

# 新零售-无人智能售货机商务数据分析报告

## 目录

- 一、背景介绍..... 2
  - 1、行业背景..... 2
  - 2、分析目标..... 2
- 二、数据说明..... 3
  - 1、数据来源..... 3
  - 2、数据量 ..... 3
- 三、销售数据分析 ..... 3
  - 1、数据预处理与分析 ..... 3
    - 任务 1.1 ..... 3
    - 任务 1.2 ..... 4
    - 任务 1.3 ..... 4
  - 2、数据可视化..... 6
    - 任务 2.1 ..... 6
    - 任务 2.2 ..... 6
    - 任务 2.3 ..... 9
    - 任务 2.4 ..... 9
    - 任务 2.5 ..... 10
  - 3、售货机画像的探索 ..... 11
    - 任务 3.1 ..... 11
    - 任务 3.2 ..... 11
  - 4.销量预测..... 14

# 一、背景介绍

## 1、行业背景

随着科学技术的发展、信息时代的到来和经济的迅速发展，人们的消费追求已逐渐倾向于便利与多元化。与此同时，互联网催生了一系列“自助购物”的现象，各种新式零售方式逐渐融入人们的生活，而自动售货机购物更是受年轻一代的喜爱。与传统的购物相比，“自助购物”让我们可以做到随时随地便可选择到自己喜欢的商品，你需要的只是“买单”，然后就可以便捷快速的获取心仪的商品，这种购物方式大大的满足了当今快节奏的生活方式。

相比之下，“线下零售”方式在“自助购物”的方式面前就显得比较被动了，单一的“线下零售”方式将无法适应多元化的消费结构。因此，人们一直在寻求一种更好的销售方式，寻求一种可以带来更高收益高效的营销方式。自 1998 年起，电子商务进入高速发展阶段，为人们的日常生活带来诸多便利，实业零售面临着巨大威胁，曾经的零售巨头“沃尔玛”陷入多家实体店倒闭的窘境；可现观如今的电子商务，其格局也逐渐趋于稳定，流量尚未有新的突破，发展似乎已经进入瓶颈。零售业如今陷入的困境，从外部供给端看来，一方面传统零售主要基于实体店面进行经营活动，实业零售人力和地面租金高昂，在电商经济的冲击下持续走低；而另一方面，后兴起的电商经济尽管发展势头迅猛，但因其实际用户体验不佳，渠道成本较高，低效率等难题，近年格局和流量也基本稳定，似乎已经触顶。从内部需求端看来，如今大数据时代到来、消费群体改变、消费观念更新和购买力不断提升，向零售行业提出了更高的要求，消费模式亟待转型创新。但是，传统零售面临着“渠道分散”、客户体验不一、成本上升、利润空间压缩等多个困局。如此窘境下，“新零售”应运而生。

作为以网络技术发展、消费升级转型为基础的产物，“新零售”向整个零售业展现出了它全新的活力。“新零售”实质是创新其运营模式，它将自助零售和线下叠加融合，顾客既能够在线下亲身进行用户体验，也能够直接在自动售货机便利购买。这样的结合模式使这一新型的零售模式兼具实业零售的实际具体和电子商务的快捷高效。

## 2、分析目标

本次选取了“新零售”的代表产物自动售货机，获取不同的位置的五台自动售货机 A、B、C、D、E 的 2017 年的销售数据，对售货机的销售数据进行商务数据分析，帮助经营者了解用户需求，

掌握商品需求量，给出合理的营销方案。具体进行以下几方面的分析：

- (1) 根据项目需求对数据进行预处理，分别提取每台售货机的销售数据，计算每台售货机每个月的平均交易额与日均订单量
- (2) 对销售数据进行可视化，并根据可视化后的结果给出分析结论
- (3) 根据热销商品分别绘制 5 台售货机画像，根据画像及分析制定合适的营销策略
- (4) 预测售货机未来一个月的每个大类商品的销量，并根据预测值给出 5 台售货机未来一个月货品的补充计划

## 二、数据说明

### 1、数据来源

数据来源：五台售货机销售数据 csv 文件

数据获取时间：2017 年 1 月 1 日—2017 年 12 月 31 日

### 2、数据量

drop\_duplicates 方法去重之后商品总数为： 303

去重之前订单详情表的形状为： (70680, 9)

依照订单号去重之后订单详情表大小为: (70680, 9)

删除原始时间错误前的长度为： 70680

删除后的列索引为： 70679

## 三、销售数据分析

### 1、数据预处理与分析

#### 任务 1.1

提取每台售货机对应的销售数据，分别保存为 5 个 csv 文件

## 任务 1.2

图表 1 2017 年 5 月份每台售货机的订单量、交易额  
与订单总量、交易总额

	A 售货机	B 售货机	C 售货机	D 售货机	E 售货机	Sum
5 月订单量	756	869	789	564	1292	4270
5 月交易额	3385.1	3681.2	3729.4	2392.1	5699	18886.8

