

目录

一、	背景介绍.....	3
1、	行业背景.....	3
2、	分析目标.....	3
二、	数据说明.....	4
1、	数据来源.....	4
2、	数据内容及示例数据.....	4
3、	数据量.....	4
三、	数据简易统计.....	5
1、	售货机 5 月份交易情况.....	5
2、	售货机全年总体交易情况.....	5
3、	售货机各月销售情况统计.....	6
四、	数据分析.....	8
1、	六月销量 top5 及其销售情况(任务 1.1).....	8
2、	交易额变化分析(任务 2.2).....	9
3、	毛利润分析(任务 2.3).....	11
4、	商品在不同月份出售情况分析(任务 2.4).....	13
5、	6、7、8 月订单随时间分析.....	15
五、	售货机画像.....	18
1、	售货机饮料类产品销售情况.....	18
2、	售货机画像及营销意见.....	19
六、	预测.....	20

1、	预测原理.....	20
2、	可预测性判断及解释.....	20
七、	小结.....	23
附、	项目完整的实现过程.....	25

一、 背景介绍

1、行业背景

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、分析目标

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，并预测每台售货机的销售额。

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。

二、 数据说明

1、数据来源

销售数据来自于某商场在不同地点安放的 5 台标号分别为 A, B, C, D, E 的自动售货机, 售货机数据的时间跨度为从 2017 年 1 月 1 日到 2017 年 12 月 31 日。

2、数据内容及示例数据

销售数据包括了全部 5 台售货机的每笔订单的具体情况 (包括订单号, 设备 ID, 应付金额, 实际金额, 商品, 支付时间, 地点, 状态, 提现情况)。

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
DD20170816	E43A6E078	4.5	4.5	68g好丽友巧克力派2枚	2017/1/1 0:53	D	已出货未退款	已提现
DD20170816	E43A6E078	3	3	40g双汇玉米热狗肠	2017/1/1 1:33	A	已出货未退款	已提现
DD20170816	E43A6E078	5.5	5.5	430g泰奇八宝粥	2017/1/1 8:45	E	已出货未退款	已提现
DD20170816	E43A6E078	5	5	48g好丽友薯愿香烤原味	2017/1/1 9:05	C	已出货未退款	已提现
DD20170816	E43A6E078	3	3	600ml可口可乐	2017/1/1 9:41	B	已出货未退款	已提现

分类数据提供的则是售货机出售的每种商品对应的大类 (饮料类、非饮料类) 和二级类 (包括乳制品、饼干糕点、肉干/豆制品/蛋等在内)。

商品	大类	二级类
100g*5瓶益力多	饮料	乳制品
100g越南LIPO奶味面包干	非饮料	饼干糕点
10g卫龙亲嘴烧香辣味	非饮料	肉干/豆制品/蛋
10g越南LIPO奶味面包干	非饮料	饼干糕点
110g顺宝九制话梅	非饮料	蜜饯/果干

3、数据量

销售数据 (附件 1) 提供了 70680 条数据

分类数据 (附件 2) 提供了 315 条数据

三、 数据简易统计

1、售货机 5 月份交易情况

id	交易额	订单量
A	3385.1	756
B	3681.2	869
C	3729.4	789
D	2392.1	564
E	5699	1292
总和	18886.8	4270

通过对这五台售货机在五月份销售情况的提取和数值计算，我们发现在这个月当中，E 售货机的交易额和订单量明显较其它售货机更多，而 D 售货机相比于其它售货机则交易额和订单量明显较少。

2、售货机全年总体交易情况

销售机	交易总额	订单总量	每笔订单交易额
A	42542.6	10486	4.06
B	53970.3	13482	4
C	61572.1	14494	4.25
D	33243.3	8713	3.82
E	95655.4	23505	4.07
全部	286983.7	70680	4.06

当我们将统计的范围放大到全年时，我们发现 E 售货机依旧拥有着高交易总额和高订单量，而 D 售货机各项数值依然较低。又由于在上面的分析中，我们统计了一个自然月的销售情况，我们可以发现，全年的销售结果与 5 月份一个月的销售结果反映在五台销售机的交易总额和订单总额的排序上基本一致，所以，很显然，我们可以认为 E 在大多数情况下都是一个比较受欢迎的购物地点，而 D 则在很多时候都显得很冷清，所以，我们在营销上应该采取**在 E 地点适当增加售货机**以尽量缓解排队情况，进而提升顾客消费热情并**适当提高 E 地点货品投放量**以防止在热门售货区域出现产品脱销的情况；自然，我们应在**D 区域适当减少货品投放量**以减少滞销情况。同时我们发现每笔交易的金额在 A, B, E 三个地点无很大差异，而在 C, D 两个地点却出现了较为明显的偏高和偏低，我们将在后续对销售机画像的分析中，深入分析这一点。

3、售货机各月销售情况统计

首先展示各售货机每月的每单平均交易额：

月份	A	B	C	D	E
1	4.506567	3.753005	4.328496	3.692664	4.680226
2	3.864035	3.255676	3.826923	3.088652	3.638372
3	3.58549	3.614717	3.769962	4.305729	4.305714
4	4.036913	4.07529	4.403678	3.790293	4.159888
5	4.477646	4.236133	4.726743	4.241312	4.410991
6	4.047394	4.06805	4.5017	4.025962	3.817856
7	4.097689	4.401739	3.988351	4.229653	3.919311
8	3.358709	3.5842	3.913582	3.316503	3.804471
9	4.307212	4.130258	4.427294	3.89939	4.125375
10	4.020703	4.11234	4.27333	3.884233	3.676125
11	4.471552	4.268784	4.352393	3.862314	4.283227
12	3.787868	3.667014	3.943043	3.57258	4.168973

通过对上面表格的观察，我们可以较为直观地发现，在 2 月，8 月，12 月三个月中，五台售货机的平均每单交易额相较于其它月份偏低，1 月，5 月，6 月，7 月，9 月，11 月则要偏高一点。

