目录

[一、 背景介绍 2](#_bookmark0)

[1、 项目背景 2](#_bookmark1)

[2、 项目目标 2](#_bookmark2)

[3、 自动售货机简介 3](#_bookmark3)

[二、 数据说明 4](#_bookmark4)

[1、 数据来源 4](#_bookmark5)

[2、 数据采集 4](#_bookmark6)

[3、 数据量 4](#_bookmark7)

[三、 数据预处理与分析 5](#_bookmark8)

[1、 提取每台售货机销售数据 5](#_bookmark9)

[2、 计算5月份交易额、订单量等数据 5](#_bookmark10)

3、 计算每月的每单平均交易额与日均订单量 6

四、 数据的分析与可视化 7

1、 绘制柱状图并进行分析 8

2、 绘制折线图、柱状图并进行分析 8

3、 绘制饼图并进行分析 9

4、 绘制气泡图并进行分析 10

5、 绘制热力图并进行分析 10

五、 生成各自动售货机画像 11

1、 给出每台售货机饮料类商品的标签 11

2、 生成各售货机画像并进行分析 11

六、 业务预测 15

1、 预测原理 15

2、 预测结果原因分析 15

六、 小结 15

# 背景介绍

## **1、 项目背景**

自阿里巴巴创始人马云提出了新零售这一概念以后，很多的线上企业纷纷开始布局线下，许多没有有效的进行线下布局的传统电商，都开始不同程度显示出了疲软状态，其中一些对物流要求比较高的电商企业压力更加明显。在消费市场风向快速变化的同时，若是仅仅依靠线上流量是很难把握消费者的。而作为零售新的服务终端之一的自动售货机，无疑给这些企业带来了福音。

自动售货机以线上经营的理念，提供线下的便利服务，以小巧、自助的经营模式节省人工成本，让实惠、高品质的商品触手可及，成为当下零售经营的又一主流模式。自动售货机内的商品的供给频率、种类选择、供给量、站点选择等是自动售货机运营者需要重点关注的问题。因此，科学的商业数据分析能够帮助经营者了解用户需求，掌握商品需求量，为用户提供精准贴心的服务，是掌握经营方向的重要手段，对自动售货机这一营销模式的发展有着非常重要的意义。

## **2、 项目目标**

1. 根据自动售货机的经营特点，对经营指标数据、商品营销数据及市场需求进行分析，完成对销量、库存、盈利三个方面各项指标的计算，按要求绘制对应图表，分析每台售货机2018年1月商品销量的预测问题。
2. 为每台售货机所销售的商品贴上标签，使其能够很好地展现销售商品的特征。

## **3、 自动售货机简介**

自动售货机和实体店铺的高昂成本相比较，自动售货机的优势其实十分明显，能够有效的避免企业资产过重和人员管理的问。另外就是自动售货机智能化程度高，不但能够实现常规的购买，也能支持线上支付，线下取货。

自动售货机的大小尺寸可以根据消费者的需求进行定制，所以部署的机动性更加强，能够让销售的范围推进到消费群体的百米范围内，这一点对于追求效率和用户体验的生鲜企业来说绝对是超值的回馈。相对于管理门店的营业员和开展标准化服务，对机器的管理比对人员的管理更加容易，而且现在自动售货机设备都是非常的智能化，这样使得企业轻松的对每个网点进行标准管理。

智能化的自动售货机最大的特点是兼具了基础零售与经营数据化。基础零售功能除可支持线上线下的互动，同时可以兼顾传统零售，而在此过程中，自动售货机基本就成为了一种线下引流终端。重要的事还不仅于此，由此产生的消费数据，借助云平台和数据分析系统，使得企业可以对消费者的行为偏好进行更精准的掌控，借此优化采购、物流、销售服务等整个体系，让企业运转效率更高，更为贴近客户需求。

# 

# 数据说明

## **1、 数据来源**

数据来源：某商场5台自动售货机销售平台

数据截止时间：2017 年 12 月 31 日

## **2、 数据采集**

商品销售信息采集：通过八爪鱼采集器，获取到商场5台自动售货机在2017年1月1日至2017年12月31日期间的订单号、设备ID、应付金额、实际金额、商品、支付时间、地点、状态、提现等商品销售信息。

商品分类信息采集：通过八爪鱼采集器，获取不同商品的大类、二级类信息。

## **3、 数据量**

商品销售信息：7068条 商品分类信息：315条

# 数据预处理与分析

单纯从数据分析和机器学习建模的角度，数据的预处理是极为重要的一个环节。对于数据预处理，包括清洗、排序、缺失值的填充、异常值的删补，平稳性检测、统计量分析、相关性分析、主成分分析、滤波处理、降维、归一化等。在此次项目数据的预处理与分析过程中我们发现了一个脏数据，即2017-2-29号的数据信息，做法是把它删除后进行其他操作。