## 任务 1.3

图表 2 2017 年每台售货机每个月份的日均订单量

月份	A 售货机	B 售货机	C 售货机	D 售货机	E 售货机
1	10.806452	11.806452	12.225806	8.354839	11.419355
2	4.071429	6.607143	7.392857	5.035714	9.214286
3	8.225806	8.548387	8.483871	6.193548	11.290323
4	14.900000	20.100000	24.466667	14.766667	29.833333
5	24.387097	28.032258	25.451613	18.193548	41.677419
6	55.633333	61.866667	62.733333	34.666667	86.433333
7	15.354839	11.129032	24.645161	10.225806	26.225806
8	21.483871	31.645161	40.612903	23.064516	57.000000
9	34.666667	58.166667	55.933333	32.766667	137.800000
10	50.483871	65.354839	71.483871	38.258065	89.580645
11	38.666667	67.700000	64.766667	40.333333	167.333333

12	64.612903	71.290323	76.741935	53.645161	104.903226
----	-----------	-----------	-----------	-----------	------------

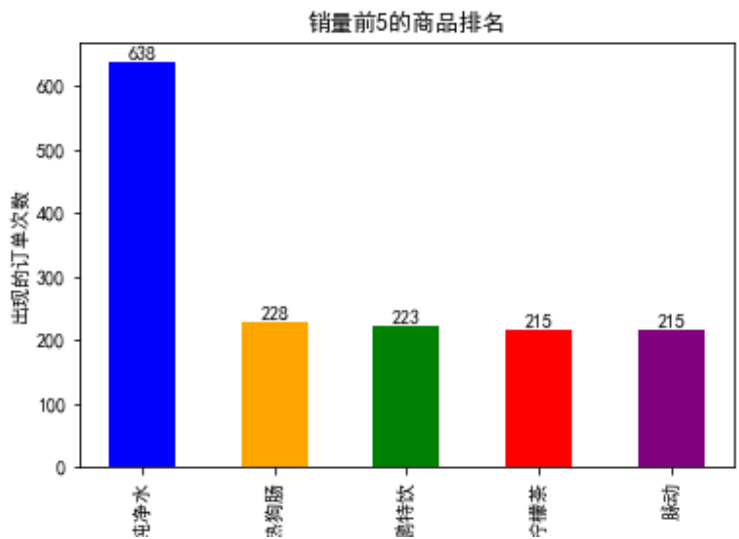
图表 3 2017 年每台售货机每个月份的每单平均交易额

月份	A 售货机	B 售货机	C 售货机	D 售货机	E 售货机
1	4.506567	3.753005	4.328496	3.692664	4.680226
2	3.864035	3.255676	3.826087	3.088652	3.638372
3	3.585490	3.614717	3.769962	4.305729	4.305714
4	4.036913	4.075290	4.403678	3.790293	4.159888
5	4.477646	4.236133	4.726743	4.241312	4.410991
6	4.047394	4.068050	4.501700	4.025962	3.817856
7	4.097689	4.401739	3.988351	4.229653	3.919311
8	3.358709	3.584200	3.913582	3.316503	3.804471
9	4.307212	4.130258	4.427294	3.899390	4.125375
10	4.020703	4.112340	4.273330	3.884233	3.676125
11	4.471552	4.268784	4.352393	3.862314	4.283227
12	3.787868	3.667014	3.943043	3.572580	4.168973

