

一、背景介绍

1、行业背景

即时便利性消费的需求在增加。随着收入水平的提升，人们对即时便利性消费的需求在增加，自动售货机的消费环境日益成熟。

房租上涨压缩便利店利润。随着房价的上涨，国内整体租金不断增长，便利店的利润被急剧压缩，而自助售卖机占地体积小，租金成本低，这为自动售货机市场发展创造了条件。

人口红利消失，用工成本上涨。劳动力人口呈现减少的趋势，老龄化问题日益严重，使得国内用工成本不断增长，而自动售货机很少需要人维护，用工成本相对较少。

技术驱动良好交互体验。自动售货机应用了移动支付的技术，支付宝和微信可以瞬间实现快捷支付，消费者短时间内就可以完成消费，非常方便。此外智能化的云管理系统为运营商提供了科学、智能化的管理工具，可以及时掌握售货机的商品情况，并且能够对消费时段进行分析。

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，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。

二、数据说明

1、数据来源

数据来源：泰迪云课堂平台

数据截止时间：2017 年 12 月 31 日

2、数据量

销售信息：70680 条记录，9 个字段

商品分类：5 台售货机（A、B、C、D、E），商品分饮料类和非饮料类两大类

三、数据预处理与分析

1、计算每台售货机 2017 年 5 月份的交易额、订单量及所有售货机交易总额和订单总量，结果如表格 1 所示：

表格 1：各售货机 5 月份交易额、订单量展示

	售货机 A	售货机 B	售货机 C	售货机 D	售货机 E
交易额	3385.1	3681.2	3733.4	2392.1	5699.0
订单量	756	869	790	564	1292

通过计算可知：所有售货机 2017 年的交易额为 286983.69（元），订单量为 70680

2、计算每台售货机每月的每单平均交易额与日均交易量，结果如表格 2 所示：

（其中位于上方的为每月的每单平均交易额数据，位于下方的为日均交易量数据）

表格 2：每台售货机每月的每单平均交易额与日均交易量

交易额、交易量	售货机 A	售货机 B	售货机 C	售货机 D	售货机 E
1 月	4.507	3.753	4.329	3.693	4.680
	10	11	12	8	11
2 月	3.864	3.256	3.826	3.089	3.638
	3	6	7	4	8
3 月	3.586	3.615	3.770	4.306	4.306
	8	8	8	6	11
4 月	4.037	4.075	4.404	3.790	4.160
	14	20	24	14	29
5 月	4.478	4.236	4.727	4.241	4.411
	24	28	25	18	41
6 月	4.047	4.068	4.502	4.026	3.818
	55	61	62	34	86
7 月	4.098	4.402	3.988	4.230	3.919
	15	11	24	10	26
8 月	3.359	3.584	3.914	3.317	3.804
	21	31	40	23	57
9 月	4.307	4.130	4.427	3.899	4.125
	34	58	55	32	137
10 月	4.021	4.112	4.273	3.884	3.676
	50	65	71	38	89
11 月	4.472	4.269	4.352	3.862	4.283
	38	67	64	40	167
12 月	3.788	3.667	3.943	3.573	4.169
	64	71	76	53	104

