一、数据读取与清洗

|  |
| --- |
| *#读取数据并处理数据* **def** readcvs(path):  data=pd.read\_csv(path,encoding=**'gbk'**)  data=data.drop\_duplicates() *#去重* columns=[**'商品规格'**,**'选购商品'**,**'支付方式'**, **'订单提交时间'**]  data[columns] = data[columns].applymap(**lambda** x: x.replace(**'\t'**, **''**)) *#删除字符串'\t'* data.loc[data[**'流量来源'**] == **'-'**, **'流量来源'**] = **'其他'** *#将‘-’字符替换成其他* **return** data |

二、功能模块

2.1）有效订单的数量以及有效占比

|  |
| --- |
| **def** OCR(path): *#有效订单率* data=readcvs(path)  x = (data[**'订单状态'**] != **'已关闭'**).sum()  **return f'有效订单总数为{**x**}\t'** \  **f'有效订单率为{**x/data.shape[0]\*100**:.2f}%\n'** \  **f'无效订单总数为{**data.shape[0]-x**}\t'**\  **f'无效订单率为{**100-x/data.shape[0]\*100**:.2f}%'** |

2.2）日均销售额：

|  |
| --- |
| *#日均销售额* **def** self\_volume(path):  data = pd.read\_csv(path, encoding=**'gbk'**)  self\_volume = data.loc[data[**'支付完成时间'**].notnull(), **'订单应付金额'**].sum()  data[**'支付完成时间'**] = pd.to\_datetime(data[**'支付完成时间'**]) *# 处理缺失值* self\_volume = self\_volume / data[**'支付完成时间'**].dt.date.nunique()  **return f'天均销售额为：{**self\_volume**:.2f}'** |

2.3）日均销售订单量：

|  |
| --- |
| *#日均销售订单量* **def** date\_self(path):  data = pd.read\_csv(path, encoding=**'gbk'**)  data[**'支付完成时间'**] = pd.to\_datetime(data[**'支付完成时间'**]) *# 处理缺失值* data\_self= data[**'支付完成时间'**].isna().count()/data[**'支付完成时间'**].dt.date.nunique()  **return f'天均订单销售量为：{**data\_self**}'** |

2.4）不同商品的销售量

|  |
| --- |
| **def** good(path):  data = readcvs(path)  goods=data.groupby(**'商品ID'**).agg(总订单量=(**'商品ID'**,**'count'**),  订单有效数量=(**'售后状态'**, **lambda** x: (x != **'已关闭'**).sum()))\  .sort\_values(**'总订单量'**,ascending=**False**)  goods[**'有效订单率'**]=goods[**'订单有效数量'**]/goods[**'总订单量'**]  goods[**'有效订单率'**]=goods[**'有效订单率'**].apply(**lambda** x : **f'{**x\*100**:.2f}%'**)  **return** goods.head() |

2.5）总体退还率：

|  |
| --- |
| **def** refund(path): *#总体退款率* data=readcvs(path)  refund\_num=((data[**'售后状态'**] == **'退款成功'**) | (data[**'售后状态'**] == **'待退货'**) | (data[**'售后状态'**] == **'代收退货'**)).sum()  x=(data[**'订单状态'**]!= **'已关闭'**).sum()  **return f'总退款数为{**refund\_num**}\t退款率为{**refund\_num/x\*100**:.2f}%'** |

2.6）支付方式数量及占比：

|  |
| --- |
| **def** pay(path): *#支付方式* data=readcvs(path)  pay=data.groupby(**'支付方式'**).agg(数量=(**'支付方式'**,**'count'**)).sort\_values(**'数量'**,ascending=**False**)  pay[**'占比'**]=pay[**'数量'**]/pay[**'数量'**].sum()\*100  pay[**'占比'**]=pay[**'占比'**].apply(**lambda** x : **f'{**x**:.2f}%'**)  **return** pay.head() |

