

拼便宜系统上线两年多，已经积累了大量的用户，本次主要针对用户数据进行分析。

一、用户总览

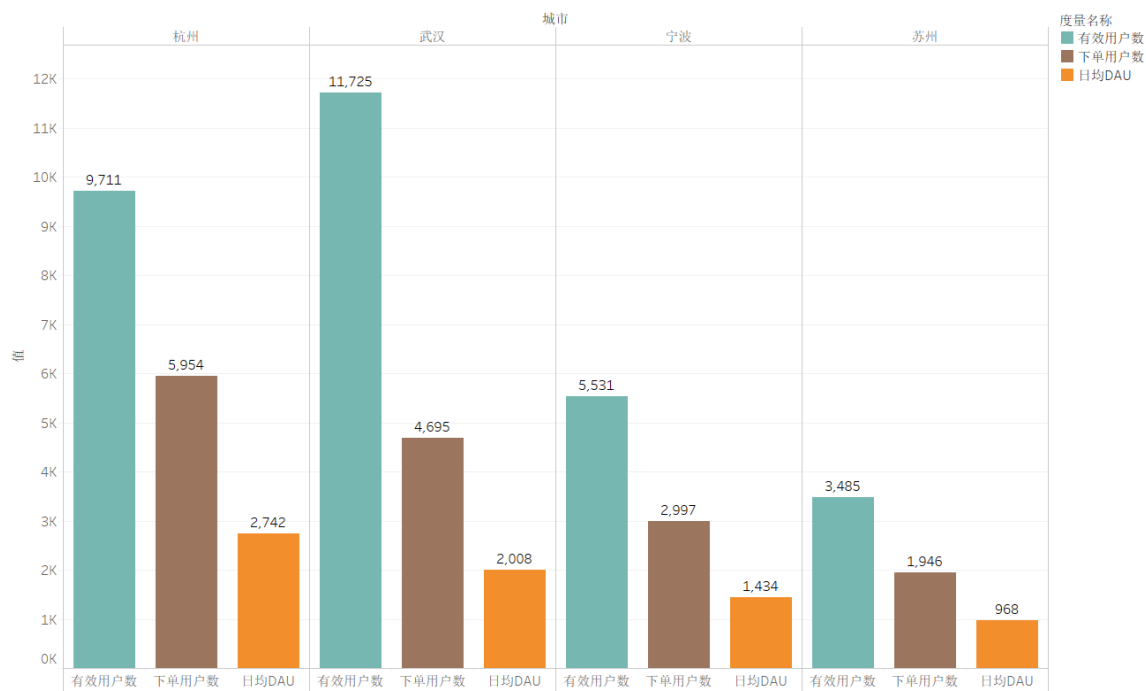
数据选取自2019年8月份的用户交易相关数据。

下图展示了各个城市用户基数、下单人数、日均DAU等一些指标的对比。

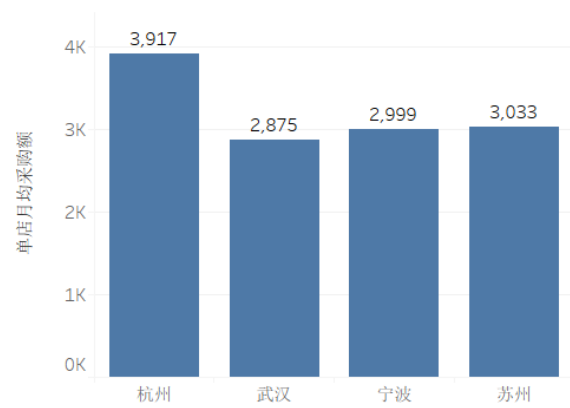
从图中可以看出：

- 1、武汉的用户基数最多，杭州次之。但是杭州下单人数、日均DAU最高，说明杭州的整体用户留存情况最好。
- 2、杭州的月店均采购额，平均单店交易sku数、平均下单频次、客单价均最高，说明杭州经过两年多的运营，已经发展的比较成熟稳定了。

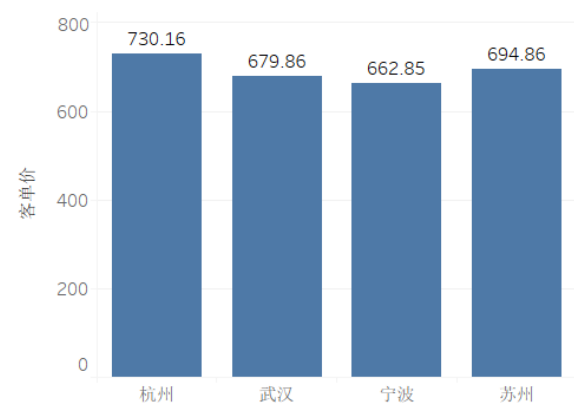
用户基数



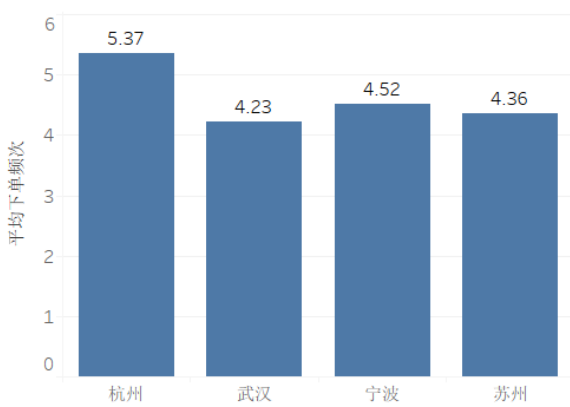
月店均采购额



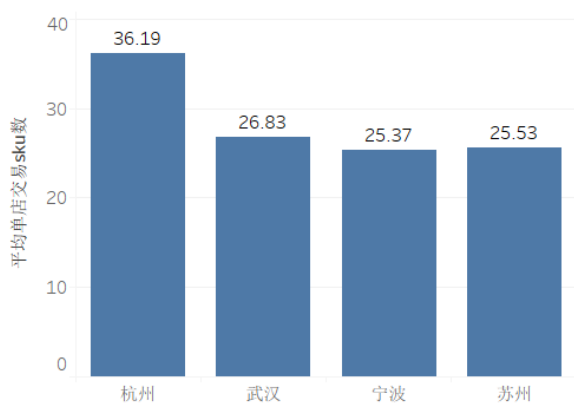
客单价



平均下单频次



单店平均交易sku数



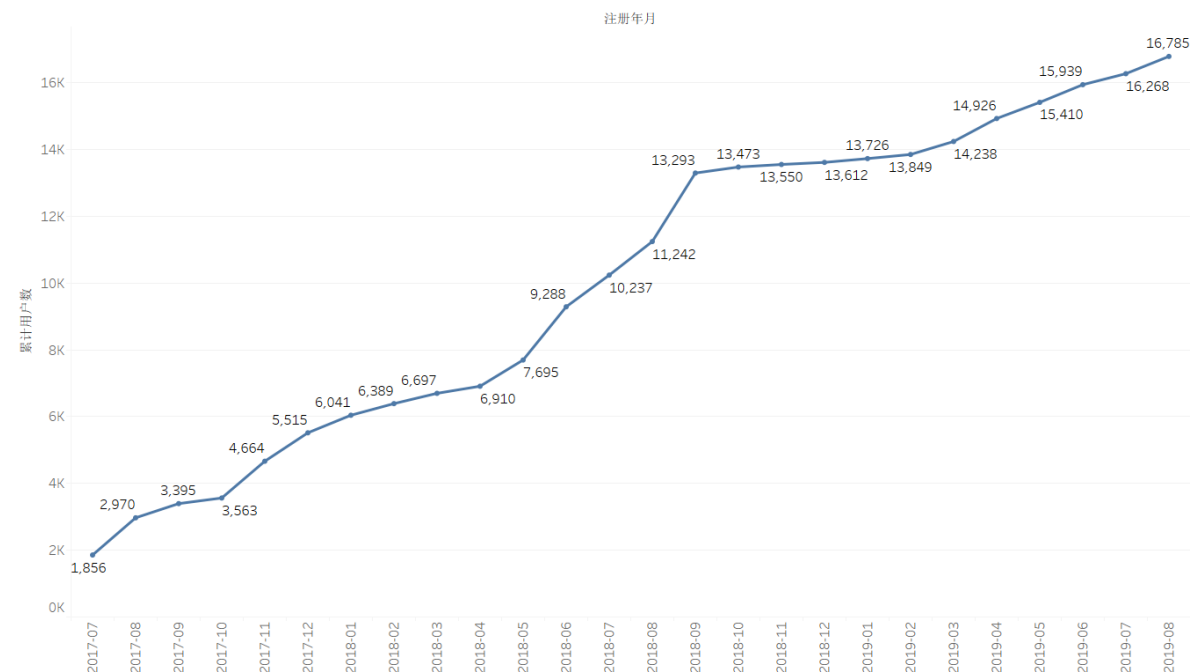
二、用户增长趋势

下图展示了四个城市的用户增长趋势图，数据选自每个城市从开城到2019年8月份每个月的累计注册用户数。

从图中可以看出：杭州、武汉和宁波在2019年4月开始用户增长率都有明显提升。

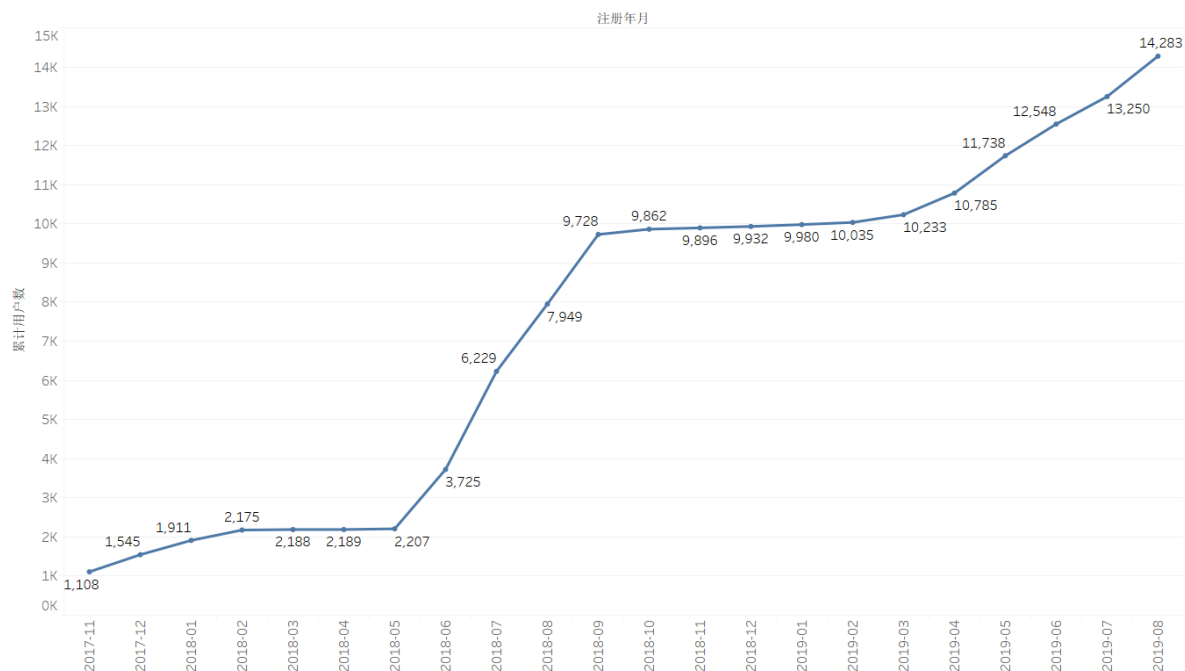
1、杭州：用户增长最快的时期是2018-05至2018-09，后面缓慢增长，2019-07、2019-08由于冲锋月的缘故，又挖掘了一批新用户。

