



抖音商城数据统计及分析



背景与目标

背景与目标

背景

近年来，随着社交媒体的普及，大数据营销也得到了迅速的发展。企业营销正在从传统的单向活动转向以用户为中心的双向互动模式。抖音小店可以更好地满足用户的需求，不断推出新的产品、新的服务，满足用户的需求。



背景与目标

目标

根据上月订单数据，上半月订单数据和近七天的订单数据对商品，价格进行分析，判断用户消费趋向，拓展宣传渠道，优化选品策略，从而获取更多的消费者，提升营销效。



上月订单数据.csv



上半月订单数据.csv



近7天订单数据.csv



数据处理

数据处理

一

将上月订单数据.csv, 上半月订单数据.csv和近七天订单数据.csv文件中的数据依次导入, 再将三个表格中的数据拼接为一个表格。进行统一处理

导入数据

二

去除重复数据, 防止数据的不准确性影响分析结果。

数据去重

三

由于数据复杂, 为方便数据分析, 将无意义的删除, 再对数据进行标准化处理。

截取重点信息

数据处理

具体步骤

导入数据后，使用pandas库中的concat方法将三个表格按行拼接为一个表格，代码如下：

```
all_data = pd.concat([last_data, half_data, recent_data], axis=0)
```

然后删除多余的列，再对数据进行去重处理后，发现表格中仍存在不规则数据例如
'----' '\t' 等符号，具体处理方法如下

```
all_data.drop_duplicates(inplace=True) # 去除重复数据
all_data.replace({'售后状态': {'-': np.nan}}, inplace=True) #将售后状态列中的 '—' 改为空值
for i in ['选购商品', '支付方式', '订单提交时间']:
    all_data[i]=all_data[i].str.strip() #利用列表推导式将数据中的 '\t' 删除
```