2、数据可视化

任务 2.1

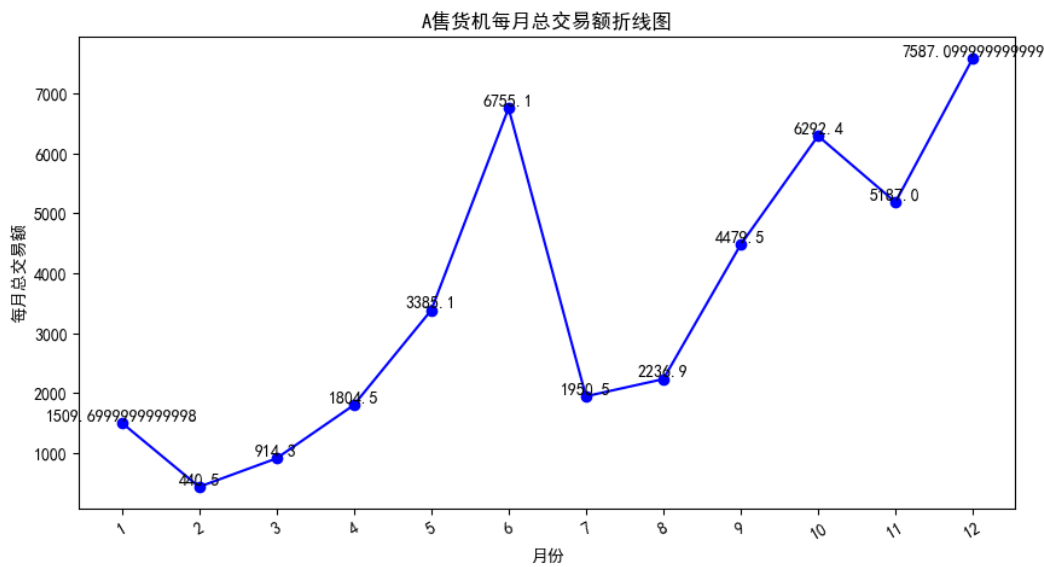
图表 4 2017 年 5 月销量前 5 的商品销量柱状图

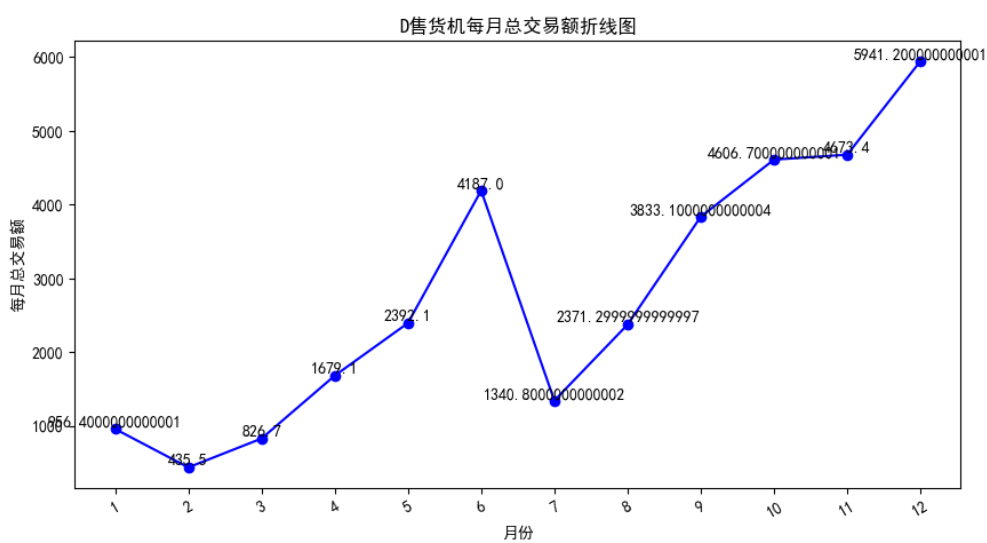
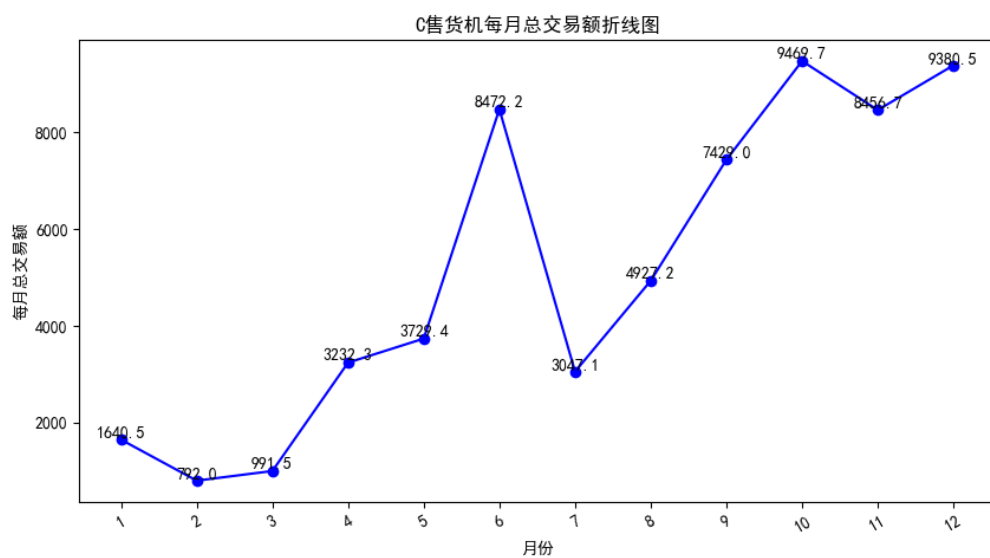
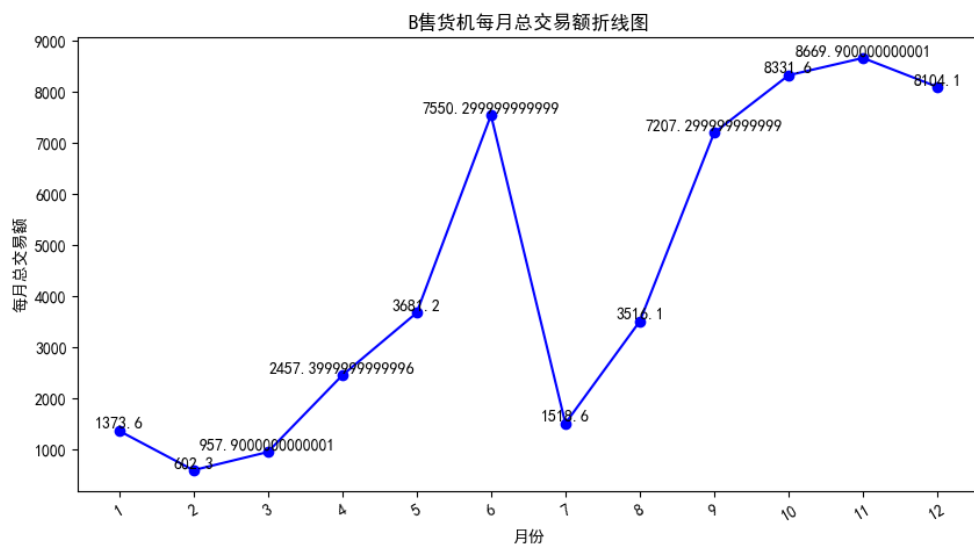