用户增长趋势图-杭州



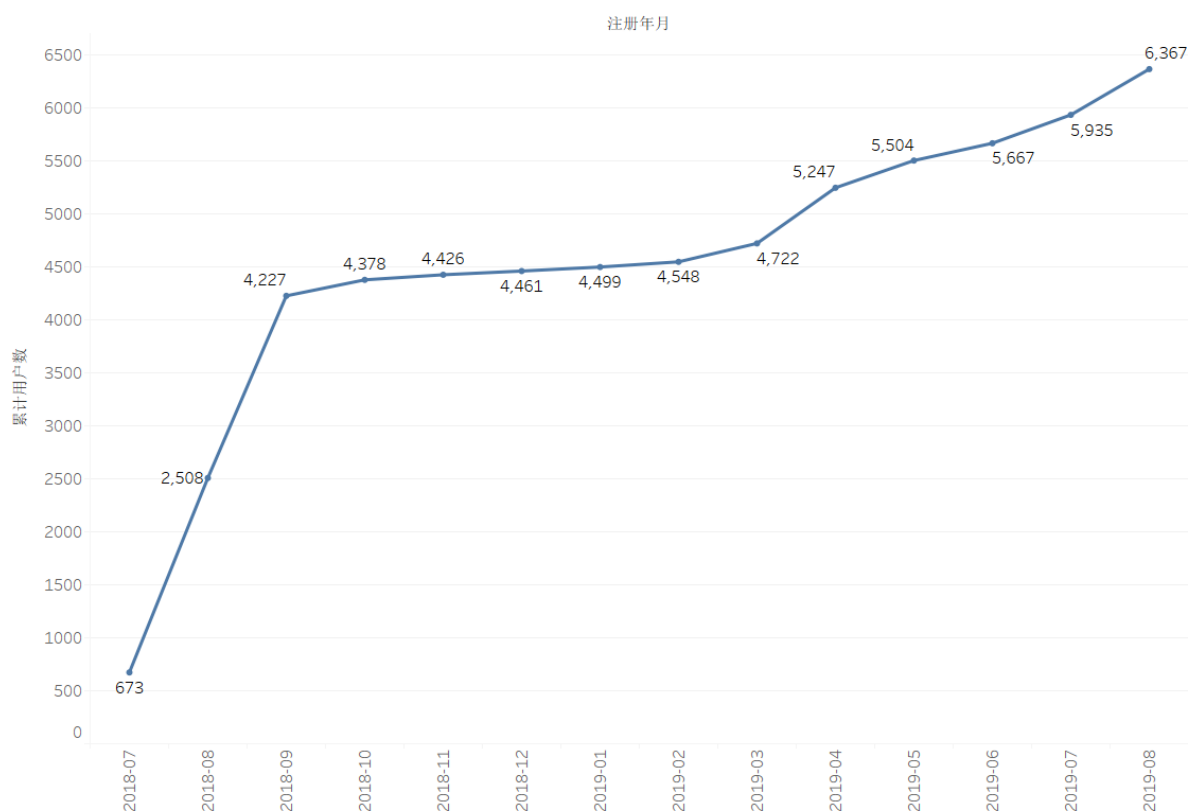
2、武汉：开城后的半年用户增长缓慢，在2018-06至2018-09这四个月增长最快，用户数增长了7500人左右；在2018-10至2019-03用户增长基本趋于平稳，在2019-04之后，用户又有所增长。

用户增长趋势图-武汉



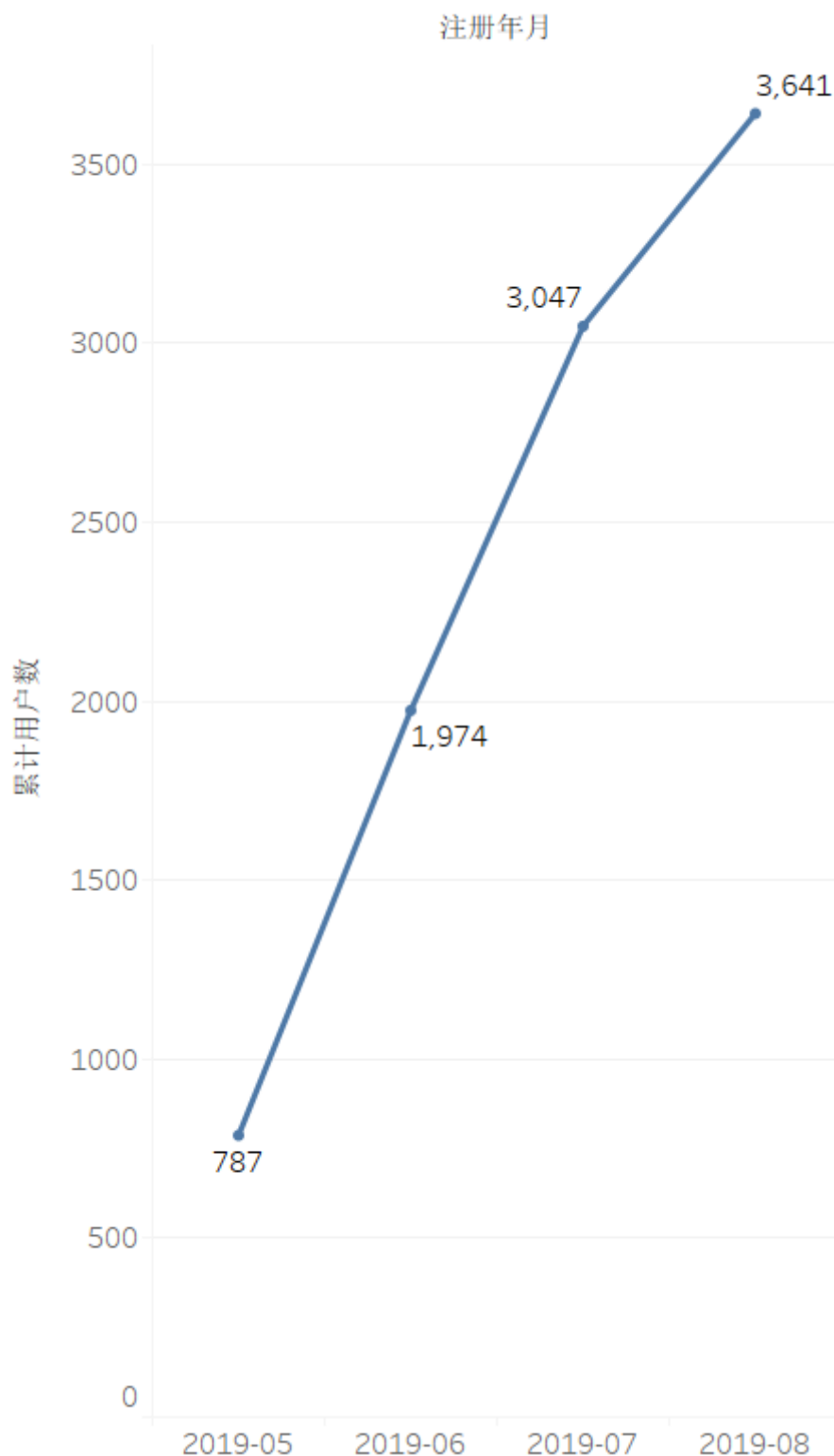
3、宁波：用户增长最快的时期是2018-08至2018-09这两个月，在2018-10至2019-03这段时间用户增长很缓慢，半年时间仅仅增长了500人左右，2019-04之后，增长趋势变快。

用户增长趋势图-宁波



4、苏州于2019年5月份开城，用户增长趋势很猛。在未来一段时间，应该会保持原有的增长趋势。

用户增长趋势图-苏州



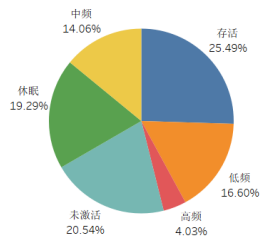
三、用户结构分布

如下是各个城市不同用户结构分布图，数据选自截止2019年08月底的所有有效注册用户数。

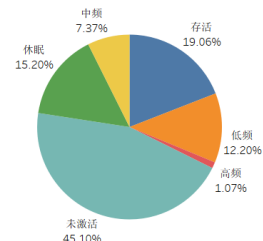
- 用户分层

根据用户历史下单情况对用户进行分类，可以将用户分成高频、中频、低频、存活、休眠和未激活6个等级。从用户分层结构可以看出，武汉有45%的用户未激活，如果能将这些用户激活，将会带来很大的收益。其他城市各个分层的比例差不多。

