

目录

新零售-无人智能售货机商务数据分析报告.....	1
一、 背景介绍.....	1
二、 数据预处理与分析.....	2
1、 数据的分类.....	2
2、 各售货机 2017 年 5 月的销售情况探究.....	2
3、 2017 年售货机总体销售情况探究.....	3
4、 各售货机每单平均交易额与日均订单量探究.....	3
三、 数据分析与可视化.....	5
1、 2017 年 6 月销量前五的商品销量柱状图.....	5
2、 每台售货机每月总交易额折线图以及交易额月环比增长率柱状图.....	6
3、 每台售货机毛利润占比饼图.....	11
4、 每月交易额均值气泡图.....	12
四、 自动售货机画像绘制.....	13
1、 商品“滞销”、“正常”、“热销”标签张贴.....	13
2、 标签拓展与自动售货机画像的生成.....	14
五、 业务预测.....	16
1、 预测未来销售额的原理.....	16
2、 预测问题.....	16

新零售-无人智能售货机商务数据分析报告

一、背景介绍

伴随收入水平的提高，人们追求灵活、方便、快捷以及自助式服务，这使得自动售货机得以迅猛发展。自动售货机具有技术含量高、销售方式新、市场潜力大、商品货真价实等优点，决定了它广阔的发展前景。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意义重大。

某商场在不同的位置安放了 5 台自动售货机，编号分别为 A、B、C、D、E，先对 2017 年每台售货机的商品销售数据以及商品分类信息进行处理分析，从而为消费者提高用户体验，经营者提供科学有效的建议。

二、数据预处理与分析

1、数据的分类

将 2017 年整体数据按照售货机进行分类，也就是提取各个售货机 2017 年数据，其部分数据如下所示：

	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
1	DD2017081(E43A6E078#		3	3	40g双汇玉米	2017/1/1 1:33 A		已出货未退	已提现
16	DD2017081(E43A6E078#		5.8	5.8	100g卫龙点	2017/1/1 19:29 A		已出货未退	已提现
17	DD2017081(E43A6E078#		0.8	0.8	咪咪虾条马	2017/1/1 19:33 A		已出货未退	已提现
20	DD2017081(E43A6E078#		3	3	250ml维他	2017/1/1 19:55 A		已出货未退	已提现
22	DD2017081(E43A6E078#		3.5	3.5	东鹏特饮	2017/1/1 20:18 A		已出货未退	已提现
24	DD2017081(E43A6E078#		3	3	250ml维他	2017/1/1 20:25 A		已出货未退	已提现
25	DD2017081(E43A6E078#		3	3	卫龙大面筋	2017/1/1 20:25 A		已出货未退	已提现
31	DD2017081(E43A6E078#		3.5	3.5	东鹏特饮	2017/1/2 8:50 A		已出货未退	已提现
34	DD2017081(E43A6E078#		1.5	1.5	卫龙亲嘴烧	2017/1/2 12:02 A		已出货未退	已提现
37	DD2017052(E43A6E078#		4.5	4.5	阿萨姆奶茶	2017/1/2 13:34 A		已出货未退	已提现

2、各售货机 2017 年 5 月的销售情况探究

首先提取 2017 年 5 月各售货机的销售数据，其部分数据如下：

支付时间	Unnamed: (订单号	设备ID	应付金额	实际金额	商品	地点	状态	提现
2017/5/1 2:11	7053	DD2017081(E43A6E078/	0.8	0.8	咪咪虾条	马A	已出货未退	已提现
2017/5/1 2:12	7054	DD2017081(E43A6E078/	3	3	250ml维他	A	已出货未退	已提现
2017/5/1 2:30	7057	DD2017071(E43A6E078/	3	3	可口可乐	(A	已出货未退	已提现
2017/5/1 7:03	7078	DD2017081(E43A6E078/	4	4	阿萨姆奶茶	A	已出货未退	已提现
2017/5/1 7:04	7080	DD2017081(E43A6E078/	2.5	2.5	伊利纯牛奶	A	已出货未退	已提现
2017/5/1 7:05	7082	DD2017081(E43A6E078/	4.5	4.5	40g呀土豆	A	已出货未退	已提现
2017/5/1 7:06	7083	DD2017081(E43A6E078/	2.5	2.5	250ml王老	A	已出货未退	已提现
2017/5/1 8:02	7087	DD2017081(E43A6E078/	3	3	维他奶	A	已出货未退	已提现
2017/5/1 11:40	7092	DD2017081(E43A6E078/	4.5	4.5	55g奥利奥	A	已出货未退	已提现