2.7）不同省份的销售订单数量以及占比

|  |
| --- |
| **def** area(path): *#地区占比* data = readcvs(path)  area=data.groupby(**'省'**).agg(数量=(**'省'**,**'count'**)).sort\_values(**'数量'**,ascending=**False**)  area[**'占比'**]=area[**'数量'**]/area[**'数量'**].sum()  area[**'占比'**]=area[**'占比'**].apply(**lambda** x : **f'{**x\*100**:.2f}%'** )  **return** area.head() |

2.8）不同app渠道的订单数量以及占比

|  |
| --- |
| **def** way(path): *#app渠道* data = readcvs(path)  way=data.groupby(**'APP渠道'**).agg(数量=(**'APP渠道'**,**'count'**)).sort\_values(**'数量'**,ascending=**False**)  way[**'占比'**]=way[**'数量'**]/way[**'数量'**].sum()\*100  way[**'占比'**]=way[**'占比'**].apply(**lambda** x : **f'{**x**:.2f}%'**)  **return** way.head() |

2.9）不同流量体裁的订单数量以及占比

|  |
| --- |
| **def** try\_way(path): *#流量体裁* data = readcvs(path)  trf\_way=data.groupby(**'流量体裁'**).agg(数量=(**'流量体裁'**,**'count'**)).sort\_values(**'数量'**,ascending=**False**)  **try**:  trf\_way.drop(**'数据将于第二天更新'**,axis=0,inplace=**True**)  **except**:  **pass** trf\_way[**'占比'**]=trf\_way[**'数量'**]/trf\_way[**'数量'**].sum()\*100  trf\_way[**'占比'**]=trf\_way[**'占比'**].apply(**lambda** x : **f'{**x**:.2f}%'**)  **return** trf\_way.head() |

2.10）不同流量来源的订单数量以及占比

|  |
| --- |
| **def** ad\_source(path):  data=readcvs(path)  ad\_source=data.groupby(**'广告渠道'**).agg(数量=(**'广告渠道'**,**'count'**)).sort\_values(**'数量'**,ascending=**False**)  ad\_source[**'占比'**]=ad\_source[**'数量'**]/ad\_source[**'数量'**].sum()\*100  ad\_source[**'占比'**]=ad\_source[**'占比'**].apply(**lambda** x : **f'{**x**:.2f}%'**)  **return** ad\_source.head() |

2.11）购买降价商品的订单数量以及占比

|  |
| --- |
| **def** low(path): *#已完成订单中降价占比* data = readcvs(path)  low=data[(data[**'订单状态'**] != **'已关闭'**) &  (data[**'降价类优惠'**]!=0)].shape[0] / data.loc[data[**'订单状态'**] != **'已关闭'**].shape[0]  **return f'购买降价商品的订单比例：{**low\*100**:.2f}%'** |

2.12）选择安心购的订单数量以及占比

|  |
| --- |
| **def** ax\_shop(path): *#安心购* data = readcvs(path)  ax\_shop=(data[**'是否安心购'**]==**'是'**).sum()/data.shape[0]\*100  **return f'选择安心购的比例为:{**ax\_shop**:.2f}%\n未选择安心购的比例为:{**100-ax\_shop**:.2f}%'** |

2.13）一天中不同时间的订单数量以及占比

|  |
| --- |
| **def** hour(path): *#时间占比* data = readcvs(path)  data.dropna(subset=[**'支付完成时间'**], inplace=**True**) *#处理缺失值* data[**'支付完成时间'**]=pd.to\_datetime(data[**'支付完成时间'**])  data[**'时间段'**]=data[**'支付完成时间'**].dt.hour.astype(int)  hour=data.groupby(**'时间段'**).agg(订单数量=(**'时间段'**,**'count'**))  hour[**'占比'**]=hour[**'订单数量'**]/hour[**'订单数量'**].sum()\*100  hour[**'占比'**]=hour[**'占比'**].apply(**lambda** x : **f'{**x**:.2f}%'**)  **return** hour |

