

# 项目成果报告

## 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.数据采集 .....	3
3.数据量 .....	3
三.任务分析 .....	4
1.任务一：数据预处理与分析 .....	4
2.任务二：数据分析与可视化 .....	5
3.任务三：生成各自动售货机画像 .....	10
4.任务四：业务预测 .....	13
四.小结 .....	13

# 一.背景介绍

## 1.行业背景

1925 年美国研制出售香烟的自动售货机，此后又出现了出售邮票、车票的各种现代自动售货机。现代自动售货机的种类、结构和功能依出售的物品而异，主要有食品，香烟，美容等自动售货机。而如今，自动售货机在国内也是相当普及，大街小巷的都能看到自动售货机的声影。但相比较外国的自动售货机，国内的自动售货机销售的品种非常单一，主要以瓶饮料售货机为主，技术质量还未完全过关，生产和运营成本较高，收益也不高。目前，中国的自动售货机拥有量约为 19 万台，相对于 13 亿的人口大国，这个数字微不足道。而在日本，平均每 25 人就拥有一台自动售货机，美国是每 50 到 60 人一台，欧洲则是每 50 到 120 人一台。我国许多大中城市已经成为自动售货机的新兴市场，未来潜力及发展方向必将日益为有关厂商关注。按照每 500 人拥有一台自动售货机来算，中国就有上百万台的市场缺口。

## 2.分析目标

这次我们的研究分析的主体就是自动售货机，而我们的分析目标就是根据某地 5 台售货机在 2017 年全年的销售数据，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，按要求绘制相应的图表，并预测每台售货机的销售额，为每台售货机所销售的商品贴上标签，是其能够很好地展现销售商品的特征。

## 二.数据说明

### 1.数据来源

数据来源：某地 5 台自动售货机

数据时间：2017 年 1 月 1 日到 2017 年 12 月 31 日

### 2.数据采集

各台自动售货机的数据都是由商家直接给出的，2017 年这一年的每一笔交易的详情都在附件中给出了，包括卖出商品的售货机在哪里，交易的时间还有交易的价格等等，都在附件一中被全部呈现出来。另外在附件二中也有各种商品的一级分类和二级分类详情。这些数据都是由对方直接给出的。

### 3.数据量

5 台售货机全年总销售量：70680

售货机所卖的商品分类：2 个大类，315 个小类

### 三.任务分析

#### 1.任务一：数据预处理与分析

在任务一中，主要就是进行一系列的数据预处理，包括把每台售货机的销售数据分开，并对应保存在不同的 CSV 文件中。另外还要计算，每台售货机在 2017 年 5 月的交易额，订单量及 5 台售货机的交易额和订单量总数。除此之外，我还会去计算每台售货机每月的每单平均交易额和日均订单量。

任务 1-2 图表：

	A	B	C	D	E	F	G	H
1		售货机E43A6E078A04134	售货机E43A6E078A04172	售货机E43A6E078A04228	售货机E43A6E078A06874	售货机E43A6E078A07631	总计	
2	5月的交易额	3681.2	3385.1	3729.4	5699	2281.1	18775.8	
3	5月的订单量	869	756	789	1292	564	4270	
4								