再展示各售货机每月的日均订单量：

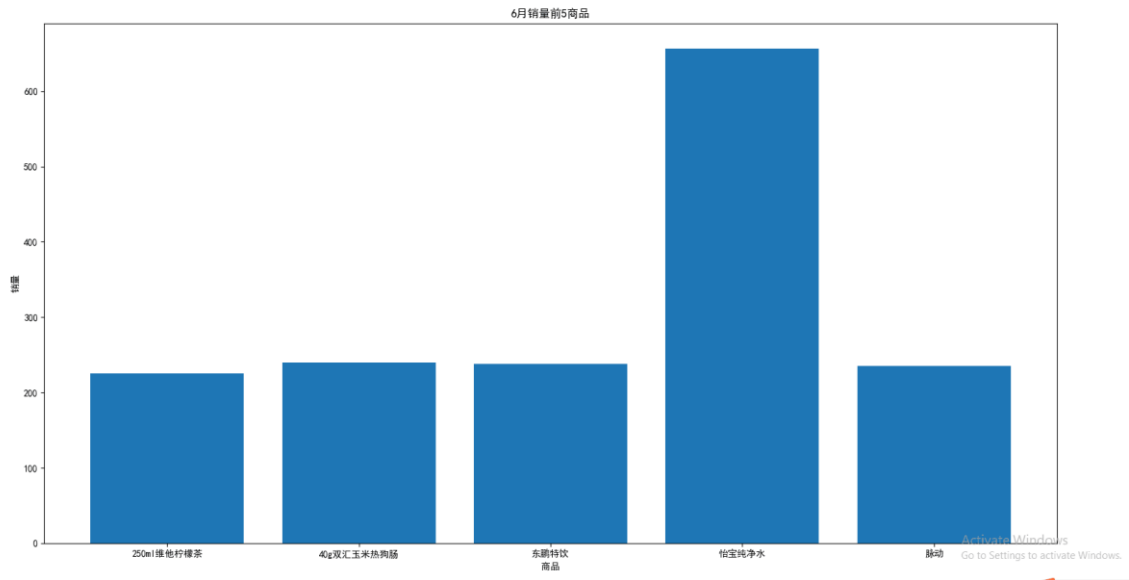
月份	A	B	C	D	E
1	13.4	15.91304	14.57692	11.26087	13.11111
2	4.56	7.708333	7.703704	5.875	9.923077
3	51	44.16667	52.6	48	50
4	14.9	20.1	24.46667	14.76667	29.83333
5	25.2	28.96667	26.3	18.8	43.06667
6	55.63333	61.86667	62.73333	34.66667	86.43333
7	17.62963	11.89655	26.34483	11.74074	27.1
8	21.48387	31.64516	40.6129	23.06452	57
9	34.66667	58.16667	55.93333	32.76667	137.8
10	50.48387	65.35484	71.48387	38.25806	89.58065
11	38.66667	70.03448	64.76667	40.33333	167.3333
12	64.6129	71.29032	76.74194	53.64516	104.9032

同样的，我们通过观察，可以发现 3 月，6 月，10 月，11 月，12 月的销售量均偏高，同时我们发现在 9, 10, 12 月，E 的销售量要明显高于其它地点，而在 D 地点，6 月和 10 月这两个热销时间段却明显的出现了日均销售量严重偏低的情况。