2.14）将三个不同时间段的结果合并到一个dataframe对象中

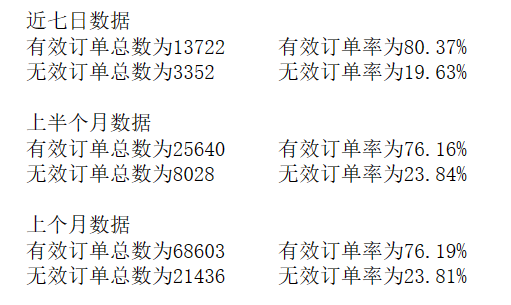
|  |
| --- |
| *#合并三个时间段的datafram对象* **def** concat(week, half\_month, month):  merged\_df = pd.concat([week, half\_month, month], axis=1, sort=**False**, keys=[**'近7日'**, **'上半个月'**, **'上个月'**])  **return** merged\_df |

三、运行

|  |
| --- |
| **import** pandas **as** pd **import** shop **as** s path=[**'E:/泰迪/资料/数据分析/抖音商城运营-总助数据统计及分析/抖音商城运营-总助数据统计及分析/近7天订单数据.csv'**,  **'E:/泰迪/资料/数据分析/抖音商城运营-总助数据统计及分析/抖音商城运营-总助数据统计及分析/上半月订单数据.csv'**,  **'E:/泰迪/资料/数据分析/抖音商城运营-总助数据统计及分析/抖音商城运营-总助数据统计及分析/上月订单数据.csv'**] pd.set\_option(**'display.unicode.east\_asian\_width'**, **True**)  *#统计有效订单率* print(**f'近七日数据\n{**s.OCR(path[0])**}\n\n上半个月数据\n{**s.OCR(path[1])**}\n\n上个月数据\n{**s.OCR(path[2])**}\n'**)  *#日均销售额*  print(**f'近七日数据\n{**s.self\_volume(path[0])**}\n\n上半个月数据\n{**s.self\_volume(path[1])**}\n\n上个月数据\n{**s.self\_volume(path[2])**}\n'**)  *#日均销售订单量*  print(**f'近七日数据\n{**s.date\_self(path[0])**}\n\n上半个月数据\n{**s.date\_self(path[1])**}\n\n上个月数据\n{**s.date\_self(path[2])**}\n'**)  *#按商品分类，统计三个时间段的数量以及比例* print(s.concat(s.good(path[0]), s.good(path[1]), s.good(path[2])).stack(),**'\n'**)  *#总退款率* print(**f'近七日数据\n{**s.refund(path[0])**}\n\n上半个月数据\n{**s.refund(path[1])**}\n\n上个月数据\n{**s.refund(path[2])**}\n'**)  *#三个时间段的支付方式数量以及占比* print(s.concat(s.pay(path[0]),s.pay(path[1]),s.pay(path[2])),**'\n'**)  *#三个时间段的地区占比* print(s.concat(s.area(path[0]),s.area(path[1]),s.area(path[2])),**'\n'**)  *#app渠道* print(s.concat(s.way(path[0]),s.way(path[1]),s.way(path[2])),**'\n'**)  *#流量体裁* print(s.concat(s.try\_way(path[0]),s.try\_way(path[1]),s.try\_way(path[2])),**'\n'**)  *#流量来源* print(s.concat(s.source(path[0]),s.source(path[1]),s.source(path[2])),**'\n'**)  *#购买降价商品的订单比例* print(**f'近七日数据\n{**s.low(path[0])**}\n\n上半个月数据\n{**s.low(path[1])**}\n\n上个月数据\n{**s.low(path[2])**}\n'**)  *#选择安心购* print(**f'近七日数据\n{**s.ax\_shop(path[0])**}\n\n上半个月数据\n{**s.ax\_shop(path[1])**}\n\n上个月数据\n{**s.ax\_shop(path[2])**}\n'**)  *#各时间段比例* print(s.concat(s.hour(path[0]),s.hour(path[1]),s.hour(path[2])),**'\n'**) |

