**学生校园消费行为分析报告**

目录

[一、 背景 1](#_Toc5586)

[二、 分析目标 1](#_Toc13879)

[三、 数据处理 1](#_Toc2326)

[四、 对食堂的就餐情况分析并进行数据可视化 1](#_Toc3970)

[五、对学生消费行为分析并建立模型 5](#_Toc28052)

[六、小结 8](#_Toc29708)

1. **背景**

某项目针对贫困本科生的“温饱问题”进行援助。在学校一卡通中心教育基金会的工作人员来找了全校一万六千余名在校本科生9月中旬到11月中旬的刷卡记录，对所有的记录进行了大数据分析。最终固定500余名“准援助对象”。可见，对学生校园消费行为分析具有重要意义。在这背景下，我们为获得某校的18级学生2019年四月份的消费情况来进行目标分析。

1. **分析目标**
2. 分析学生的消费行为和食堂的运营状况，为食堂运营提供建议。
3. 构建学生消费细分模型，为学校判定学生的经济状况提供参考意见。
4. **数据处理**

第一步：

data1.csv、data2.csv对应学生ID表、消费记录表。使用data1.csv、data2.csv，

根据实际项目需求对数据进行清洗，去除缺失值所在列的数据TerSerNo、ConOperNo对项目无影响可直接删除，对数据清洗为保证数据的有效性。进行描述统计，转换时间类型为datatime，删除异常时间点（这里我们认为5点以前为异常时间点，该时间段内学校不可能有消费记录等），保存清洗后数据为task1\_X1,task1\_X2。

第二步：

对学生个人信息表和消费记录表进行关联。由数据说明表中可知每个学生的校园卡号CardNo唯一，根据CardNo关联task1\_X1,task1\_X2合并为task1\_X3。关联后的task1\_X3的数据正是我们主要探究的18级学生四月份的消费情况。

1. **对食堂的就餐情况分析并进行数据可视化**
2. 绘制各食堂就餐人次的占比饼图，分析学生早中晚餐的就餐地点是否有显著差别。以时间间隔接近的多次刷卡记录视为一次就餐行为。