杭州



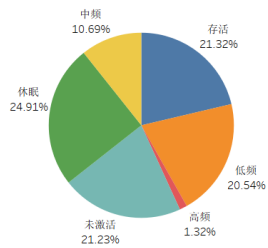
武汉



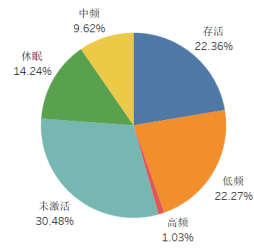
用户分层

- 存活
- 低频
- 高频
- 未激活
- 休眠
- 中频

宁波

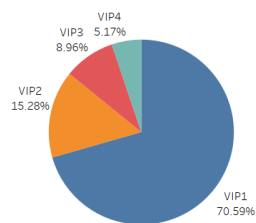


苏州

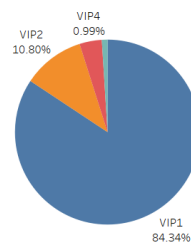


• VIP等级

杭州



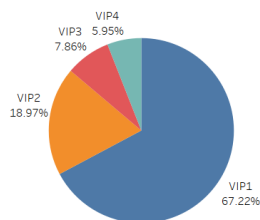
武汉



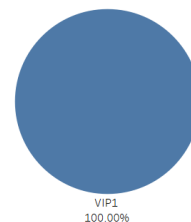
VIP等级

- VIP1
- VIP2
- VIP3
- VIP4

宁波



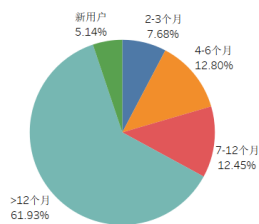
苏州



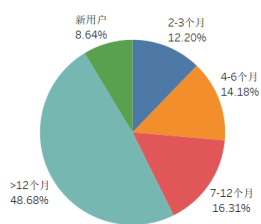
• 注册时长

根据注册时长对用户进行分类，可以将用户分成新用户、2-3个月、4-6个月、7-12个月、12个月以上5个等级。由于杭州开城最早，所以杭州注册时长12个月以上的用户最多，其次是武汉、宁波。苏州5月份刚开城，从图中可以看出，注册时长2-3个月的用户最多。

杭州



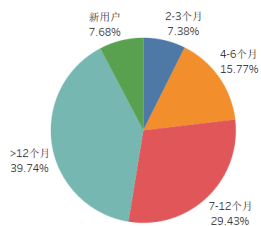
武汉



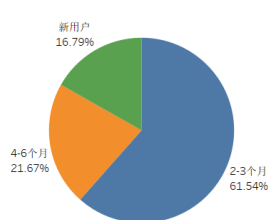
注册时长

- 2-3个月
- 4-6个月
- 7-12个月
- >12个月
- 新用户

宁波



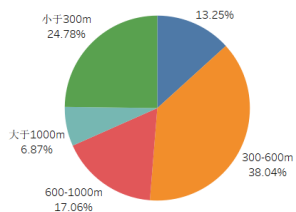
苏州



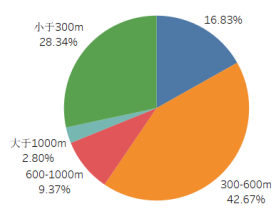
• 店铺规模

根据店铺规模对用户进行分类，可以将用户分成小于300m、300-600m、600-1000m、大于1000m四个等级。可以看到300-600m规模的用户最多，约占35%-40%不等，小于300m规模的用户次之，约占26%左右。

杭州



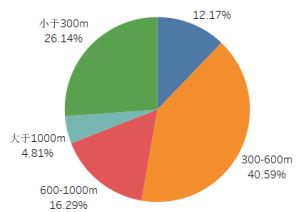
武汉



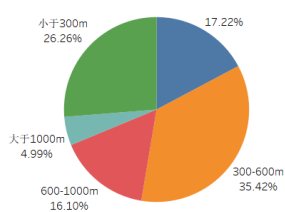
店铺规模

- Null
- 300-600m
- 600-1000m
- 大于1000m
- 小于300m

宁波



苏州



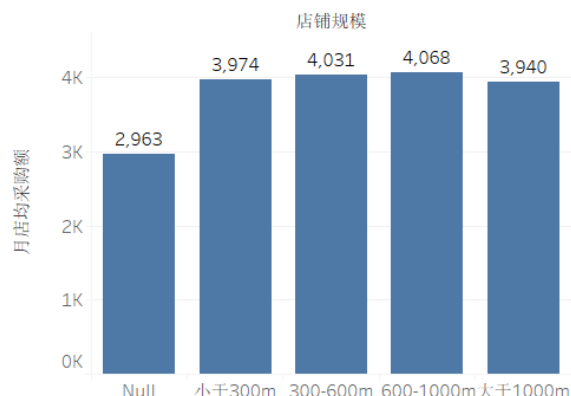
四、各城市各类型用户指标对比

数据选取自2019年8月份的用户交易相关数据。

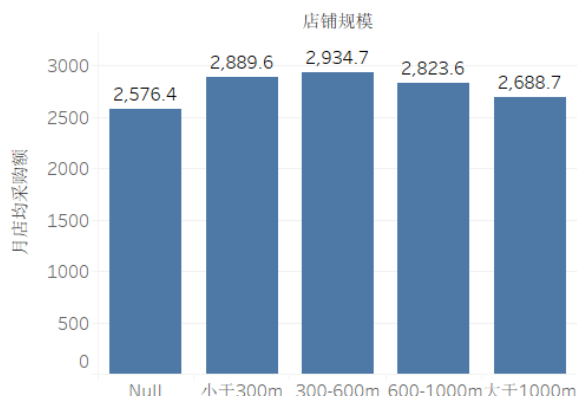
1、不同店铺规模下的月店均采购额对比

除了苏州以外，其他城市不同店铺规模的月店均采购额基本上差不多，且店铺规模大的店铺，月店均采购额反而小。建议后期可以对店铺规模大的店铺投入更大的关注，或许可以带来不错的收益。

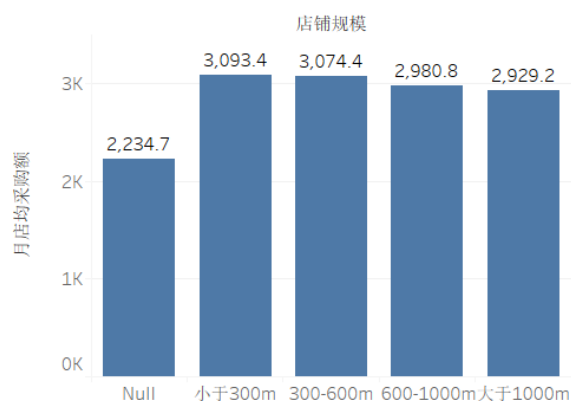
杭州



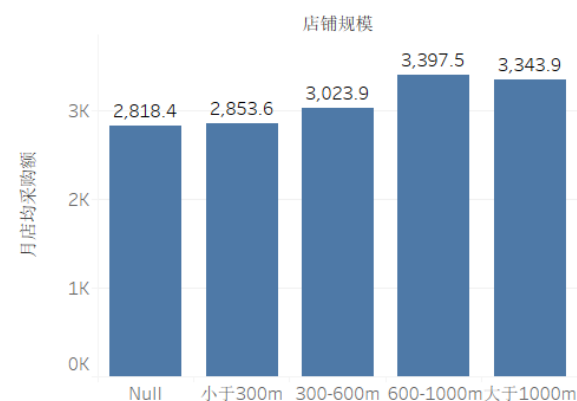
武汉



宁波



苏州



2、不同注册时长下的指标对比

杭州、武汉和苏州统计数据显示，注册时长在6个月以上的老用户对于核心指标贡献更大，忠诚度最高，注册时长在7-12个月的用户留存率都较低，可出推出相应的激活手段。

- 1) 杭州：新用户注册不多，留存率较高，8月份77%的GMV是由注册6个月以上的老用户贡献，新老用户对于核心指标的贡献差异明显，老用户的月均采购额、下单频次和采购SKU数都明显高于新用户；
- 2) 武汉：新用户留存率最低，注册半年内用户下单占比高于杭州和宁波，各类注册时长用户的平均采购频次都略低于其他三城；
- 3) 宁波：各类注册时长用户的留存率差异不大，老用户的月均采购额差异相对较小；
- 4) 苏州：基本符合城市历史规律，注册时间越长的用户对于指标的贡献更大，各类型用户月均采购额、下单频次和平均SKU数都高于武汉和宁波，发展势头很好。

杭州

	新用户	注册时长			
		2-3个月	4-6个月	7-12个月	>12个月
注册用户数	517	829	1,564	2,567	11,308
下单用户数	380	522	780	510	3,762
留存率	73.50%	62.97%	49.87%	19.87%	33.27%
8月份GMV占比	3.17%	7.32%	12.63%	7.72%	69.15%
平均月采购额	1,949	3,269	3,778	3,533	4,287
平均下单频次	2.7	4.4	5.1	4.8	5.9
平均sku数	21.7	32.9	35.3	33.2	38.7

武汉

	新用户	注册时长			
		2-3个月	4-6个月	7-12个月	>12个月
注册用户数	1,033	1,488	1,722	2,041	7,998
下单用户数	572	762	752	372	2,237
留存率	55.37%	51.21%	43.67%	18.23%	27.97%
8月份GMV占比	6.46%	15.59%	15.79%	7.73%	54.42%
平均月采购额	1,526	2,761	2,835	2,806	3,284
平均下单频次	2.3	4.0	4.1	4.1	4.8
平均sku数	17.1	26.2	27.9	26.6	29.2

宁波

	新用户	注册时长			
		2-3个月	4-6个月	7-12个月	>12个月
注册用户数	432	427	955	1,951	2,603
下单用户数	257	231	421	729	1,359
留存率	59.49%	54.10%	44.08%	37.37%	52.21%
8月份GMV占比	4.16%	7.54%	13.57%	25.51%	49.23%
平均月采购额	1,454	2,933	2,897	3,145	3,255
平均下单频次	2.3	4.5	4.4	4.7	4.9
平均sku数	14.9	27.1	24.9	25.9	26.9

苏州

	新用户	注册时长	
		2-3个月	4-6个月
注册用户数	594	2,243	806
下单用户数	385	1,138	423
留存率	64.81%	50.74%	52.48%
8月份GMV占比	13.28%	60.45%	26.27%
平均月采购额	2,035	3,135	3,665
平均下单频次	3.1	4.5	5.2
平均sku数	18.9	25.9	30.6