任务 1-3 图表：

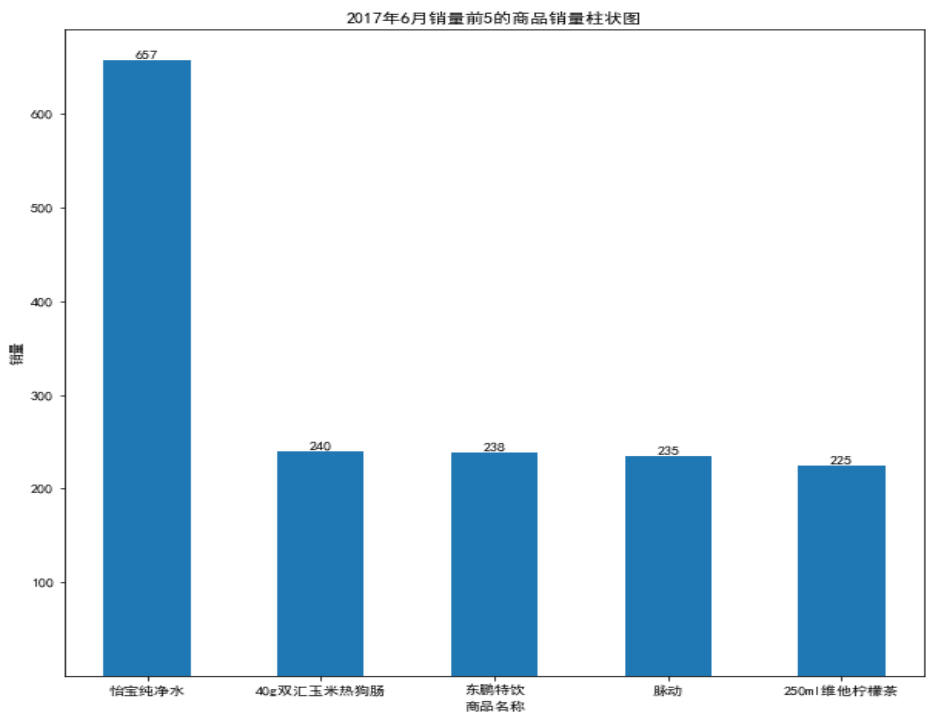
	A	B	C	D	E	F	G
1		售货机E43A6E078A04134	售货机E43A6E078A04172	售货机E43A6E078A04228	售货机E43A6E078A06874	售货机E43A6E078A07631	
2	1月每单的平均交易额	3.75	4.48	4.33	4.68	3.69	
3	2月每单的平均交易额	3.26	3.3	3.83	3.64	3.09	
4	3月每单的平均交易额	3.61	3.55	3.77	4.31	4.31	
5	4月每单的平均交易额	4.08	3.9	4.4	4.16	3.79	
6	5月每单的平均交易额	4.24	4.35	4.73	4.41	4.24	
7	6月每单的平均交易额	4.07	4.12	4.5	3.82	4.03	
8	7月每单的平均交易额	4.4	4	3.99	3.92	4.23	
9	8月每单的平均交易额	3.6	3.35	3.92	3.81	3.32	
10	9月每单的平均交易额	4.13	4.14	4.43	4.13	3.9	
11	10月每单的平均交易额	4.11	3.91	4.27	3.68	3.88	
12	11月每单的平均交易额	4.27	4.17	4.35	4.28	3.86	
13	12月每单的平均交易额	3.67	3.63	3.94	4.17	3.58	
14							
15	1月日均订单量	11.81	5.16	12.23	11.42	8.35	
16	2月日均订单量	6.61	3.39	7.39	9.21	5.04	
17	3月日均订单量	8.55	5.06	8.48	11.29	6.19	
18	4月日均订单量	20.1	8.03	24.47	29.83	14.77	
19	5月日均订单量	28.03	10.03	25.45	41.68	18.19	
20	6月日均订单量	61.83	30.37	62.7	86.36	34.67	
21	7月日均订单量	11.13	8.94	24.65	26.23	10.23	
22	8月日均订单量	31.52	15.29	40.55	56.94	23.06	
23	9月日均订单量	58.17	22.4	55.93	137.8	32.77	
24	10月日均订单量	65.35	27.61	71.48	89.58	38.26	
25	11月日均订单量	67.7	24.8	64.7	167.33	40.33	
26	12月日均订单量	71.29	39.35	76.74	104.87	53.55	
27							

从图表中不难看出，售货机 E43A6E078A06874 无论是销量还是平均交易额都在 5 台售货机中处于领先地位，同期的日均订单量更是遥遥领先其他 5 台售货机。而从总的销售情况来看，年末的日均订单量普遍要比年初的时候要好，也就是说，随着时间的发展，各台售货机的销售情况都在变好，日均订单量和月总订单量较之前都有一定程度的提升。