以上是粗略展示统计数据的环节，接下来我们将对包括以上数据在内并联系其他数据的方式进行进一步的分析。

四、 数据分析

1、六月销量 top5 及其销售情况



图片横轴从左至右依次为 250ml 维他柠檬茶，40g 双汇玉米热狗肠，东鹏特饮商品，怡宝矿泉水，脉动。

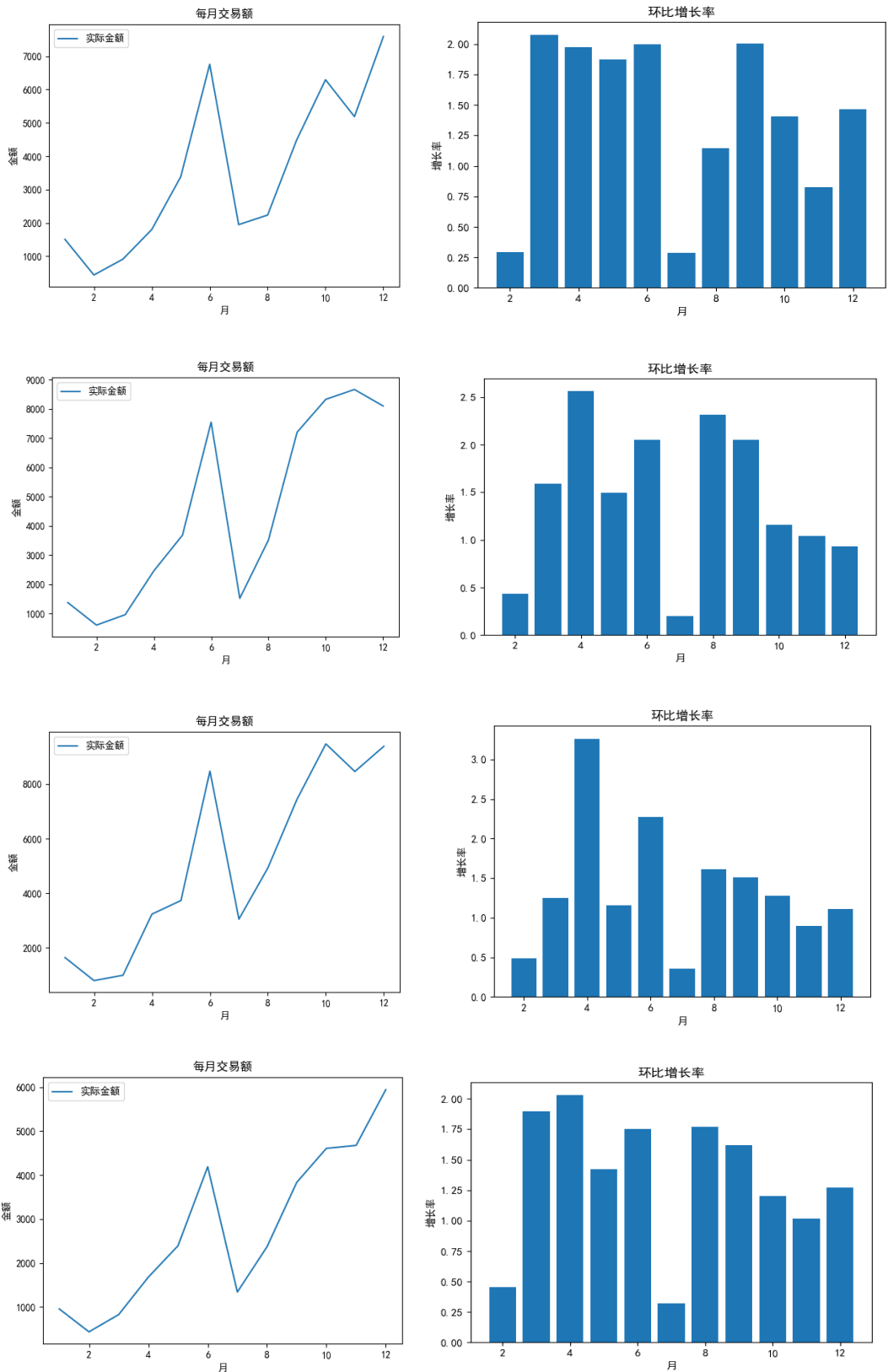
通过对六月份销量 top5 产品统计图的分析，我们可以比较清楚地看到：

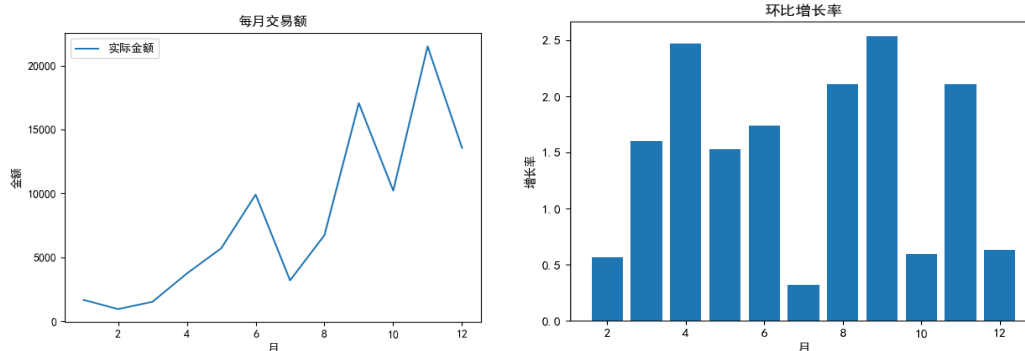
- 1) 在整个六月 top5 中，饮料类产品占据四席；
- 2) 怡宝矿泉水的销售量突破六百，销量超过其他产品一倍以上。

对这些现象进行分析以及营销方法的探索，我们可以看到，六月份饮料类饮品销售状况火热，依据常识，我们认为天气开始变得炎热可能是饮料类产品销售量较高的原因之一，并且在即将到来的 7,8 月，**饮料产品的销量可能进一步提高**；同时，我们猜测，怡宝矿泉水可能具有以下优点使得其在饮品竞争中占据优势：**较低的价格和相比于其它含糖饮料更具有单纯的补充水分的特征**，所以，我认为，在售货机中，**我们可以投放少量价格和利润要比怡宝品牌更高的纯净水/矿泉水产品，同时适当减少怡宝矿泉水的投放**，除此之外，考虑到纯净水对应的消费人群（诸如需要运动后大量补水或者喜好无糖饮料），**我们可以在部分地点尝试投放一些利润更高的无糖类饮料或者诸如 top5 榜单中适宜运动后饮用的运动饮料东鹏特饮等**，在经过一段时间的考察后，再行决定是否推广。对于其他销量较好的产品，我们应继续保持较高的投放量，同时可以宣传推广与以上四种产品相近的商品，促进同类产品的销售。也可以尝试与其他销量一般的产品进行捆绑销售。

2、交易额变化分析

以下是绘制得到的每台售货机的每月交易额与环比增长率(顺序依次为 A, B, C, D, E):





（为方便制图且使图像美观，各售货机“环比增长率”的统计图中，我们的纵轴“增长率”实际为：

本月销售额/上月销售额*100%，理论增长率应为纵坐标减去 1）

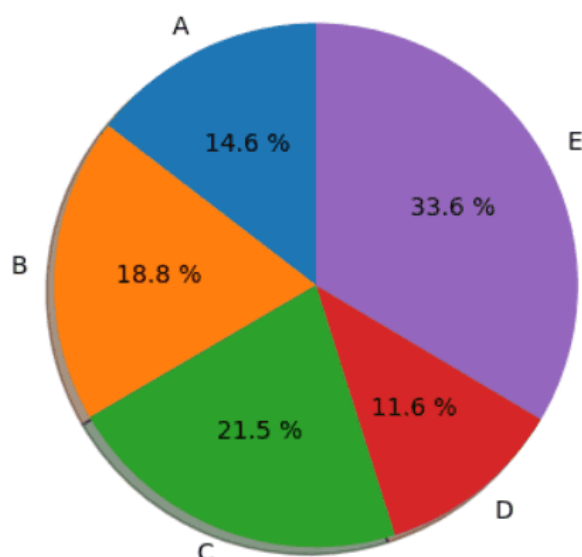
通过观察上面的各售货机每月交易额和环比增长率，我们可以明显发现，在 2 月和 7 月两个月份，我们的各售货机交易额均出现了非常明显的大幅度下降，考虑原因并提出解决措施，我们认为有以下几点：1）有可能是 2、7 两个月正值假期，受到假期人流量影响，2、7 两月购物的人较少，因为售货机位置固定，所以我们没有较好的提高人流量的解决办法，但是**我们可以减少投放保质期短的商品**，防止产品滞销导致出现货物损失进而导致成本增加；2）有可能是因为在 2、7 两月我们采取了不恰当的促销方式或者做出了其他不适当的对销售机售货的改变，我们应当**回顾我们之前在 2、7 月采取的措施，进行调整和反省**，以避免下一年度再出现类似情况。

