

参赛编号：	DAC23000045
赛题题号：	A

用户行为价值分析与挖掘

摘 要

随着互联网的快速发展，线下实体店铺也逐渐在线上进行销售。电商是当今用户最大的交易市场之一，电商行业也逐渐成熟，并且在网络和疫情的影响下，在线上的消费行为满足全年龄段用户。有效的分析用户行为可以帮助了解用户的习惯和喜好，做出精准的营销策略促进用户的后续消费。

用户行文分析结果的准确大部分来自于数据的准确，数据清洗是重要步骤。采用 Python 编程语言，对数据的缺失值，重复值，异常值等噪声数据进行清洗并做出可视化。通过数据特征之间的关系创建新特征来更好的表示用户行为，为后续建立分类模型做好数据基础。

对用户是否采取特定措施之前，需要分析用户的习惯，爱好和特定的外在特征。通过用户消费的时间维度，地区维度，消费能力维度等多维度分析用户的购买习惯并做出相应的可视化。为了进一步获取用户在店铺的视角下是哪一类客户，构建 RFM 模型对用户进行客户价值分析，并对用户进行了用户画像分析，得到最终的用户表示结果可视化。

平台推出不同商家的代金券，是否为用户发放商家代金券也成为一个重要操作，将用户使用代金券作为正样本，未领取和未使用作为负样本，由于正负样本差距过大，采取将负样本复制的方法来保持平衡。然后通过所得出的特征指标进行建模，建立朴素贝叶斯，决策树，逻辑回归，K 近邻模型，结果令人满意并根据结果给出最终的报告。

关键词：用户行为分析，RFM 模型，用户画像，朴素贝叶斯，决策树，逻辑回归

目 录

一. 背景与挖掘目标.....1

二. 特征工程.....1

 2.1 数据预处理1

 2.1.1 文件读取1

 2.1.2 文件数据预处理2

三. 用户分析.....4

四. 客户价值分析.....10

 4.1 RFM 模型10

 4.2 用户画像11

五. 用户优惠券发放模型.....13

 1. 数据采样.....13

 2. 模型预测.....13

六. 优惠券发放策略.....15

（在此页开头）

一. 背景与挖掘目标

电商是当今用户最大的交易市场之一，电商行业也逐渐成熟，所有市场中可售卖的商品全都在平台中存在，并且在网络和疫情的影响下，在线上的消费行为满足全年龄段用户。用户的交易行为推动了电商的发展。对于电商平台来说，通过用户的交易行为、交易金额、交易频率、交易时间等信息，分析用户的行为，可促进二次营销，精准营销，促活等。需要挖掘到用户的购买行为所存在的规律，用户的生活习惯，店铺的火热程度。更深层次的分析，要得到用户在店铺角度的价值，分析出用户属于哪一类客户，并作出相应的决策。对用户进行特定的操作，需要根据已有的数据做出预测，通过分类，回归或聚类等机器学习模型，可以得到想要的结果，并根据结果对用户采取行动以便于促进用户消费。

二. 特征工程

2.1 数据预处理

2.1.1 文件读取

首先读取保存到本地的用户数据，并查看数据，如图 1 所示。

	订单ID	商家ID	用户ID	付款日期	使用状态	实付金额	邮费	省份	城市	购买数量
0	S10816	100110	笑嘻嘻	2022-02-17 08:09:10.000	1.0	172	0	江苏省	无锡市	1
1	S19000	103840	芽张	2022-04-19 00:24:09.000	1.0	79	0	上海	上海市	1
2	S13184	100017	2天天向上7403	2022-02-22 16:45:20.000	1.0	124	0	辽宁省	盘锦市	1
3	S17539	100180	9chs05	2022-04-13 09:09:07.000	1.0	46	0	福建省	厦门市	2
4	S26017	102464	ejaneinmecn	2022-06-16 21:27:39.000	1.0	186	0	北京	北京市	1
...
20178	S28637	105725	0amethystjing27	2022-06-29 09:23:03.588	1.0	74	0	上海	上海市	1
20179	S17731	101642	0萝卜咸菜1	2022-04-18 00:03:58.000	1.0	95	0	北京	北京市	2
20180	S28031	105614	0peiya7	2022-06-24 19:08:46.000	1.0	49	0	上海	上海市	1
20181	S15726	100144	0whh101529	2022-03-21 22:25:12.000	1.0	165	0	河南省	郑州市	1
20182	S19967	100017	8用心刺痛太阳17	2022-05-01 11:32:43.000	1.0	150	0	湖北省	十堰市	1

