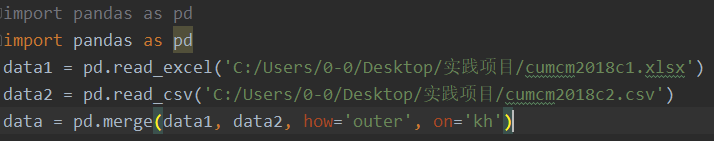
项目分析报告

1 数据准备

本报告所用数据为某商场2015年-2018年1月的销售情况与会员信息数据，项目优先对销售情况和会员信息进行一一匹配，利用python进行外连接式合并如下



合并后的数据情况为（2041027，15）。

2 数据处理

本报告所研究的内容是会员消费情况，故在将会员信息表和销售状况表合并后剔除非会员数据，其中有价值的数据是在销售时间内具有消费记录的会员，故对合并后数据中无消费记录者数据进行剔除。



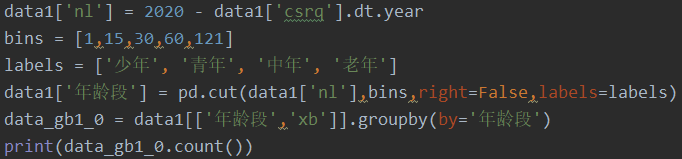
而后所的数据形状为（876147，15）。

3 数据可视化

3.1 会员特征分析

3.1.1 会员年龄构成

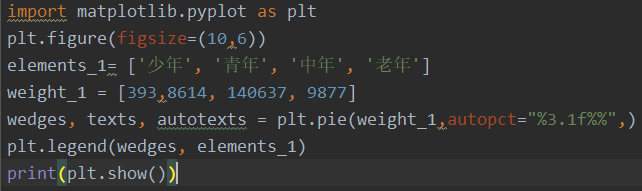
本报告对所有会员所处年龄段进行同进，对小于15岁者为少年，15至30岁者为青年，30至60岁者为中年，60岁以上则是老年。



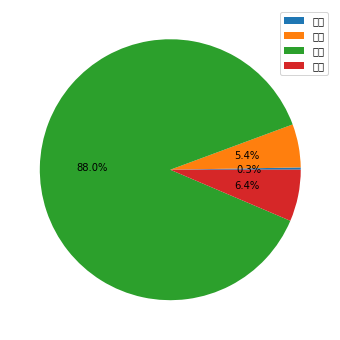
先对会员信息中的会员出生日期进行年份提取，再用2020将之减去，利用pandas模块中的cut函数进行会员年龄的分组，可得统计数据



以此数据为依据，可有代码如下：



将其数据可视化有



不同年龄段会员分布图

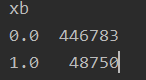
以上分布结果显示在该百货商场中，中年会员占据绝大多数，为88%，老年人其次，为6.4%，少年最少，不足1%。

