熟悉货架，目标明确，我们建议于每月 5-8 天前的最后一个周日调整货架摆放位置，增加消费者在店内挑选商品所需时间，增加销售额。

（2）由于节假日中客流量较少，可以进行节假日促销，或是周末促销，对部分商品予以折扣，吸引消费者周末或是节假日购买，而不是在节假日前购买，改变其消费习惯，增加进入便利店消费次数。

## 第6章 数据问题汇总

### 6.1 订单列表

因为在校园内，所以所有订单皆打九折优惠，可是在 11 月 8 日出现几个小时的所有订单都没有打折，原价购买的情况。

订单列表中由六个退款订单记录无法找到并关联对应用户购买记录。

### 6.2 用户标签表

在闸机流水表中，有部分员工类型的用户 id 也具有购买行为，理应也被视作顾客，但是在用户标签表里并没有这类既是员工又是顾客的用户标签，希望在之后的数据里能够添加上这一类特殊的顾客的用户标签。

### 6.3 闸机流水表

存在一些站在闸机口没进店但是误判进店的情况，存在无人进店却有出

---

店记录的情况。进店和扫码可能由于光线等原因产生误判，出店由于店内光线等环境固定所以误判几率低。希望能调整算法和灵敏度参数，提高判定进出店的准确率。