四、数据分析与可视化

1、绘制 2017 年 6 月销量前 5 的商品柱状图，结果如图 1 所示：

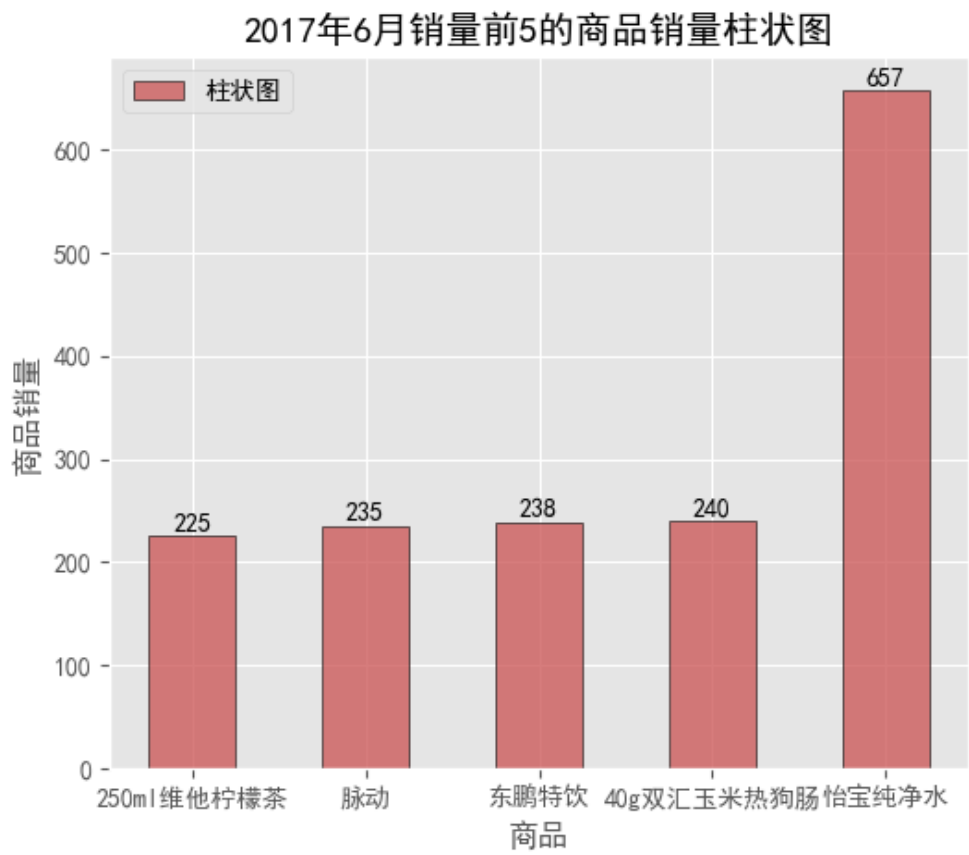


图 1：2017 年 6 月销量前 5 的商品柱状图

由图 1 可知，2017 年 6 月销量前 5 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，其中怡宝纯净水的销量表现较为突出，其它四种商品的销量几近相同。

2、绘制每台售货机每月总交易额折线图及交易额月环比增长率柱状图，结果如图 2、图 3 所示：

（为节省篇幅，图 3 展示的仅是售货机 C 的交易额月环比增长率柱状图，其它售货机的交易额月环比增长率柱状图可在文件夹“任务 2”中找到）

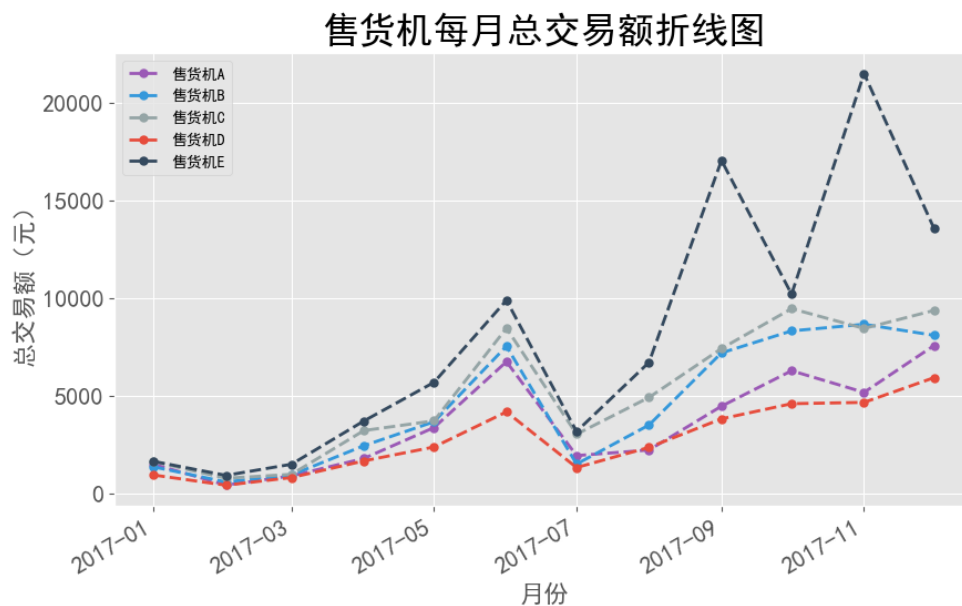


图 2: 售货机每月总交易额折线图

由图 2 知: 各售货机每月总交易额总体呈上升趋势, 在 2017 上半年, 其增长趋势一致; 在 2017 下半年, 售货机 E 每月总交易额增长迅猛且有较大的波动, 而其它售货机表现相近。