结果如下：

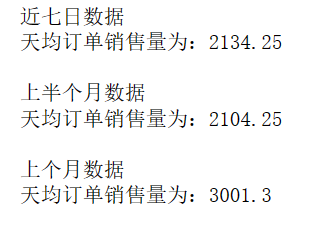
1、有效订单数以及占比



2、日均销售额



3、日均销售订单量



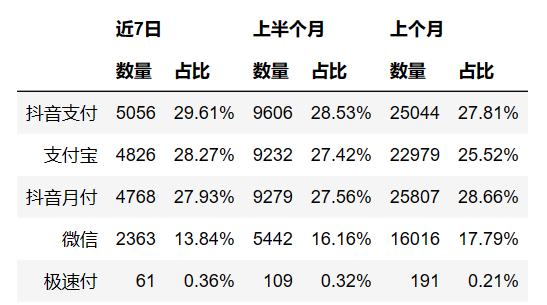
4、不同商品的销售数量



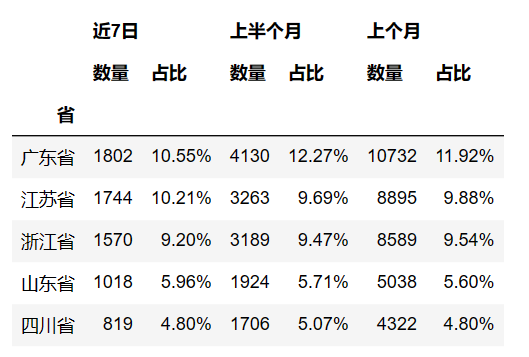
5、总退款率



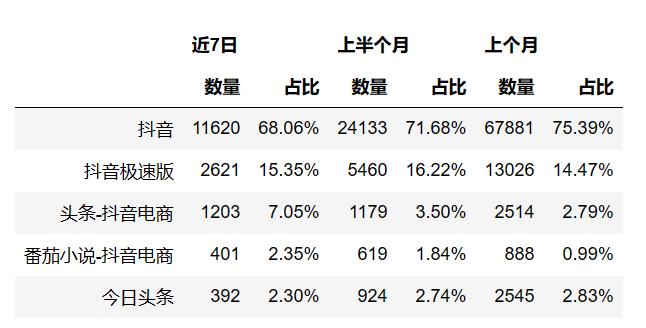
6、三个时间段的支付方式数量以及占比



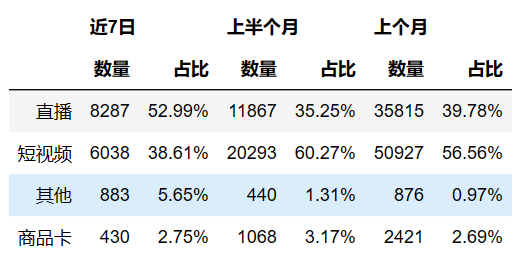
7、三个时间段的地区占比



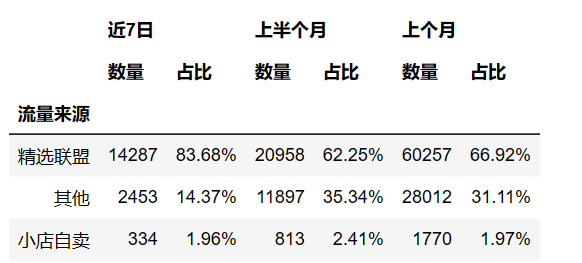
8、不用app渠道的订单数量以及占比



9、不同流量体裁的订单数量以及占比



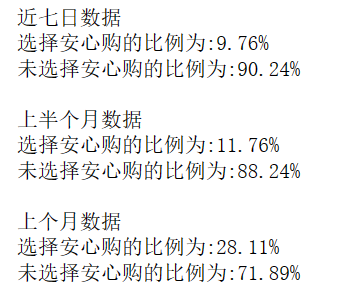
10、不同流量来源的订单数量以及占比



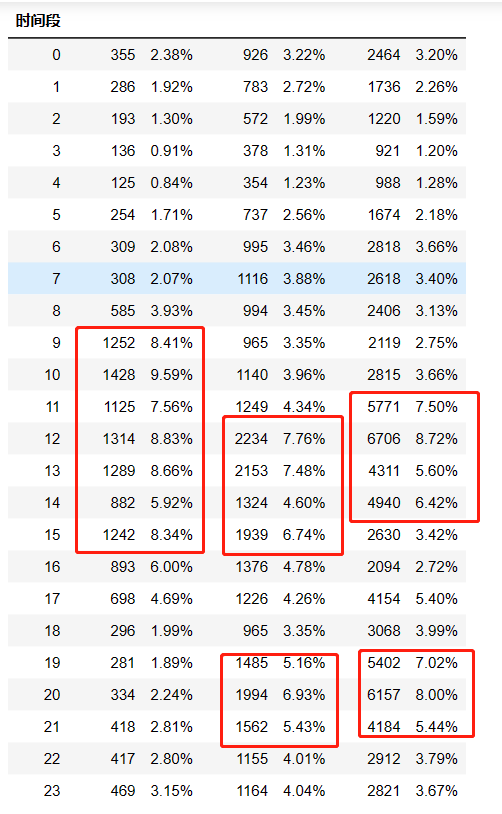
11、购买降价商品的订单数量以及占比



12、选择安心购的订单数量以及占比



13、一天中不同时间的订单数量以及占比



四、分析

4,1）独立分析

1、订单情况：近7日有效订单率与上半个月和上个月相比有所提升，有效订单率保持在较高水平。继续保持良好的产品质量和服务准确性，以提高客户满意度和忠诚度。

2、销售额：近七日的天均销售额略高于上半个月，但远低于上个月。考虑一些促销活动来刺激销售额，例如限时折扣、买一赠一或捆绑销售等。同时，通过精确广告投放和推广活动，吸引更多潜在客户。

3、订单销售量：近七日的天均订单销售量与上半个月相差不大，均比上个月低，尝试增加产品种类，提供更多样化的选择，以吸引更多不同需求的顾客。