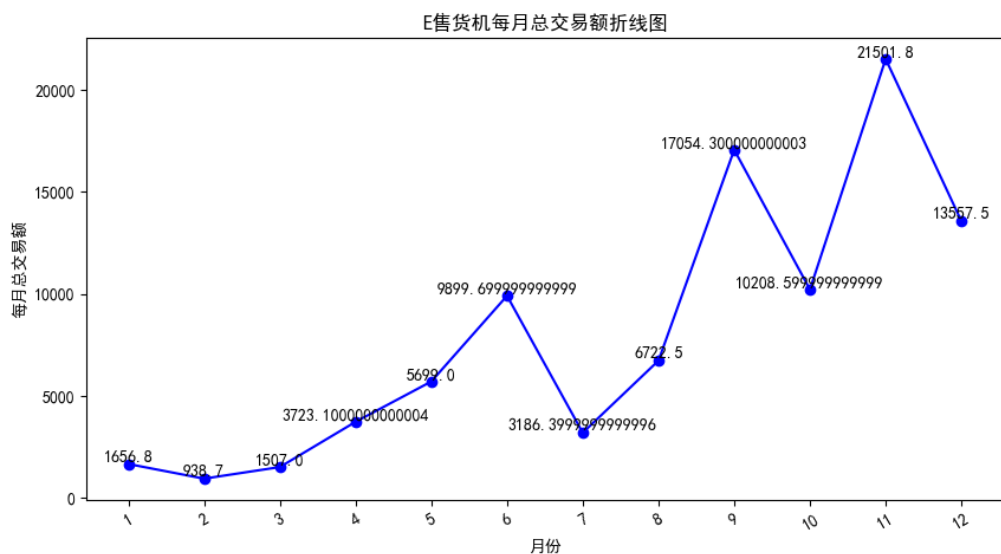
由图表 4 可知 2017 年 6 月销量前 5 的商品分别为纯净水、热狗肠、东鹏特饮、柠檬茶、脉动，其中纯净水的销量尤为突出。