## **1、提取每台售货机销售数据**

附件1的数据已经以excel形式给出，直接使用函数搜索获取关键信息即可。

搜索 函数函数对信息提取即可

优点：提取过程简洁，速度较快。

缺点：提取结果准确性与信息内容相关。

使用h函数提取数据，read\_csv提取，to\_csv生成文件

## **计算5月份交易额、订单量等数据**

表1每台售货机2017年5月份对交易额、订单量

|  |  |  |
| --- | --- | --- |
|  | 交易额 | 订单量 |
| A | 2816.8 | 756 |
| B | 3223.8 | 869 |
| C | 2993.8 | 789 |
| D | 2047.9 | 564 |
| E | 4996.1 | 1292 |
| 合计 | 16078.4 | 4270 |

分析：由表1可以看出，某商场中不同地点的五台无人智能售货机中，地点E5月份的交易额和订单量最多，地点D5月份的交易额和订单量最少，但总体来看有3个地点的交易额和订单量较为平均。

## **3、计算每月的每单平均交易额与日均订单量**

表2 每台售货机每月的每单平均交易额

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 每台售货机每月的每单平均交易额 | | | | | |
| 地点  月份 | A | B | C | D | E |
| 1 | 3.600597 | 3.416667 | 3.646966 | 3.427799 | 3.773729 |
| 2 | 3.280702 | 3.139459 | 3.567633 | 2.975177 | 3.404264 |
| 3 | 3.177255 | 3.346415 | 3.526616 | 3.479167 | 4.134286 |
| 4 | 3.620134 | 3.502156 | 3.754087 | 3.529571 | 3.586927 |
| 5 | 3.725926 | 3.709781 | 3.794423 | 3.631028 | 3.866950 |
| 6 | 3.505333 | 3.550108 | 3.580553 | 3.631442 | 3.489780 |
| 7 | 3.576261 | 3.617971 | 3.533246 | 3.497792 | 3.493358 |
| 8 | 3.120270 | 3.399796 | 3.621922 | 3.179720 | 3.612960 |
| 9 | 3.664423 | 3.489341 | 3.641776 | 3.454934 | 3.668191 |
| 10 | 3.487987 | 3.599062 | 3.692554 | 3.476560 | 3.506986 |
| 11 | 3.737759 | 3.493255 | 3.646629 | 3.463306 | 3.771972 |
| 12 | 3.398952 | 3.386697 | 3.566877 | 3.306494 | 3.703690 |

分析：从表2中看出，每台售货机每月的每单平均交易额都为3点多，相差不大，只有地点D2月份的每单平均交易额较低和地点E4月份的每单平均交易额较高。

表3 每台售货机每月的日均订单量

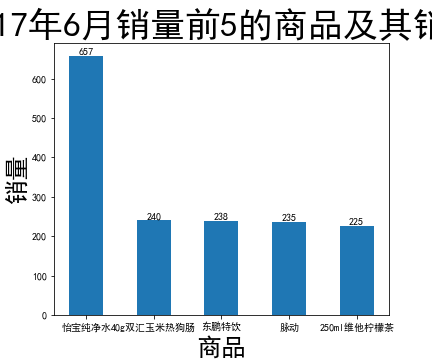
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 地点  月份 | A | B | C | D | E |
| 1 | 10.806452 | 11.806452 | 12.225806 | 8.354839 | 11.419355 |
| 2 | 4.071429 | 6.607143 | 7.392857 | 5.035714 | 9.214286 |
| 3 | 8.225806 | 8.548387 | 8.483871 | 6.193548 | 11.290323 |
| 4 | 14.900000 | 20.100000 | 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 | 57.00000 |
| 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 |

分析：由表3可以看出每台售货机的日均订单量在4月-6月、10月-12月比较多，而在1月-3月、7月-9月比较少，说明了购买者在春秋两季在售货机购买量多，在冬夏两季购买量少，这可能跟季节和气候有关。