然后计算 5 月各售货机的销售总额与订单量，统计汇总后的结果如下表所示：

	A	B	C	D	E
订单量（个）	756	869	789	564	1292
销售额（元）	3385.1	3681.2	3729.4	2392.1	5699

数据显示，2017 年 5 月，售货机的订单量从多到少的排名为：E、B、C、A、D；而售货机销售额从多到少排名为：E、C、B、A、D。

3、2017 年售货机总体销售情况探究

所有售货机	
总销售额（元）	286979.7
总订单量（个）	70679

数据显示，2017 年所有售货机总订单量为 70679 个，总销售额为 286979.7 元。

4、各售货机每单平均交易额与日均订单量探究

将已经提取好的各个售货机销售情况表分别按月分类，再对各个售货机各个月份中的“实际金额”求和，对“订单号”计数，可得到各个售货机每个月的销售总额与订单数，再进一步得到各个售货机每个月的每单平均交易而，汇总情况如下：

A 售货机		
月份	每单平均交易额(元)	订单数
1 月	4.507	335
2 月	3.864	114
3 月	3.585	255
4 月	4.037	447
5 月	4.478	756
6 月	4.074	1669

7 月	4.098	476
8 月	3.359	666
9 月	4.307	1040
10 月	4.021	1565
11 月	4.472	1160
12 月	3.788	2003

B 售货机		
月份	每单平均交易额(元)	订单数
1 月	3.753	366
2 月	3.256	185
3 月	3.615	265
4 月	4.075	603
5 月	4.236	869
6 月	4.068	1856
7 月	4.402	345
8 月	3.584	981
9 月	4.13	1745
10 月	4.112	2026
11 月	4.269	2031
12 月	3.667	2210
D 售货机		

C 售货机		
月份	每单平均交易额(元)	订单数
1 月	4.328	379
2 月	3.826	207
3 月	3.77	263
4 月	4.404	734
5 月	4.727	789
6 月	4.502	1882
7 月	3.988	764
8 月	3.914	1259
9 月	4.427	1678
10 月	4.273	2216
11 月	4.352	1943
12 月	3.943	2379
月份	每单平均交易额(元)	订单数
1 月	3.693	259
2 月	3.089	141

3 月	4.306	192
4 月	3.79	443
5 月	4.241	564
6 月	4.026	1040
7 月	4.23	317
8 月	3.317	715
9 月	3.899	983
10 月	3.884	1186
11 月	3.862	1210
12 月	3.573	1663

E 售货机		
月份	每单平均交易额(元)	订单数
1 月	4.68	354
2 月	3.638	258
3 月	4.306	350
4 月	4.16	895
5 月	4.411	1292
6 月	3.818	2593
7 月	3.919	813
8 月	3.804	1767
9 月	4.125	4134
10 月	3.676	2777
11 月	4.283	5020
12 月	4.169	3252

三、数据分析与可视化

1、2017 年 6 月销量前五的商品销量柱状图

首先，将 2017 年 6 月总体销售情况数据提取出来。然后根据商品进行分类，并对商品销量进行统计，得到结果。数据表明，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、脉动与 250ml 维他柠檬茶，其销售量如下表所示：

商品	销量
怡宝纯净水	657
40g 双汇玉米热狗肠	240
东鹏特饮	238
脉动	235
250ml 维他柠檬茶	225

用柱状图将其表示出来：



2、每台售货机每月总交易额折线图以及交易额月环比增长率柱状图

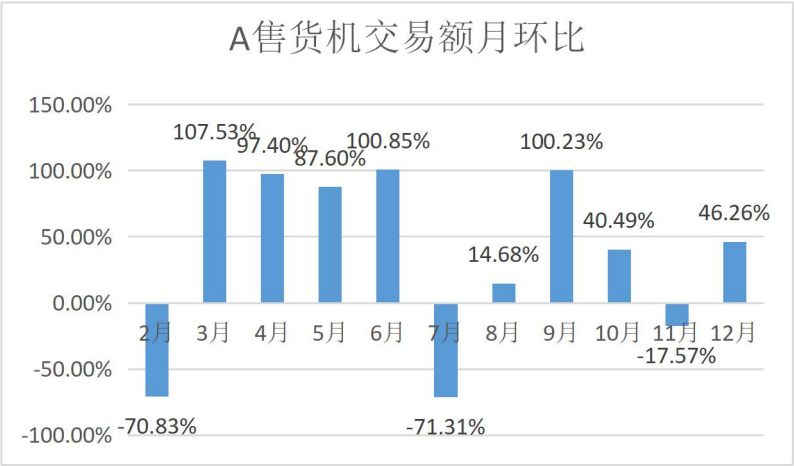
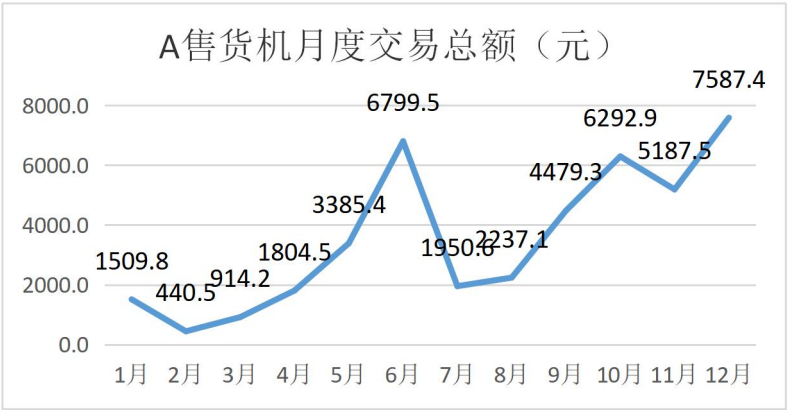
首先，将已经提取好的的各售货机销售数据根据月份分类，再对其交易额进行求和，得到每台售货机每月总交易额。再对数据进一步分析，求出交易额月环比。以下用这折线图的形式展示出每台售货机每月总交易额情况，用柱状图的形式展示出每台售货机每月交易额月环比（由于缺乏 2016 年各个售货机的销售数据，因此月环比数据 2017 年 2 月开始，至 2017 年 12 月）。