## 2.任务二：数据分析与可视化

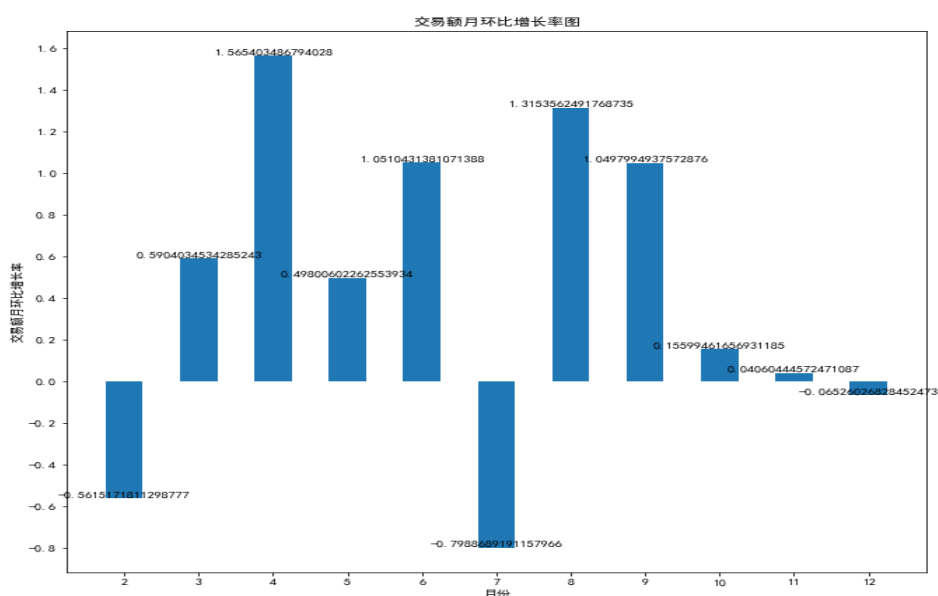
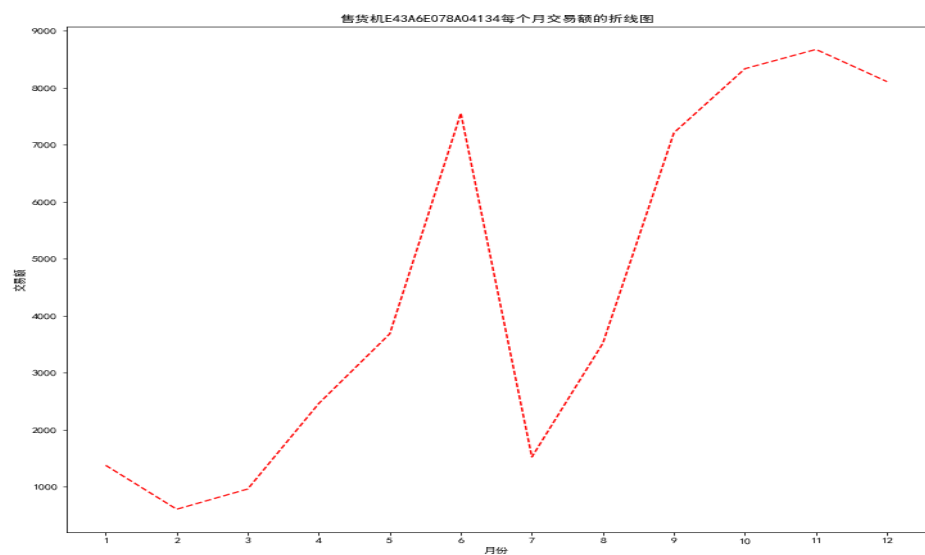
任务二主要就是要根据要求，对数据进行可视化处理，并且画出各种的图表。

任务 2-1 图表：



从图表中我们很容易能够看出来，6月销售量前5的商品分别是怡宝矿泉水，40g 双汇玉米热肉肠，东鹏特饮，脉动和250ml 维他柠檬茶。而在这5个数据里面，怡宝矿泉水又以月销量657瓶遥遥领先其他的商品，成为6月销量最好的商品。