3.1.2 会员性别构成

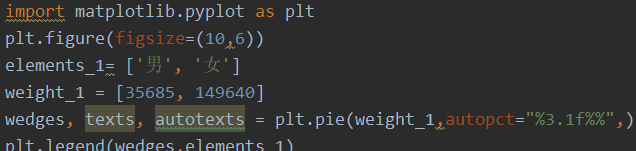
对会员信息表中的xb一列进行统计，其中“0”为女性，“1”为男性，则可有



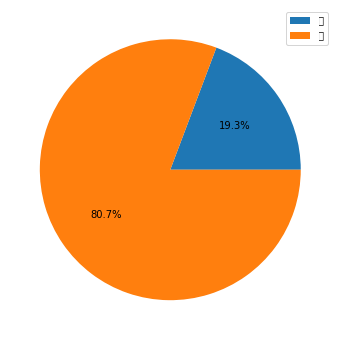
有其结果为



则以此数据为基础，按照以下代码



可得会员性别的分布情况如下：



会员性别分布图

以上分布结果显示在该百货商场中，女性会员占80.7%为该百货商场的主要群体，这与我国当下国情相一致。

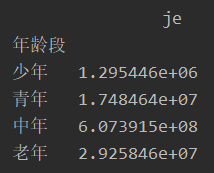
3.2 特征顾客消费分析

3.2.1 会员年龄消费金额分析

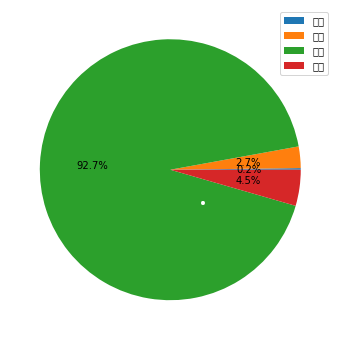
在前文合并处理后的数据中，按年龄段分组运行相似代码，并加入消费金额的求和，即有：



则有各年龄段2015-2018年的消费金额如下：



可视化结果为：

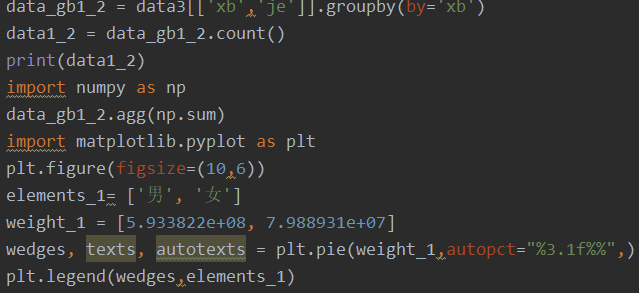


各年龄段消费金额饼状图

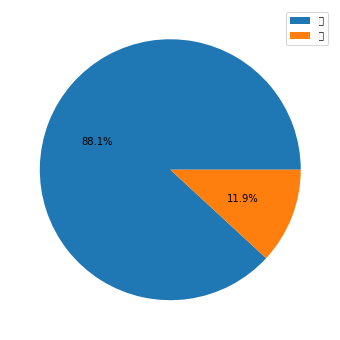
以上分布结果显示在该百货商场中，占会员大多数的中年群体，其消费金额也在所有会员的消费金额占92.7%，相较而言是最具价值的群体。

3.2.2 会员性别消费金额分析

与上文所述不同年龄的会员消费金额分析相似，本位以下列代码可将不同性别的会员消费金额进行分析。



即可视化结果为

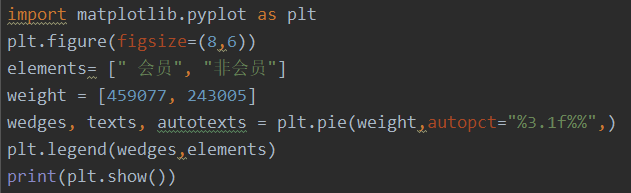


男女消费金额饼状图

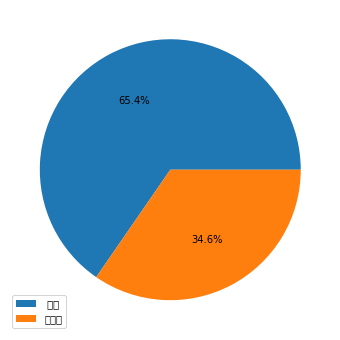
在上述分析中，女性的消费金额占88.1%，男性占11.9%，与会员中男女比例相类似，即说明男女消费力度相似。