各售货机月度销售数据					
	A	B	C	D	E
1 月	1509.8	1373.6	1640.3	956.5	1656.7
2 月	440.5	602.4	792.0	435.5	938.6
3 月	914.2	958.0	991.5	826.8	1507.1
4 月	1804.5	2457.2	3232.5	1679.0	3723.2
5 月	3385.4	3681.1	3729.6	2391.9	5699.0
6 月	6799.5	7550.2	8472.8	4187.0	9900.1
7 月	1950.6	1518.7	3046.8	1340.9	3186.1
8 月	2237.1	3515.9	4927.7	2371.7	6721.7
9 月	4479.3	7206.9	7428.5	3832.7	17052.8
10 月	6292.9	8330.9	9469.0	4606.4	10208.3
11 月	5187.5	8670.3	8455.9	4673.0	21500.7
12 月	7587.4	8104.1	9380.4	5941.9	13557.6

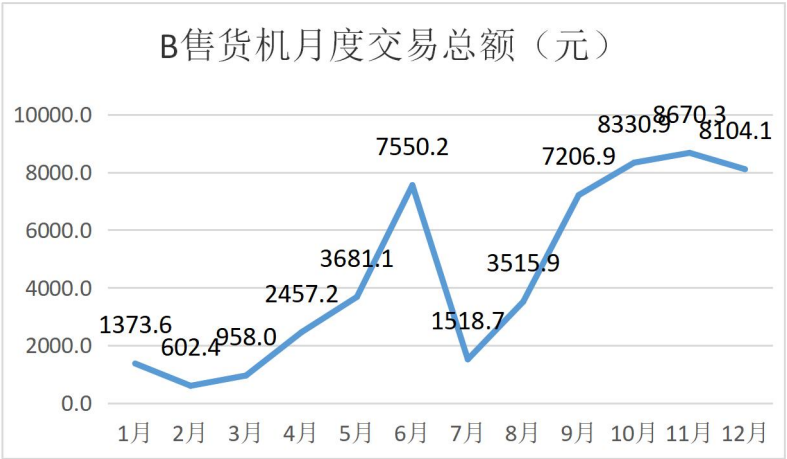
各售货机交易额月环比					
	A	B	C	D	E
1 月	/	/	/	/	/
2 月	-70.83%	-56.15%	-51.72%	-54.46%	-43.35%
3 月	107.53%	59.04%	25.19%	89.82%	60.57%
4 月	97.40%	156.50%	226.02%	103.08%	147.04%
5 月	87.60%	49.81%	15.38%	42.46%	53.07%
6 月	100.85%	105.11%	127.18%	75.05%	73.72%
7 月	-71.31%	-79.89%	-64.04%	-67.97%	-67.82%
8 月	14.68%	131.51%	61.73%	76.87%	110.97%
9 月	100.23%	104.98%	50.75%	61.61%	153.70%
10 月	40.49%	15.60%	27.47%	20.19%	-40.14%
11 月	-17.57%	4.07%	-10.70%	1.45%	110.62%
12 月	46.26%	-6.53%	10.93%	27.15%	-36.94%

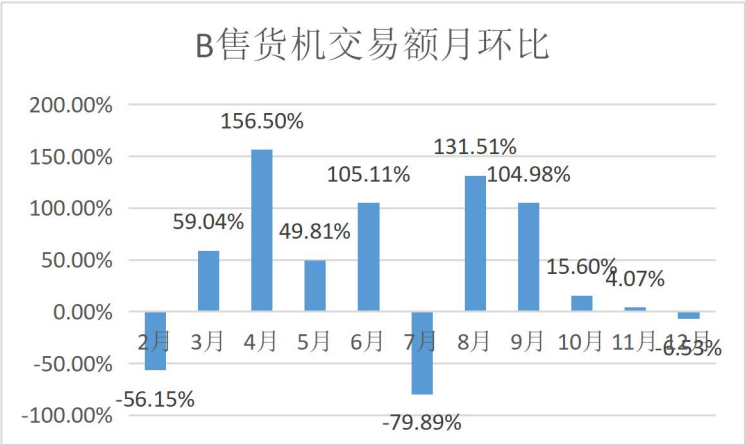
每台售货机每月交易额折线图及月交易额环比增长率柱状图如下图所示：

A:

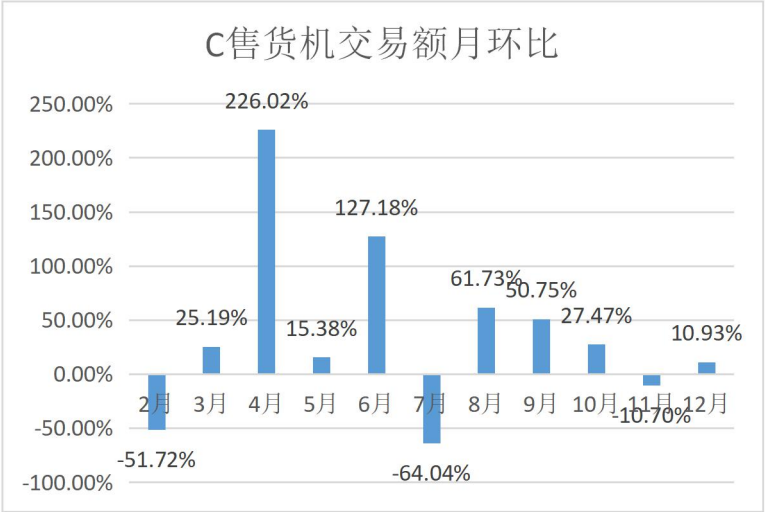
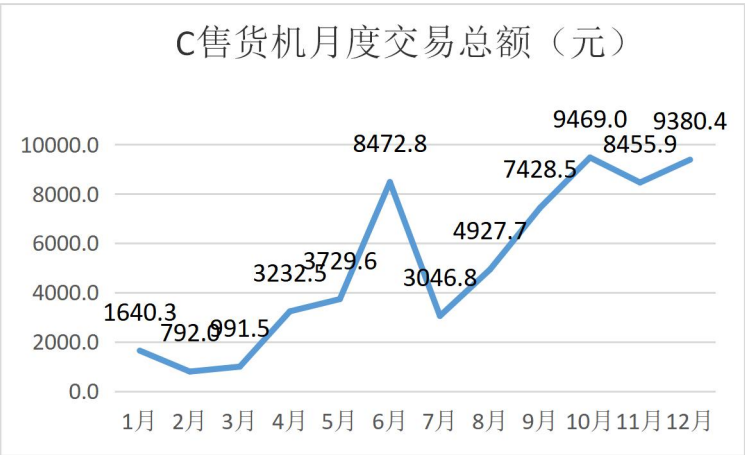


B:

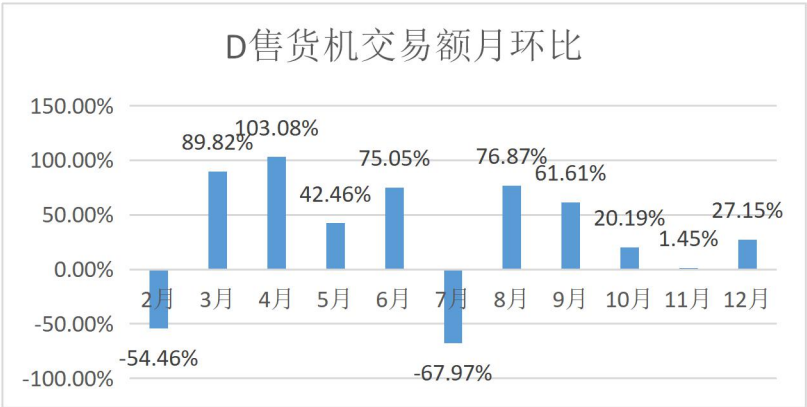
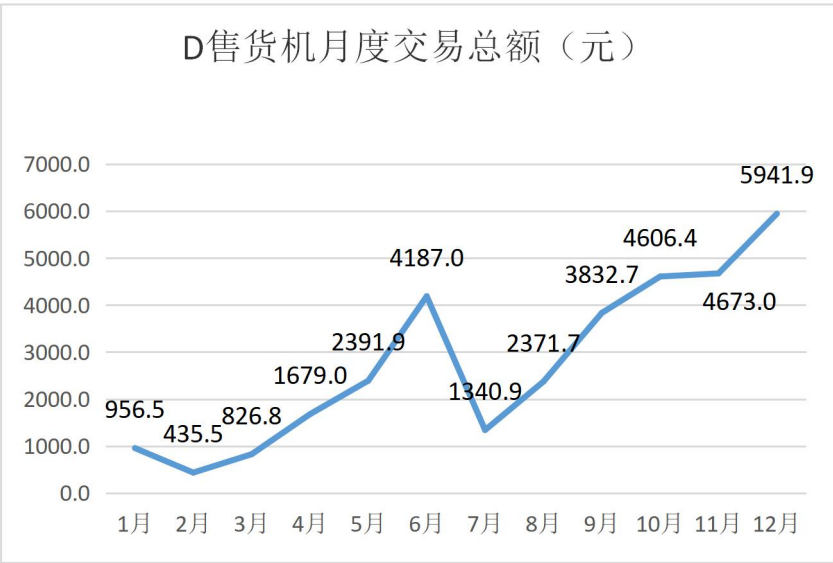