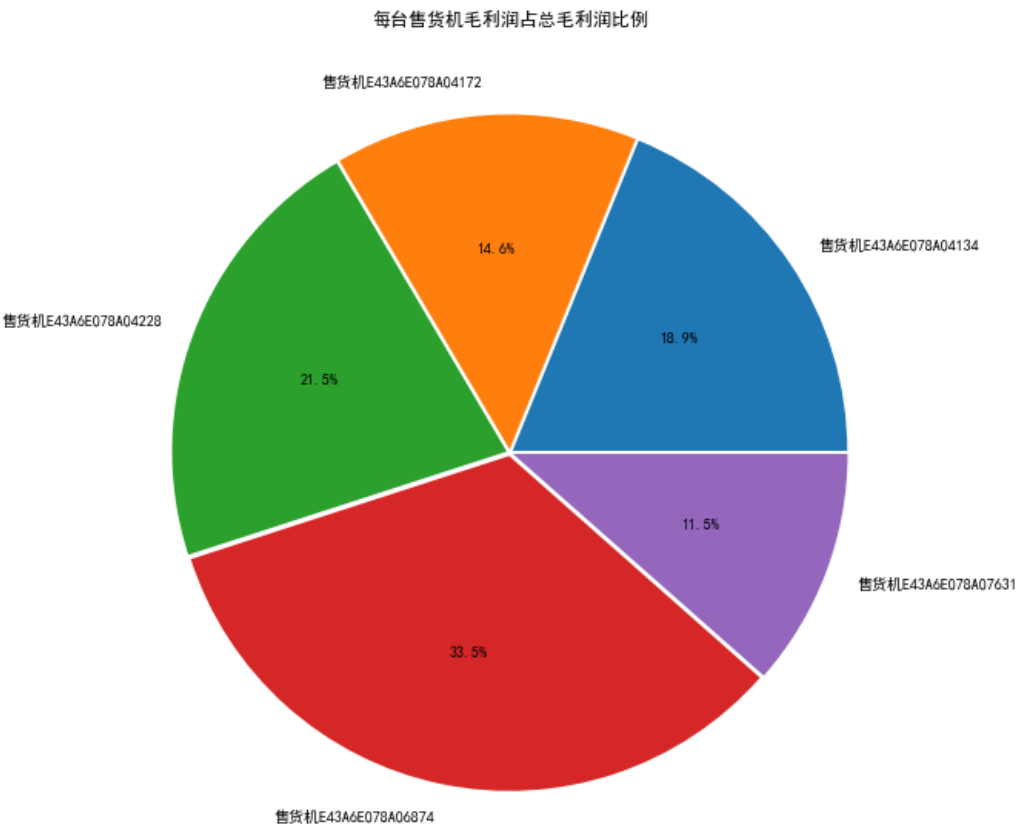
任务 2-2 图表：



任务 2-2 就是要我们探讨每台售货机每月总交易额和交易额月环比增长率，并在此基础上，完成每台售货机每月总交易额的折线图和交易额月环比增长率的柱状图。由于每台售货机的每月交易额的变化情况相近，所以我在这里就用售货机 E43A6E078A04134 的数据来进行一个近似代替分析，其他的售货机折线图和柱状图与售货机 E43A6E078A04134 的图相当接近，就不在这里一一列出分析了。从上图中，我们不难看出，2 月和 7 月是两个交易额下跌较大的月份，交易额月环比增长率分别为-56%和-79%，交易额较之前的月份都有比较明显的下滑。另一方面，4 月，6 月，8 月和 9 月是四个月交易额环比增长超过 100%的月份，也就是说这四个月的月交易额都比前一个月要多出至少一倍，而 4 月是表现最好的月份，月销售额的增长超过了 150%。而在这么多个月份里，售货机 E43A6E078A04134 月交易额最好的月份就是 11 月，总的交易额达到了 8669.9 元。