3.2.3 会员与非会员订单分析

经过对消费状况表的统合，以总订单数减去会员订单数则为非会员的订单数，故而可运行以下代码



即有可视化图形为：

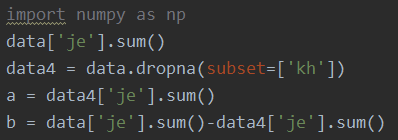


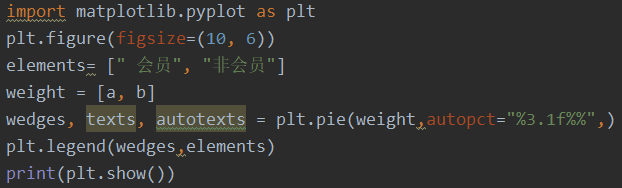
会员与非会员订单数之比

其中，会员的订单数占较大部分，达到65.4%，约为非会员所构成的订单数的两倍。这与会员多为该百货商场附近的居民，可多次到该百货商场消费一致。

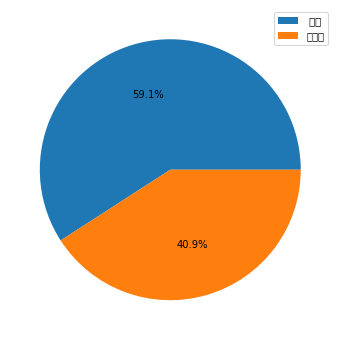
3.2.4 会员与非会员消费金额之比

与研究订单数的方法相一致，在消费状况中将所产生的总消费金额减去会员消费金额则为非会员所产生的消费金额，运行一下代码。





及其结果可视化为



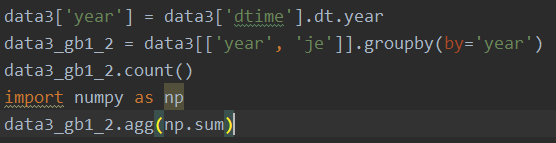
会员与非会员消费金额之比

在上图中，会员所产生的消费金额超过一半，占总体的59.1%，但相较其所产生的订单数来说，就稍显不足。由此图可知，会员所产生的消费与非会员所产生的消费相比，其每单消费金额的均值较小。

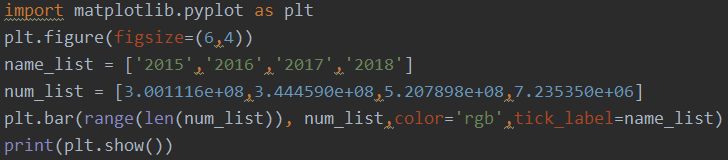
3.3 时间分析

3.3.1 年增长

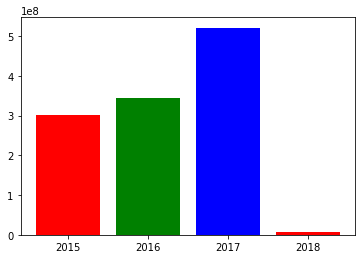
本报告对该商场三年的销售额进行统计



对所的结果运行以下代码



则有每年的销售额如下：

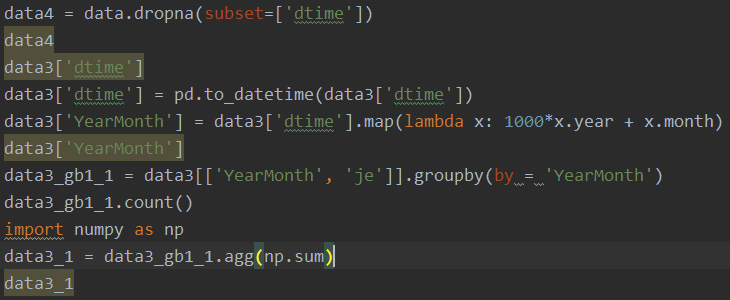


各年销售额柱状图

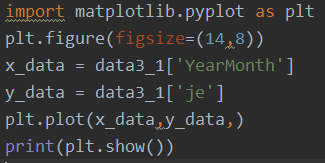
在上图中，2018年的销售额为其一月份部分时间所产生，非一年所得。故而从图中可知，该百货商场的销售额在2015、2016和2017三年间都呈现上升态势，发展良好。