五、用户偏好

数据选取自2019年8月份的用户交易相关数据。

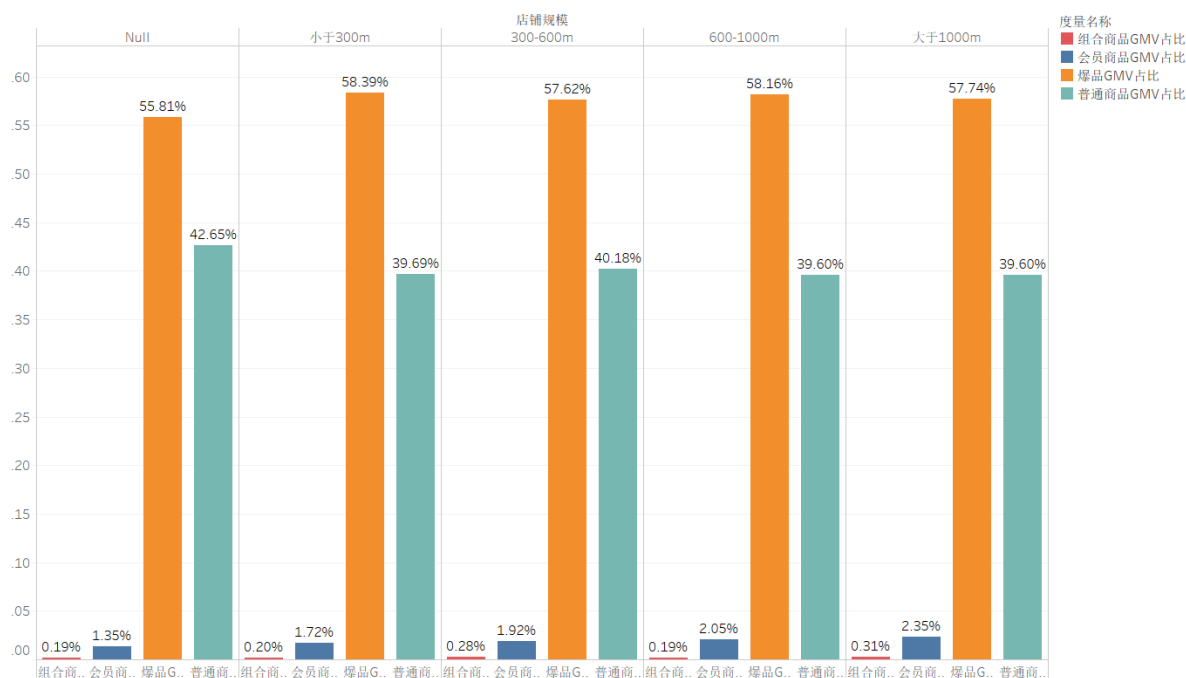
一、各类型用户的活动商品GMV占比

• 店铺规模

不同规模店铺对于活动商品的偏好基本一致，其中苏州的普通商品GMV占比最低在20%左右，武汉的组合商GMV占比最高，杭州的普通商品GMV占比最高，会员商品GMV占比高于组合商品。

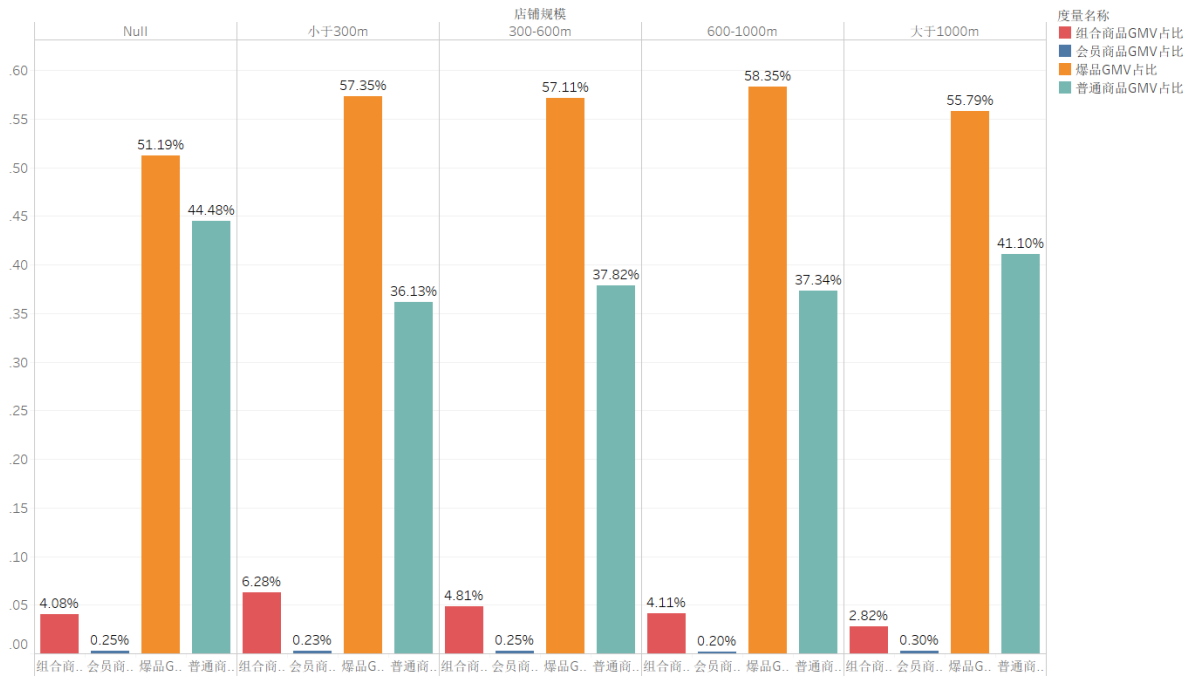
1) 杭州：不同规模店铺对于活动的偏好差异不大，但从下图可以看出，随着店铺规模增大会员商品GMV占比略有升高；

杭州



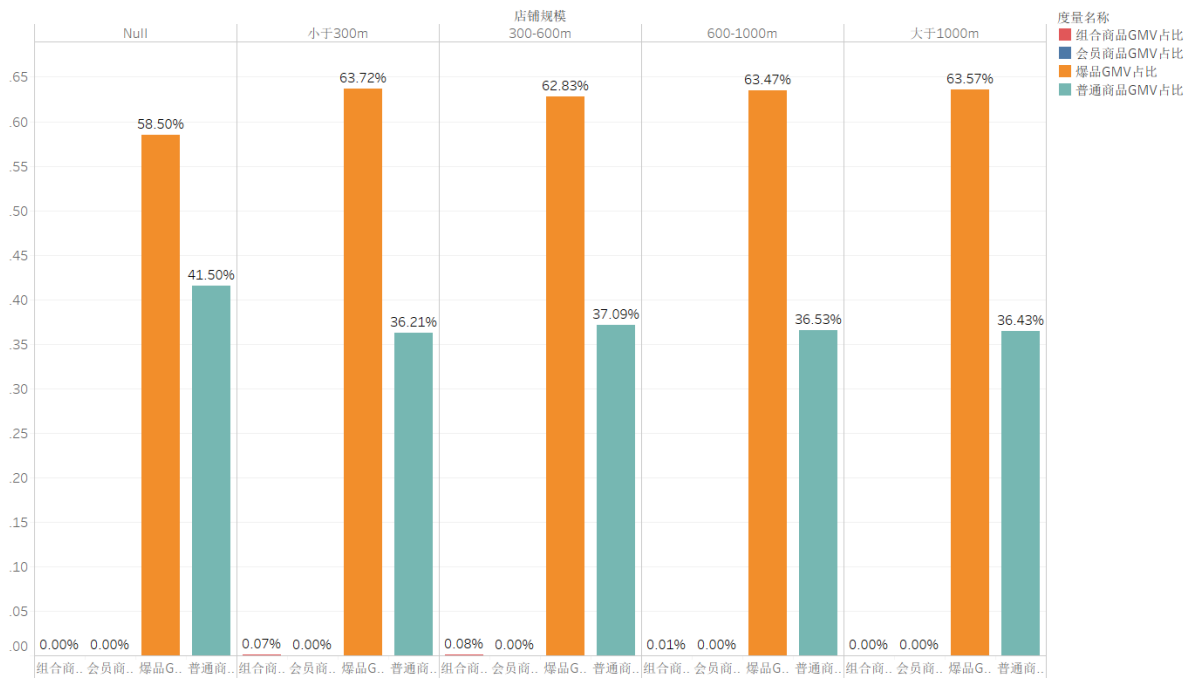
2) 武汉：随着店铺规模增大组合商品的GMV占比下降；

武汉



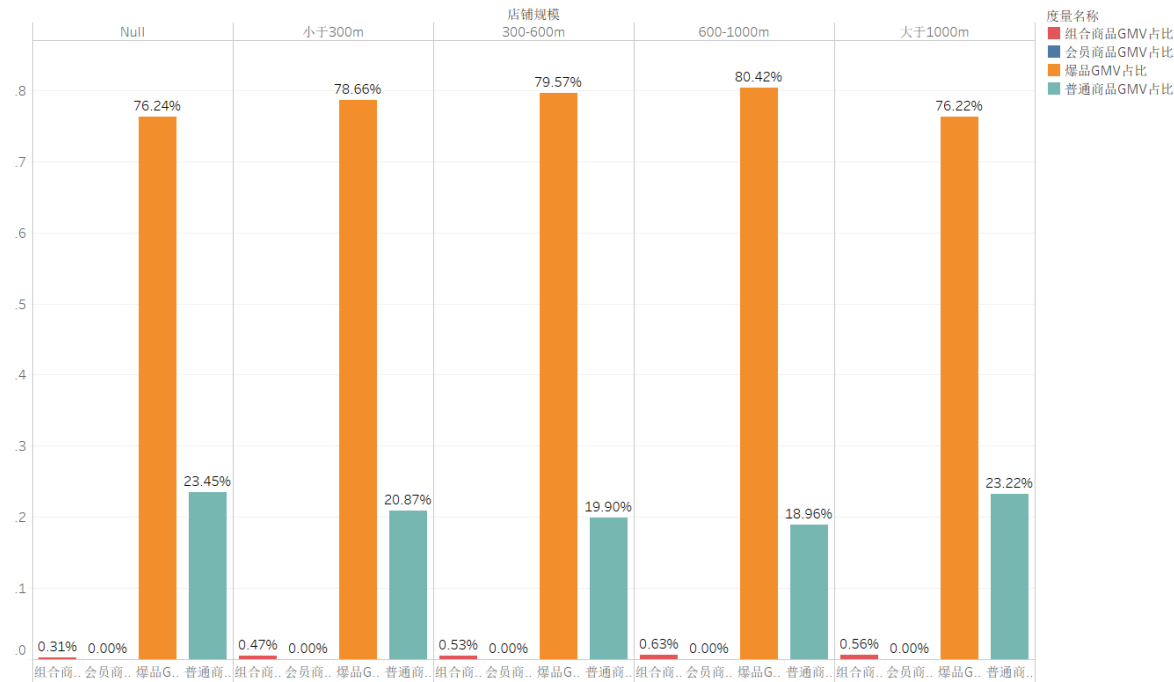
3) 宁波：各规模店铺对于活动的偏好差异不大，63%左右都是爆品，可能是因为宁波其他活动商品较少或无明显吸引力；