数据处理

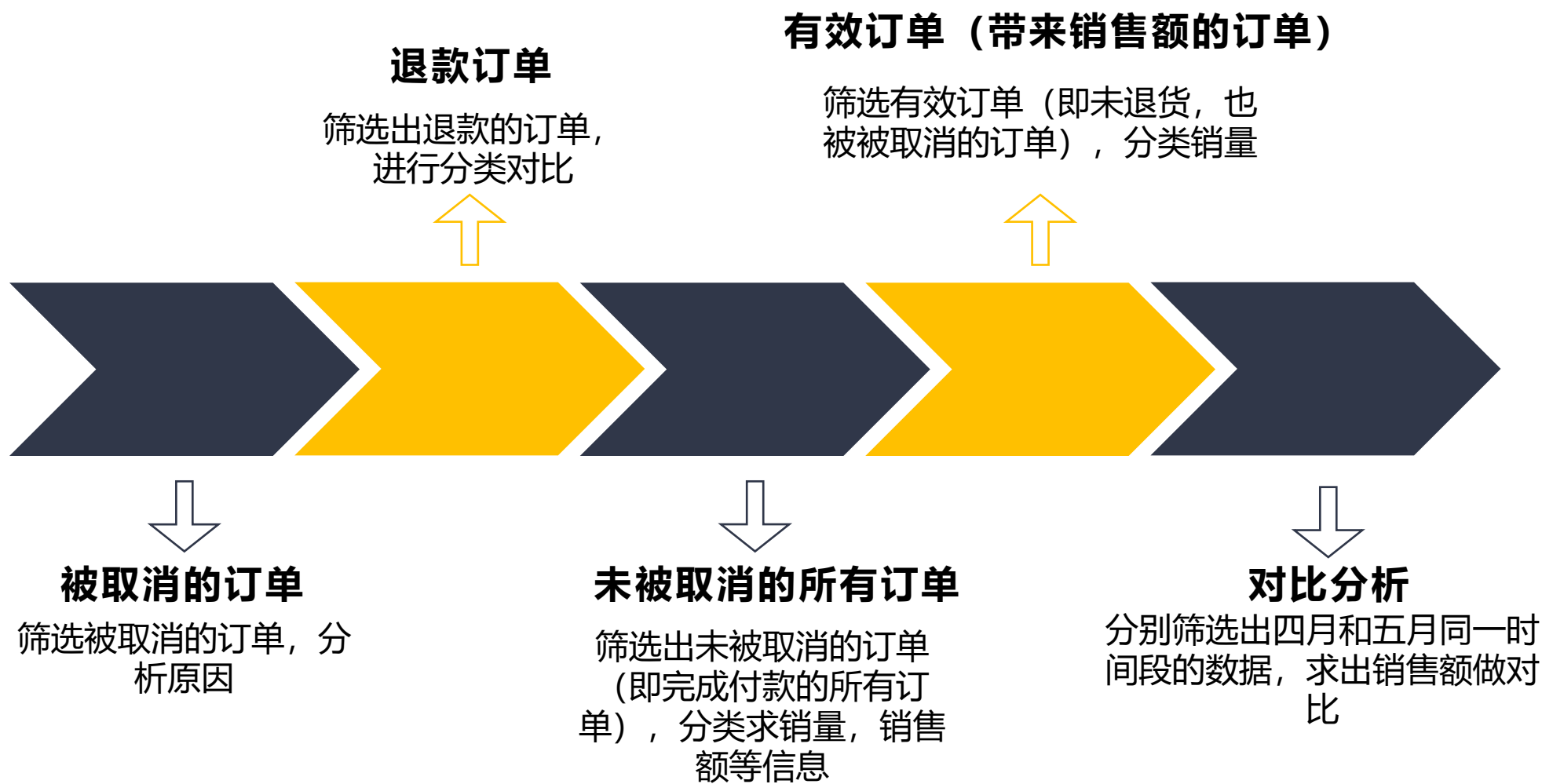
最终得到的数据如下：

all_data											
t[2]:											
	选购商品	商品数量	订单应付金额	支付方式	省	订单提交时间	旗帜颜色	APP渠道	售后状态	取消原因	流量渠道
0	【买一送二】新一代保暖护腰，双面磁石保暖，四根钢板支撑减压	1	69.0	微信	上海市	2023-04-30 23:59:57	灰色	抖音	NaN	用户超时未支付自动取消	其他
1	LIARJAN/莱尔健【标准】家庭综合训练拉力绳，一套练全身肌肉	1	69.0	支付宝	云南省	2023-04-30 23:59:32	灰色	抖音	退款成功	NaN	其他
2	LIARJAN/莱尔健【标准】家庭综合训练拉力绳，一套练全身肌肉	1	69.0	抖音支付	山东省	2023-04-30 23:59:13	灰色	抖音	NaN	NaN	其他



数据分析

数据分析



数据分析

筛选出被取消的订单，并以支付方式分类求出被取消的订单数量，代码实现如下：

```
undone_data=all_data[all_data['取消原因'].notnull()] #筛选被取消的订单
undone_data['取消数量']=undone_data['商品数量']
undone_data=undone_data.groupby(['支付方式']).agg({'取消数量':'sum'})
undone_data.sort_values(by='取消数量',ascending=False)
```

取消数量

支付方式

抖音月付	29409
抖音支付	10301
支付宝	9243
微信	2206
极速付	345

数据分析

筛选出退款的订单，并以商品类型，支付方式，商品颜色分类求出退款的订单数量，代码实现如下：

```
delete_data=all_data[all_data['售后状态'].notnull()]\n# 计算不同商品的退款量\ndelete_data['退款量']=delete_data['商品数量']\nlost_sales=delete_data.groupby(by=['选购商品']).agg({'退款量':'sum'})\nlost_sales.sort_values(by='退款量',ascending=False).head(10)
```