# 数据的分析与可视化

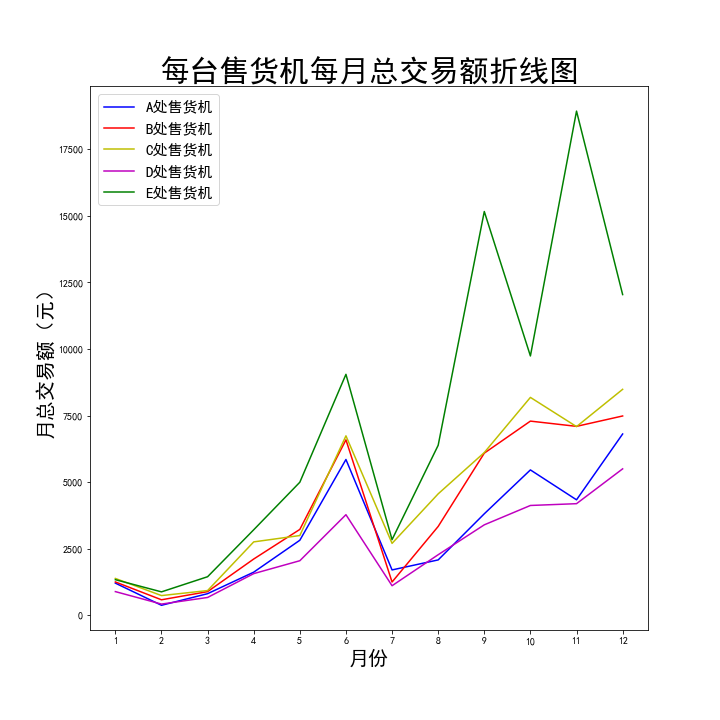
数据可视化是数据分析很重要的一部分，它能帮助我们更好的从繁杂的数据中更直观更有效的获取信息。   
其中，matplotlib是用来创建图表的工具包之一。其目的是为Python构建一个Matlab式的绘图接口，结果除了图形界面显示，还可以把图片保存为pdf、svg、jpg、png、gif等形式。matplotlib是可视化必须要掌握的基础。

## **1、 绘制柱状图并进行分析**

****

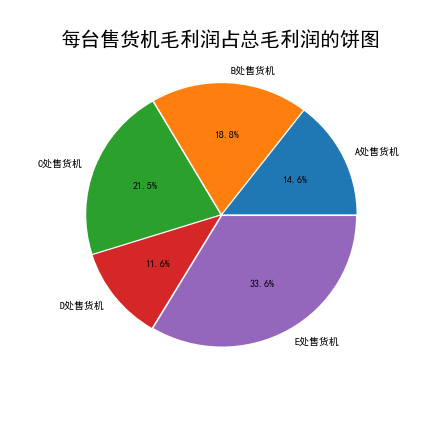
分析：由柱状图可以看出，2017年6月销量前五的商品分别为：怡宝纯净水40g、双汇玉米热狗肠、东鹏特饮、脉动、250ml维他柠檬茶。其中怡宝纯净水40g销量最高且远高于其他四个，说明人们更多得喜欢在无人智能售货机购买纯净水。另外，前五的商品中有四个是饮料类的，说明了购买者多为觉得在无人机购买饮料类的东西方便些。