图 1 数据读取展示

2.1.2 文件数据预处理

首先根据 `describe()` 函数查看数据分布，并做出箱型图查看异常数据，如图 2，图 3 所示。

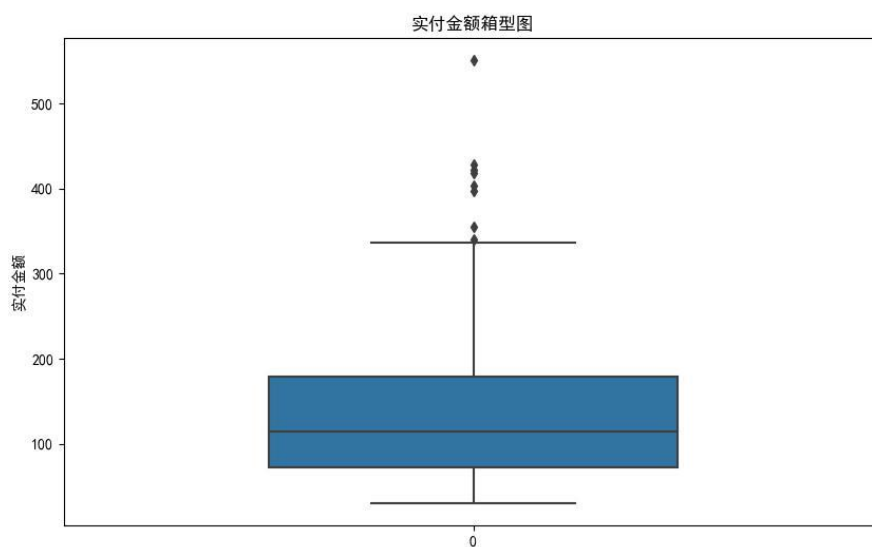


图 2 实付金额箱型图

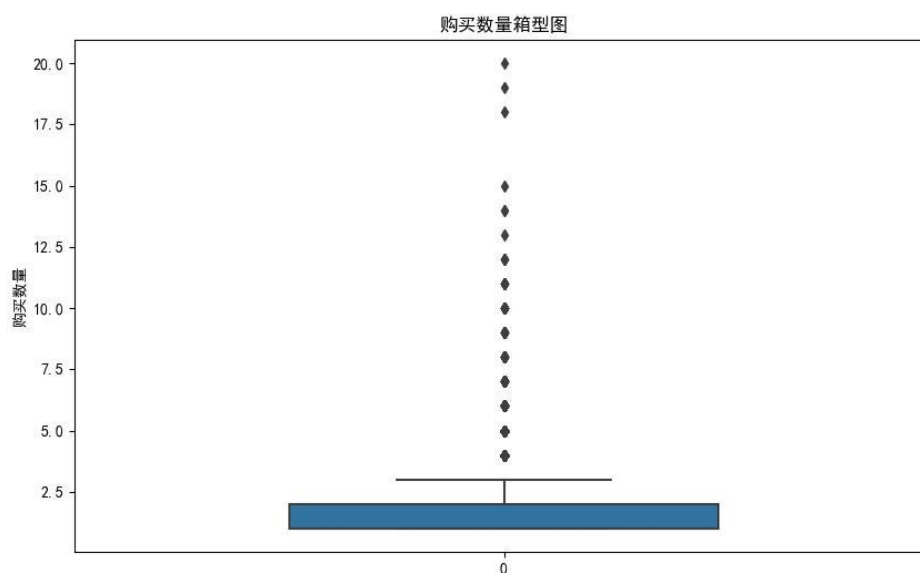


图 3 购买数量箱型图

根据图中结果可以看出在实付金额存在大金额数据，在购买数量中也存在大量购买记录，通过代码查看这些疑似异常值的信息。

查看数据结果如图 4 所示。

	订单ID	商家ID	用户ID	付款日期	使用状态	实付金额	邮费	省份	城市	购买数量
2471	S2091	100158	多没有乌龟多	2022-01-11 10:00:26.000	1.0	260	0	北京	北京市	19
4959	S21328	104273	小kk铺	2022-05-16 08:19:06.000	1.0	155	0	上海	上海市	20
5619	S21327	104272	小kk铺	2022-05-16 08:18:23.000	1.0	155	0	上海	上海市	20
6594	S21325	104270	小kk铺	2022-05-16 08:14:14.000	1.0	427	0	上海	上海市	23
14022	S21322	104268	小kk铺	2022-05-16 08:12:14.000	1.0	422	0	上海	上海市	23
14686	S14913	103027	小kk铺	2022-03-16 10:55:03.000	1.0	415	0	上海	上海市	50
14966	S8300	100123	0jessie3504	2022-01-29 18:45:46.098	1.0	340	0	福建省	福州市	18
15417	S16163	103297	小kk铺	2022-03-27 13:54:50.000	1.0	1028	0	上海	上海市	82
15558	S21329	104274	小kk铺	2022-05-16 08:19:42.000	1.0	155	0	上海	上海市	20
15746	S27052	100021	8刘美丽158251989	2022-06-20 10:53:05.427	1.0	265	0	云南省	红河哈尼族彝族自治州	20
17641	S14936	103033	小kk铺	2022-03-16 16:51:01.000	1.0	1000	0	上海	上海市	50
17831	S14915	103028	小kk铺	2022-03-16 10:59:28.000	1.0	425	0	上海	上海市	50
19183	S21330	104275	小kk铺	2022-05-16 08:21:25.000	1.0	160	0	上海	上海市	20
20126	S21324	104269	小kk铺	2022-05-16 08:13:11.000	1.0	427	0	上海	上海市	23

图 4 异常数据统计图

由图可看出大部分异常数据来自于名为小 kk 铺的用户，明显可以看出很多数据购买时间间隔很接近，所以认为此用户是刷单的行为，将此用户的数据删除。后续，对付款时间是否超过了指定范围，邮费等数据是否出现异常等进行查看，均为正常值。

对数据中的缺失值进行统计，通过 `df.info()` 查看缺失值，结果显示使用状态存在缺失，因为未领取优惠券的用户在此特征中不存在值，为确保有效的统计这些用户，通过 `df.fillna()` 函数将未领取优惠券的用户的使用状态设置为-1。

针对重复值，用户不可能在同一时间购买同种商品，如若出现此类现象，购买数量会做出显示，所以对数据通过 `df.drop_duplicates()` 进行去重复，结果不存在重复值。

构造新特征，将用户的时间特征拆分成月份、天数、小时，构造各个商家优惠券发放比率、用户使用、领取未使用、未领取占比。并且去除了针对上述指标的计数，仅留下了比率，对商家特征建立的公式如下。

$$C_{\text{优惠券发放总数}} = C_{\text{用户使用优惠券}} + C_{\text{用户领取未使用优惠券}} \quad \text{公式 (2.1)}$$

$$R_{\text{优惠券发放使用率}} = \frac{C_{\text{用户使用优惠券}}}{C_{\text{优惠券发放总数}}} \quad \text{公式 (2.2)}$$

$$R_{\text{未领取优惠券占比}} = \frac{C_{\text{未领取优惠券用户}}}{\text{SUM}(C_{\text{未领取优惠券用户}} + C_{\text{优惠券发放总数}})} \quad \text{公式 (2.3)}$$