	退款量
选购商品	
【限量特惠专属】LIARJAN/莱尔健【标准】家庭综合训练拉力绳，一套练全身肌肉	4380
LIARJAN/莱尔健【标准】家庭综合训练拉力绳，一套练全身肌肉	3741
【5根金属支撑条】支撑透气护腰带，四季通用，支撑减压 束腰收腹	3033
【限量优惠专属】LIARJAN/莱尔健【标准】家庭综合训练拉力绳，一套练全身肌肉	1897
【买一送二】新一代保暖护腰，双面磁石保暖，四根钢板支撑减压	1029
【居家健身必备】11件综合训练拉力绳,一套搞定全身肌肉,4根80磅	172
LIARJAN/莱尔健【升级】家庭综合训练拉力绳，一套搞定全身肌肉	47
LIARJAN/莱尔健计数握力器 不同握力阶段人群的锻炼神器 臂肌	40
专业级全能运动护膝 男女膝盖篮球装备 关节半月板保护套跑步	18
LIARJAN/莱尔健「麒麟臂专用」拉力绳+握力器 强化力量 智能计数	17

数据分析

#统计不同支付方式订单的退款数量

```
lost_sales1=delete_data.groupby(by='支付方式').agg({'退款量':'sum'})  
lost_sales1
```

退款量

支付方式

微信	2102
抖音支付	3736
抖音月付	4807
支付宝	3768

```
lost_sales2=delete_data.groupby(by='旗帜颜色').agg({'退款量':'sum'})  
lost_sales2
```

退款量

旗帜颜色

橙色	1
灰色	13373
紫色	335
红色	685
绿色	18
蓝色	1

数据分析

筛选出未被取消的所有订单，并以商品类型分类求出订单数量，销售额及退款率，得到结果如下：

选购商品	总销量	销售额	退款量	退款率
LIARJAN/莱尔健【标准】家庭综合训练拉力绳，一套练全身肌肉	40399	2365889.37	3741.0	9.26%
【限量特惠专属】LIARJAN/莱尔健【标准】家庭综合训练拉力绳，一套练全身肌肉	31546	1936935.99	4380.0	13.88%
【5根金属支撑条】支撑透气护腰带，四季通用，支撑减压 束腰收腹	21968	1668993.94	3033.0	13.81%
【限量优惠专属】LIARJAN/莱尔健【标准】家庭综合训练拉力绳，一套练全身肌肉	15846	1016190.21	1897.0	11.97%
【买一送二】新一代保暖护腰，双面磁石保暖，四根钢板支撑减压	8436	587431.96	1029.0	12.20%
【居家健身必备】11件综合训练拉力绳,一套搞定全身肌肉,4根80磅	2049	141546.26	172.0	8.39%
LIARJAN/莱尔健计数握力器 不同握力阶段人群的锻炼神器 臂肌	570	12479.46	40.0	7.02%
LIARJAN/莱尔健【升级】家庭综合训练拉力绳，一套搞定全身肌肉	353	41499.43	47.0	13.31%
专业级全能运动护膝 男女膝盖篮球装备 关节半月板保护套跑步	197	13701.00	18.0	9.14%
LIARJAN/莱尔健「麒麟臂专用」拉力绳+握力器 强化力量 智能计数	173	20736.50	17.0	9.83%

数据分析

以支付方式分类，算出不同支付方式对应的退款率如下：

	总销量	退款量	退款率
支付方式			
抖音月付	35586	4807.0	13.51%
支付宝	34331	3768.0	10.98%
抖音支付	30147	3736.0	12.39%
微信	21776	2102.0	9.65%
极速付	17	0.0	0.00%