另外，E 售货机在 10、12 两月同样出现了大幅度的交易额下滑情况，我们可以参考上述第二点（回顾、调整、反省）来做出一些更好的举措。

除这两个月以外，其他月份均有着较高的交易额或者良好的交易势头（高增长率）。

3、毛利润分析

以下为每台售货机毛利润占占总毛利润比例的饼图（假设饮料类毛利润为25%，非饮料类为20%）：



通过上图很明显可以看出 E 售货机的利润额极高，占到总额的三分之一，而 D 售货机的销售额却仅仅只有 E 的三分之一，A, B, C 三台售货机表现一般。这一结果可以佐证我们之前在“售货机全年总体交易情况”中提出的营销要求：**在 E 增加售货机和提高商品投放量，减少 D 处商品投放量。**

同时我们对比“售货机全年总体交易情况”的统计表，并生成一个新的比例表时可以发现：

售货机	交易总额比（E为基准）	利润对比（E为基准）
A	0.445	0.435
B	0.564	0.56
C	0.644	0.64
D	0.348	0.345
E	1	1

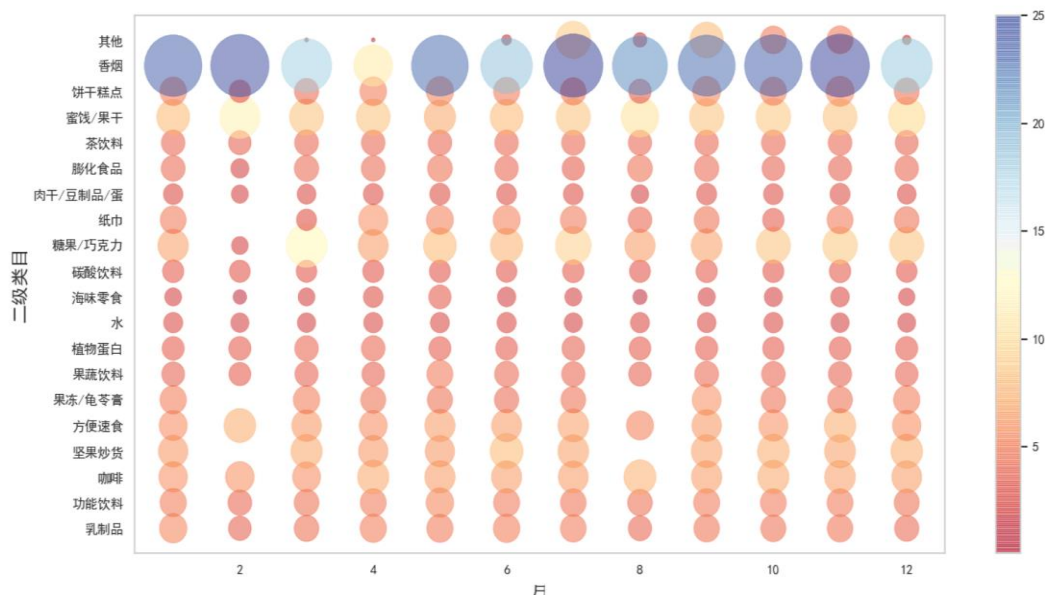
（交易总额比是用各售货机的总交易额除以 E 的，利润对比是用每台售货机毛利润占占总毛利润比例除以 E 的）

通过上表，我们可以发现，五台售货机的交易总额和利润比例大致相当，反

映出每台售货机出售饮品和非饮料类的产品比例大致相当,即五台售货机的数据没有表明某台售货机对某种大类产品（饮料类/非饮料类）具有明显的偏好,所以我们可以维持这些售货机投放商品的大类比（饮料类产品与非饮料类的产品）大致相同。当然, E 售货机由上表可以推断出容易出售更大比例的饮料类,但是也仅仅是稍大,没有特别大的必要性去考察是否需要对 E 售货机做出不同的投放方案。