宁波



4) 苏州：爆品GMV占比高达75%以上，各类用户对于不同活动类型偏好差异不明显，主要还是受城市自身运营的活动类型影响。

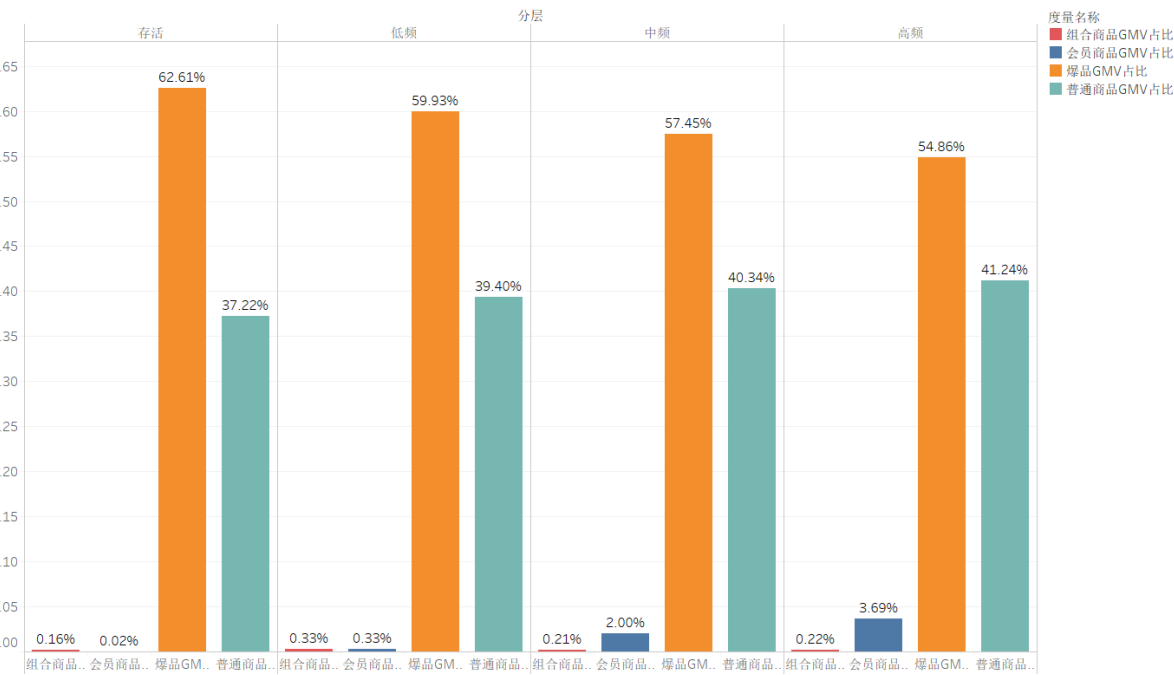
苏州



• 用户分层

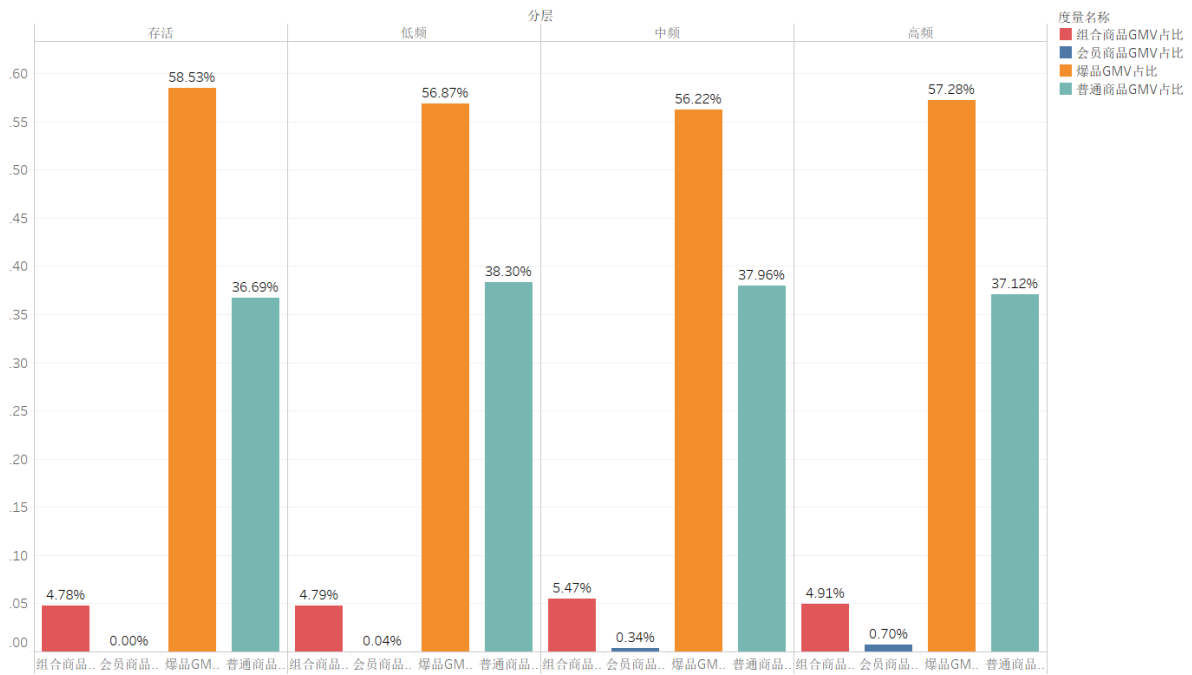
1) 杭州：中高频用户会员商品的GMV占比最高，随着用户下单频次的增加可享受更多会员优惠，故爆品GMV占比越小，各类型用户对于组合商品的偏好差异不大；

杭州



2) 武汉：各分层用户的活动偏好稳定几乎无差异；

武汉

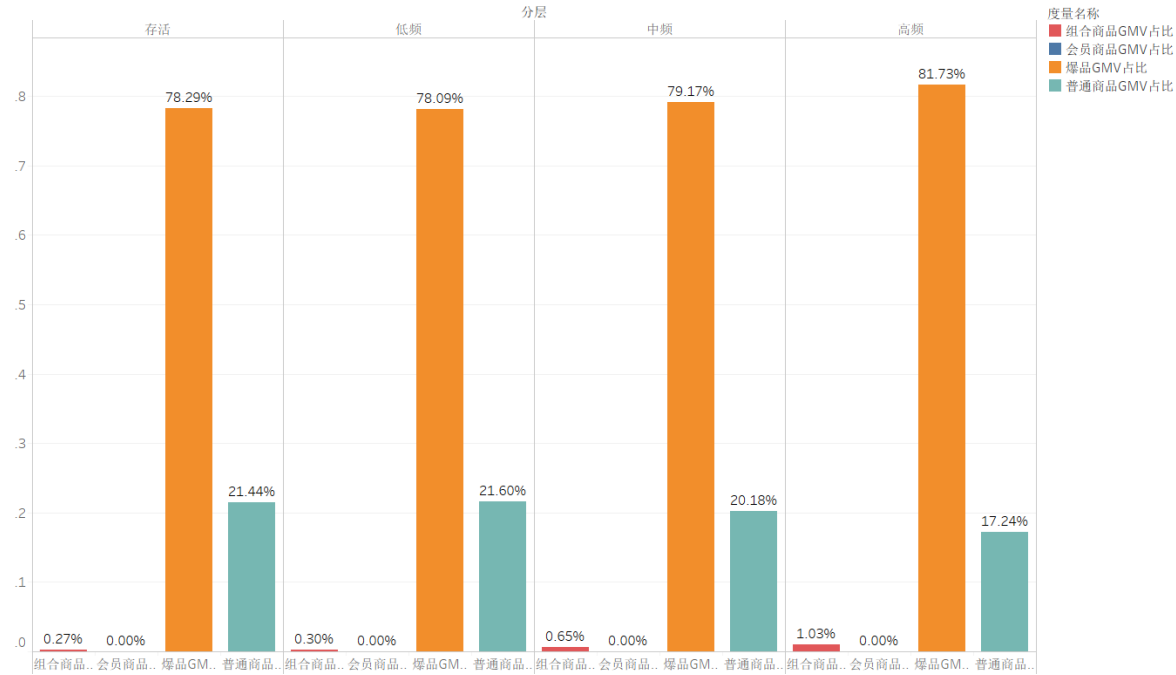


3) 宁波：存活用户爆品GMV占比最低，高中低频用户对于爆品的依赖相对更强；

宁波



4) 苏州：高频用户的普通商品GMV占比最低仅为17%，苏州用户对于活动商品的依赖很高。



二、分层用户红包使用情况

数据选自8月份用户下单红包使用情况。

红包使用率 = 本月使用的红包 / 本月发放的红包，由于赔付红包使用期限较长，存在本月使用历史发放的情况较多，故存在使用率大于100%的情况；系统发放红包包含很多红包种类，发放数量最大使用率偏低，使用该类红包的订单GMV占比最多，一般用户下单频次越高各类红包的使用率越高，费率越高。

1、杭州：随着用户下单频次增多，各类型红包的使用率都升高，拼福利红包使用率升高且红包费率升高，高频用户的拼福利红包费率高达2.11%，明显超出红包平均费率；

杭州

分层		红包分类					
		红包口令兑换	会员红包	赔付红包	拼福利红包	签到红包	系统发放
存活	发放数量	1	21	58	220	1,922	99,766
	使用数量		4	43	12	104	2,712
	使用人数		4	41	12	97	2,212
	使用率		19.05%	74.14%	5.45%	5.41%	2.72%
	GMV占比		0.15%	1.42%	0.61%	3.16%	86.92%
	红包使用金额占比		0.08%	0.81%	0.28%	1.63%	88.02%
	红包费率		0.75%	0.83%	0.67%	0.75%	1.47%
低频	发放数量		358	175	586	6,259	65,885
	使用数量		87	278	109	727	4,595
	使用人数		77	243	104	453	1,558
	使用率		24.30%	158.86%	18.60%	11.62%	6.97%
	GMV占比		1.68%	4.12%	2.12%	9.89%	79.64%
	红包使用金额占比		1.32%	2.78%	1.58%	7.80%	82.87%
	红包费率		0.91%	0.78%	0.86%	0.91%	1.20%
中频	发放数量		1,701	309	2,058	14,058	58,555
	使用数量		608	504	493	2,851	7,829
	使用人数		373	388	397	846	1,390
	使用率		35.74%	163.11%	23.96%	20.28%	13.37%
	GMV占比		5.65%	3.30%	3.82%	18.48%	68.00%
	红包使用金额占比		5.34%	2.43%	5.99%	17.04%	68.26%
	红包费率		1.13%	0.89%	1.88%	1.11%	1.21%
高频	发放数量		1,369	224	1,188	6,740	18,468
	使用数量		743	365	339	2,070	4,155
	使用人数		224	195	246	357	429
	使用率		54.27%	162.95%	28.54%	30.71%	22.50%
	GMV占比		11.25%	3.68%	4.16%	21.90%	58.92%
	红包使用金额占比		11.57%	2.71%	6.93%	19.67%	59.02%
	红包费率		1.30%	0.93%	2.11%	1.14%	1.27%