以商品颜色分类，算出不同颜色的商品对应的退款率如下：

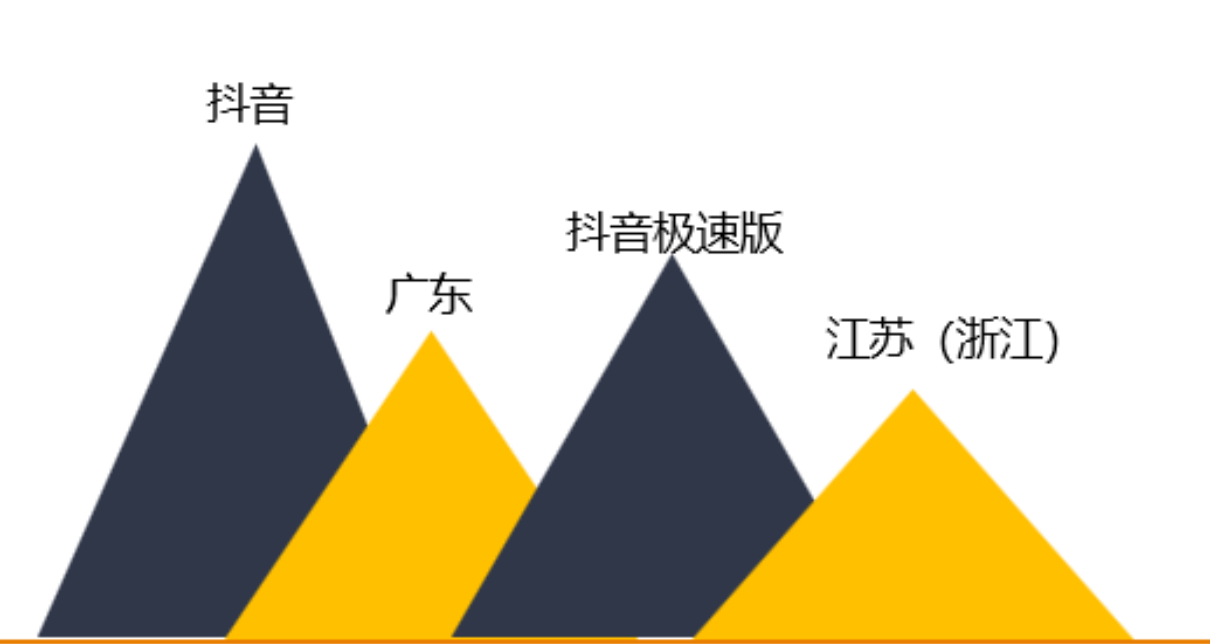
	总销量	退款量	退款率
旗帜颜色			
灰色	118779	13373	11.26%
红色	2265	685	30.24%
绿色	412	18	4.37%
紫色	396	335	84.60%
蓝色	3	1	33.33%
橙色	2	1	50.00%

根据用户支付方式的退款率可以看出，支付方式对用户退货有一定影响，比如：使用抖音月付支付的用户退款率比支付宝支付的用户退款率大一些。

根据商品颜色的退款率可以看出，不同的商品颜色对用户退款具有很强的关联性，红色和紫色商品的退款率很大，该现象可能是由商品色差因素造成的，需要改进。

数据分析

筛选出有效订单，并以APP渠道以及地区分类求出成交量top10，得到结果如下：



APP渠道	成交量	省	成交量
抖音	79407	广东省	12074
抖音极速版	15374	江苏省	10856
头条-抖音电商	3811	浙江省	10497
今日头条	3185	山东省	6061
番茄小说-抖音电商	1553	四川省	5366
头条lite-抖音电商	1090	福建省	4771
西瓜-抖音电商	1045	安徽省	4743
头条极速版	947	河南省	4436
番茄畅听-抖音电商	303	湖北省	4246
西瓜视频	266	上海市	3811

由结果可知，该商品的主要销售渠道是抖音及抖音的附属APP，在其他APP则需要加大宣传力度，扩大销售覆盖面。成交量与地区经济水平无太大关联性，而跟地区人口密集度有较强的关联性，应该对人口密集城市的用户加大宣传。

数据分析

由于数据只有4月1日—5月23日的，故截取4月1日—4月22日和5月1日—5月22日的数据，对分别求出两组数据中不同APP渠道的销售总额，并做对比分析，代码如下：

```
# 将'订单提交时间'列转换为日期时间类型:
```

```
valid_data['订单提交时间'] = pd.to_datetime(valid_data['订单提交时间'])
```

```
# 截取4月1日到4月22日数据
```

```
april_data = valid_data[
```

```
    (valid_data['订单提交时间'] >= pd.to_datetime('2023-04-01')) #设置起始日期
```

```
    & (valid_data['订单提交时间'] <= pd.to_datetime('2023-04-23'))] #设置截止日期
```

```
# 截取5月1日到5月22日数据
```

```
may_data = valid_data[
```

```
    (valid_data['订单提交时间'] >= pd.to_datetime('2023-05-01')) #设置起始日期
```

```
    & (valid_data['订单提交时间'] <= pd.to_datetime('2023-05-23'))] #设置截止日期
```