任务 2.2

图表 5 每台售货机每月总交易额折线图

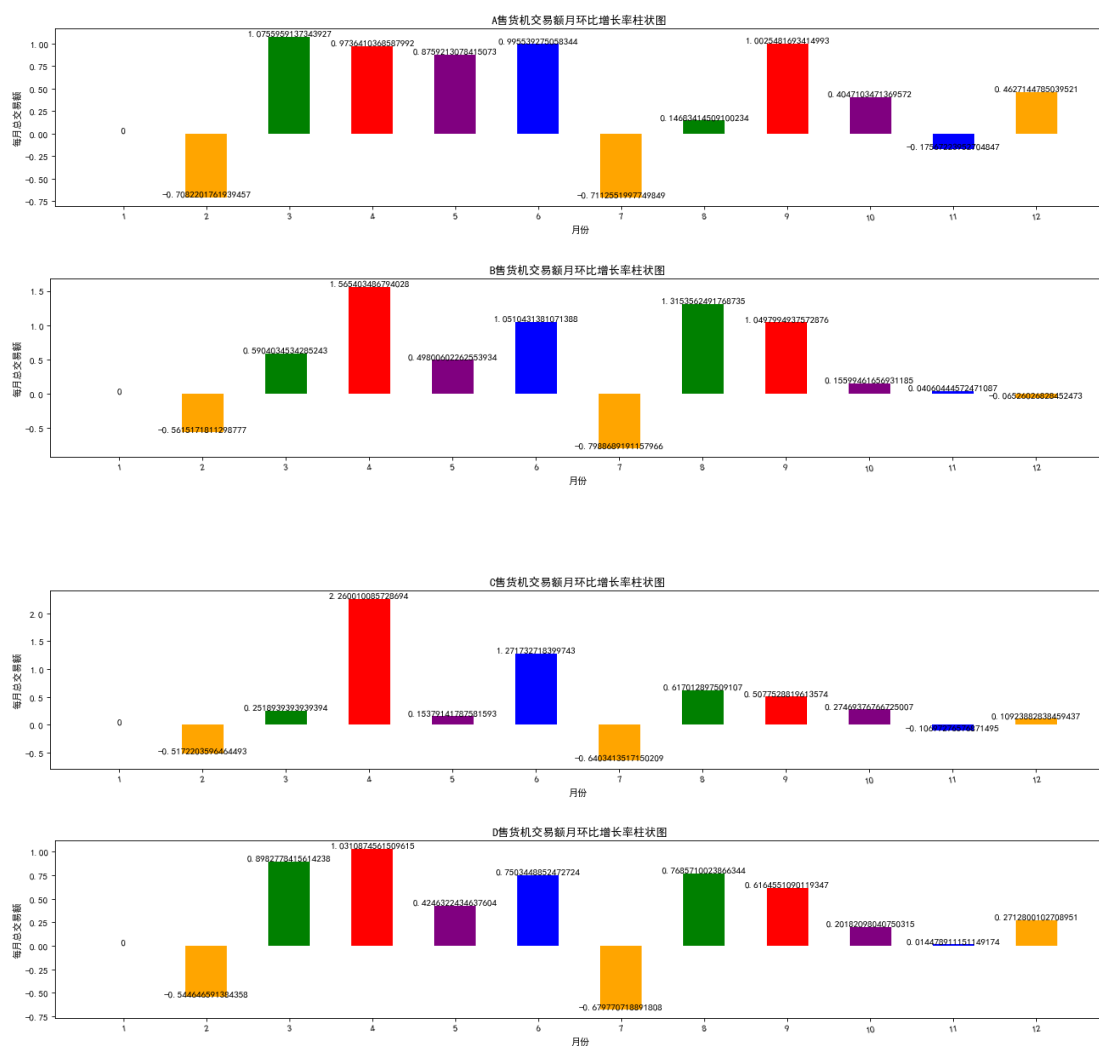




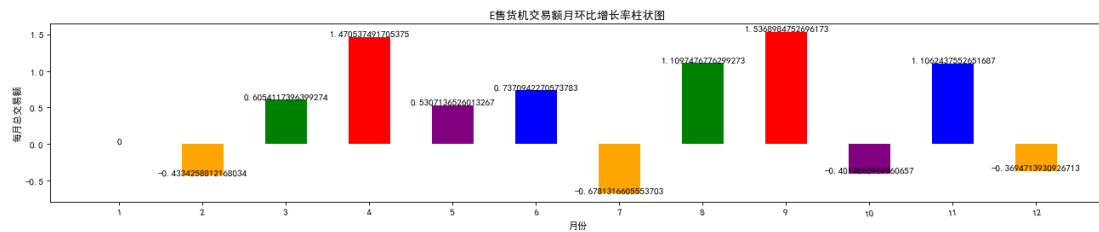


从图表 5 可以清晰的看出每台售货机的每个月的总交易额及 2017 年各个月份的交易额变化趋势，同时将交易额月环比增长率通过下面的图表 6 进行可视化。

图表 6 每台售货机交易额月环比增长率柱状图