2、武汉：口令红包费率较高，但发放数量不多，其他红包类型的红包费率差异不大，拼福利红包的使用率很低，签到红包的使用率较其他城市略高，新人红包使用率在四城中最高；

武汉

		红包分类						
分层		红包口令兑换	会员红包	赔付红包	拼福利红包	签到红包	系统发放	新人礼包领取
存活	发放数量	28	16	11	332	2,243	66,861	2,666
	使用数量	16	4	32	2	517	1,533	595
	使用人数	16	4	31	2	440	1,329	460
	使用率	57.14%	25.00%	290.91%	0.60%	23.05%	2.29%	22.32%
	GMV占比	0.61%	0.19%	1.01%	0.08%	16.64%	60.25%	21.22%
	红包使用金额占比	0.93%	0.11%	0.62%	0.05%	9.11%	64.89%	24.30%
	红包费率	2.39%	0.91%	0.97%	1.00%	0.86%	1.70%	1.81%
低频	发放数量	78	219	44	870	6,194	45,682	583
	使用数量	60	66	65	32	1,842	2,474	547
	使用人数	59	58	62	30	823	1,135	186
	使用率	76.92%	30.14%	147.73%	3.68%	29.74%	5.42%	93.83%
	GMV占比	1.05%	1.40%	1.14%	0.68%	29.84%	55.92%	9.98%
	红包使用金额占比	2.08%	1.05%	0.93%	0.45%	23.01%	59.08%	13.40%
	红包费率	2.66%	1.01%	1.10%	0.90%	1.04%	1.42%	1.81%
中频	发放数量	104	766	45	1,183	8,449	29,140	101
	使用数量	82	286	65	53	3,847	2,844	230
	使用人数	78	194	61	50	730	831	56
	使用率	78.85%	37.34%	144.44%	4.48%	45.53%	9.76%	227.72%
	GMV占比	0.95%	4.31%	0.76%	0.81%	43.62%	46.52%	3.04%
	红包使用金额占比	2.03%	3.98%	0.73%	0.56%	39.03%	49.63%	4.04%
	红包费率	2.89%	1.24%	1.30%	0.93%	1.21%	1.44%	1.79%
高频	发放数量	29	437	22	315	1,999	4,662	
	使用数量	26	239	25	24	1,371	723	
	使用人数	26	91	21	20	137	142	
	使用率	89.66%	54.69%	113.64%	7.62%	68.58%	15.51%	
	GMV占比	0.97%	11.56%	0.84%	1.11%	47.78%	37.74%	
	红包使用金额占比	2.04%	12.02%	0.85%	0.79%	47.07%	37.22%	
	红包费率	2.78%	1.38%	1.34%	0.94%	1.30%	1.31%	

3、宁波：高频用户的拼福利红包费率高达2.15%；

宁波

分层		红包分类			系统发放	新人礼包领取
		赔付红包	拼福利红包	签到红包		
存活	发放数量	67	514	2,231	61,155	1,076
	使用数量	12	15	148	1,246	174
	使用人数	12	15	142	980	145
	使用率	17.91%	2.92%	6.63%	2.04%	16.17%
	GMV占比	0.66%	0.82%	7.84%	80.65%	10.03%
	红包使用金额占比	0.41%	0.68%	5.25%	81.41%	12.25%
	红包费率	0.93%	1.24%	1.00%	1.51%	1.83%
低频	发放数量	120	1,654	6,773	61,618	201
	使用数量	78	160	1,142	2,970	135
	使用人数	77	148	682	1,098	70
	使用率	65.00%	9.67%	16.86%	4.82%	67.16%
	GMV占比	1.48%	3.15%	21.48%	71.19%	2.70%
	红包使用金额占比	1.17%	3.01%	17.67%	74.33%	3.82%
	红包费率	0.99%	1.20%	1.04%	1.32%	1.78%
中频	发放数量	188	1,848	6,425	31,415	31
	使用数量	184	355	1,773	2,848	51
	使用人数	157	245	537	601	26
	使用率	97.87%	19.21%	27.60%	9.07%	164.52%
	GMV占比	3.00%	6.04%	29.70%	60.41%	0.85%
	红包使用金额占比	2.43%	7.25%	26.37%	62.73%	1.22%
	红包费率	1.00%	1.48%	1.09%	1.28%	1.75%
高频	发放数量	36	660	1,320	4,200	1
	使用数量	49	259	529	612	1
	使用人数	41	67	83	84	1
	使用率	136.11%	39.24%	40.08%	14.57%	100.00%
	GMV占比	2.70%	16.16%	33.28%	47.80%	0.06%
	红包使用金额占比	1.91%	25.48%	28.03%	44.49%	0.08%
	红包费率	0.97%	2.15%	1.15%	1.27%	2.04%

4、苏州：拼福利红包的使用高费率较其他城市偏低，各类红包的使用率和费率相对稳定。

苏州

		红包分类					
分层		红包口令兑换	赔付红包	拼福利红包	签到红包	系统发放	新人礼包领取
存活	发放数量	8	7	154	1,312	64,020	1,284
	使用数量	4	1	5	59	730	242
	使用人数	4	1	5	49	554	211
	使用率	50.00%	14.29%	3.25%	4.50%	1.14%	18.85%
	GMV占比	0.73%	0.08%	0.75%	4.71%	72.12%	21.61%
	红包使用金额占比	0.76%	0.08%	0.66%	3.65%	70.00%	24.84%
	红包费率	1.61%	1.57%	1.37%	1.20%	1.50%	1.77%
低频	发放数量	25	14	545	3,518	68,578	718
	使用数量	11	7	50	426	2,220	407
	使用人数	11	7	43	255	741	200
	使用率	44.00%	50.00%	9.17%	12.11%	3.24%	56.69%
	GMV占比	0.31%	0.18%	1.95%	10.84%	75.48%	11.25%
	红包使用金额占比	0.35%	0.18%	1.75%	10.13%	72.69%	14.90%
	红包费率	1.62%	1.45%	1.28%	1.34%	1.38%	1.89%
中频	发放数量	30	21	817	2,698	30,007	137
	使用数量	18	15	203	779	1,905	150
	使用人数	18	12	110	223	346	53
	使用率	60.00%	71.43%	24.85%	28.87%	6.35%	109.49%
	GMV占比	0.54%	0.41%	6.68%	20.63%	67.48%	4.27%
	红包使用金额占比	0.60%	0.45%	6.58%	19.76%	66.67%	5.92%
	红包费率	1.53%	1.54%	1.36%	1.32%	1.36%	1.91%
高频	发放数量	6	2	385	536	3,658	4
	使用数量	5	1	195	256	366	4
	使用人数	5	1	35	41	45	1
	使用率	83.33%	50.00%	50.65%	47.76%	10.01%	100.00%
	GMV占比	0.59%	0.10%	21.25%	25.21%	52.20%	0.66%
	红包使用金额占比	0.63%	0.13%	22.61%	24.92%	51.14%	0.58%
	红包费率	1.51%	1.82%	1.51%	1.40%	1.39%	1.24%

六、用户留存率

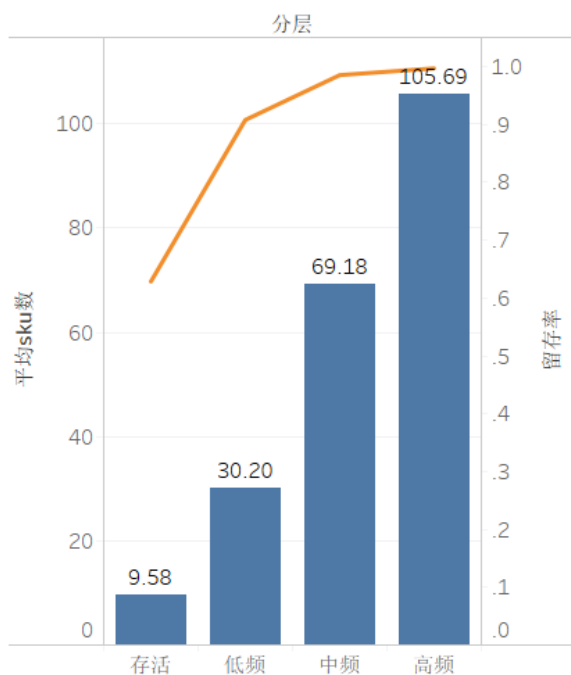
数据选自2019年07月份的所有下单用户在2019年08月份的留存情况。

从下图可以看出：

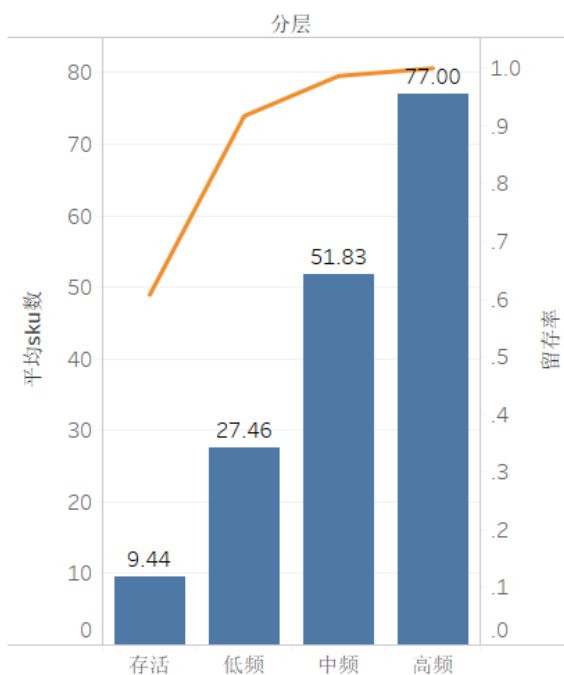
- 1、留存率：高频>中频>低频>存活,说明下单频次多的用户，越不容易流失；
- 2、随着平均交易sku数的增大，留存率也越来越来大；