总的来说，每台售货机的情况都很相近，每个月的销售额几乎都能获得提升，而到了年末，每个月的销售额几乎都能娶到全年的高点，而少数的月份（例如售货机 E43A6E078A04134 在 2 月与 7 月）录得销售额的下降很大的原因也是因为学校放假等等客观原因造成的。这些销售额随着月份逐渐攀升的数据也能一部分地说明人们越来越喜欢到这些售货机那里购物，售货机的销量与人们对售货机的认可程度越来越高了。

任务 2-3 图表：

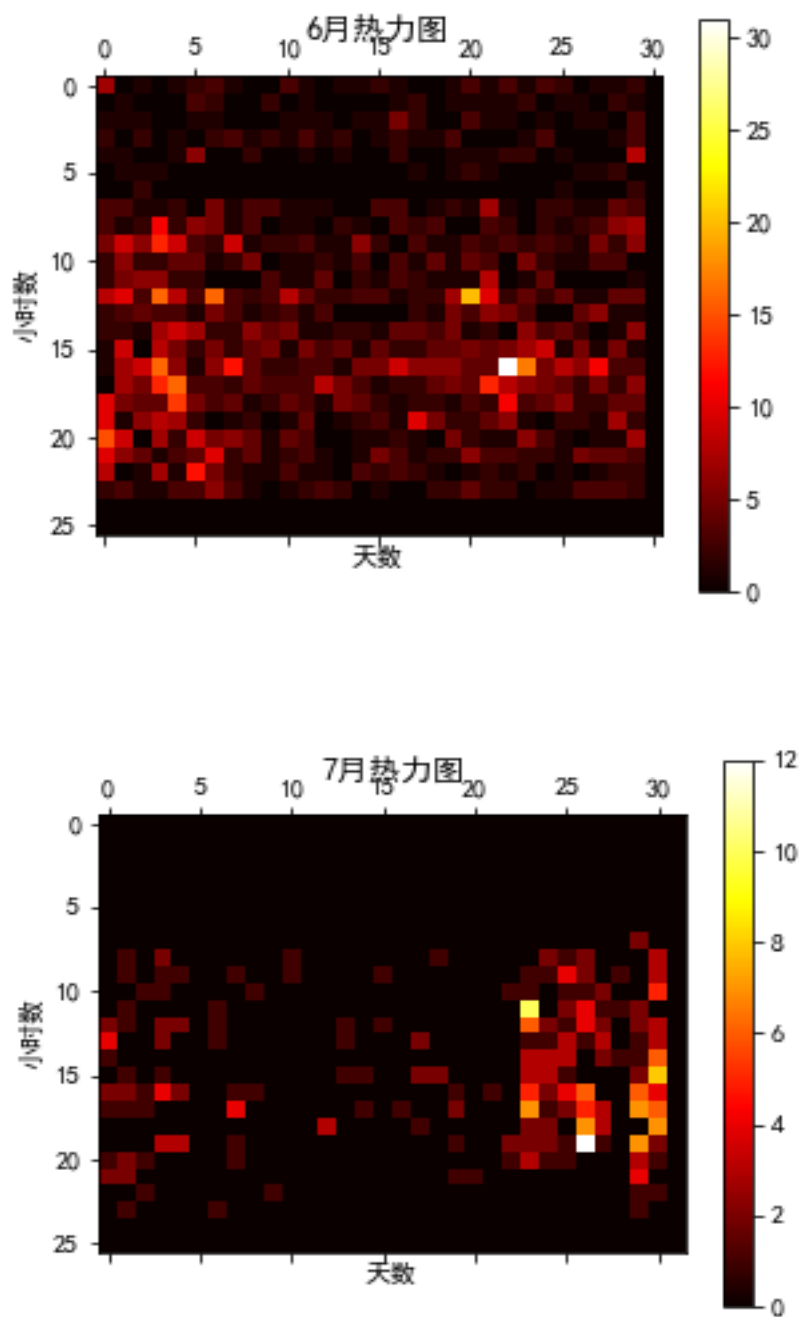


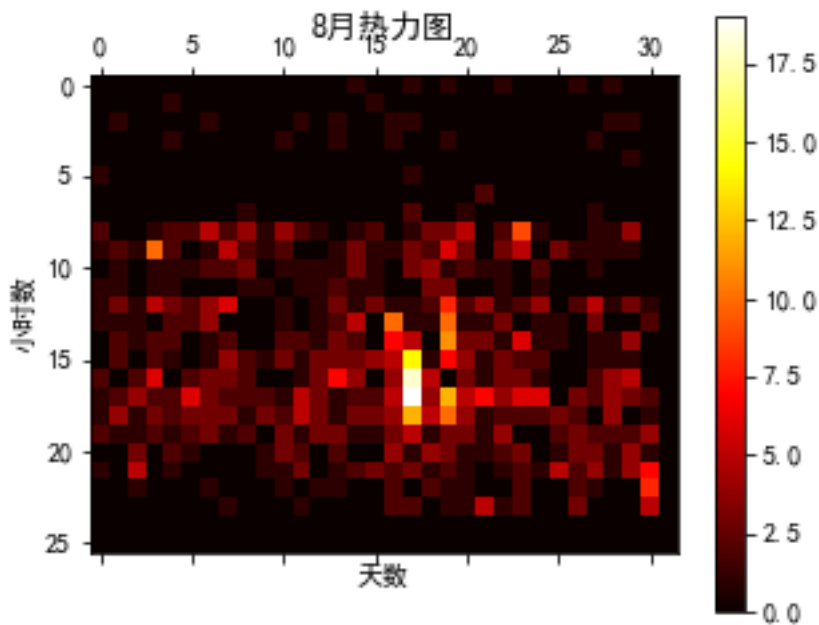
任务 2-3 就是要计算出每台售货机毛利润占总毛利润的百分比，并将这些数据变成饼图呈现出来，具体的结果看上图。在上面这幅利润占比的饼图中，我们能清晰地看到售货机 E43A6E078A06874 的毛利润占总毛利润的比重最大，也就是说，最大一部分的利润是从售货机 