对用户特征建立的公式如下。

$$C_{\text{单个用户优惠券使用率}} = \frac{C_{\text{单个用户使用优惠券}}}{C_{\text{单个用户购买总数}}} \quad \text{公式 (2.4)}$$

$$C_{\text{单个用户优惠券未使用率}} = \frac{C_{\text{单个用户领取未使用优惠券}}}{C_{\text{单个用户购买总数}}} \quad \text{公式 (2.5)}$$

$$C_{\text{单个用户未领取优惠券占比}} = \frac{C_{\text{单个用户未领取优惠券}}}{C_{\text{单个用户购买总数}}} \quad \text{公式 (2.6)}$$

结果如图 5 所示。

	订单ID	商家ID	用户ID	付款日期	使用状态	实付金额	邮费	省份	城市	购买数量	日期_月	日期_天	日期_小时	优惠券发放使用率	未领取优惠券占比	用户优惠券使用率	用户优惠券未使用率	用户未领取优惠券占比
0	S10816	100110	笑囡佳	2022-02-17 08:09:10.000	1.0	172	0	江苏省	无锡市	1	2	17	8	1.0	0.0	1.0	0.0	0.0
1	S12525	100244	笑囡佳	2022-02-19 00:07:55.000	1.0	192	0	江苏省	苏州市	1	2	19	0	1.0	0.0	1.0	0.0	0.0
2	S26273	100286	笑囡佳	2022-06-17 09:50:14.078	1.0	42	0	江苏省	无锡市	3	6	17	9	1.0	0.0	1.0	0.0	0.0
3	S4612	100110	ynej7	2022-01-14 16:44:10.613	1.0	253	0	浙江省	杭州市	1	1	14	16	1.0	0.0	1.0	0.0	0.0
4	S13542	100110	加breez菲	2022-02-28 11:49:48.000	1.0	138	0	黑龙江省	哈尔滨市	2	2	28	11	1.0	0.0	1.0	0.0	0.0
...
20165	S17748	103616	ooh笑h	2022-04-18 00:04:50.000	1.0	87	0	上海市	上海市	1	4	18	0	1.0	0.0	1.0	0.0	0.0
20166	S22753	104547	2tb66429	2022-05-27 17:25:57.000	1.0	99	0	上海市	上海市	1	5	27	17	1.0	0.0	1.0	0.0	0.0
20167	S20528	104127	0duckduck17	2022-05-08 13:58:20.000	1.0	123	0	上海市	上海市	3	5	8	13	1.0	0.0	1.0	0.0	0.0
20168	S28637	105725	0amethystjing27	2022-06-29 09:23:03.588	1.0	74	0	上海市	上海市	1	6	29	9	1.0	0.0	1.0	0.0	0.0
20169	S28031	105614	0peiya7	2022-06-24 19:08:46.000	1.0	49	0	上海市	上海市	1	6	24	19	1.0	0.0	1.0	0.0	0.0

图 5 构造特征展示图

三. 用户分析

为探究用户在时间上的消费习惯，对用户从三个不同的时间维度做出了分析。由图中数据显示可以看出，1、2、6 月份销售额明显高于其他三个月，特别是 1 月份。对此现象，考虑 1、2 月份是春节期间，所以用户的需求会大大增加，6 月份销售额增加，可能是由于电商活动 6.18 所导致，用户可以通过低价格购入想购买的商品。针对此三个月，商家或电商 APP 可以推出活动、加大优惠力度、推出新商品等。具体如图 6 所示。

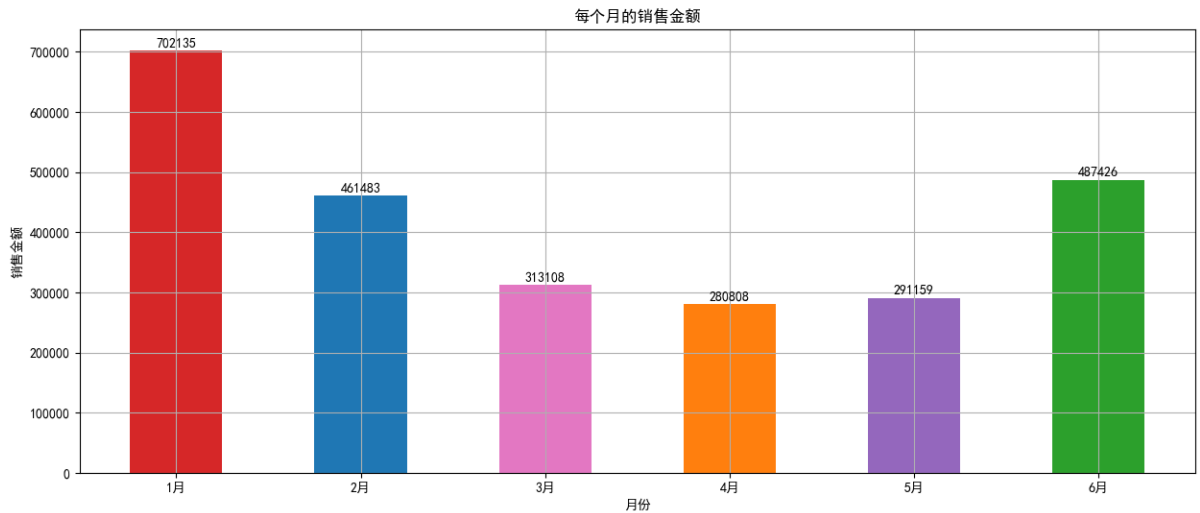


图 6 用户每月习惯

其次是用户在每天的习惯,通过图中呈现结果可知,用户在每个月的 11 号, 16 号-20 号, 消费水平有明显的增加。首先, 针对 11 号, 认为在 6 个月的数据中, 11 号会有鼓励消费因素, 因为在 12-15 号消费水平又回归正常水平, 所以 11 号可以不认为是用户的月份消费习惯。在 16 号-20 号, 可能由于用户在 15 号进行薪资发放, 用户会在发放后的 5 天进行消费, 这是用户的消费习惯。具体如图 7 所示。