4、商品订单情况:近7日各商品有效订单率都保持在相对较高的水平，其中商品ID为‘3595154472139107812’以及‘3573991726941080892’卖的最热销，主推热门商品，广告投放时优先展示热销商品，同时重点关注销量低的产品，分析是否有市场需求，产品质量等问题，根据过去的订单情况，调整定价、促销活动和产品策略，以提高订单数量和有效订单率

5、退款情况：退款率呈逐渐减少的趋势，在这段时间内客户对商品的满意度较高，退款情况相对较少，同时了解退款原因：分析退款原因，了解客户退款的主要问题，例如产品质量、物流问题或者客户服务不佳等。这将有助于你定位问题并采取相应的改进措施。

6、支付方式：抖音支付和支付宝是主要的支付方式，占据较大比例，其他支付方式的比例相对较低。确保支付通道的稳定性和安全性，同时考虑添加其他热门支付方式以满足不同用户的需求，或者考虑删除极速付的支付方式。

7、地域分布：广东、江苏、浙江是主要的消费省份，订单数量较多。重点关注广东、江苏和浙江等主要消费省份，加大市场推广力度，提高品牌知名度和产品曝光度。同时，根据地域差异，了解当地消费者喜好和需求，进行产品优化和定价策略调整。

8、平台：根据提供的数据，在近7日、上半个月和上个月中，抖音平台的订单数量占比都是最高的，其他平台的订单数量占比相对较低。注重抖音平台：由于抖音平台订单数量最高，建议将更多资源和精力投入到抖音平台的营销活动中。可以考虑增加抖音上的广告投放、提升产品在抖音上的曝光度，以吸引更多用户下单购买。利用其他平台的潜力：虽然其他平台的订单数量占比较低，但仍具有一定潜力。可以加大对抖音极速版、头条-抖音电商、番茄小说-抖音电商和今日头条等平台的关注度，在这些平台上进行更有针对性的营销活动，提升订单数量和占比。定期分析订单数据，观察不同平台的订单分布情况是否发生变化，并及时调整营销策略。可以根据订单高峰时段和平台偏好等因素，优化广告投放时段和内容，提高订单转化率。