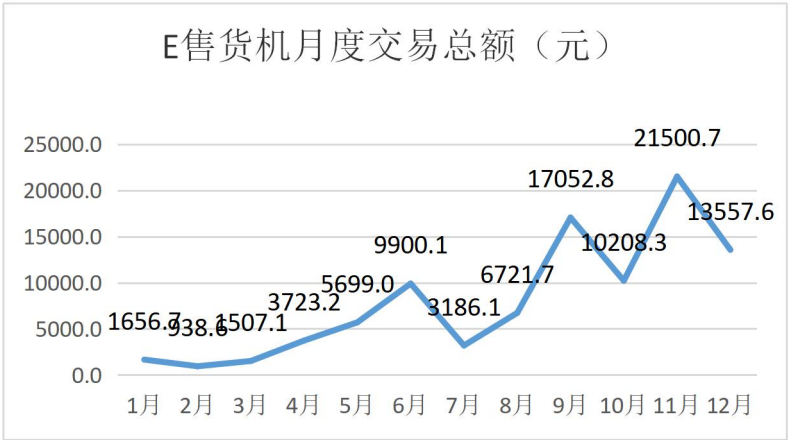
C:

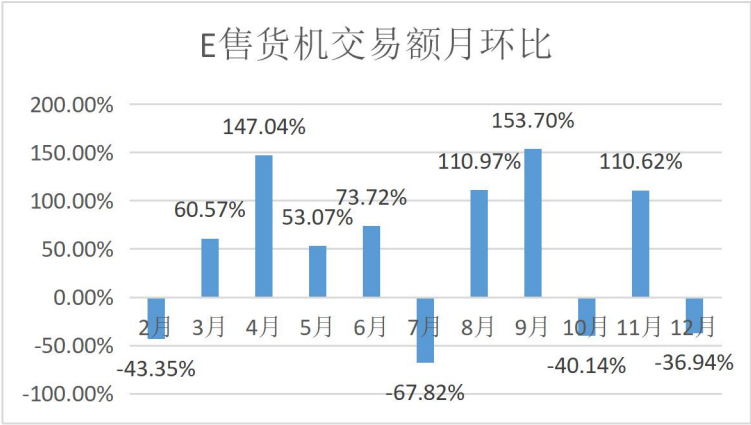


D:



E:





3、每台售货机毛利润占比饼图

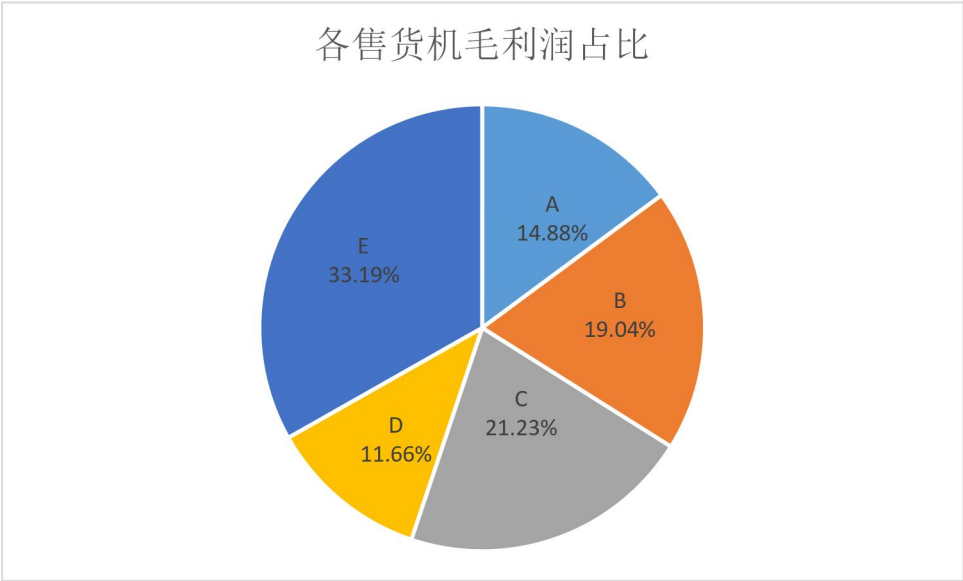
已知毛利润情况是：饮料类 25%，非饮料类 20%。首先，在已经提取好的各个售货机销售情况表中，根据商品分类信息表为商品与其是否属于饮料类的信息进行匹配，部分数据如下：