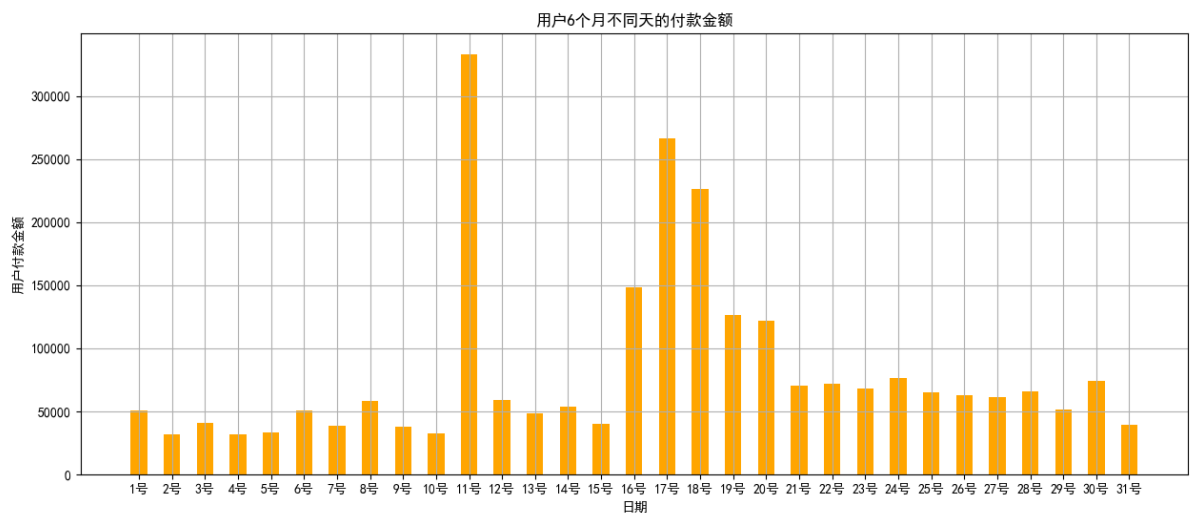


图 7 用户每天习惯

最后, 用户在每天 24 小时的消费习惯, 根据用户在每天 24 小时的消费习惯来看, 部分用户会在 0 点后选择睡眠, 部分会选择继续购物直至 1 点。直到第二天早上 6 点后呈现上升趋势, 并且在 10 点达到新的高峰期, 推测这是因为商品在抢购的时候是在第二天 10 点开始, 所以此时间段消费额达到顶峰。10 点中后回归正常值, 并且在 19 点后呈现上升趋势。根据用户的小时习惯来看, 我们可以重点将活动放在 19 点或 20 点开始, 持续 2-4 个小时的热卖时间。具体如图 8 所示。

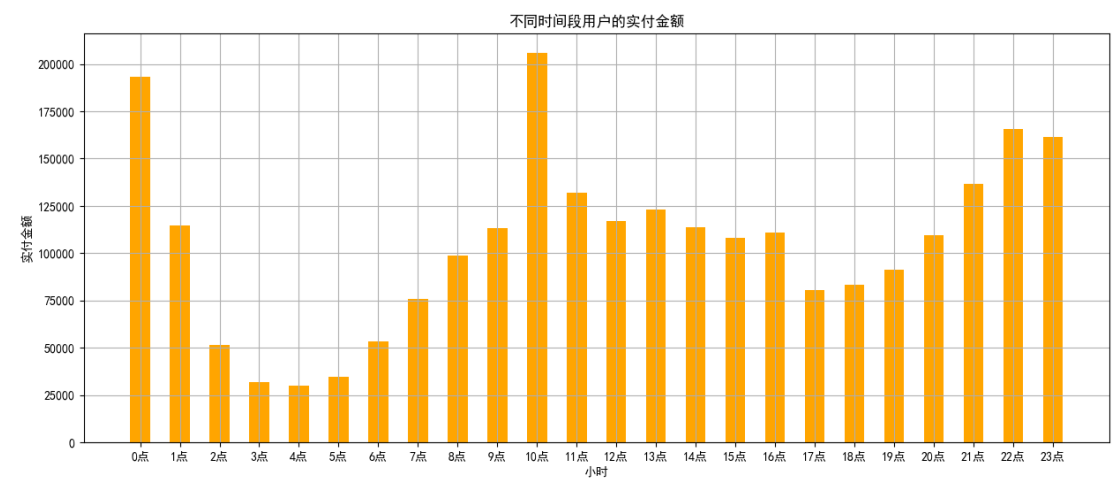


图 8 用户 24 小时习惯

我们针对地理纬度，统计出在这半年内不同地区的消费总金额前十名，我们将深圳市从广东省分离出来，原因是因为重庆市也没算在四川省内。前十省份消费最高的是上海，要高于其他省份 2-4 倍，前二到前五，消费金额水平相当。通过此图我们可以针对地区进行制定策略，为了不损失公平，可以保持原价不变，优惠券等额优惠，在高消费地区售卖时多货源，低消费区少货源。并且为了鼓励低消费区消费，采取购买多件免运费，消费满某金额抽奖等活动。具体如图 9 所示。