3.3.2月增长情况

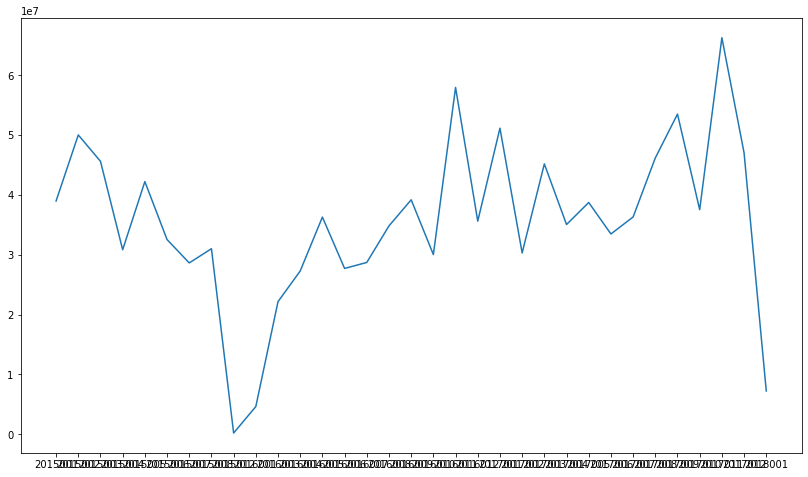
由于年份较短，故对三年间的36个月所产生的销售额进行统计，与上文同理，可运行下列代码



以其结果可视化，有



则有各月份销售额所绘制的折线图如下：

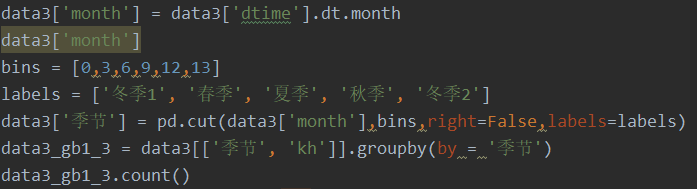


各月销售额折线图

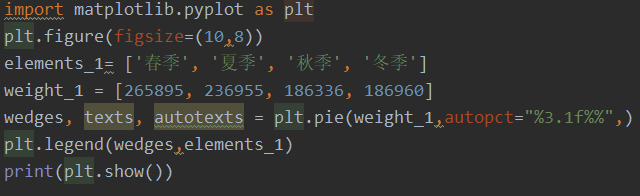
上图中，销售额整体呈现上升态势，不过增长趋势较为平缓，特别在每年5月和12月时又有明显下滑现象。

3.3.3 季节消费人数

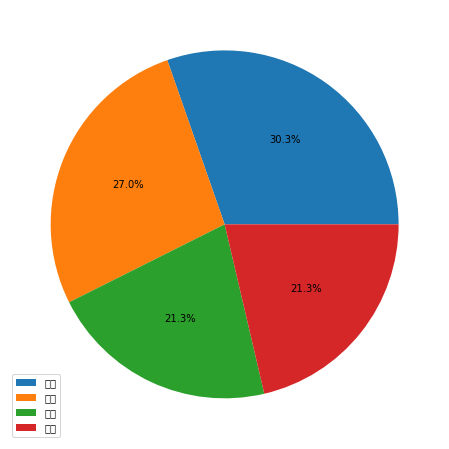
基于上述分析，在对各个季节的消费人数进行探究，使用cut函数对用户消费时间进行分组，即可运行



对所的结果中冬季1和冬季2相加则为冬季的人数值，基于此结果数据，对数据进行可视化过程，运行



可得下列饼状图

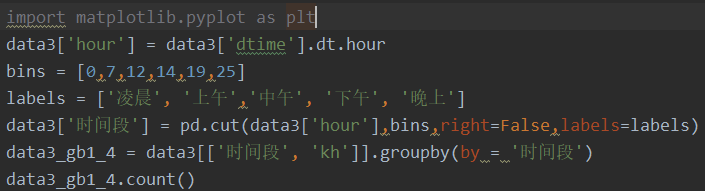


各季节消费人数饼状图

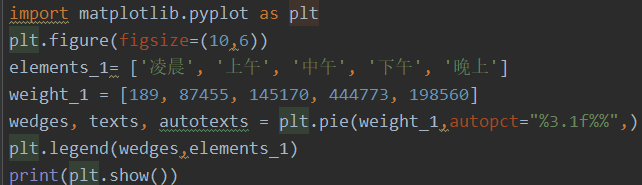
从上图可知，该百货商场在春夏两季人流量更多，都超过了总人流量四分之一，其中又以春季为最，大30.3%。而秋冬两季相差不大，都低于四分之一，且其数值都在21.3%左右。