E43A6E078A06874 那里得来的，占比达到了 33.5%，而利润占总利润最小的售货机是售货机 E43A6E078A07631，产生的利润仅占总利润的 11.6%。



任务 2-4 和任务 2-5 分别是画交易额均值的气泡图和售货机 E43A6E078A04228（以下简称为售货机 C）在 6, 7, 8 月订单量的热力图，都是用来研究交易情况的一类图表。

任务 2-5 图表：





### 3.任务三：生成各自动售货机画像

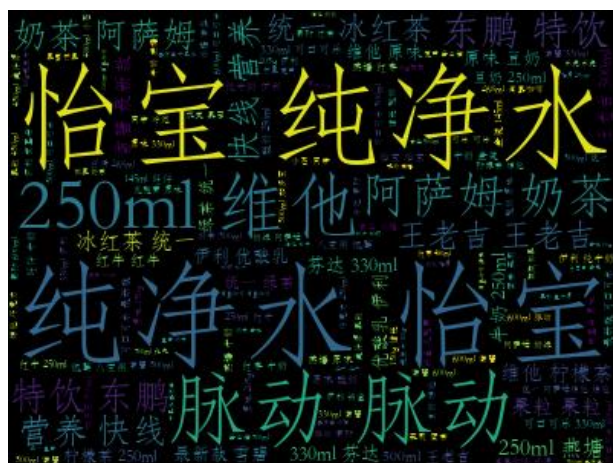
具体的情况可以参考提交的报告结果，这里我只是简单地总结一下结果。

任务 3-1 是要求我们去标各类饮料的标签。首先我在这里用来判断是畅销，正常还是滞销的标准是这台售货机饮料类的平均销量，超过平均销量 50%即为畅销，低于平均销量 50%的即为滞销，其他的即为正常。