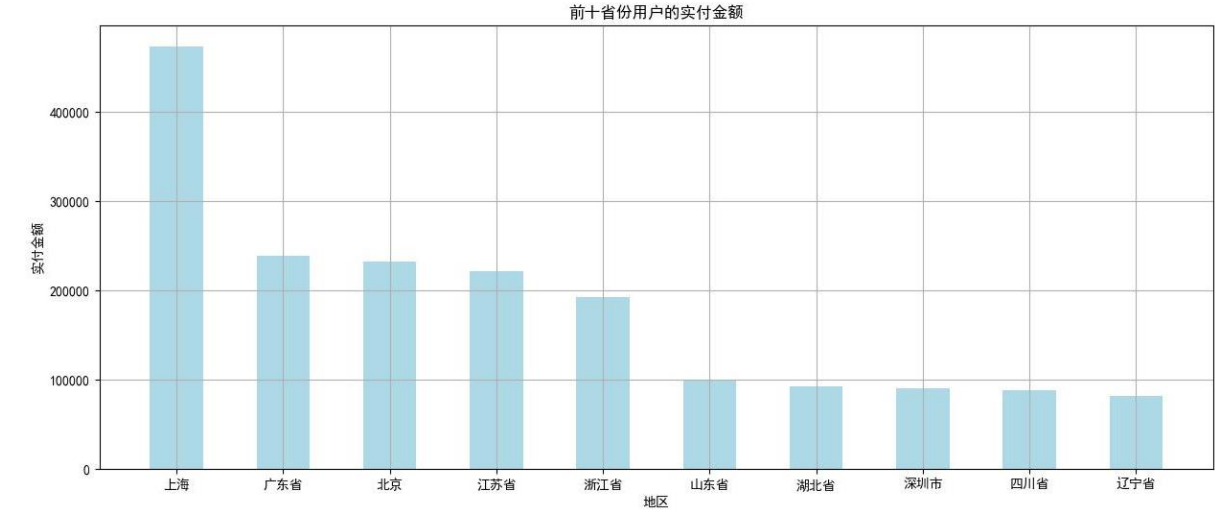


图 9 前十省份用户的实付金额

此外，统计用户使用优惠券情况，做出优惠券发放策略。大部分用户都是会使用优惠券的，未领取和领取未使用的用户占比相近，领取未使用优惠券的用户可能是系统自动领取、领取后忘记使用、优惠券面额太小等原因导致未使用。为了促销销售，可以推出折扣优惠券，第二件打折或优惠券面额提高一些，适当的从高额 1 位数变成小额 2 位数，从心理上表现出优惠差距。用户领取优惠券后，店铺后台可以给用户发送消息表明用户已经领取优惠券，做到适当的提醒，来提高用户对店铺的优惠印象。具体如图 10 所示。

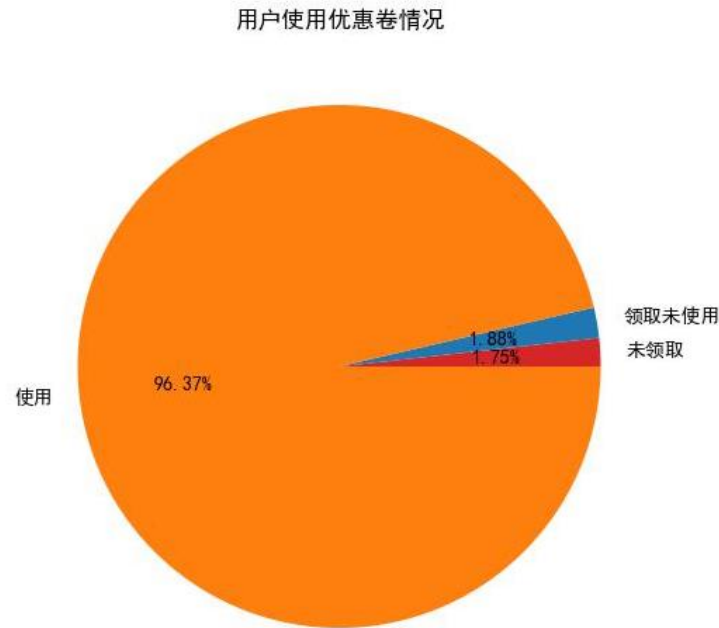


图 10 用户使用优惠券情况

对用户单次购买数量做出了分析，从图中可以看出，大多数用户单次购买仅购买一件商品，少部分购买 2 件，对此可以看出用户对购买多件商品欲望很低或者用户认为下次需要再购买即可，不需要购买很多，毕竟价钱都是一样的。所以可以实施第二件折扣，第三件折上折等活动鼓励用户消费。统计结果如图 11 所示。

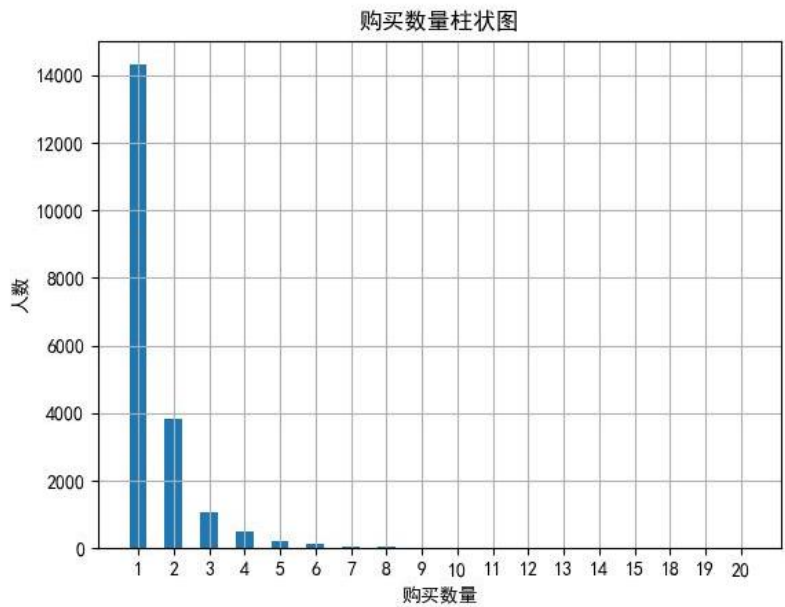


图 11 用户购买数量统计图

用户在同一家店铺多次购买说明用户对此店铺具有依赖，此店铺的商品和优惠能够满足顾客的需求，商铺需要维持此类用户的存在。经过数据的统计，用户

仅在相同店铺消费 1 次或 2 次。我们考虑到大部分用户的消费数据只有一条，我们统计了不同用户的消费频率，结合在同家店铺的消费次数来进行分析。大部分用户仅仅消费了一次，并且绝大部分用户也在同一店铺购买一次。通过购买频率图分析得到存在 6.28% 用户购买了 2 次及以上，但是用户在同一店铺购买仅有 0.08%，说明大部分用户不会在同一家店铺购买商品或用户购买的商品不是同一类型。针对此现象，店铺要对多次购买的用户做出留存策略，对只消费了一次的用户做出商品评价调查。具体如图 12，图 13 所示。