4、商品在不同月份出售情况分析

商品（以二级类目为分类标准）在不同月份的出售情况如下表：



观察上图并作出分析，我们非常容易得到如下结论：

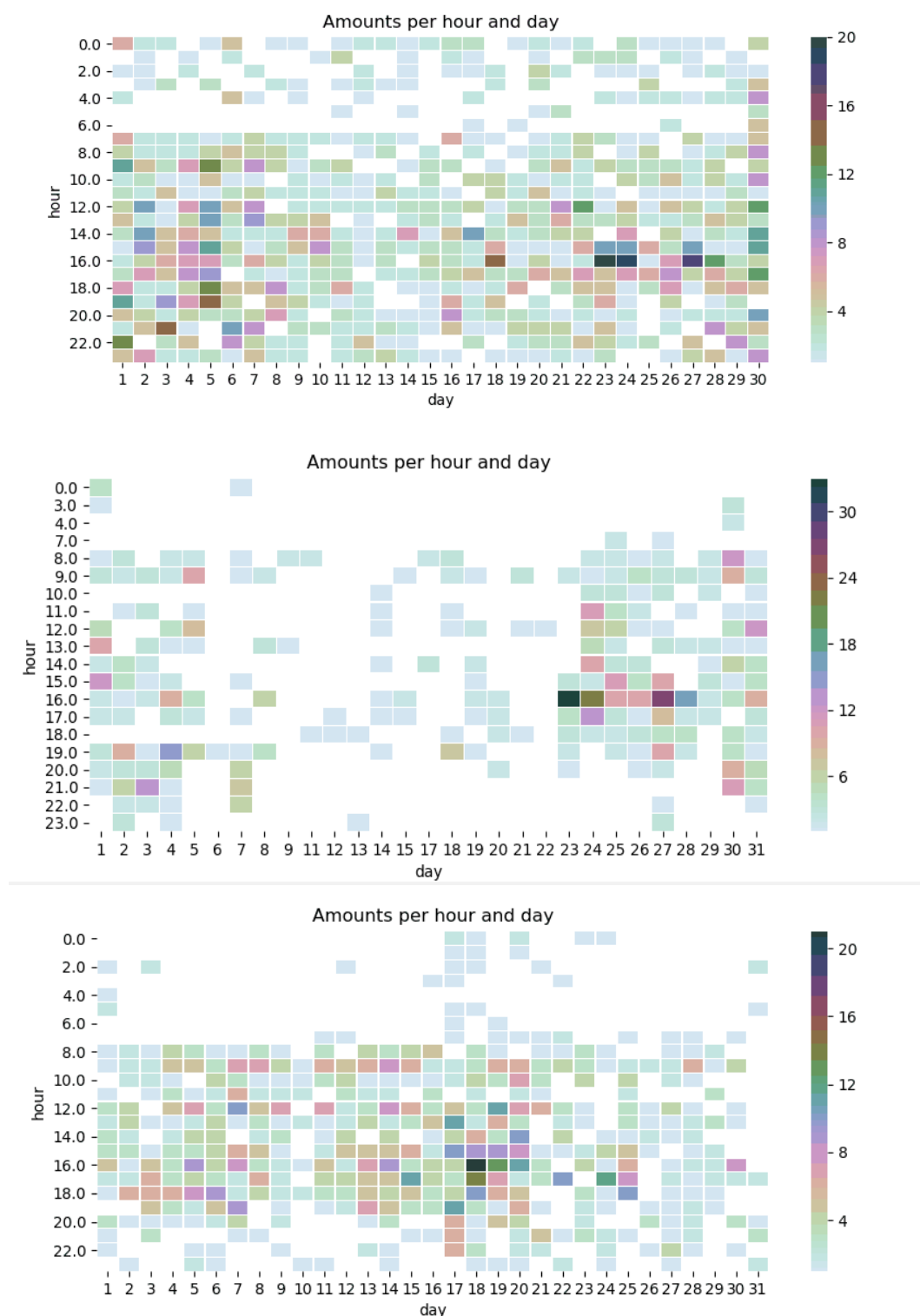
- 1) 几乎在任何月份中，香烟类的产品出售情况都十分优秀，月月占据榜单首位；
- 2) 并无任何商品出现非常大幅度明显的季节性变化或月份性变化，各商品的销售情况都可以说是十分稳定；但是在较为稳定的同时却也出现了极少的个例，比如 2、8 两个月份均出现了坚果炒货和果冻/龟苓膏的断销，2 月出现了纸巾的断销。在营销层面我认为，需要调查回顾我们在 17 年的 2、8 两个月是否在售货机中投放了坚果炒货和果冻/龟苓膏两种产品，如果我们没有投放，那么我建议保持一定量的对上述产品的投放，因为在上图中，我们可以发现，这两种产品的出售量其实是较为稳定的，可以稍作适量投放。如果我们投放了但是并未售出，则我们可以尝试在之后的 2~3 年内的 2、8 月继续坚持投放，因为坚果炒货的保质期一般为 4~12 月，果冻和龟苓膏能达到 6~12 月以上，较少量的投放即便在当年 2、8 月未能出售，也可以在之后的几个月卖出，并不会造成损失，在 2~3 年之后，我们可以根据之后的样本再行考察是否应进行这两种货品的投放。纸巾保质期同样较长，它的投放方式可以参考上述

对坚果炒货和果冻龟苓膏的分析。

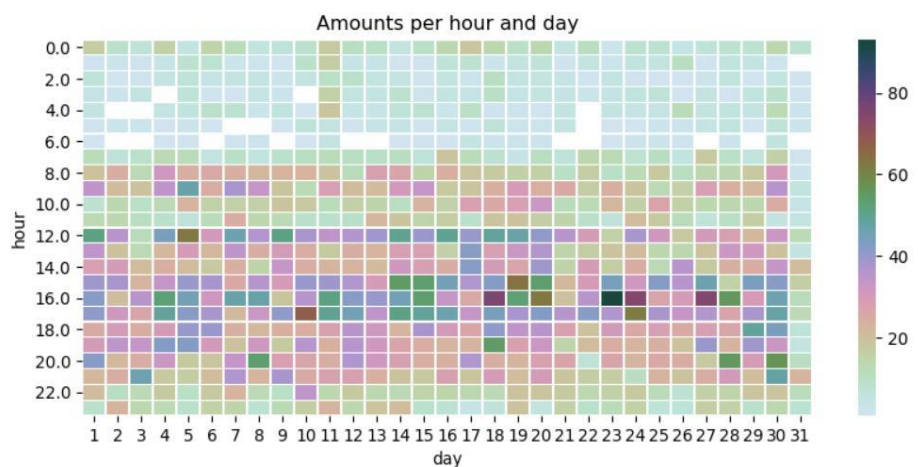
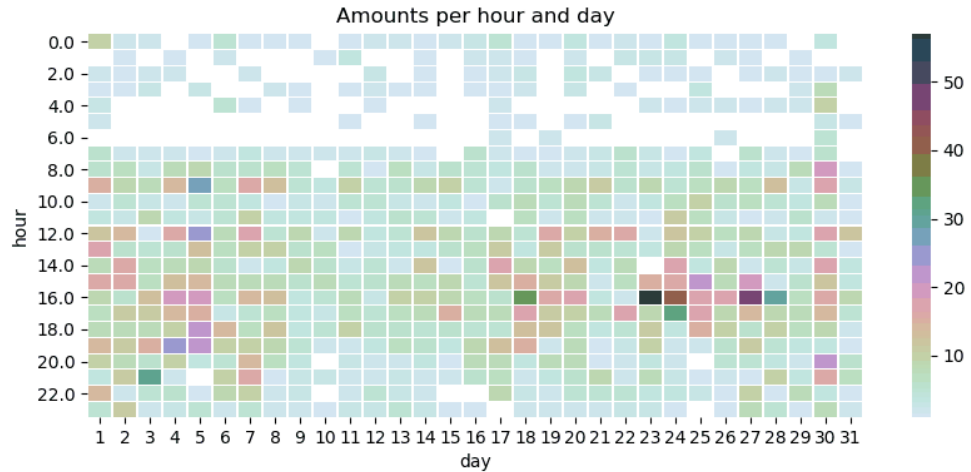
- 3) 我们可以依据上图略微改变商品投放的比重。
- 4) 与之前分析的不同，在夏季，各种饮料的出售量并没有出现想象中的小高潮，我希望在销售端做出一些改变，比如提供一些冷藏的饮料或者冷冻水来促进销售。