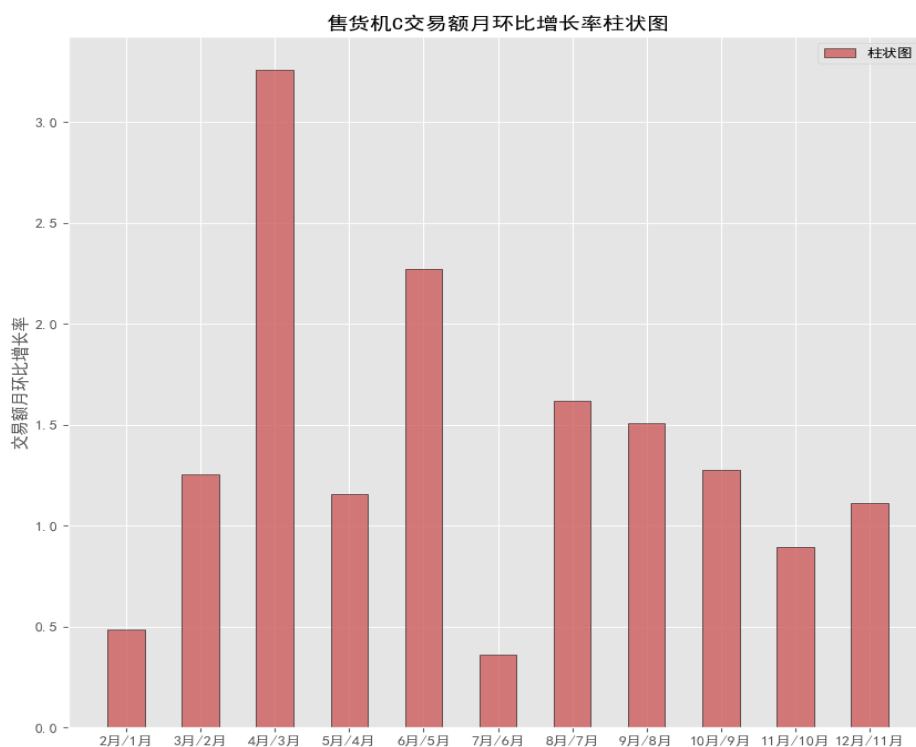


图 3: 售货机 C 交易额月环比增长率柱状图

3、绘制每台售货机毛利润占总毛利润比例的饼图，其中饮料类毛利率为 25%，非饮料类为 20%，结果如图 4 所示：

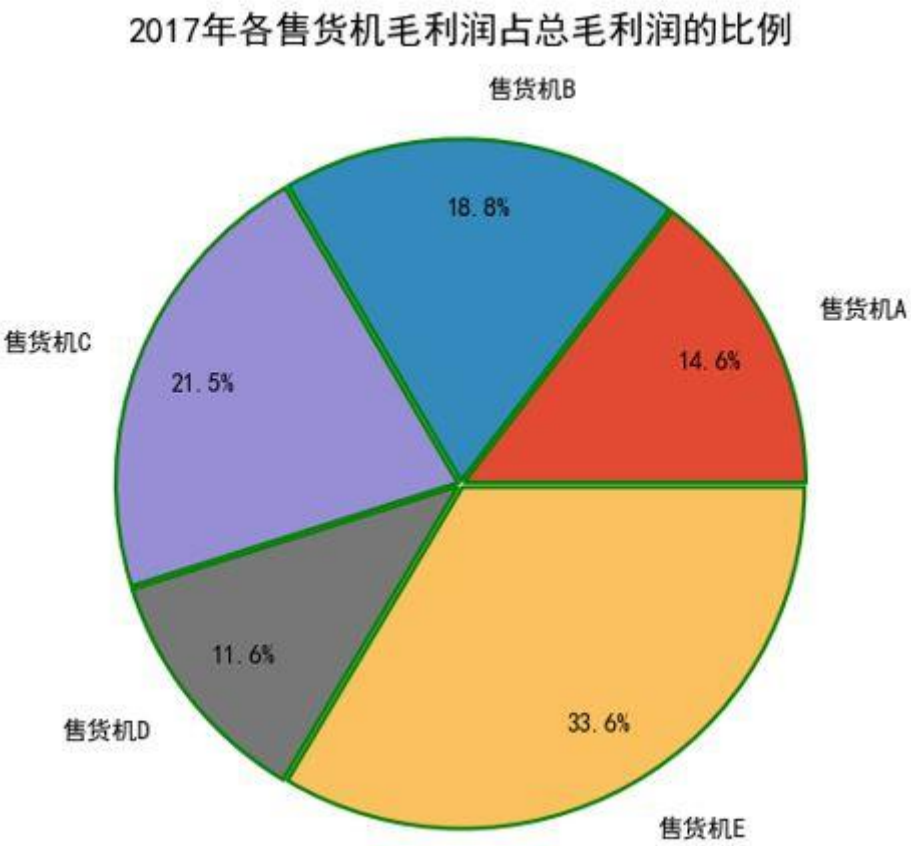


图 4：每台售货机毛利润占总毛利润比例的饼图

由图 5 可知，售货机 E 毛利润占总毛利润比例最高（33.6%），而售货机 D 