由此可以看出留存率跟用户下单频次和交易sku数有很大关系，下单频次多和交易sku数多的用户，对我们平台越忠诚，越不容易流失。

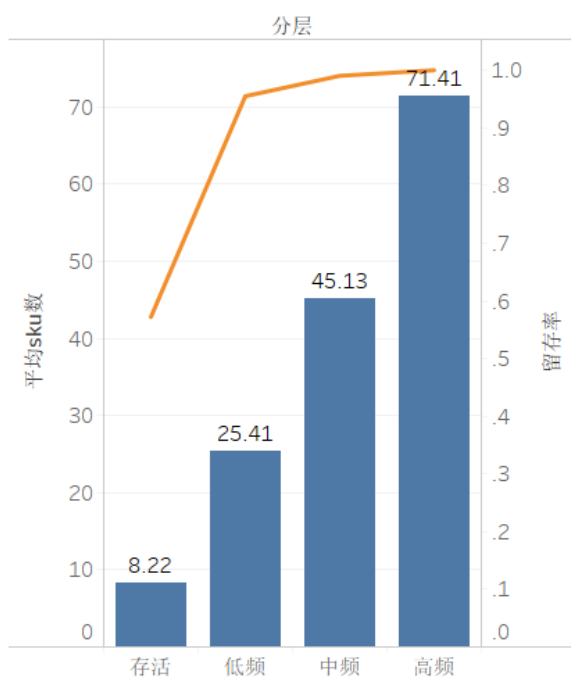
各分层留存率-杭州



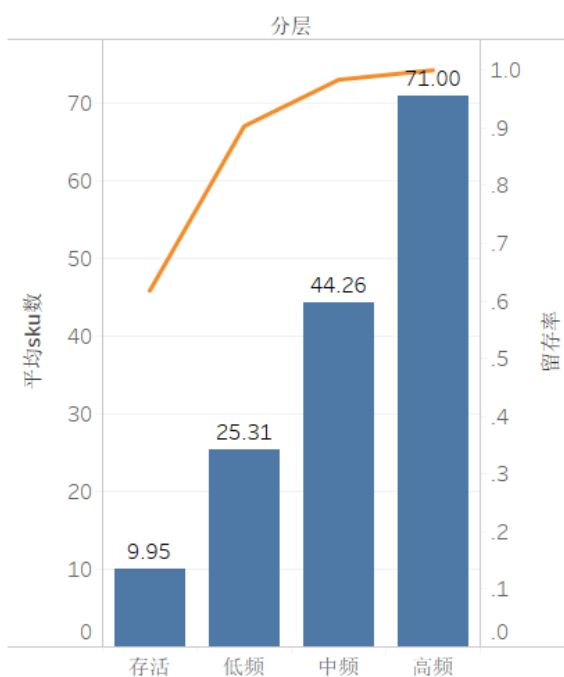
各分层留存率-武汉



各分层留存率-宁波



各分层留存率-苏州

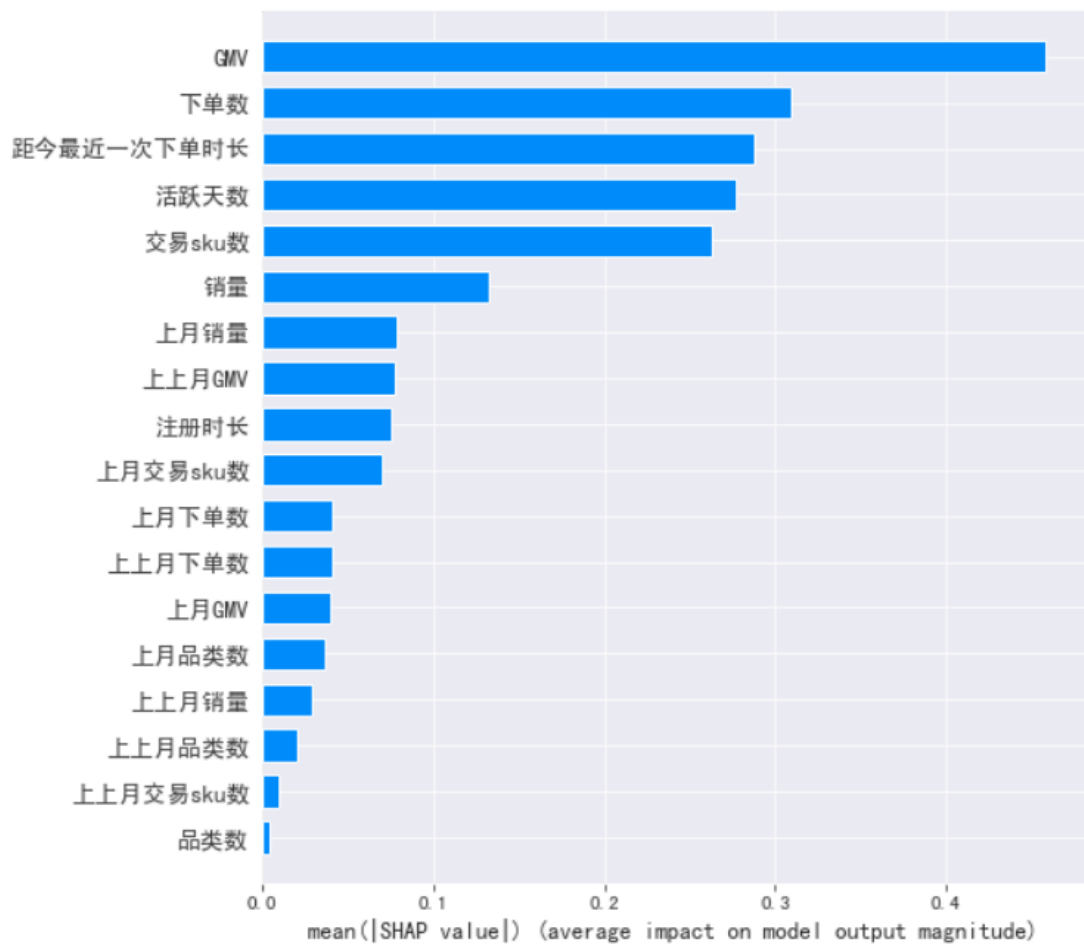


七、用户流失预测模型

定义：基于当月的下单用户，根据一些指标特征，去预测这些用户在未来30天内的流失（未下单）情况。

1、特征重要性

根据用户流失预测模型，可以看出影响用户流失（或留存）最重要的特征是GMV、下单数、距今最近一次下单时长、活跃天数、交易sku数等。

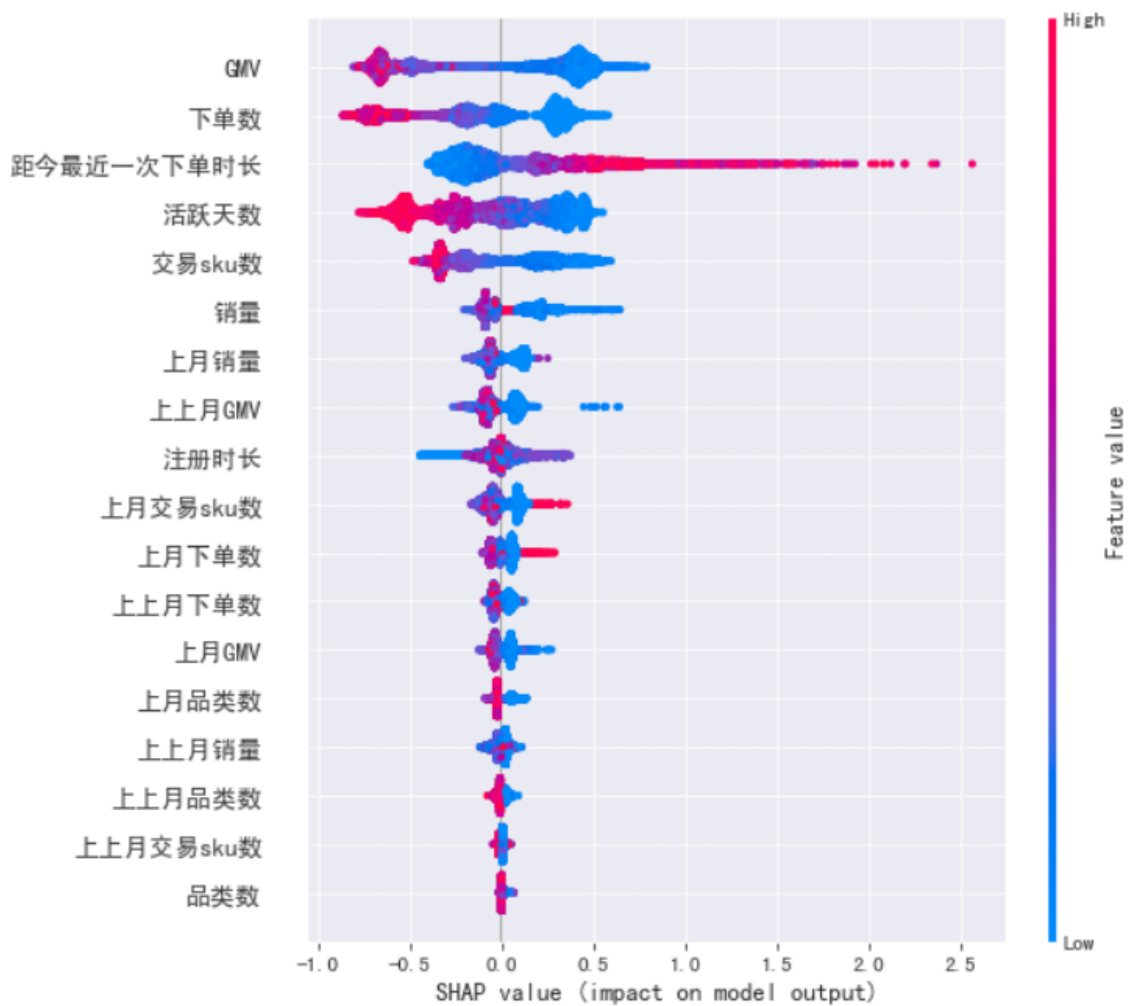


2、SHAP_value

SHAP_value最大的优势是能反映出特征的影响力，而且还表现出影响的正负性。

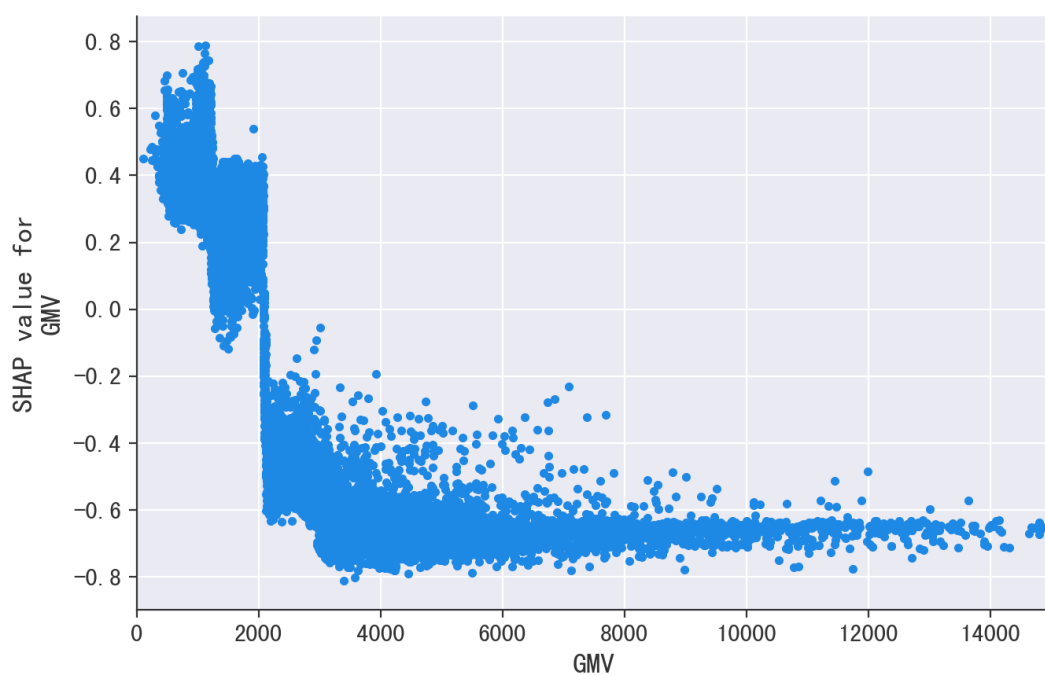
下图中每一行代表一个特征，横坐标为SHAP_value。SHAP_value>0,表示对流失有正向影响力，SHAP_value<0,表示对流失有负向影响力。一个点代表一个样本，颜色越红说明特征本身数值越大，颜色越蓝说明特征本身数值越小。