3.3.4 各时间段消费人数情况

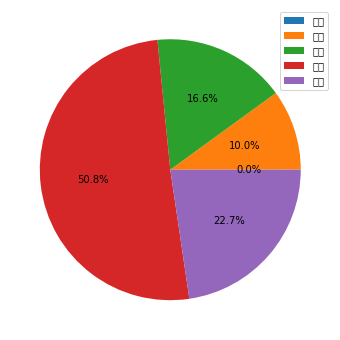
本文将一天分为五个时间段，其中零点到6点为凌晨，6点-11点为上午，11点-13点为中午，13点-18点为下午，18点以后为晚上。同样利用cut函数进行分组，即



以其结果进行可视化，运行



则有：



各时间段人数分布饼状图

有上图可知，在一天之中，凌晨的客户最少，下午的客户最多，占总数的一半，其次为晚上，占22.7%，再就是中午和下午，分别占16.6%和10%。

4 客户分类

本文在传统RFM模型上进行改进，对百货商场三年来会员销售状况进行会员的分类。

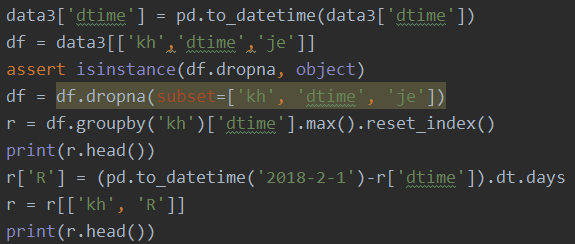
4.1 RFM模型

RFM模型是衡量客户价值和客户创利能力的重要工具和手段。在RFM模型中设定，客户数据库中有3个要素，这3个要素构成了数据分析最好的指标，即最近一次消费，消费频率和消费金额三项。

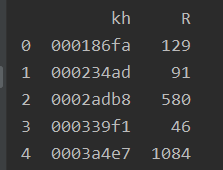
4.2 模型建立

4.2.1 最近消费R

最近一次消费意指上一次购买的时候，理论上，上一次消费时间越近的顾客应该是比较好的顾客，对提供即时的商品或是服务也最有可能会有反应。营销人员若想业绩有所成长，只能靠偷取竞争对手的市场占有率，而如果要密切地注意消费者的购买行为，那么最近的一次消费就是营销人员第一个要利用的工具。历史显示，如果我们能让消费者购买，他们就会持续购买。这也就是为什么，0至3个月的顾客收到营销人员的沟通信息多于3至6个月的顾客。该商场销售数据到2018年1月，依照RFM模型可令时间“2018-2-1”减去每个会员消费时间的最大值，并令之为“R”，运行

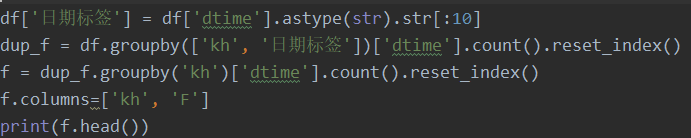


可得到每个会员的R，有

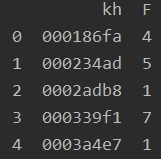


4.2.2 消费频率F

消费频率是顾客在限定的期间内所购买的次数。亦可以称为最常购买的顾客，也是满意度最高的顾客。如果相信品牌及商店忠诚度的话，最常购买的消费者，忠诚度也就最高。增加顾客购买的次数意味着从竞争对手处偷取市场占有率，由别人的手中赚取营业额。则在三年中的消费次数即为消费频率，运行

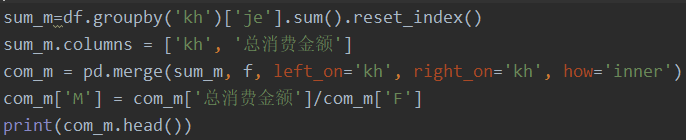


有各个会员的消费频率，如：



4.2.3 消费金额M

消费金额是所有数据库报告的支柱，理论上M值和F值是一样的，都带有时间范围，指的是一段时间（通常是1年）内的消费金额，在工作中我认为对于一般店铺的类目而言，产品的价格带都是比较单一的，比如：同一品牌美妆类，价格浮动范围基本在某个特定消费群的可接受范围内，加上单一品类购买频次不高，所以对于一般店铺而言，M值对客户细分的作用相对较弱。本文按会员卡号进行分组对各个会员进行此阶段内的消费金额统计，并令之为M，运行