5、 6、7、8 月订单随时间分析

任务要求我完成热力图如下，三幅图依次为 6、7、8 三个月：



而为了便于分析和得到一些更好的规律，我还绘制了 6、7、8 三个月的总图和全年的总图如下（依次为 6、7、8 三个月的总图和全年总图）：



下面，我将依照这 5 幅图展开相关分析：

- 1) 由图 1,2,3,4，我们可以发现，在夏季，在凌晨两点之后，早上七点之前，商品出售量极低；早上八点到中午十二点以及晚上七点后到凌晨一点这一段时间中，货物销售稳定但并不是很高，而从中午十二点到晚上七点，商品销售量极高。推广到图 5，我们容易发现，全年的销售与当日时间的关系与夏季情况相同。
- 2) 我们通过联系当日的销售量和星期的关系，发现周末的销售量（星期五、六、日）比其它时候出售更多的货物的概率更大，但是这种情况出现的概率大的程度也并不是十分明显。
- 3) 七月中旬，销售机 C 出现了严重的断销现象，我们应当回顾找到原因，指导下一步的决策。
- 4) 在营销方面，我们应当本着维持 12~19 时较高的商品销售量和利润额，适当提升 7~12 及 19~次日 1 时两个时间段的利润和销量的原则，所以我提

供以下策略：首先，我们可以采取分时间打折的营销策略，比如在早晨或者晚间，提供一些促销活动，吸引顾客；第二，我们应该在早上七点之前补充足够多的货物，在 12 时到 19 时这个时段，尽量保证售货机可以正常使用且有足够多的货品，在 19 时之后我们也应少量补充以弥补上一时间段大量的消耗；第三，在 12 到 19 时的这一时间段，我们可以制造一些供不应求的情况（比如减少投放低价物品，投放一些价格略高的同类商品）来提高利润。

五、销售机画像

1、售货机饮料类产品销售情况

我们通过整理获得销售机饮料类产品的销售情况，我们通过分析，决定将销售量大于等于 100 且排名该售货机前 30 的产品标记为热销，将销售量小于 10 的商品标记为滞销，其它产品标记为正常，可以得到以下结果（仅展示各售货机的热销产品，且五个表格依次为 A, B, C, D, E）：

怡宝纯净水	531	热销	怡宝纯净水	1342	热销	怡宝纯净水	999	热销
东鹏特饮	448	热销	东鹏特饮	488	热销	脉动	759	热销
阿萨姆奶茶	326	热销	阿萨姆奶茶	457	热销	东鹏特饮	525	热销
脉动	309	热销	脉动	431	热销	阿萨姆奶茶	456	热销
营养快线	255	热销	营养快线	425	热销	营养快线	435	热销
统一冰红茶	228	热销	统一冰红茶	334	热销	王老吉（罐）	402	热销
250ml维他柠檬茶	181	热销	王老吉（500ml）	310	热销	统一冰红茶	394	热销
250ml维他原味豆奶	176	热销	果粒橙	245	热销	果粒橙	328	热销
红牛	166	热销	250ml维他柠檬茶	236	热销	芬达330ml	271	热销
统一绿茶	151	热销	250ml维他原味豆奶	229	热销	雪碧最新款	259	热销
果粒橙	145	热销	可口可乐330ml	212	热销	伊利优酸乳	258	热销
伊利优酸乳	143	热销	统一绿茶	209	热销	250ml维他柠檬茶	255	热销
145ml旺仔牛奶盒装	131	热销	王老吉（罐）	201	热销	统一绿茶	246	热销
250ml燕塘红枣牛奶	128	热销	红牛	198	热销	王老吉（500ml）	238	热销
250ml燕塘甜牛奶	126	热销	芬达330ml	198	热销	红牛	231	热销
雪碧（500ml）	123	热销	伊利优酸乳	179	热销	可口可乐330ml	219	热销
250ml燕塘原味酸奶	112	热销	伊利纯牛奶	175	热销	250ml维他原味豆奶	217	热销
王老吉（500ml）	108	热销	145ml旺仔牛奶盒装	172	热销	怡宝	182	热销
银鹭八宝粥	103	热销	雪碧最新款	139	热销	银鹭八宝粥	166	热销
可口可乐（500ml）	101	热销	雪碧（500ml）	137	热销	250ml燕塘原味酸奶	134	热销
			250ml燕塘甜牛奶	131	热销	250ml燕塘甜牛奶	129	热销
			250ml燕塘原味酸奶	130	热销	250ml燕塘红枣牛奶	129	热销
			银鹭八宝粥	120	热销	145ml旺仔牛奶盒装	127	热销
			可口可乐最新款	116	热销	伊利纯牛奶	127	热销
			250ml燕塘红枣牛奶	114	热销	可口可乐最新款	125	热销
			雪碧330ml	103	热销	250ml红牛	119	热销
			百事可乐500ml	102	热销			