	订单号	设备ID	应付金额	实际金额	商品	大类	支付时间	地点	状态	提现
1	DD2017081	E43A6E078	3	3	40g双汇玉米肠	非饮料	2017/1/1 1:33 A		已出货未退	已提现
16	DD2017081	E43A6E078	5.8	5.8	100g卫龙点	饮料	2017/1/1 19:29 A		已出货未退	已提现
17	DD2017081	E43A6E078	0.8	0.8	咪咪虾条	非饮料	2017/1/1 19:33 A		已出货未退	已提现
20	DD2017081	E43A6E078	3	3	250ml维他	饮料	2017/1/1 19:55 A		已出货未退	已提现
22	DD2017081	E43A6E078	3.5	3.5	东鹏特饮	饮料	2017/1/1 20:18 A		已出货未退	已提现
24	DD2017081	E43A6E078	3	3	250ml维他	饮料	2017/1/1 20:25 A		已出货未退	已提现
25	DD2017081	E43A6E078	3	3	卫龙大面筋	饮料	2017/1/1 20:25 A		已出货未退	已提现
31	DD2017081	E43A6E078	3.5	3.5	东鹏特饮	饮料	2017/1/2 8:50 A		已出货未退	已提现
34	DD2017081	E43A6E078	1.5	1.5	卫龙亲嘴烧	饮料	2017/1/2 12:02 A		已出货未退	已提现

再将各个售货机的饮料类与非饮料类交易总额进行汇总，计算出各个售货机的毛利润，再求各个售货机毛利润占比，汇总情况如下：

	A	B	C	D	E
非饮料	12344.3	13537.7	22017.2	9245.4	31452.4
饮料	30198.3	40429.6	39554.9	23991.9	64198
毛利润（元）	10018.435	12814.94	14292.165	7847.055	22339.98
所有售货机总利润（元）	67312.575				
毛利润占比（%）	14.88%	19.04%	21.23%	11.66%	33.19%

用饼图的形式将每台售货机毛利润占总毛利润的比例直观地体现出来：

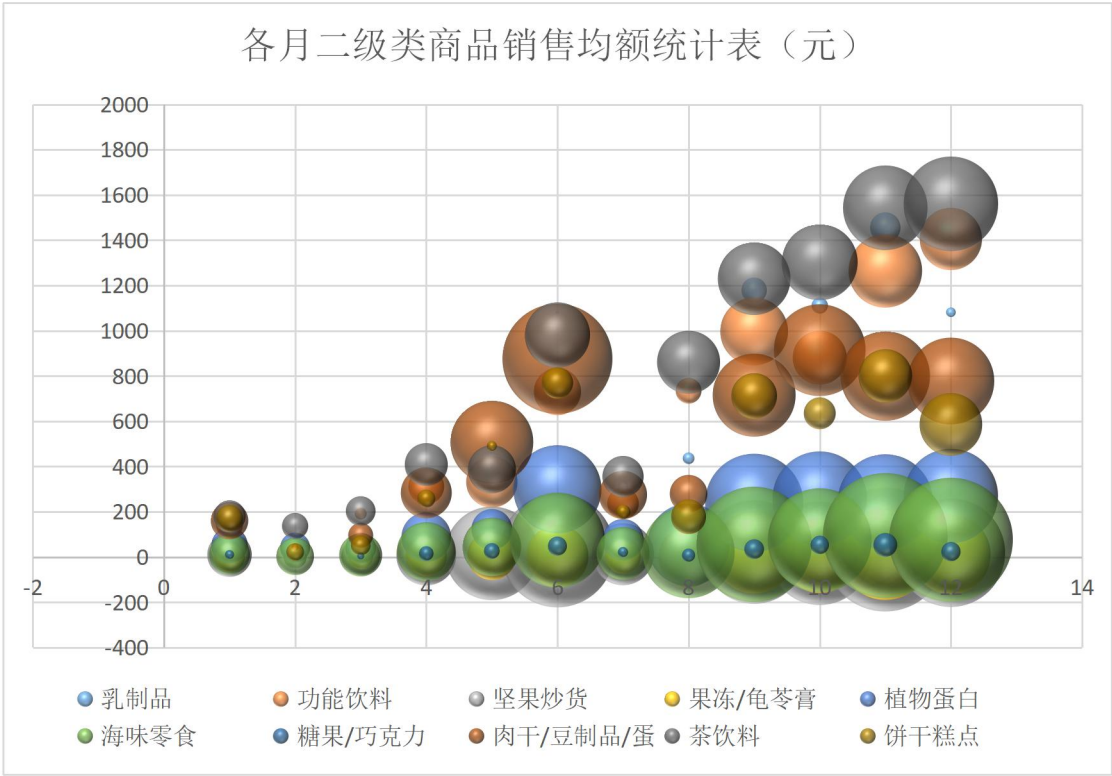


饼图显示，售货机 E 对总毛利润贡献最大，占到 33.19%，售货机 D 对总利润贡献最先，占比 11.66%。

4、每月交易额均值气泡图

将附件一中的数据按照月份分隔出来，并对商品进行二级类名称的匹配，再对匹配后的销售数据根据“二级类”对“实际金额”进行求和，得到每月二级类商品销售额均值数据，汇总如下：