我们可以直观地看出GMV是一个很重要的特征，GMV越小，会给流失带来正向影响力。距今最近一次下单时长也会影响流失，红色点主要集中在SHAP_value大于0的区域，可见距今最近一次下单时长越长越容易流失。

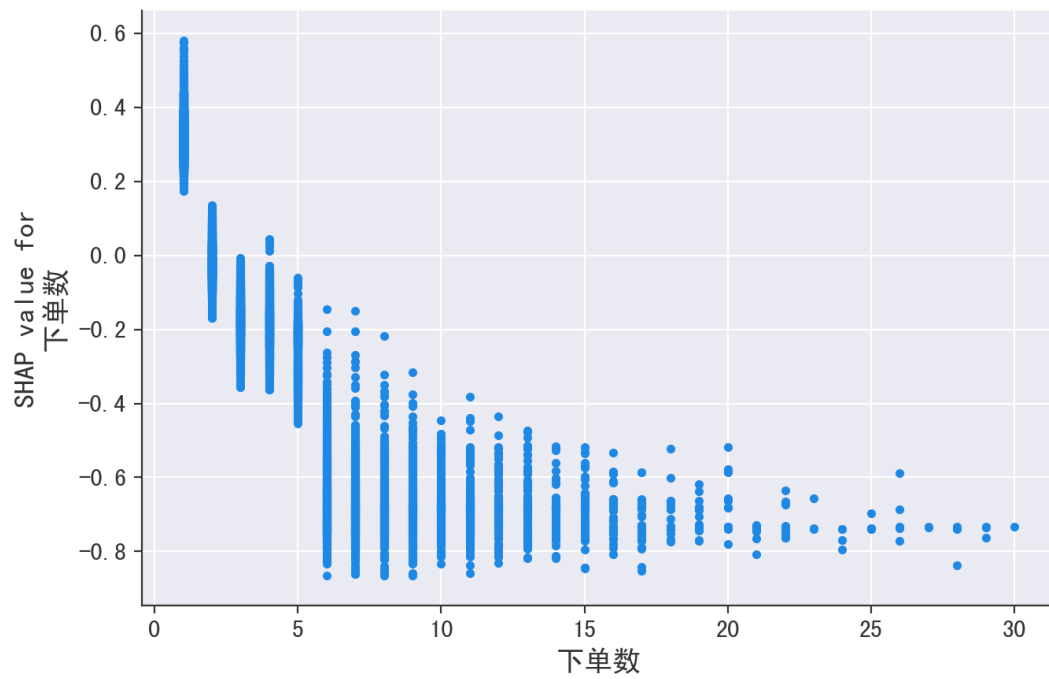


以下绘制了特征值对流失的影响程度。横坐标是特征值，纵坐标是shap_value值，值越大，影响力越大。

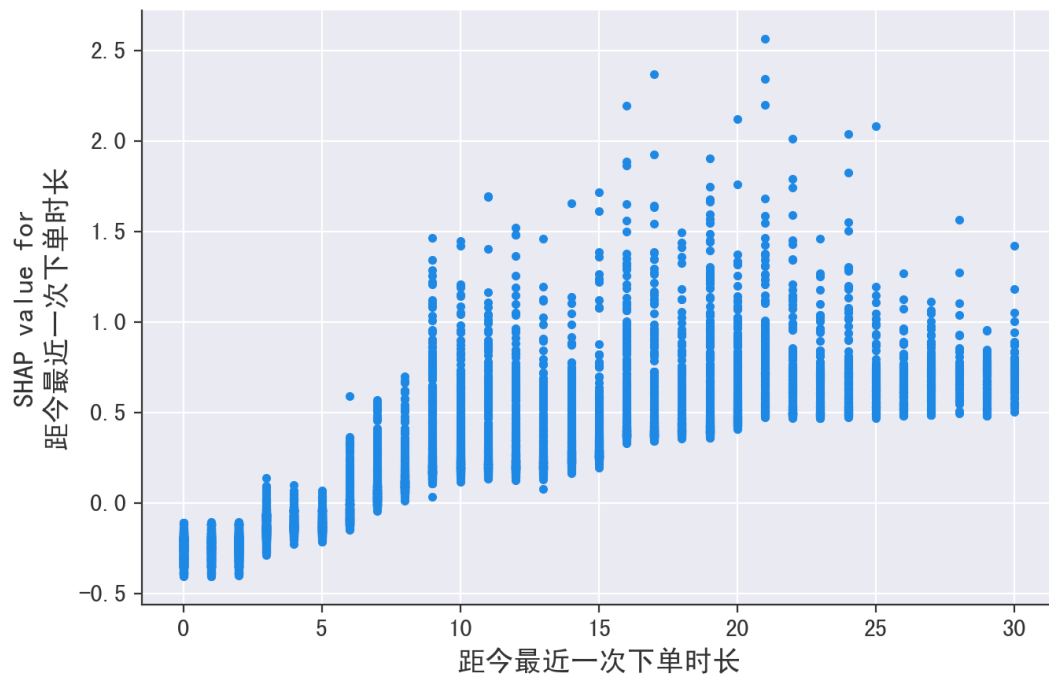
1、当GMV<2000时，shap_value>0,说明当GMV<2000时，对流失有着正向的影响力。



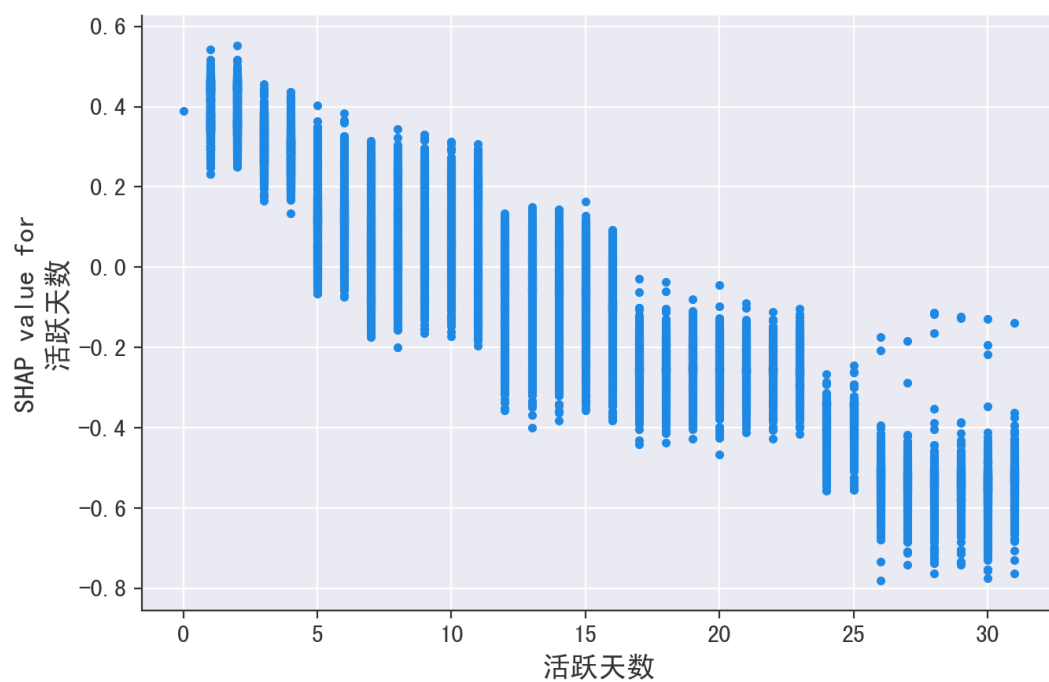
2、当下单数<=2时，shap_value>0,说明当下单数<=2时，对流失有着正向的影响力。



3、当距今最近一次下单时长>5时，shap_value>0,说明当距今最近一次下单时长>5时，对流失有着正向的影响力。



4、当活跃天数<=5时，shap_value>0,说明当活跃天数<=5时，对流失有着正向的影响力。



5、当交易sku数 ≤ 15 时， $\text{shap_value} > 0$ ，说明当交易sku数 ≤ 15 时，对流失有着正向的影响力。