东鹏特饮	411	热销	怡宝纯净水	1705	热销
怡宝纯净水	387	热销	脉动	996	热销
阿萨姆奶茶	365	热销	营养快线	805	热销
营养快线	319	热销	阿萨姆奶茶	792	热销
统一冰红茶	309	热销	东鹏特饮	709	热销
脉动	283	热销	统一冰红茶	626	热销
统一绿茶	188	热销	统一绿茶	490	热销
红牛	180	热销	果粒橙	470	热销
雪碧最新款	178	热销	王老吉（500ml）	417	热销
王老吉（罐）	177	热销	王老吉（罐）	410	热销
可口可乐（500ml）	176	热销	250ml维他柠檬茶	397	热销
果粒橙	168	热销	250ml维他原味豆奶	391	热销
可口可乐最新款	144	热销	芬达330ml	382	热销
雪碧（500ml）	132	热销	伊利优酸乳	375	热销
250ml维他柠檬茶	109	热销	雪碧最新款	309	热销
伊利优酸乳	102	热销	红牛	289	热销
			250ml燕塘原味酸奶	249	热销
			可口可乐330ml	244	热销
			伊利纯牛奶	241	热销
			250ml燕塘甜牛奶	232	热销
			145ml旺仔牛奶盒装	228	热销
			银鹭八宝粥	220	热销
			250ml燕塘红枣牛奶	206	热销
			可口可乐（500ml）	202	热销
			268ml雀巢咖啡丝滑拿铁	197	热销
			雪碧330ml	173	热销
			250ml红牛	172	热销
			雪碧（500ml）	172	热销
			维他柠檬茶	169	热销
			500ml统一阿萨姆奶茶	158	热销

2、 售货机画像及营销意见

在上一步，我已经通过任务要求完成了一部分画像，接下来我将在上面的基础上，利用非饮料类商品出售的数据和两类商品出售的总数据添加新的标签，并以此得到销售机的画像。在画像中，我将通过提取并标明售货机热销产品 top3 类目和总种数，滞销产品种数，以及更热销的产品大类（饮料/非饮料）以及总销量来使每个售货机的画像能反映出能指导营销策略的结果。

售货机	热销产品（饮料类）			热销产品（非饮料类）			
A	怡宝纯净水	东鹏特饮	阿萨姆奶茶	30g无穷农场盐_鸡蛋	13g无穷烤鸡小腿蜂蜜	40g双汇玉米热狗肠	
B	怡宝纯净水	东鹏特饮	阿萨姆奶茶	30g无穷农场盐_鸡蛋	鸭翅	香豆干	
C	怡宝纯净水	脉动	东鹏特饮	30g无穷农场盐_鸡蛋	卫龙大面筋	20g咪咪虾条马来西亚风味	
D	东鹏特饮	怡宝纯净水	阿萨姆奶茶	卫龙大面筋	30g无穷农场盐_鸡蛋	香豆干	
E	怡宝纯净水	脉动	营养快线	40g双汇玉米热狗肠	30g无穷农场盐_鸡蛋	鸭翅	
热销产品类数（饮料+非饮料）			滞销产品类数（饮料+非饮料）			热销大类	总销量
32	20+12		114	40+74		饮料	10486
40	27+13		103	44+59		饮料	13482
40	26+14		90	34+56		饮料	14494
25	16+9		112	46+66		饮料	8713
52	29+23		65	24+41		饮料	23505

由以上结果，我们可以进行分析并得到以下结论及营销策略：

- 1) 毫无疑问，从销售结果上可以反映出来，饮料类产品的热销程度要大于非饮料类，而且相比于非饮料类，更不容易滞销，所以我们在之后投放新品时，选择饮料类的风险相比之下要小一些。
- 2) 我们可以发现，怡宝矿泉水，东鹏特饮，阿萨姆奶茶，脉动等饮料十分受欢迎，鸡蛋，玉米热狗肠，卫龙大面筋，鸭翅这些产品是非饮料类的佼佼者，所以，我们一方面是要加大投放量来防止断货，同时，我们也可以增加与其类别相近的新品、竞品（饮用水类，其他品牌的运动饮料，其他品牌鸡蛋、辣条）来满足消费者的更多需求。
- 3) 我们可以发现，处于滞销状态（销量在 10 以下）的产品种类很多，我们可以通过停止相关产品的投放以降低成本，同时若我们已买进大量该类产品，我们可以采取捆绑销售、优惠促销等方式去库存。

六、 预测

1、预测原理

在我看来，预测的原理应该是指，我们可以基于我们之前的售货机营收情况以及售货机做优惠活动等行为的历史资料，并且考虑售货机未来相比于现在进行的发展以及其他外界事物的变化，以此进行分析来得到未来某段时间内销售机的营收情况。

2、可预测性判断及解释

首先，我的观点是：利用我手中现在的已知数据（附件 1，附件 2 以及由这两个表格衍生出来的统计图、表）是很难进行预测的，即便预测也只能进行非常不精准的预测。

以下是我关于不可预测原因的解释：

- 1) 首先，任务要求我对 2018 年 1 月的销售情况进行预测，但是如果要进行该预测的话最简单有效的方法有二：通过统计分析近几年 1 月的销售情况来预测或者利用上一年的数据，将 2018 年 1 月看做 2017 年 12 月进行预测。
- 2) 对于第一种方法，很显然我们缺少 2016 年及以前售货机 1 月的数据。
- 3) 对于第二种方法，我们可以从“交易额变化分析”这一节的统计图中很明显的看出，在过去的一年里，售货机的销售额非常不稳定，变化非常的频繁，并无明显的规律，所以对 2018 年 1 月的预测很难通过规律得出。
- 4) 除了上述第三点原因以外，我认为，即便是过去的一年里，销售机的销售额有较为稳定的变化趋势，也难以进行预测。首先，我们不能把每个月份都平等而孤立的看待，每个月都有其自身的特点，并且在不同地区的售货机销售额会受这个特点的影响，因为我们要预测 2018 年 1 月的销售情况，但是，1 月却具有其他月份不具有的特点，比如，1 月份时学校已经放寒假，可能导致学校附近商场的售货机销售额降低，1 月份临近年关，很多人去选择购置年货，更大量的人流势必会提升在商业街的商场的售货机销售额；1 月学子和少部分上班族返乡，可能带动乡镇中商场售货机的销量，而城市售货机销量下降。从这个角度看，没有哪个月份有和 1 月相同的特点，而 A, B, C, D, E 又未告知明确的地点，用其他月份的数据明显太过死板化，