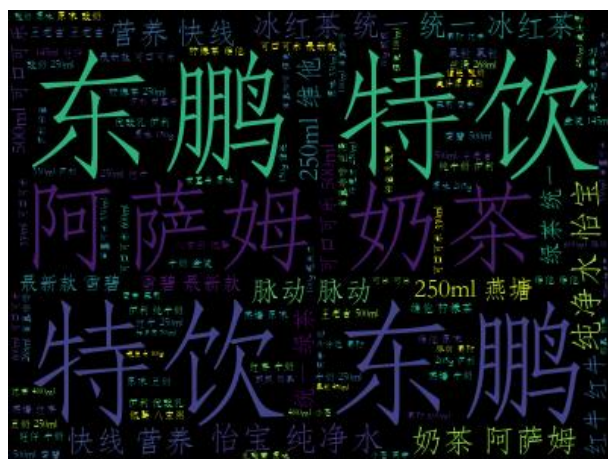
而在统计出来的数据中，我发现极端情况非常明显，也就是说畅销的非常畅销，例如说怡宝纯净水或者东鹏特饮这类特别畅销的商品，平均销量可以是总平均销量的几倍，怡宝纯净水甚至还能卖到上千瓶。而另一边滞销的也非常滞销，一台售货机一年可能就卖几瓶，更有甚



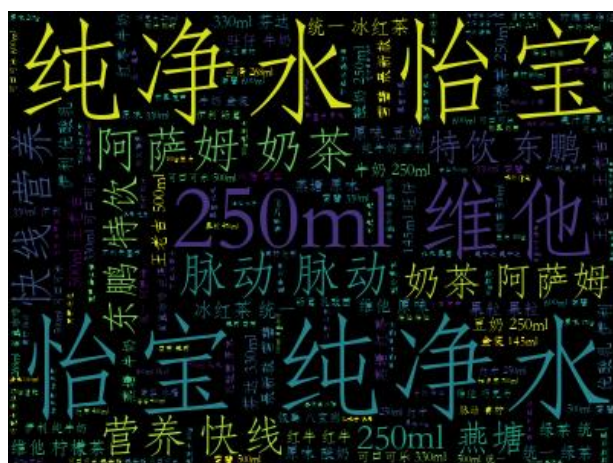
售货机 C 的词云图:



售货机 D 的词云图:



售货机 E 的词云图:



## 4.任务四：业务预测

预测未来销售额的原理是根据过去的数据来分析销售额的规律与走向，以此来预测未来的销售额数据。常用的预测模型有趋势外推预测方法，回归预测方法，卡尔曼滤波预测模型，组合预测模型，BP神经网络预测模型。而这次我采用的预测方法则是时间序列方法，按时间的间隔来做分隔，预测销售额的规律，从而预计出 2018 年 1 月的销售额。由于时间和数据的关系，我只预测了 A, B, C, D 四台售货机的销售额，E 售货机的销售额预测方法其实是一样的，可以按一样的方法进行预测。

A 售货机: 3851.772536

B 售货机: 9253.290620

C 售货机: 10330.085169

D 售货机: 5498.155761

## 四.小结

总的来说，总共 5 台的售货机在 2017 年的表现是越来越好的，月销售量和销售额在年初的时候都不太理想，但到了年末的时候都有比较大的升幅，每台售货机的毛利都有 1 万元左右，月销量也有了明显的涨幅。总体来说，5 台售货机在 2017 年的总利润达到了可观的 54065.75 元。而在 5 台售货机中，售货机 E43A6E078A06874（也就是售货机 E）表现得最好，2017 年一年的利润超过了 18000 元，占据总

利润的 33.5%。而在商品方面，最畅销的商品莫过于怡宝纯净水，总销量达到了 4964 瓶，远远领先于其他的商品，另外一些表现比较好的商品还有东鹏特饮和 250ml 维他柠檬茶等。除此之外，在分析中我发现有些商品的销量非常堪忧，甚至可能会出现一年才卖了 1 件的情况出现，所以我认为要继续发展自动售货机就要清理掉这些滞销商品，特别是那些一年才卖几件的超级滞销货。另外想要继续发展自动售货机，还能引进一些新的销售物品，例如口红等的一些小物品，这些在国外的售货机中已经出现了，而且在售卖了一段时间后，发现购买者给予的反馈还不错，大家也都接受售货机出售这一类商品。