毛利润占比最低且远低于 20%，该数据表明需要重点对售货机 D、E 销售数据进行分析，以探索数据背后的原因。

4、绘制每月交易额均值气泡图，横轴为时间，纵轴为商品的二级类目，结果如图 5 所示：

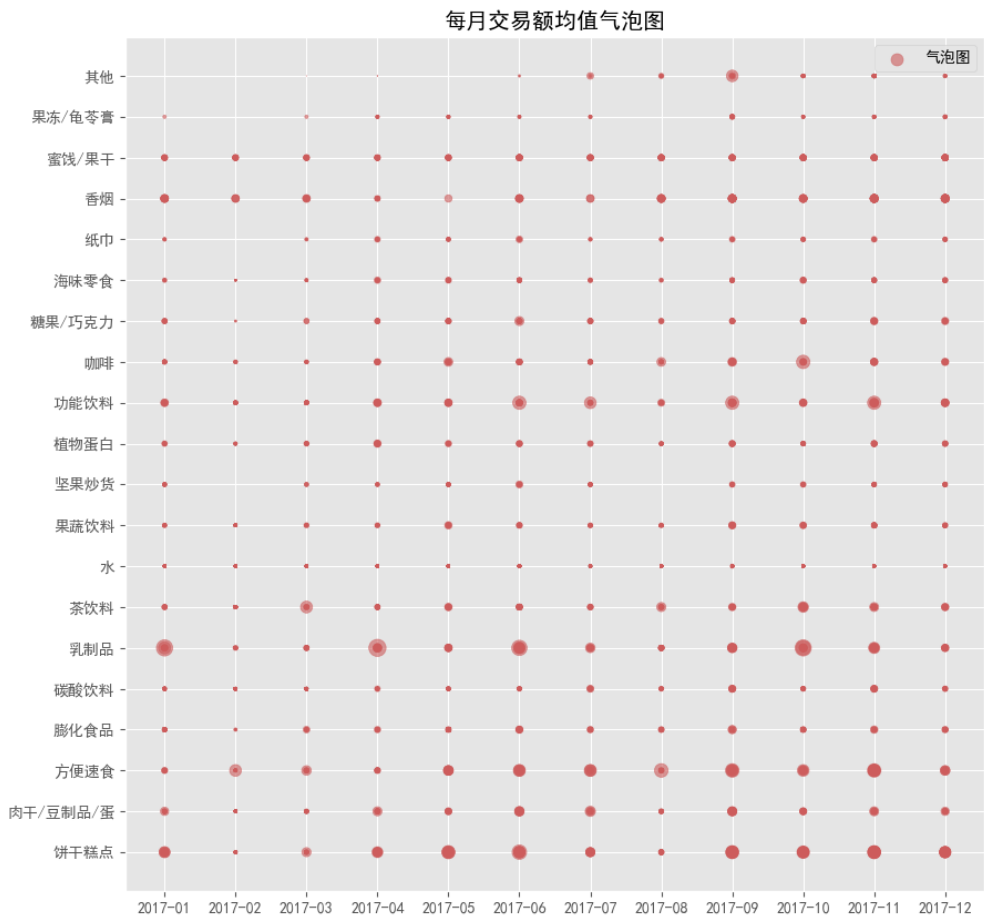


图 5：每月交易额均值气泡图

由图 5 可知，大多数商品在 2017 年不同月份的交易额分布均匀，并没有随时间发生较大变化，个别商品（如乳制品、饼干糕点、方便速食）在某些月份的订单量出现了明显波动。