用户在同一店铺购买次数占比

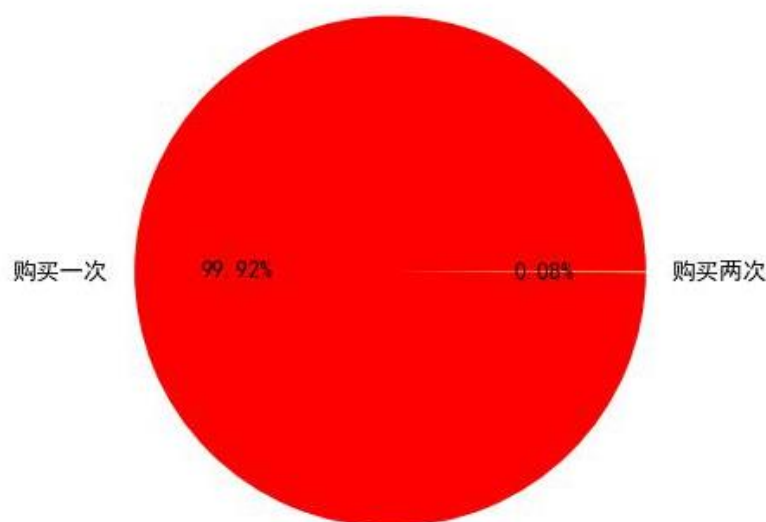


图 12 用户在同一店铺购买比率

用户购买频率比率图

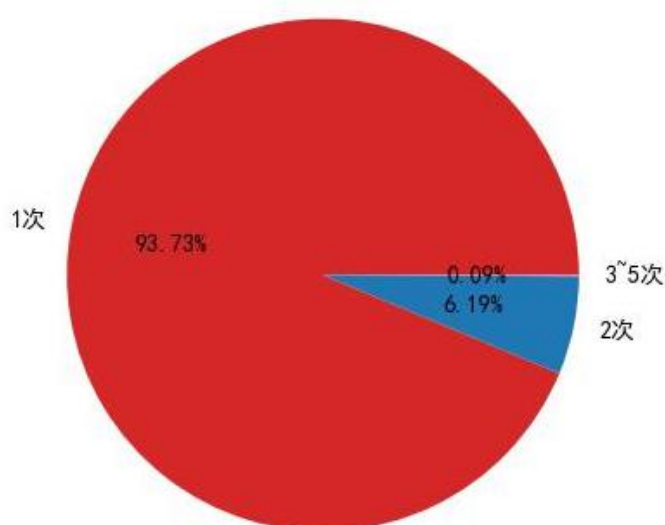


图 13 用户购买频率比率

了解用户的消费水平是做出针对性策略的重要信息，我们统计不同消费水平用户的数量。由于数据中消费金额不算大，所以我们设置消费等级表，详见表 1。统计结果如图 14 所示，可以看出低消费用户占大多数，是高消费的二倍多，所以在对用户做出营销策略时可以根据用户的消费等级来进行优惠，店铺的用户如果大多数都是中低消费水平可以适当的增加优惠券的面额，提高优惠力度。大多数用户是高消费时，可以适当的做出调研，表示商品的优点，留住用户。

表 1 消费等级表

消费金额	消费等级
0-100	低消费
100-200	中消费
>200	高消费

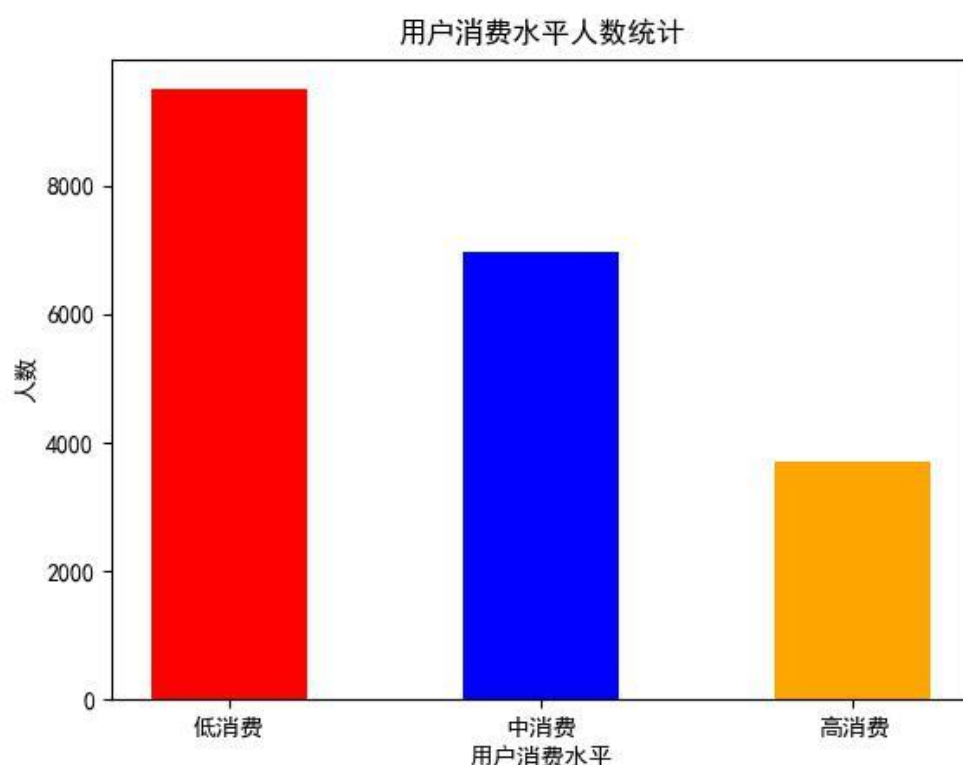


图 14 用户消费水平统计

由于数据的性质，大多数用户仅仅登录一次，如果按照此类数据来计算留存率等特定指标数据，计算后的结果会大幅度偏向于 1 或 0 这两个极值，所以对用户进行推荐的时候并没有太大的参考意义。