9、渠道来源：精选联盟是主要的流量来源，其他渠道的贡献相对较小，且有增加趋势，可以进一步与合作伙伴合作，共同开展促销活动，扩大品牌影响力。另外，考虑拓展其他渠道来源，例如社交媒体、搜索引擎广告等，以增加多样化的流量来源，避免来源过于单一。

10、流量来源：可以看出精选联盟在近7日、上半个月和上个月的订单数量占比都是最高的。其他流量来源的订单数量占比相对较低。注重精选联盟，由于精选联盟在订单分布中占比最高，建议加大对精选联盟的关注度，与联盟合作伙伴密切合作，共同提升业绩。提升其他流量来源，尽管其他流量来源的订单占比较低，但仍有一定潜力。可以通过优化产品推广策略，增加品牌曝光度，吸引更多的流量来源下单购买，避免流量来源渠道单一

11、购买降价商品的订单比例：近七日购买降价商品的订单比例明显高于上半个月和上个月。可以继续进行定期促销活动，吸引更多顾客购买降价商品，并通过提高销售量来弥补降价带来的利润损失。

12、安心购选择比例：选择安心购的比例逐渐降低，未选择安心购的比例逐渐增加。增加更多的保障服务，降低价格或提供额外的特殊权益，以吸引更多顾客选择安心购。

13、近7日、上半个月和上个月的订单数量分布情况显示，不同时间段存在订单数量差异。在近7日中，第10小时的订单数量最高；在上半个月和上个月中，第12小时的订单数量最高。这些高峰时段可以作为调整营销策略的参考，例如在订单高峰时段进行促销活动或增加客户服务人员。然而，还应综合考虑其他因素，如产品热度和广告宣传，以做出更全面的营销决策。

综合分析：

在近7日的数据中，客户购买降价商品比例明显增加，这应该是商家的优惠促销活动，但是在日均销售量中以及日均销售额中，销售量与销售额并没有增加，相比起上个月的数据反而还下降了，再结合流量体裁以及流量来源分析进行分析，这很可能是因为商家的营销策略有所改变，加入了精选联盟，提高了直播带货的比例，给于了很大的优惠力度，但是短视频权重和其他流量来源不足导致曝光度不够，这意味着虽然进行了大力促销优惠和直播带货的一系列措施，但由于流量来源不足，带货主播的影响力不大，顾客对产品的关注度和购买意愿有所下降，因此利润反而下降。

基于以上分析，商家可以考虑以下策略来提升销售量：

1.提高短视频的曝光度：商家可以增加短视频的制作和推广力度，通过各大短视频平台提升曝光度。可以考虑制作有趣、吸引人的短视频内容，以吸引更多用户的关注。

2.多渠道投放广告：商家应考虑通过多种渠道进行广告投放，以增加流量来源。可以选择合适的广告平台、社交媒体、线上线下渠道等，提高产品的曝光率和知名度。

3.优化精选联盟策略：商家应仔细评估精选联盟的效果，并根据实际情况进行调整和优化。可以与联盟合作方进行沟通，共同制定更具吸引力的直播带货计划，提高直播的质量和吸引力。

4.数据分析与个性化推荐：商家可以通过客户数据分析，了解顾客购买偏好和需求，进行个性化推荐。通过精准营销，向顾客提供他们感兴趣的商品和优惠信息，提高购买意愿和转化率。

5.定期评估和调整策略：商家应定期评估营销策略的效果，并根据数据反馈进行相应调整。不断优化和改进营销手段，找到最适合目标顾客群体的推广方式。