计算四月有效销售额

```
april_data['四月销售额'] = april_data['商品数量'] * april_data['订单应付金额']
```

```
april = april_data.groupby(by=['APP渠道']).agg({'四月销售额': 'sum'})
```

计算五月有效销售额

```
may_data['五月销售额'] = may_data['商品数量'] * may_data['订单应付金额']
```

```
may = may_data.groupby(by=['APP渠道']).agg({'五月销售额': 'sum'})
```

将april表和may表连接

```
contrast = pd.merge(april, may['五月销售额'], on='APP渠道', how='left')
```

```
contrast['五月销售额'].fillna(0, inplace=True) #将表中的空值改为0
```

加入一行计算总销售额

```
total_row = pd.DataFrame(contrast.sum()).T
```

```
total_row.index = ['总计']
```

```
contrast = contrast.append(total_row)
```

增加差值列, 判断不同平台的销售趋势

```
contrast['差值'] = contrast['五月销售额'] - contrast['四月销售额']
```

```
contrast
```

数据分析

结果如右图所示：

由结果可知，五月的总销售额比四月的总销售额降低了很多。且不同的APP渠道在四月和五月的销售额有很大差异，因素也可能存在多样性，故将数据根据差值的正负分为两类，表示营业额上涨和营业额下降。

	四月销售额	五月销售额	差值
今日头条	118404.00	76118.00	-42286.00
头条-抖音电商	95044.57	111857.20	16812.63
头条lite-抖音电商	23402.90	34782.90	11380.00
头条极速版	38506.00	20050.00	-18456.00
学浪	207.00	375.00	168.00
悟空浏览器-抖音电商	1082.00	1034.00	-48.00
抖音	2530288.89	1684583.58	-845705.31
抖音极速版	453632.10	369077.49	-84554.61
火山	2884.00	7834.00	4950.00
火山小视频	3563.00	1478.00	-2085.00
番茄小说	6821.00	0.00	-6821.00
番茄小说-抖音电商	31996.55	50215.50	18218.95
番茄畅听	474.00	128.00	-346.00
番茄畅听-抖音电商	7160.00	9400.90	2240.90
皮皮虾	69.00	0.00	-69.00
皮皮虾-抖音电商	197.00	0.00	-197.00
穿山甲-抖音电商	741.00	888.00	147.00
站外	69.00	138.00	69.00
西瓜-抖音电商	27126.90	26696.90	-430.00
西瓜视频	12740.00	4368.00	-8372.00
总计	3354408.91	2399025.47	-955383.44

数据分析

筛选出营业额上涨的APP渠道

```
uptrend= contrast[contrast['差值'] > 0]
uptrend
```

	四月销售额	五月销售额	差值
头条-抖音电商	95044.57	111857.2	16812.63
头条lite-抖音电商	23402.90	34782.9	11380.00
学浪	207.00	375.0	168.00
火山	2884.00	7834.0	4950.00
番茄小说-抖音电商	31996.55	50215.5	18218.95
番茄畅听-抖音电商	7160.00	9400.9	2240.90
穿山甲-抖音电商	741.00	888.0	147.00
站外	69.00	138.0	69.00