5、绘制售货机 C 在 6、7、8 三个月订单量的热力图，横轴以天为单位，纵轴以小时为单位，结果如图 6 所示：

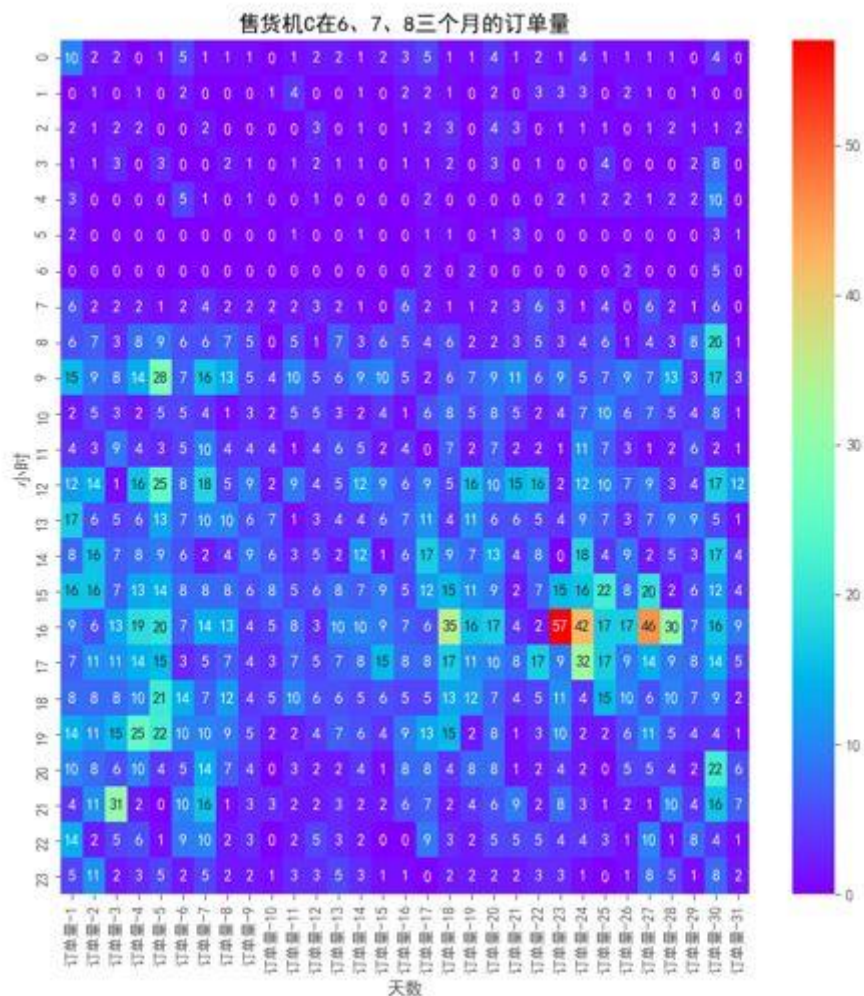


图 6: 售货机 C 在 6、7、8 三个月订单量的热力图

由图 6 可知，售货机 C 在 6、7、8 三个月订单量随时间分布较为均匀，但在部分时间段有一定的分布规律：

(1) 随天数变化：在 6、7、8 月的月初（1 号-5 号）、月末（22 号-28 号），订单量较多，其中每月的 23 号、27 号表现较为突出；

(2) 随小时变化：在每天的 0 点-12 点，订单量基本为零，大部分订单发生时间主要在 12 点-19 点。

五、生成各自动售货机画像

1、分析各售货机商品销售数据，给出每台售货机饮料类商品的标签，具体步骤如下所示：

(1) 判别依据：考虑到各售货机摆放位置、所售商品种类等不同因素，这里采用售货机各商品年销售量的平均值作为判定商品销售为“正常”或“滞销”的依据，以年销售量（>200）作为判定商品销售为“热销”的依据。

(2) 给出每台售货机饮料类商品的标签，结果如表格 3 所示：

(3) 给出每台售货机非饮料类商品的标签，结果如表格 4 所示：

表格 3：各售货机饮料类商品标签

标签 \ 售货机	售货机 A	售货机 B	售货机 C	售货机 D	售货机 E
热销	6	13	17	6	24
正常	28	22	10	25	11
滞销	78	80	87	74	78

表格 4：各售货机非饮料类商品标签

标签 \ 售货机	售货机 A	售货机 B	售货机 C	售货机 D	售货机 E
热销	2	2	2	1	7
正常	42	38	41	32	35
滞销	118	106	109	97	112