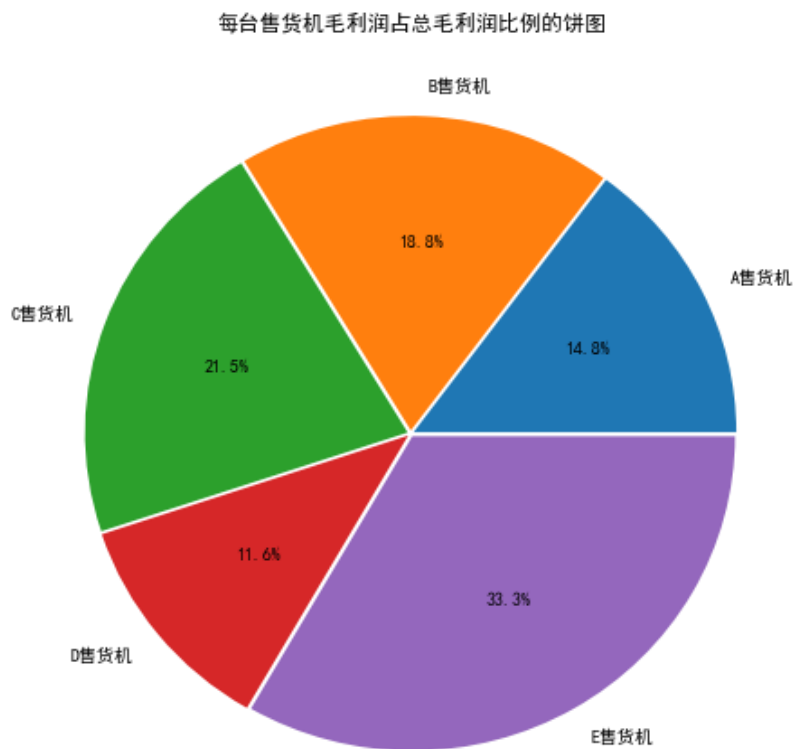






### 任务 2.3

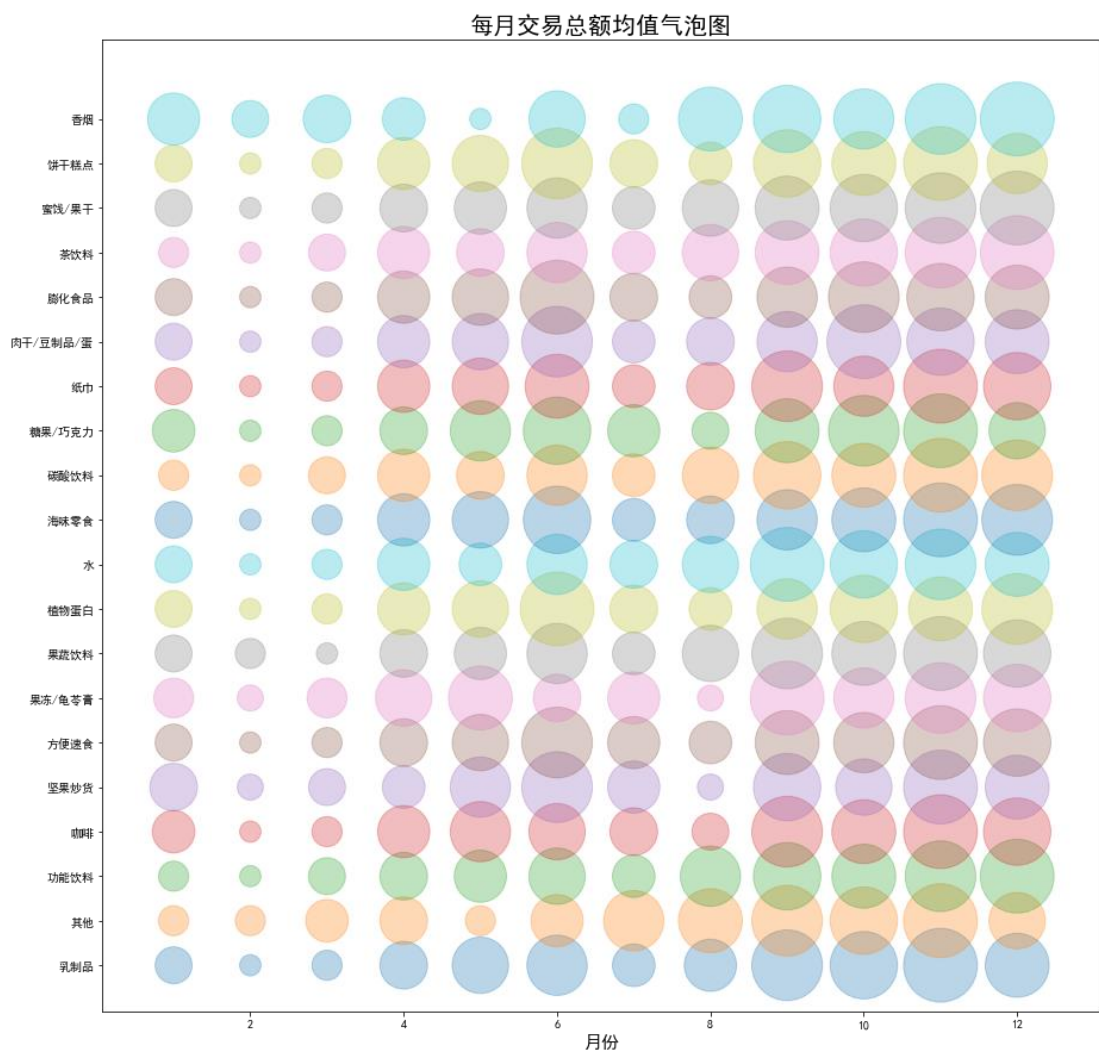
图表 7 每台售货机毛利润占总毛利润比例的饼图



由图表 7 可知，E 售货机毛利润占总毛利润的比例最高，较其他 4 台售货机领先，对于利润贡献最为突出。而 A 售货机与 D 售货机的毛利润占比较低，其中货品销量与其饮料类别占比较低均有一定的关系。