四. 客户价值分析

4.1 RFM 模型

在面向客户制定运营策略、营销策略时，希望能够针对不同的客户推行不同的策略，实现精准化运营，以期获取最大的转化率。精准化运营的前提是客户关系管理，而客户关系管理的核心是客户分类。通过客户分类，对客户群体进行细分，区别出低价值客户、高价值客户，对不同的客户群体开展不同的个性化服务，将有限的资源合理地分配给不同价值的客户，实现效益最大化。由于大多数客户仅有一条消费数据，如果采取 CLV 模型会导致大部分用户偏向同一类，导致结果出现偏差，故采取另一种经典的分类模型，RFM 模型。

将 RFM 模型中的三个维度根据数据来做出表示，具体如表 2 所示。

表 2 RFM 维度表示

维度	表示
R	最后一次付款距 2022-07-30 的天数
F	用户购买的频率
M	用户购买的总金额

基于这三个维度进行建模，得到最终的用户价值表示，统计出每个价值维度的总人数，做出柱状图来直观的查看用户价值分布，具体如图 15 所示。

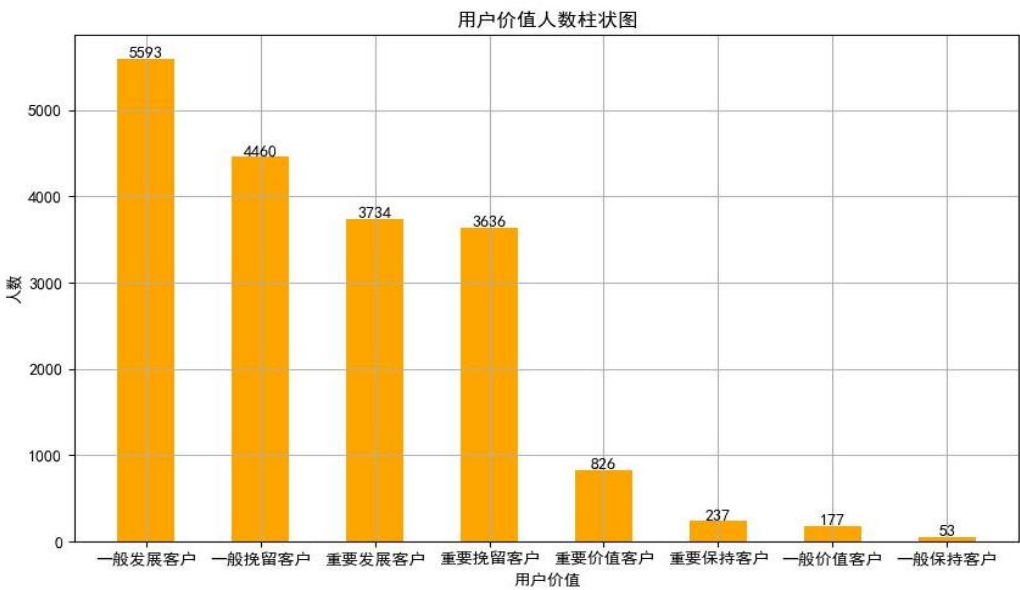


图 15 客户价值分布图

根据图中数据可知，分类人数最多的是一般发展客户，其次是一般挽留客户。一般发展客户消费频率低，消费金额少并且消费距规定时间长，这一类客户可能是趁着有折扣或者当时需要才购买商品，对商家无忠诚度。一般挽留客户消费频率低，消费金额少，但是消费时间距今短，店铺可以强化对这类用户的联系。

其次就是重要发展和重要挽留客户，这两类用户的人数较多且相近。重要发展客户相比于一般发展客户，其消费金额数值大，具有很大的发展潜力。商家应加强这类客户的满意度，使他们逐渐成为忠诚客户。

重要挽留客户同样相比一般挽留客户消费金额高，商家可以着重对此类用户增加资源，促使他们进行多次消费。

重要保持和重要价值客户是商家大多数经济的来源，根据二八规则，这两类用户虽然人数较少，不过是重中之重的用户。重要保持客户的消费频率和消费金额都比较高，消费周期也短，这类客户是忠诚客户，需要重点维护。

重要价值客户，相比重要保持客户消费时间比较久，商家应加强与这类客户的互动，召回用户，延长客户的生命周期。

最后两类客户，数量少并且不具有较高的价值，可以减少成本，节约资源，将资源利用在大部分用户和重要用户上。

4.2 用户画像

用户画像也叫用户标签，通过对用户搭建画像，做出用户喜好，用户消费水平，用户的购买习惯等。为了更好的搭建用户画像，进行了多标签构建，分别有属性标签，统计标签，规则标签，具体特征如表 3 所示。通过构建的标签，可以对用户进行精准营销，针对不同消费者给予不同类型和不同金额的优惠券，提高拉新，留存，转化等运行效率。其次洞察用户的核心需求，组织用户所需要的活动，促进用户的消费。最后，目前推荐算法已经逐渐成熟，统计消费者的相关标签也可以促进推荐，让用户得到更好的体验。最终，对用户画像的图像表示如图 16 所示，这是针对某一个用户的画像，如果需要可以输入用户的名字进行查找，如果不需要可以随机进行画像。