数据分析得到的预测值可能不会准确。

- 5) 除了上述提到的时间特点(1月份具有的与众不同的性质), 区位影响(售货机放置的地点)、已知数据外, 我们从商业角度上还要考虑很多, 包括但不限于以下几点: 我们之前出现销量波动时售货机出现了什么问题或者我们做了什么活动, 这些问题和活动是否在 2018.1 还会出现/举办; 我们放置售货机的位置是否有可能出现其他公司的售货机, 如果有, 则大概率会影响我们的销量……

综上所述, 我认为我们通过现有数据进行预测很可能只是用统计学方法、利用上面第二个办法进行销量预测, 这很可能达不到目的, 因为商业分析不单单是一个统计学的问题, 我们需要考虑到很多有可能影响到我们销量的外在因素, 这些数据是绝对不够的。

但是, 我认为给我以下数据, 我可以做出一些较为简单的预测来大致估计 2018.1 的销量:

- 1) 额外给我 2015~2016 年的销售数据, 这样我可以由这三年的数据来进行统计分析, 得到一个统计学意义上的结果, 而且我认为这个结果也可以比较准确。
- 2) 告知我这五台售货机具体的地理位置, 并且将 2017 年 10,11,12 月三个月的售货机详细记录(包括但不限于售货机故障, 售货机货品不足, 售货机举行促销活动), 有这些数据, 我们可以简单预测与这三个月天气情况相近, 市场变化不大的 1 月销售额, 并且可以通过具体地理位置, 排除一些上述提到的过年、放假等情况的影响, 我认为这样的预测可能更准确。

除此之外, 如果能给我更多的数据, 我可以做出一些更好的预测, 这些数据可以有:

- 1) 我们的销售机之前和现在竞争对手的情况。
- 2) 更多台销售机数据。
- 3) 销售机的活动记录和 1 月预计计划开展的活动内容(如促销、加入新品)。
- 4) 售货机放置的地点, 此地点是否发生变化(比如放置售货机的商场人流量减少)。

- 5) 人们对于售货机的态度改变（如果有条件的话，可以参考诸如“自动售货机如何使用”的搜索量或其他相关评论）。

以上将会非常有助于抛开单纯的数据，从更多方面进行对于销售的预测，并且相比于由数据而得到的单纯的预测，我们更有利于把握影响销售机销量的根本原因，从而及时改变我们的营销策略来提高销售额。

七、 小结

本节中，我将对上述完成的分析结论单独提取出来，便于结论的快速查找：

- 1) 关于售货机地点的分析。我们要适当的在 E 处增加售货机、提升投货量，减少在 D 处的投货量，对于 A, B, C 则暂时不进行处理；
- 2) 关于售货机时间的分析。2、7 月这两个月的销售额偏低，我们要回顾一下这两个月我们是否采取了不当的行动还是出现了其它意料以外的情况，如果没有，则我们可以适当减少投放一些保质期短的商品。E 售货机在 10、12 月也出现了上述问题，我们可以采取相同的措施进行处理；
- 3) 关于饮品。现在卖得好的产品我们可以加大投货量；我们可以投放一些具有较高利润的饮用水；防止顾客“喝腻了”，提供一些与东鹏特饮、脉动等销售火热的饮料相近的商品（运动饮料、维生素水），可能会取得一些不错的效果；做出一些改变使饮品的出售在夏季更加吸引人可能是我们提高夏季营业额的不错手段；
- 4) 关于投货的商品调整。可以提升香烟类产品的种类数，适当增加投货量，我们可以依据“四、4”中的图片简单调整投货比例，同时，我们可以简单考察坚果炒货、果冻、龟苓膏、纸巾等在特定月份的投放与销售情况，但这个操作并不应有较高优先级，因为它并不会短时间内提高太多利润；
- 5) 关于货物补充。我们要注意销售机投货的时间，我们要保证 12~19 时“够卖”，7~12 及 19~24 时“有得卖”；
- 6) 关于新品。投放新品时，选择饮料类产品的滞销风险会小一些，新品可以选择和热销产品相近的，诸如奶茶类饮料、香肠类、辣条类可能会收到不错的效果。
- 7) 关于已经产生的滞销品。滞销品种类很多，可以采取买赠、打折、甚至是“两元抽奖随机得商品”的方式进行促销去库存。不是很值得冒风险对滞销品做宣传。
- 8) 关于预测。依照现有的数据很难预测，我们想要预测的话可能需要再搜集一些往年数据，拿到售货机具体地点，最好还可以做一下市场

调研，了解一下竞争对手的售货机。

附、项目的完整实现过程

- 1) 我拿到项目数据（附件 1、2）之后，首先完成了一个表格的拆分，因为数据量较小，我采取的是一个比较简单的筛选——提取方法；
- 2) 拆分完表格之后，我发现时间数据的格式十分整齐，为了后续好处理，我运用 `python` 代码，将时间数据拆分，将货物售出时间的月、日、小时、分分别提取了出来；
- 3) 运用已完成的表格和 `python` 代码，比较轻松的获得 5 月份和全年售货机的交易情况；
- 4) 接下来我用已完成的表格和 `python` 代码，处理数据，并依照数据画出任务 2 需要的全部统计图，依照统计图完成数据分析的任务；
- 5) 生成销售机画像的步骤中，我首先利用代码完成对各售货机各饮料类商品的销量统计，之后用 `Excel` 表格进行标签；为了生成更有意义的画像，我又统计了非饮料类、总榜两项并标签，并依照该结果生成了一个包含部分销售机属性的画像；
- 6) 后续的预测过程，我进行了一些简单的回归预测和趋势外推预测，发现预测的结果在未考量现实意义时并不好，我分析了该预测偏差的原因并阐述了更多关于进行预测需要的要求；
- 7) 完成了全部任务后，我进行了小结，将之前分析的比较零散的结果放在一起，便于读者快速得到报告的结论。