### 任务 2.4

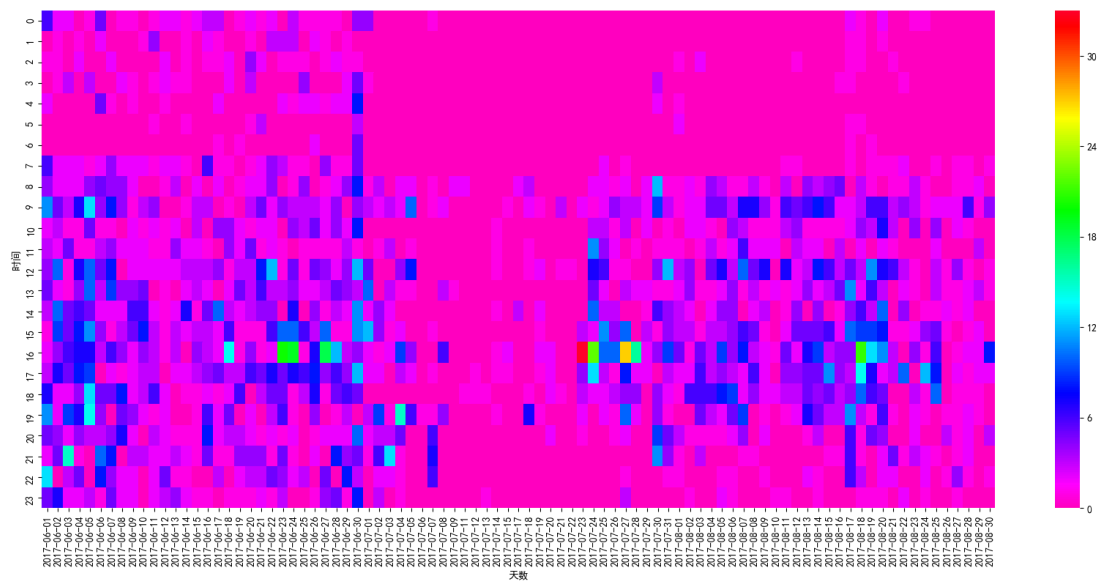
图表 8 每月交易额均值气泡图



图表 8 以时间（月份）为横轴，以商品的二级类目为纵轴，绘制了每月交易额均值的气泡图。总体来看，售货机订单主要集中在夏季与秋季，因为夏季与秋季对于货品尤其是饮料类的需求量较高，因而交易额均值较高，在气泡图中体现为气泡体积较大。而 1-3 月气候较为寒冷，7-8 月为暑期放假，对于货品尤其是饮料类的销量打击较大，因此体现为气泡较小。

## 任务 2.5

图表 9 售货机 C 6、7、8 三个月订单量热力图



横轴以天为单位，纵轴以小时为单位，绘制售货机 C 6、7、8 三个月订单量的热力图，图表 9 为所示。其中暖色调表示订单量高，冷色调表示订单量较为低。从横轴来看，6、7、8 三个月中，7 月份的订单量较其余两月较少；从纵轴来看，凌晨时间段（0 点-8 点）订单量明显较少，而下午时间段（16 点前后）较为集中。

### 3、售货机画像的探索

#### 任务 3.1

根据 5 台售货机销售数据，给出每台售货机商品的标签，结果保存为 csv 文件

#### 任务 3.2

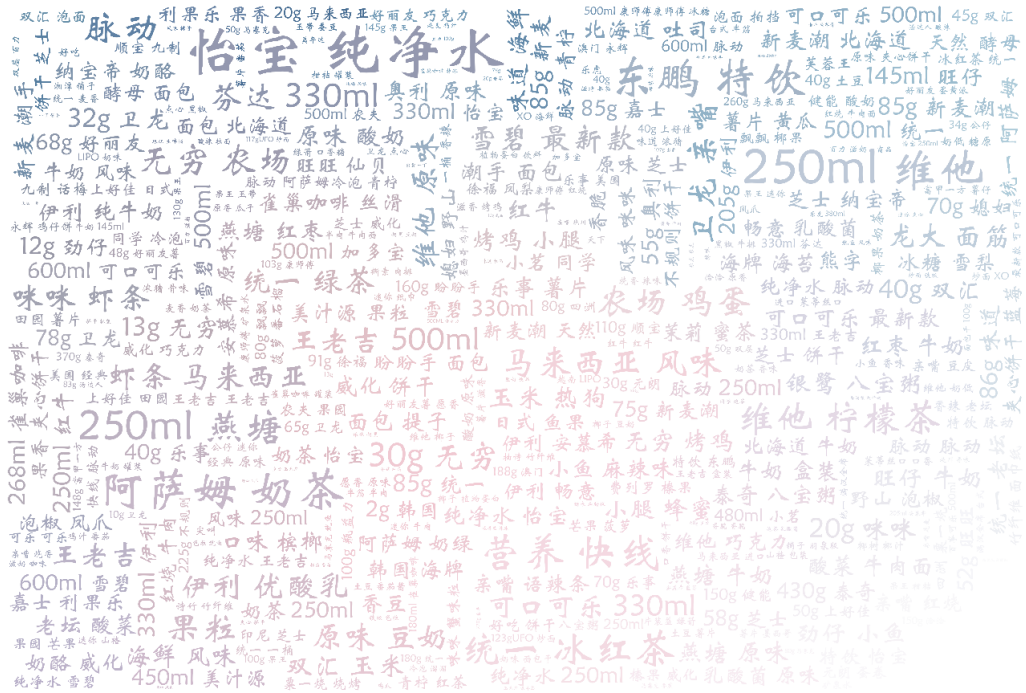
图表 10 A 售货机画像词云图



图表 11 B 售货机画像词云图



图表 12 C 售货机画像词云图



图表 13 D 售货机画像词云图





[illegible]