表 3 标签特征对应表

标签类型	用户特征
属性标签	用户名，IP 地址
统计标签	消费能力，消费频次，优惠券使用习惯
规则标签	消费时间习惯（每天，每月）

笑嘻佳 高频
喜欢在上午进行消费
喜欢在月中进行消费
江苏省无锡市 高消费用户
经常使用优惠券

图 16 某用户用户画像

五. 用户优惠券发放模型

平台推出了不同商家的代金券，用于发放给用户，不过用户之间也存在着差异，如何给那些有潜力的用户发放优惠券促进消费成为了主要问题，针对此问题，建立了四种预测模型，分别为朴素贝叶斯分类模型，Logistic 回归模型，决策树模型，K 近邻模型。在利用模型做出预测之前，需要对数据进行抽样。

1. 数据采样

模型预测的效果突出往往来自于干净的数据，为了确保建立的模型能够准确的预测结果，检查数据是否可以带入模型进行训练。

认定用户使用优惠券为 1，领取未使用和未领取认定为 0，通过 0-1 分类模型来进行预测。可以看到，正样本数据和负样本数据具有严重的不平衡，这会导致预测出来的结果严重偏向于正样本，对测试集上的预测结果产生严重的过拟合，并不准确。所以，采取将负样本复制多次达到和正样本数量接近的方法来缓解正负样本不平衡的方法。

此外，还需要将字符型数据进行量化，将其参与模型中进行运算。我们将用户消费前 10 的城市设定为 1，其余设定为 0. 具体如图 17 所示。

实付金额	邮费	购买数量	日期_月	日期_天	日期_小时	优惠券发放使用率	未领取优惠券占比	用户优惠券使用率	用户优惠券未使用率	用户未领取优惠券占比	F	M	R	用户所在地
190	0	1	4	18	7	0.965035	0.013793	0.500000	0.000000	0.5	2	443	102	0
79	0	1	6	18	5	0.977444	0.022059	0.000000	1.000000	0.0	1	79	41	0
99	0	2	2	17	4	0.983333	0.010989	0.500000	0.500000	0.0	2	278	162	1
106	0	1	1	11	18	0.973154	0.019737	0.666667	0.333333	0.0	3	343	199	1
84	0	1	1	9	23	0.977444	0.022059	0.666667	0.333333	0.0	3	152	201	1
...
87	0	1	4	18	0	1.000000	0.000000	1.000000	0.000000	0.0	1	87	102	1
99	0	1	5	27	17	1.000000	0.000000	1.000000	0.000000	0.0	1	99	63	1
123	0	3	5	8	13	1.000000	0.000000	1.000000	0.000000	0.0	1	123	82	1
74	0	1	6	29	9	1.000000	0.000000	1.000000	0.000000	0.0	1	74	30	1
49	0	1	6	24	19	1.000000	0.000000	1.000000	0.000000	0.0	1	49	35	1

图 17 模型参数

2. 模型预测

我们采用了决策树，K 近邻，Logistic 回归，朴素贝叶斯，LightGBM 集成学习模型来预测用户是否使用优惠券，以便后续对用户进行优惠券下发。我们将数据分为训练集和测试集，训练集占比 80%，测试集占比 20%。由于数据值在不同特征下的大小差异过大，所以我们采用了最大最小值方法来对数据进行归一化处理，公式如下：

$$X_{new} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (\text{式 5.1})$$

我们对处理好的训练集代入模型中进行训练，训练后将处理好的训练集和预测出的结果进行计算相似性，结果如表 4 所示。

表 4 模型精准度结果表

模型	精准度
决策树	0.998
K 近邻	0.974
Logisitic 回归	0.993
朴素贝叶斯	0.988
LightGBM 集成学习	0.996

我们利用五种机器学习模型对用户使用优惠券进行了预测，根据预测结果显示，模型能够根据用户的习惯精准的预测出用户是否会使用优惠券，为商家提供了是否为用户发放优惠券的衡量标准。接下来针对优惠券投放做出策略。

六. 优惠券发放策略

上述机器学习模型预测出了商家是否给用户发放优惠券，准确度很高，几乎可以等于使用优惠券的用户都在模型的评估结果里。所以我们可以分析使用优惠券的客户的行为。

首先我们根据用户所在地来判断使用优惠券的情况，因为用户使用优惠券和未使用优惠券（包含领取未使用和未领取）的比率为 1: 1，使用优惠券在前十省份的用户占比高达 73%，未使用优惠券的用户也达到 72%，说明再下发优惠券的时候城市仅仅可以当作二级指标，不能当作决定性指标。

根据用户的消费习惯，1，2，6 月用户购买力度大大增加，在此时可以对用户进行优惠券下发，以便促成交易。并且在每个月的 11，16，17，18 号，针对活动促销，可以选在这三天进行，因为用户在这三天购买情况远高于其他时间，并且在 2，3，4，5，7，9，10，31 号不适合搞活动促销，且不宜发优惠券。

如果商家选择发放优惠券，在每天的 21 点以后凌晨 1 点之前，还有第二天早上的 10 点，这段期间是用户购买的高峰期，可以选择在此刻进行下发优惠券。方便用户周知。

根据用户本人的购买频率下发优惠券是很准确的，用户在购买频率多的情况下，可以下发优惠券并邀请成为会员，这样能够促使用户继续购买。由于大多数用户都是购买了 1 次，并且在这六个月没有进行后续的消费，也就是说如果是新客户，在客户满足以上情况的前提下，可以进行优惠券发放，如果不满足，可以根据店铺的实际情况，比如新开店铺需要增加购买量来提升知名度，这样就需要下发优惠券，如果是老店铺，新用户第一次购买可以发放优惠券，如果第一次购买没有使用优惠券，可以考虑第二次不需要下发。如果用户多次购买，在考虑下发优惠券提高用户忠诚度。

其次，根据店铺商品的价格，如果店铺商品普遍在 80 元左右，那么店铺要根据商品来给予优惠券，如果商品本身价格很低，那么优惠券可以根据店铺目前情况进行下发。

此上，就是优惠券下发策略。