由表格 3、表格 4 可知，只有个别商品标签为“热销”，大多数商品的标签为“滞销”，标签的判别结果与商品销售者期望严重偏离。

商品年销售量的平均值被作为判定商品销售为“正常”或“滞销”的依据，这是大多数商品标签为“滞销”的原因。而数据中多数饮料类商品年销售量仅为个位数或低于 50，说明该判别依据某种程度上可以接受。

2、营销意见：

(1) 销售部需要通过市场调查对售货机所销售的商品种类进行调整，重点放在标签为“滞销”的商品；

(2) 对年销售量仅为个位数或年销售量在 20、30 前后的商品，考虑将其下架；

(3) 对标签为“热销”的商品及标签为“正常”的销量靠前的商品，考虑增加货位，以减少定期补货的物流费用。

六、商品销量预测

1、算法原理

ARIMA 模型：模型形式为：ARIMA(p,d,q)(P,D,Q)，其中 p 是指移除季节性变化后的序列所滞后的 p 期，d 是指移除季节性变化后的序列进行了 d 阶差分，q 指移除季节性变化后的序列进行了 q 次移动平均。

2、预测示例

选用数据为售货机 A 饮料类销售额数据，以 2017 年每一天的销售额进行预测，对数据进行平稳性处理，最终得到的模型为ARIMA(31,2,2)(31,2,2), 模型训练效果如图 7 所示：

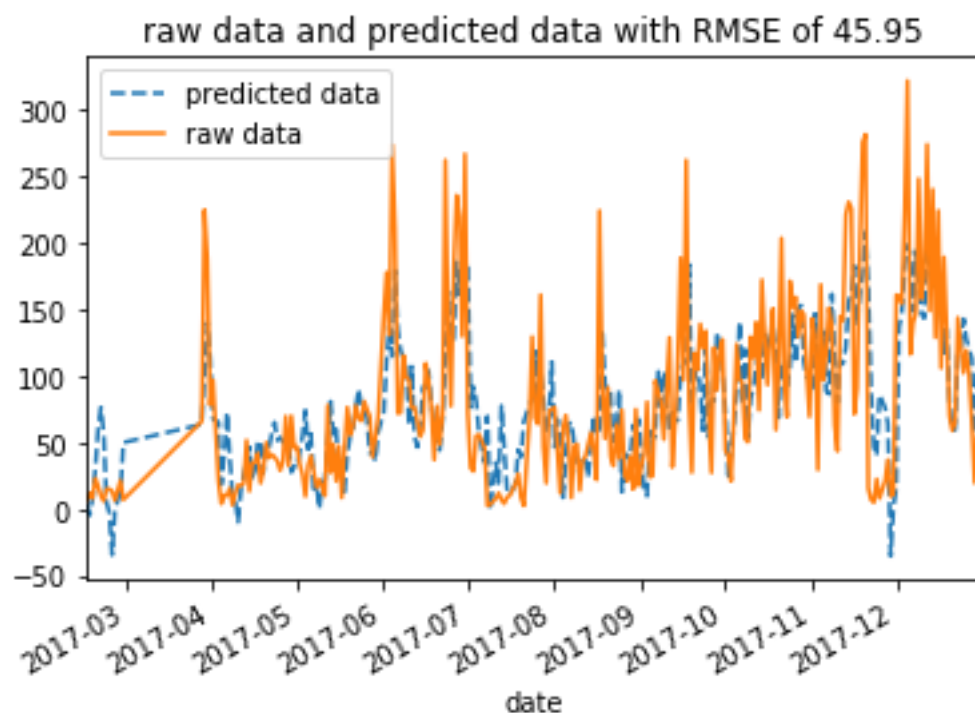


图 7：ARIMA 模型训练效果

运用该模型对售货机 A 饮料类商品在 2018 年 1 月份销售额进行预测，预测值为 3253.64（元）

3、模型评估

（1）即便模型已经通过检验并给出了预测值，但由于所给的数据仅为 2017 年的销售数据，缺少更多年份的销售数据，模型的预测结果可信度并不高。

（2）国内商业环境变化多端，很多因素会影响销售额，该模型仅考虑了数据的时间性规律。

综上，该模型可以用来预测，但可靠性不高，至少需要经营者提供更多年份的销售数据才有较高的可靠性。

4、数据要求

- （1）售货机的销售数据时间跨度需要有 2 或 3 年；
- （2）考虑其它因素对销售额的影响，如广告活动、竞争对手影响等。