从销售种类品牌来看，消费群体的品牌选择倾向较为明显，口碑好的产品如怡宝、维他等 5 台售货机均占销量前三。

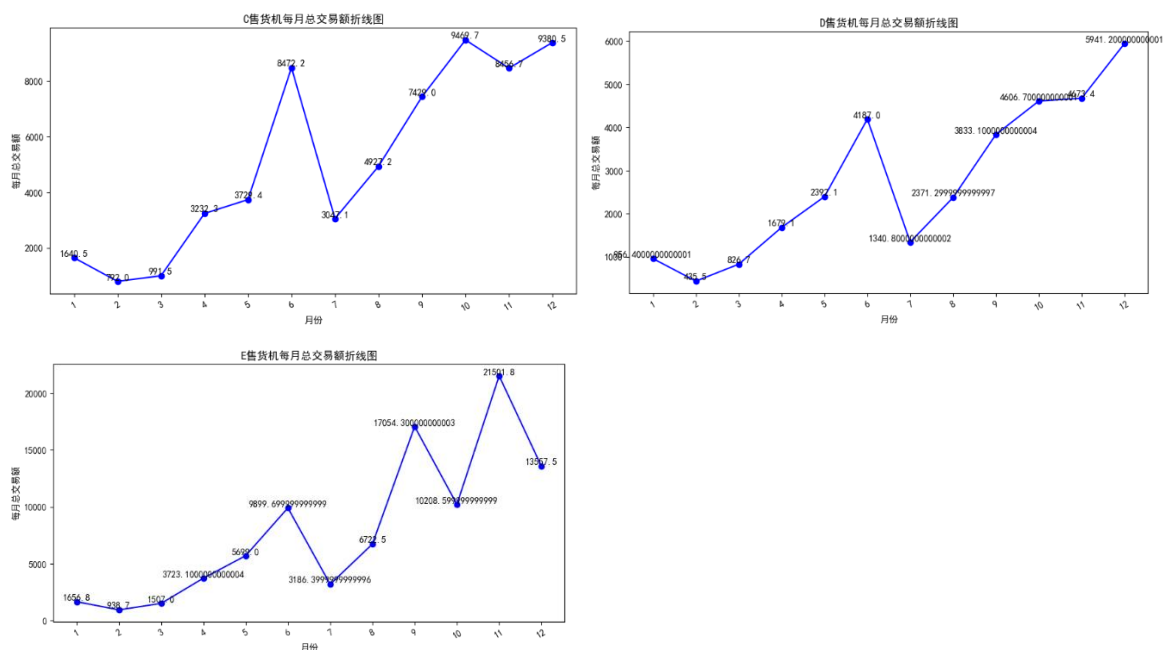
为了预测 2018 年 1 月份的销量，并给出相应的货品补充计划，本次选取 5 台售货机的数据图进行分析，销售统计时段为 2017 年 1 月 1 日到 12 月 31 日。

**A 售货机每月总交易额折线图**

月份	每月总交易额
1	1500
2	400
3	914
4	1800
5	3388
6	6755
7	1950
8	2238
9	4472
10	6292
11	5145
12	7587

**B 售货机每月总交易额折线图**

月份	每月总交易额
1	1372
2	657
3	957
4	2457
5	3694
6	7550
7	1516
8	3514
9	7207
10	8334
11	8669
12	8104



“新零售”实质是创新其运营模式，它将自助零售和线下叠加融合，顾客既能够在线下亲身进行用户体验，也能够直接在自动售货机便利购买。这样的结合模式使这一新型的零售模式兼具实业零售的实际具体和电子商务的快捷高效。

消费群体对货品的需求越来越多样化，因而科学有效地预测货品销量，并根据预测合理安排货品补货量变得至关重要。BP 神经网络是人工智能中的代表性方法，具有优良的非线性预测效果。以 BP 神经网络技术为主线，结合 ARIMA 模型和主成分分析，分别对售货机的月度销量和年度销量进行预测。在汽车月度销量预测问题中，分别使用了 ARIMA 模型和 BP 神经网络单独进行了预测。针对月度数据具有的周期波动特点，对 ARIMA 模型采用具有季节调整的形式。结果显示，这两种模型对于预测 2018 年 1 月份的销量均具有良好的预测精度。

为进一步提高预测效果，根据预测误差平方和最小的原则将两种模型进行组合，构建新的组合预测模型，结果表明在月度销量预测问题中，采用了地点、时间、季节等指标对货品销量进行预测。运用灰色理论分析得到选用的指标和货品销量的关联度，其中货品销量和季节与货品种类的关联度最高。将这些指标标准化后作为 BP 神经网络的输入值对销量进行预测，预测结果的平均相对误差为 2.34%。针对指标中可能存在的信息冗余，采用了主成分分析方法对输入数据进行降维处理。结果显示，主成分 BP 神经网络预测的平均相对误差为 1.71%，比 BP 神经网络降低 0.53%。