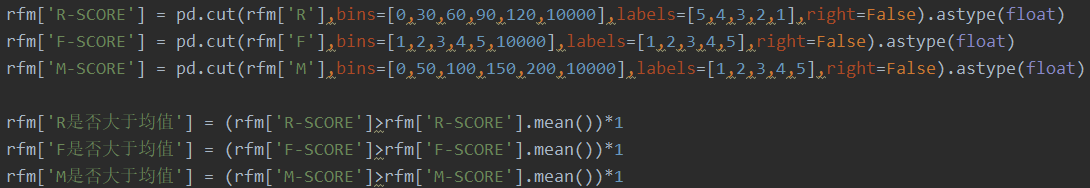


可有各会员的消费金额情况，如



4.2.4 人群数值

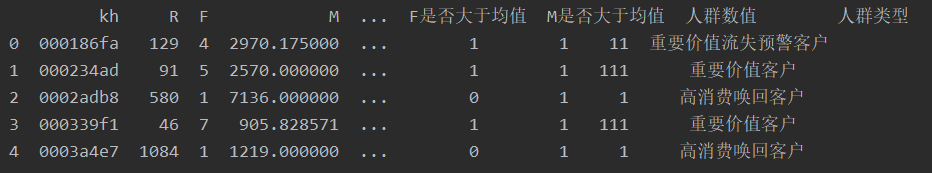
在计算得到各个会员的R、F、M值时，对三者具体的值进行分类，其中最近消费时间最小越好，而消费频率和消费金额则越大越好。而后分别使三者与其对应的均值相比较，进行1/0赋值，即