二级类	各月二级类商品销售均值统计表（元）											
	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
乳制品	229.6	68.6	118.96	371.46	530.32	1081.1	314.88	435.88	1178.52	1110.86	1455.74	1081.34
其他	0	0	0.02	0.06	0	0.68	5.2	6.4	30.08	12.42	43.14	4.34
功能饮料	143.7	0	191	311.66	331.66	730.74	236.2	734.42	996.76	880.78	1264.5	1405.16
咖啡	33.9	4.7	6.7	59.3	122.1	103	46	30.8	212.8	135	250.6	179.6
坚果炒货	10	0	4.2	6.16	15.96	29.36	12.3	0	21.7	15.92	32.9	18.72
方便速食	91.4	24.58	69.12	164.9	404.5	601.9	178.14	154.64	496.42	494.1	709.46	544.1
果冻/龟苓膏	1	36.1	1	5.4	6.4	3.4	4.4	0	21.6	6	13.2	11
果蔬饮料	43.4	12.72	28	78.6	110.7	178.2	59.3	169.8	316.6	258.5	390.7	292.5
植物蛋白	52.64	41.64	39.4	87.72	121.72	299.4	76.28	54.1	239.34	257.64	243.2	269.72
水	61.04	41.64	43.34	114.06	77.18	359.02	79.3	314.28	445.88	413.6	421.4	411.96
海味零食	12.18	2.9	8.5	23.88	44.5	74.36	15.56	21.68	55.6	73.94	95.06	76.06
碳酸饮料	76.74	65.7	85.16	160.32	158.68	417.68	129.7	393.24	615.74	499.18	721.18	710.92
糖果/巧克力	10.66	0.8	4.72	15.58	26.4	48.14	21.76	7.92	34.46	53.24	53.34	23.68
纸巾	3.84	0	2.32	9.8	11.52	16.92	4.8	8.08	18.08	15.84	26.32	17.32
肉干/豆制品/蛋	160.5	43.8	95.54	283.6	507.5	875.74	274.58	279.5	714.96	913.84	799.58	777.52
膨化食品	65.04	2.36	29	120.32	318.22	559.94	107.1	65.5	318.86	389.76	370.16	348.5
茶饮料	179.6	136.9	202.7	408.1	387.6	981	355.3	860.6	1230	1302.5	1543.2	1560.6
蜜饯/果干	47.96	32	41	85.86	107.2	197.58	78.88	184.06	244.64	266.68	330.88	413.34
饼干糕点	171.2	21.7	50.8	257.5	490.6	769.1	199	177.7	711.6	634.8	800.2	585.7
香烟	33	14.6	18	15	4.6	45.6	10	56.2	97	47.2	133	182



四、自动售货机画像绘制

标签是对某一类对象的某一类特征的概括性描述，而画像则是对某一类对象的多项特征的描述。本报告对各个商品的销售量用“贴标签”的方法进行描述，为经营者掌握商品销售情况提供支持。

1、商品“滞销”、“正常”、“热销”标签张贴

首先，对标签进行定义：

销售量	标签
0-50	滞销
51-200	正常
>200	热销

然后，将每台售货机销售数据按照商品类别分类，并统计各个商品在每台售货机的年销售总量，并根据标签的定义为每台售货机的每件商品贴上标签，部分情况如下：

A售货机商品商品	销量	标签
100g*5瓶益力多	64	正常
100g卫龙点心面黑椒牛排味	12	滞销
100g果王咸柑桔罐装	20	滞销
103g康师傅红烧牛肉面	19	滞销
107g出前一丁桶面酱香牛肉王	13	滞销
10g卫龙亲嘴烧香辣味	72	正常
10g越南LIP0奶味面包干	3	滞销
10g越南LIP0奶味面包干0015	3	滞销
10g越南LIP0奶味面包干0046	1	滞销
110g顺宝九制话梅	59	正常
117gUF0炒面铁板色拉鱿鱼风味	26	滞销

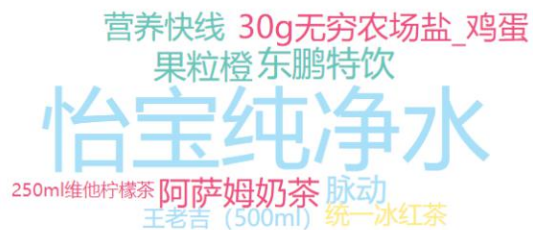
2、标签拓展与自动售货机画像的生成

标签拓展：销售量前十商品可以在一定程度上反映售货机的热销情况，因此本文
中将售货机销售前十商品作为售货机的标签，并利用词云图作为售货机的画像，
汇总情况如下：

A 售货机商品商品	销量	标签
怡宝纯净水	1705	热销
营养快线	805	热销
阿萨姆奶茶	792	热销
东鹏特饮	709	热销
统一冰红茶	626	热销
统一绿茶	490	热销
果粒橙	470	热销
王老吉（500ml）	417	热销
王老吉（罐）	410	热销
250ml 维他柠檬茶	397	热销