## **绘制折线图、柱状图并进行分析**



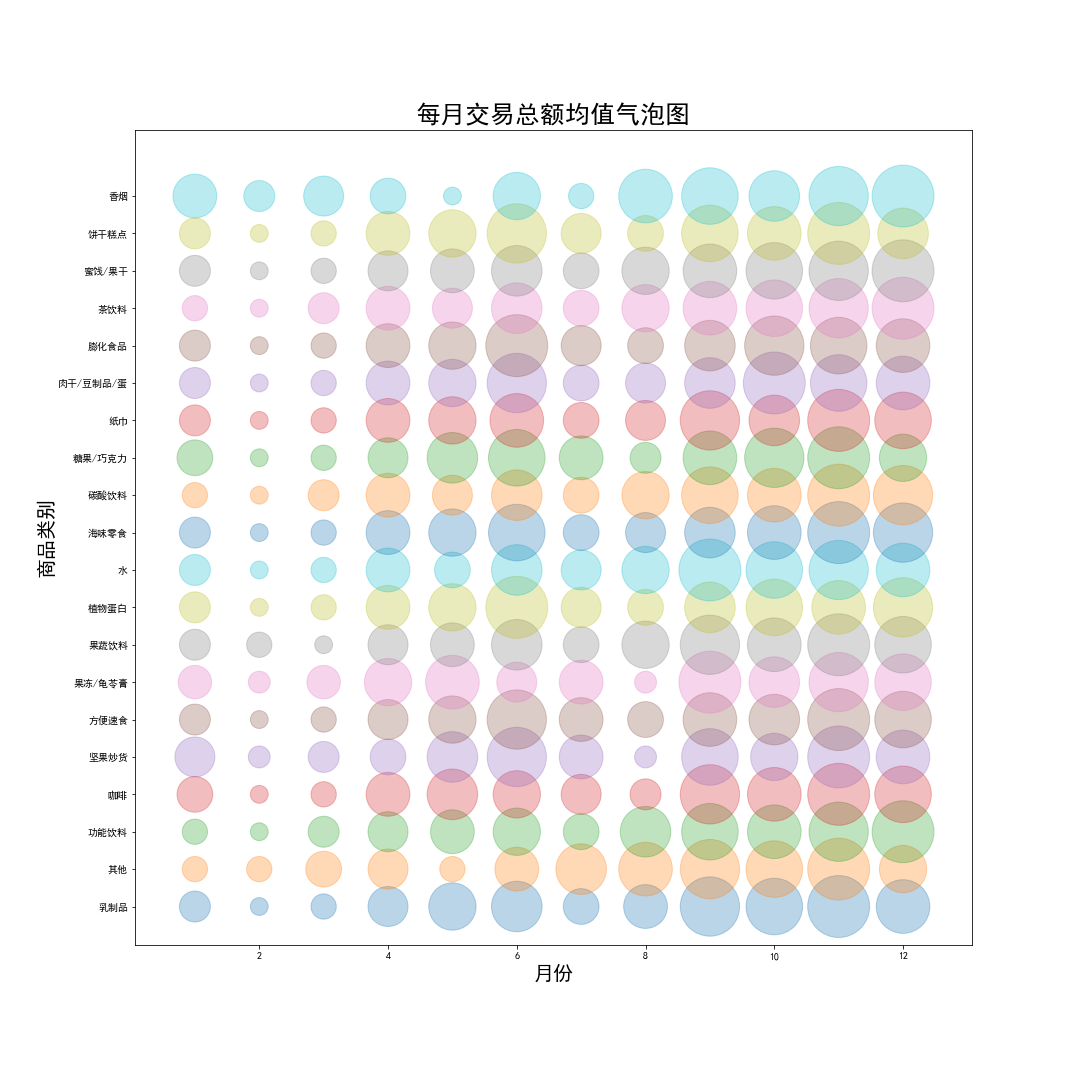
分析：由折线图可以看出，每台售货机的月总交易额呈现递增趋势，其中E处售货机递增趋势最明显且其每月总交易额都大于其他售货机的每月总交易额。体现了地点E处的售货机销售较好，其他四个地点较为平均。