再将所得的三个数值进行组合，便可有其数值所表示的人群，共计有八类，为111、110、101、100、11、10、1、0。



分别命名为重要价值客户、消费潜力客户、频次深耕客户、新客户、重要价值流失预警客户、一般客户、高消费唤回客户和流失客户。如上代码，有其结果如下图：

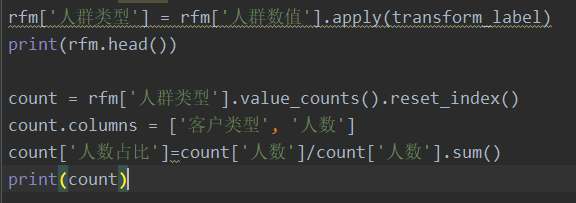


4.3 客户分类结果分析

依据RFM所分的人群类型，本文对商场会员中各个类别的会员进行人数和消费金额的对。

4.3.1 各类会员人数分析

在三年消费情况来看，各类会员的人数进行统计，有

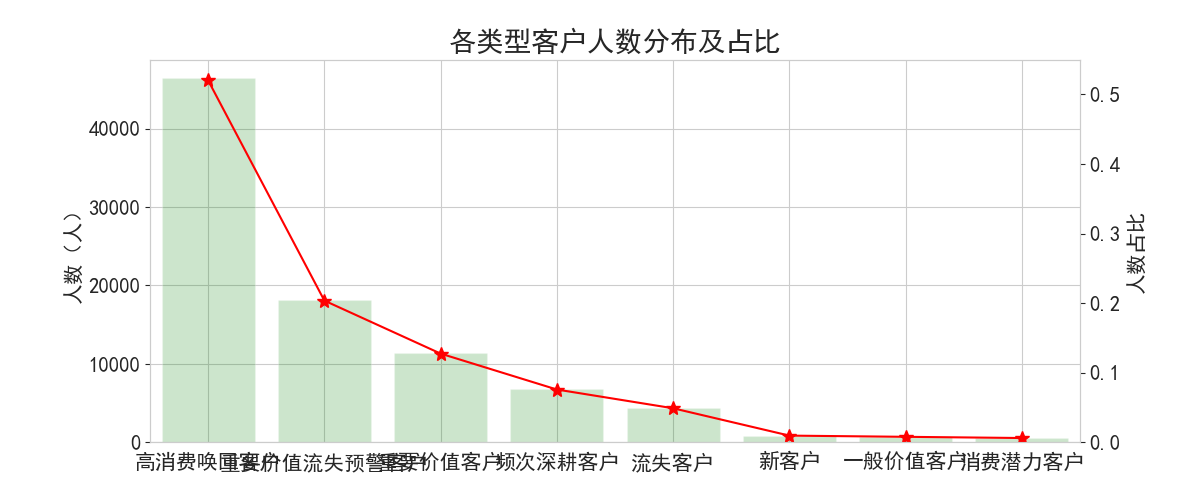


有其结果



从结果来看，重要价值客户只占12.75%，而高消费唤回客户和重要价值流失预警客户共占72%，这部分是消费高于平均水平，但其消费频次和最近消费时间来说，属于将留未留的阶段，显然该商场的会员流失将为严重。

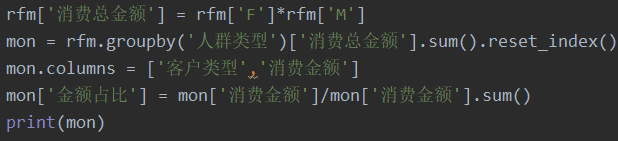
将其可视化，则有



图中显示，该商场潜力客户、一般价值客户都较少，为该百货商场会员的留存敲响警钟。

4.3.2 各类会员消费金额分析

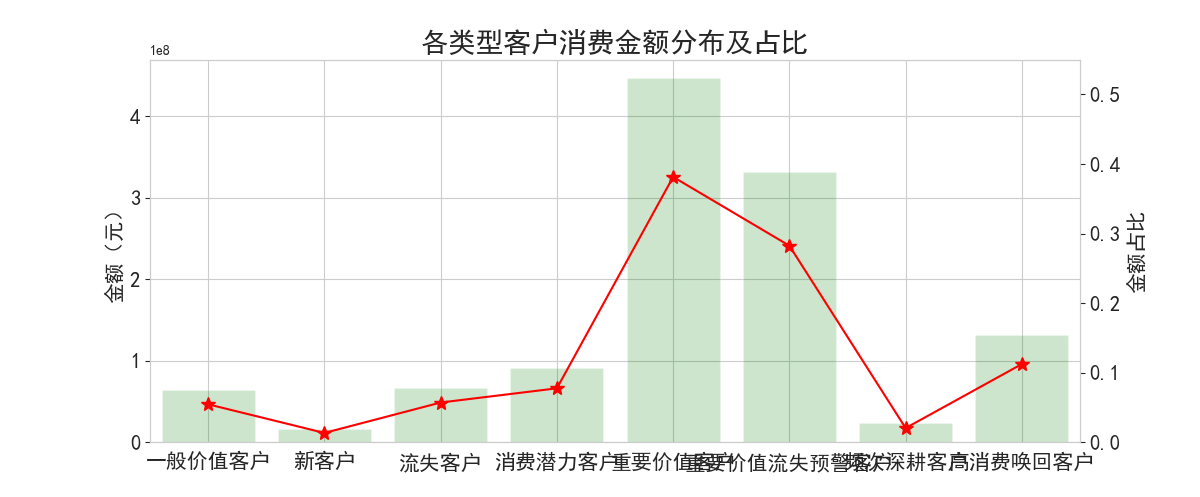
同样基于RFM模型的分类，本文对该百货商场中会员的消费金额进行分类统计，有运行



得其结果，为



则可知，与人数的分数不同，消费金额来说，重要价值客户占所有会员的38.14%，同样高消费唤回客户和高价值流失预警客户也大约占38%。



同样表示该商场的优质会员流失的情况较为严重。

5 客户画像

本文依据前文所分类别，对该百货商场会员的人群数值、年龄、性别、入会年限、消费月份、消费时间、折扣、商品售价和消费金额进行聚类，在反复实验后共聚5各类别，其情况如下雷达图：