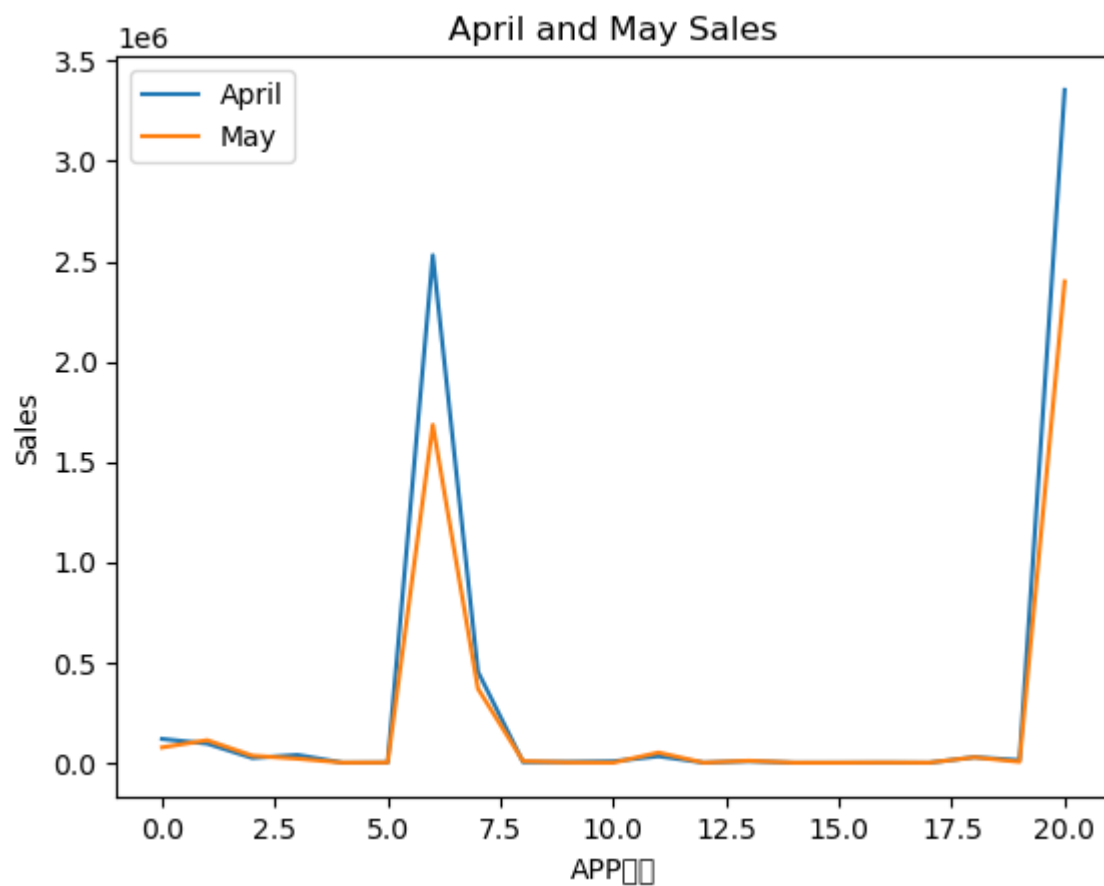
筛选出营业额下降的APP渠道

```
downturn= contrast[contrast['差值'] < 0]
downturn
```

	四月销售额	五月销售额	差值
今日头条	118404.00	76118.00	-42286.00
头条极速版	38506.00	20050.00	-18456.00
悟空浏览器-抖音电商	1082.00	1034.00	-48.00
抖音	2530288.89	1684583.58	-845705.31
抖音极速版	453632.10	369077.49	-84554.61
火山小视频	3563.00	1478.00	-2085.00
番茄小说	6821.00	0.00	-6821.00
番茄畅听	474.00	128.00	-346.00
皮皮虾	69.00	0.00	-69.00
皮皮虾-抖音电商	197.00	0.00	-197.00
西瓜-抖音电商	27126.90	26696.90	-430.00
西瓜视频	12740.00	4368.00	-8372.00
总计	3354408.91	2399025.47	-955383.44

数据分析

根据营业额绘制折线图如下：



结合销售额对比表和折线图可知，销售额下降最多的APP渠道是‘抖音’，而增加最多的‘番茄小说-抖音电商’的增加量远小于减少量，由此可以看出，消费群体在不断流失。



决策建议



决策建议

建议一：

优化用户体验，可通过加强商品质量，增加商品展示图的真实性，来减少商品退货，订单取消等事件的发生。还可以对购买过的用户实行问卷奖励机制，通过老用户的反馈意见对商品不断优化，增加商品好评率。

建议二：

拓展新用户群体，定期举办促销活动，如新用户限时折扣、买一送一或赠品活动，以吸引新用户的关注并鼓励他们尝试购买您的产品。

建议三：

加大产品宣传力度，数据显示销售额有明显的下降趋势，可能是宣传力度不到位导致的，可以通过短视频创新与商品推荐结合，激发用户购买欲。



THANK YOU