## **绘制饼图并进行分析**



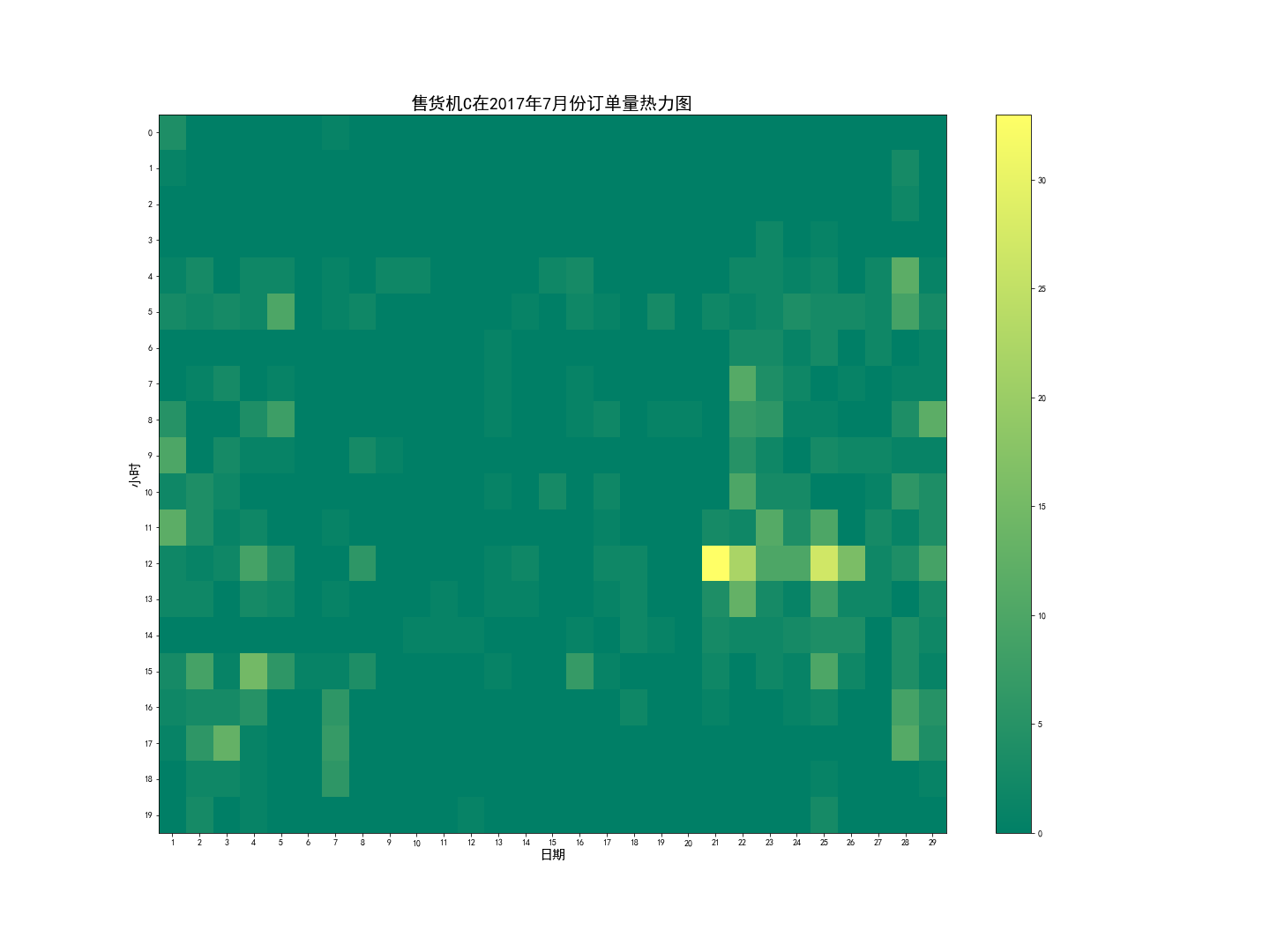
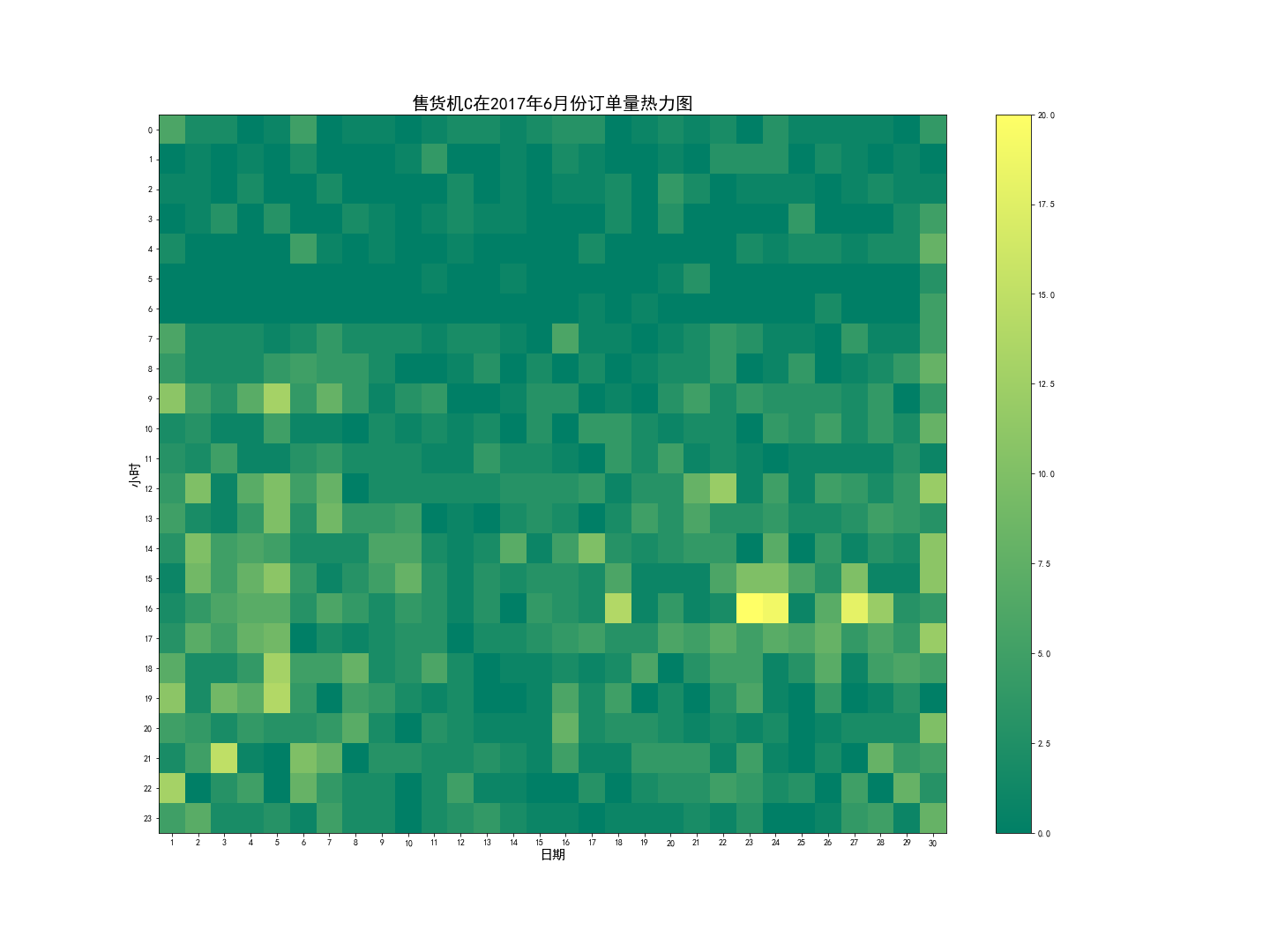
分析：由饼图可以看出，A处售货机毛利润占总利润的14.6%、B处售货机毛利润占总利润的18.8%、C处售货机毛利润占总利润的21.5%、D处售货机毛利润占总利润的11。6%、E处售货机毛利润占总利润的33.6%。D处售货机获利最少，E处售货机获利最多。