B 销售量前十商品	销售额	标签
怡宝纯净水	1342	热销
东鹏特饮	488	热销
阿萨姆奶茶	457	热销
脉动	431	热销
营养快线	425	热销
统一冰红茶	334	热销
王老吉（500ml）	310	热销
果粒橙	245	热销
30g 无穷农场盐_鸡蛋	243	热销
250ml 维他柠檬茶	236	热销



C 销售量前十商品	销量	标签
怡宝纯净水	999	热销
脉动	759	热销
东鹏特饮	525	热销
阿萨姆奶茶	456	热销
营养快线	435	热销
王老吉（罐）	402	热销
统一冰红茶	394	热销
果粒橙	328	热销
芬达 330ml	271	热销
雪碧最新款	259	热销

统一冰红茶 250ml 维他柠檬茶
营养快线 芬达 330ml
东鹏特饮
怡宝纯净水
果粒橙 阿萨姆奶茶
脉动 王老吉（罐）

D 销售量前十商品	销量	标签
东鹏特饮	411	热销
怡宝纯净水	387	热销
阿萨姆奶茶	365	热销
营养快线	319	热销
统一冰红茶	309	热销
脉动	283	热销
统一绿茶	188	正常
红牛	180	正常
雪碧最新款	178	正常
王老吉（罐）	177	正常

营养快线 脉动 统一绿茶
红牛 阿萨姆奶茶
怡宝纯净水
东鹏特饮
王老吉（罐） 统一冰红茶

E 销售量前十商品	销量	标签
怡宝纯净水	1705	热销
脉动	996	热销
营养快线	805	热销
阿萨姆奶茶	792	热销
东鹏特饮	709	热销
统一冰红茶	626	热销
统一绿茶	490	热销
果粒橙	470	热销
王老吉（500ml）	417	热销
王老吉（罐）	410	热销

东鹏特饮 统一冰红茶
营养快线 脉动
怡宝纯净水
阿萨姆奶茶 王老吉（罐）
果粒橙 王老吉（500ml） 统一绿茶

总体来说：怡宝纯净水、脉动、东鹏特饮、阿萨姆奶茶、营养快线、王老吉（罐）、统一冰红茶、果粒橙是每个售货机都具有的“爆款”，而值得关注的是，A 售货

机的两款王老吉（罐装与 500ml）都在前十中；B 售货机销售量前十中出现了 30g 无穷农场盐_鸡蛋，这也是所有售货机销售量前十中的非饮料类产品；C 售货机前十商品中出现了芬达 330ml 商品；D 的东鹏特饮销售量超过怡宝饮用水，这也是唯一一台售货机有商品销量超过怡宝，同时，在 D 商品销量前十中，出现了雪碧最新款与红牛；E 中出现了统一绿茶，并且两款王老吉（罐装与 500ml）都在前十中。因此，经营者应对售货机共同的“爆款”加以关注，适当提高存货量或者替代品以满足客户需求。同时，不同售货机中出现的“特异”标签也反映了不同地点，由于经过的客户年龄段差异等的因素带来的客户偏好的不同，经营者应该适当提供这样的商品或者替代品以满足客户需求。

五、业务预测

1、预测未来销售额的原理

①惯性原理

事物的发展都有时间上的延续性，即惯性。事物某个属性或指标的发展趋势的特征(如发展方向、发展速度、变化周期等) 在一段时间呈现延续性，而两个或多个指标间的相互关系也会按照一定的格局延续下去，这就为预测提供了基础。

②相关性原理

各种事物之间都存在着直接或间接的联系， 都与其他事物的发展存在或大或小的相互影响、相互制约、相互促进的关系，这些关系或联系就是科学研究中的相关性。许多研究指标与一些其它事物之间存在着非常紧密度的联系，这也为预测提供了依据。

2、预测问题

我认为，仅根据附件 1 提供的数据不能完成 2018 年 1 月售货机交易额预测，理由如下：

①惯性原理

上述预测原理第一条是惯性原理，也就是销售额在一段时间内会按找数据呈现的趋势延续下去。而通过对各个售货机一年销售的数据初步分析，售货机在夏季与冬季会有明显的增长，在秋季也有明显的下滑。而在一整年中，售货机整体大趋势是销售量不断增加，如果仅仅根据这一年的值，预测 2018 年 1 月的销售额会呈现较大的数值，这与实际可能有较大的偏差，因此我认为数据的覆盖时间段应该拉长。

②相关性原理

上述预测原理第二条是相关性原理，也就是被预测变量与预测变量应该有一定的相关性。通过分析附件一中的数据，我们认为与销售量有一定相关性的指标为支付时间与实际金额，但这样的预测变量是远远不够的。我们应该有其他影响销量的信息，例如天气、节假日、商城活动、客流量等等（非数值的信息将其量化），我认为应该将数据的维度扩大，来实现较为可行的预测。

因此，我认为为了实现 2018 年 1 月销售机交易额预测，我们应该掌握更长时间段、更大数据维度的信息。经营者需要提供给我们近三年的销售数据，以及与销售数据密切相关的信息，例如人流量、天气等等。