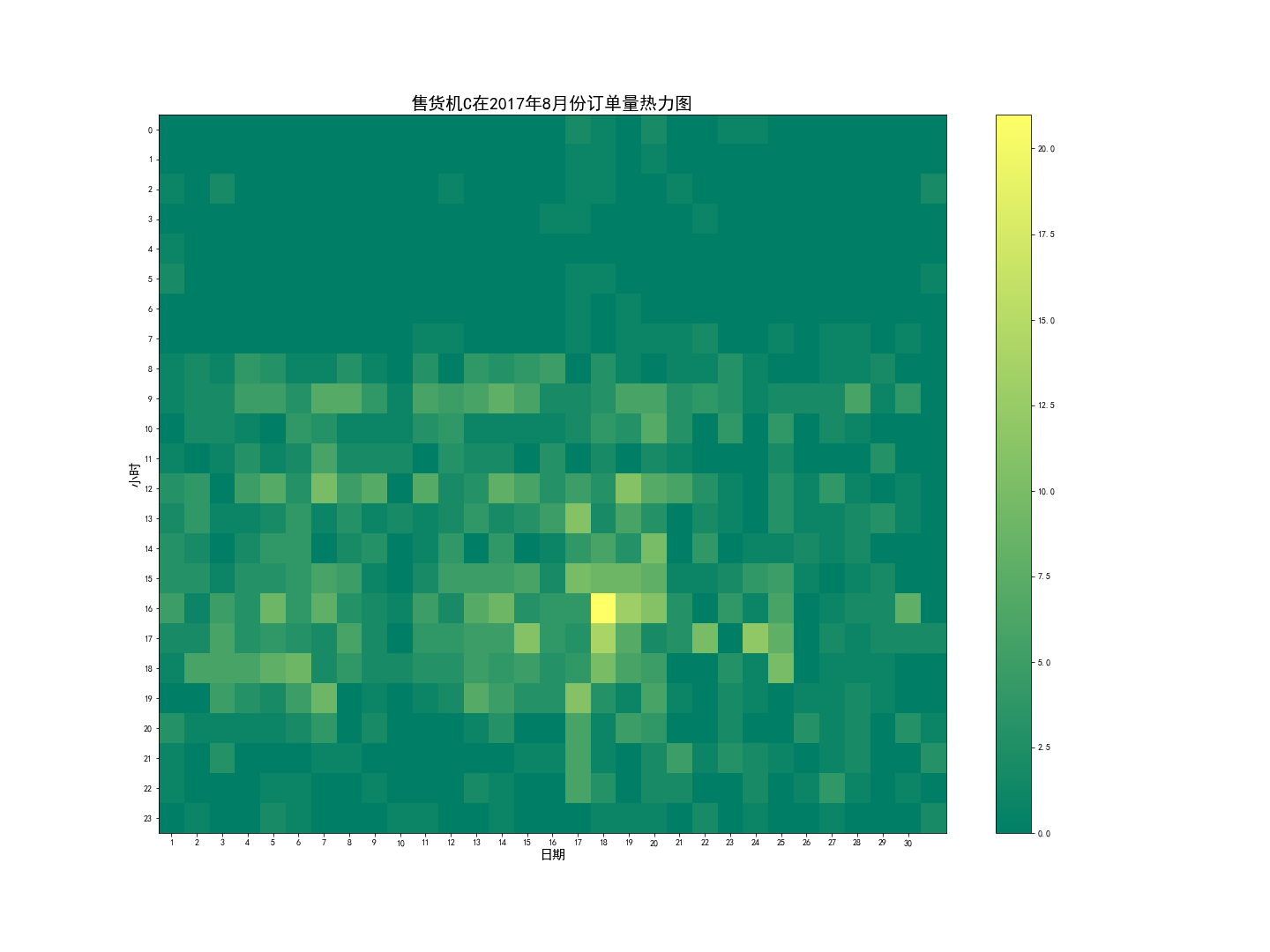
## **绘制气泡图并进行分析**



分析：由气泡图可以看出，每类商品的交易总额均值在4月-6月，10月-12月之间比较多，在1月时商品的交易总额均值总体都很少。

## **绘制热力图并进行分析**



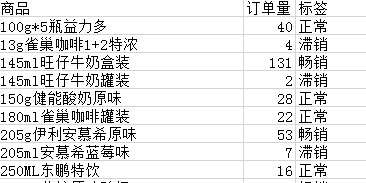


分析：由以上三个热力图可以看出，售货机C在六月份和八月份的订单量要比七月份对得多。说明七月份使用无人智能售货机购买商品的人很少，这也许跟气候有关。

# 生成各自动售货机画像

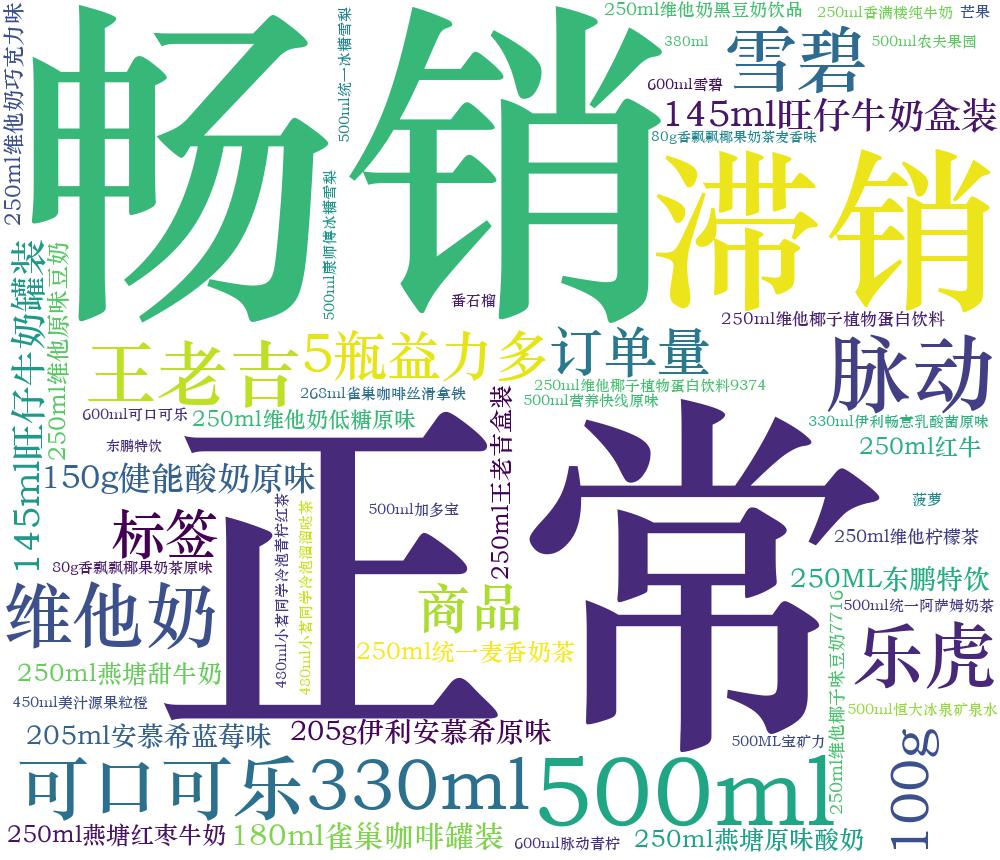
## **给出每台售货机饮料类商品的标签**

通过分析售货机商品销售数据，总结规律，给出了每台售货机饮料类商品的标签。如下表给出售货机A部分饮料类商品标签：



分析：从表中可以看出，有的商品畅销、有的商品正常、有的商品滞销，且每台售货机的商品标签可能都不太一样，因此每台售货机存放的商品应该因地制宜。

## **生成各售货机画像并进行分析**



地点A售货机画像

分析：由画像可以看出，地点A售货机大部分商品处于畅销和正常状态，只有少量是滞销的。可口可乐、雪碧、脉动等最为畅销。



地点B售货机画像

分析：由画像可以看出，地点B售货机商品处于畅销和滞销状态相当。可口可乐、小茗同学、500ml王老吉等最为畅销。



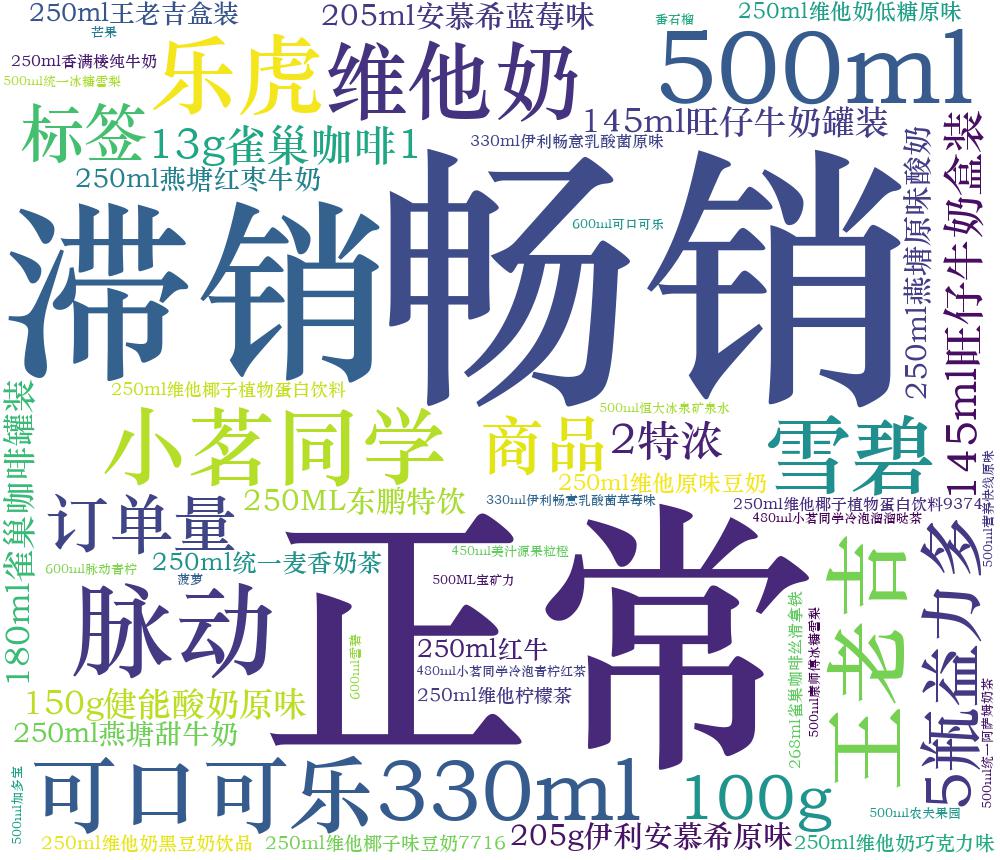
地点C售货机画像

分析：由画像可以看出，地点C售货机商品大部分处于畅销和正常状态。只有极少量处于滞销状态。维他奶、脉动、雪碧等比较畅销。



地点D售货机画像

分析：由画像可以看出，地点D售货机商品大部分处于畅销和滞销状态。只有少量处于正常状态。维他奶、脉动、500ml小茗同学等比较畅销。



地点E售货机画像

分析：由画像可以看出，地点E售货机商品处于畅销、正常、滞销状态相当。维他奶、脉动、500ml小茗同学等饮料类商品很畅销。

# 六、业务预测

## **1、 预测原理**

预测未来销售额的原理是：根据以往的销售情况以及使用系统内部内置或用户自定义的销售预测模型获得的对未来销售情况的预测。销售预测可以直接生成同类型的销售计划。销售计划的中心任务之一就是销售预测，无论企业的规模大小、销售人员的多少，销售预测影响到包括计划、预算和销售额确定在内的销售管理的各方面工作。

## **2、 预测结果原因分析**

我认为根据附件提供的数据无法对每台售货机的每个大类商品在2018年1月的交易额进行预测。因为附件只给出了每台售货机一年的销售信息且每个类别所对应的商品都只是有限的，经过数据预处理并不能找出太多异常数据，数据太少不具备代表性，因此不足以支持预测工作。经营者应在数据采集方面注意采取更广泛，更具代表性的，更大范围的数据。

1. **小结**

随着人们对智能售货机的使用更为熟悉、以及售货商品种类的多元化、支付方式的便捷化，无人智能售货机行业的增长空间巨大。但是离不开科学的商业数据分析。科学的商业数据分析能够帮助经营者了解用户需求，掌握商品需求量，为用户提供精准贴心的服务，是掌握经营方向的重要手段，对自动售货机这一营销模式的发展有着非常重要的意义