根据task1\_X2的数据提取各食堂的消费记录，我们认为6-8点为早餐时间，11-13点为午餐时间，17-19点为晚餐时间。绘制的饼图如下：

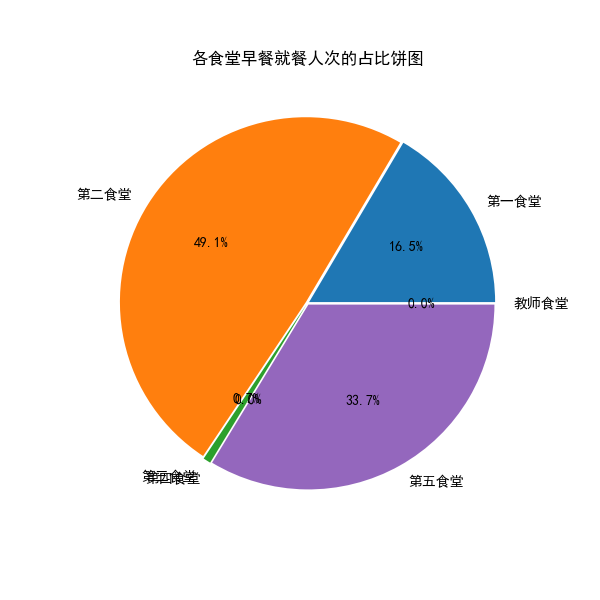


图4.1各食堂早餐就餐人次占比

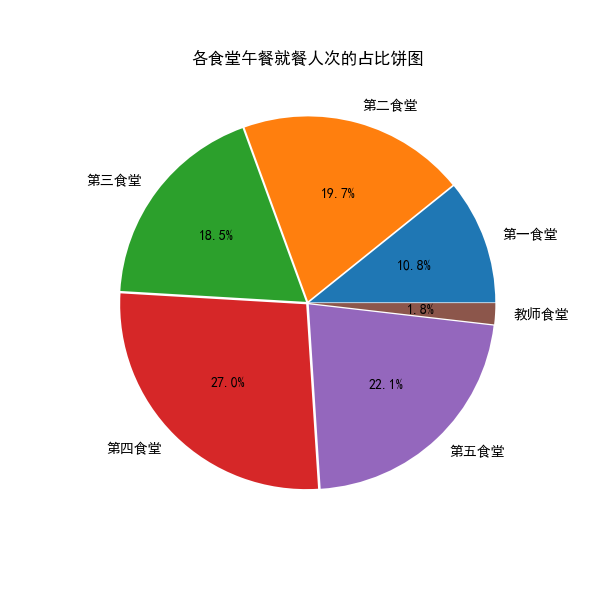


图4.2各食堂午餐就餐人次占比

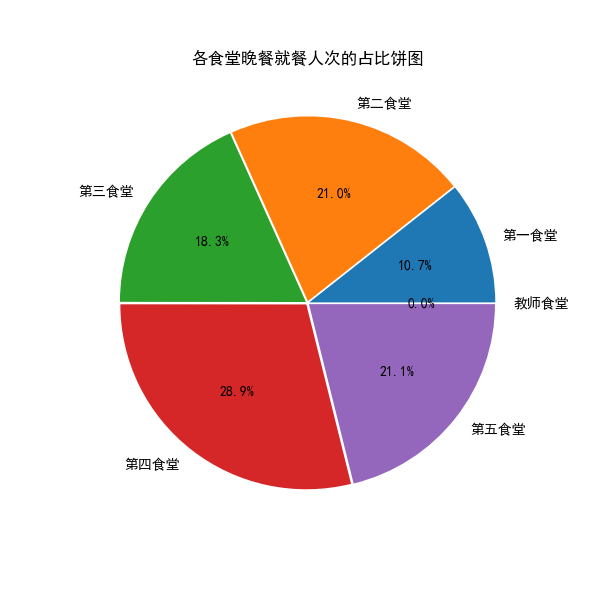


图4.3各食堂晚餐就餐人次占比

分析：由图4.1、图4.2、图4.3可直观得出：早餐就餐人次由高到低食堂排名是第二食堂、第五食堂、第一食堂、第三食堂，第四食堂，其中教师食堂无人就餐；午餐就餐人次由高到低食堂排名是第四食堂、第五食堂、第二食堂、第三食堂、第一食堂、教师食堂；晚餐就餐人次由高到低食堂排名是第四食堂、第五食堂、第二食堂、第三食堂、第一食堂，其中教师食堂无人就餐。

1. 通过食堂刷卡记录，分别绘制工作日和非工作日食堂时间曲线图（这里的工作日与非工作日按照国家节假日规定分类），分析食堂早中晚餐就餐峰值。绘制图形如下：

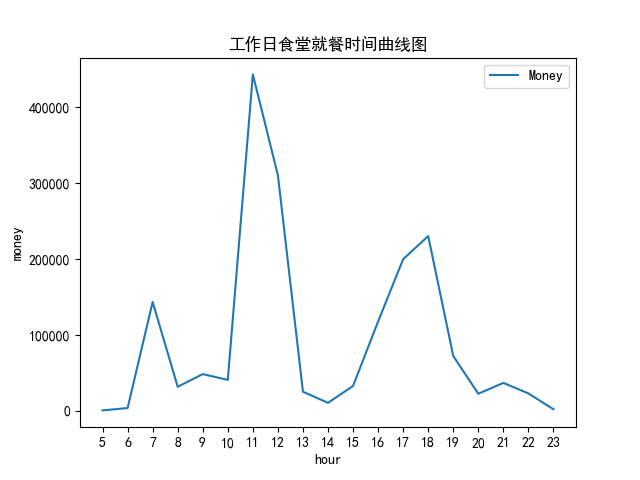


图4.4工作日食堂就餐时间曲线

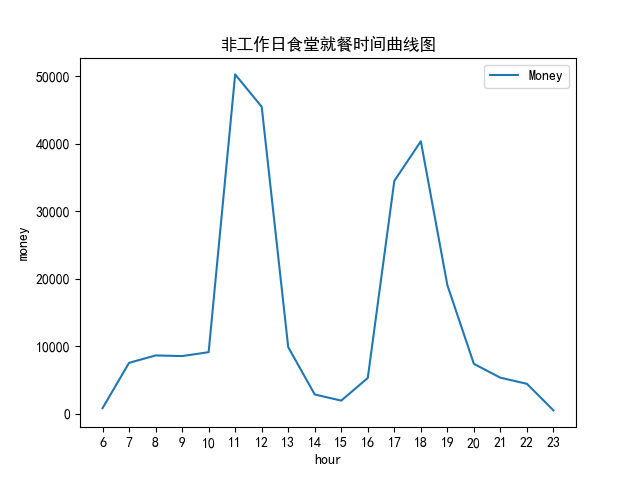


图4.5非工作日食堂就餐时间曲线

分析：由图4.4和图4.5直观看出工作日与非工作日均出现两个峰值，也都在午餐11点左右和晚餐18点左右时间。工作日的就餐量约为非工作日的8倍。

3)根据上述分析的结果，为食堂的运营提供建议。

根据各食堂的早中晚就餐人次、工作日与非工作日就餐量，可以此为依据各食堂准备相应食材量，避免食材浪费或不够；并且在人流高峰期设置多可购买窗口，避免因人流大影响就餐。

**五、对学生消费行为分析并建立模型**

1）根据18级学生的整体消费，计算本月人均刷卡频次和人均消费，分析不同性别学生群体的消费特点，分别计算各专业不同性别学生人均消费情况，并绘制条形图如下：

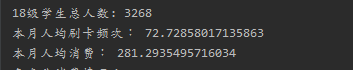


图5.1

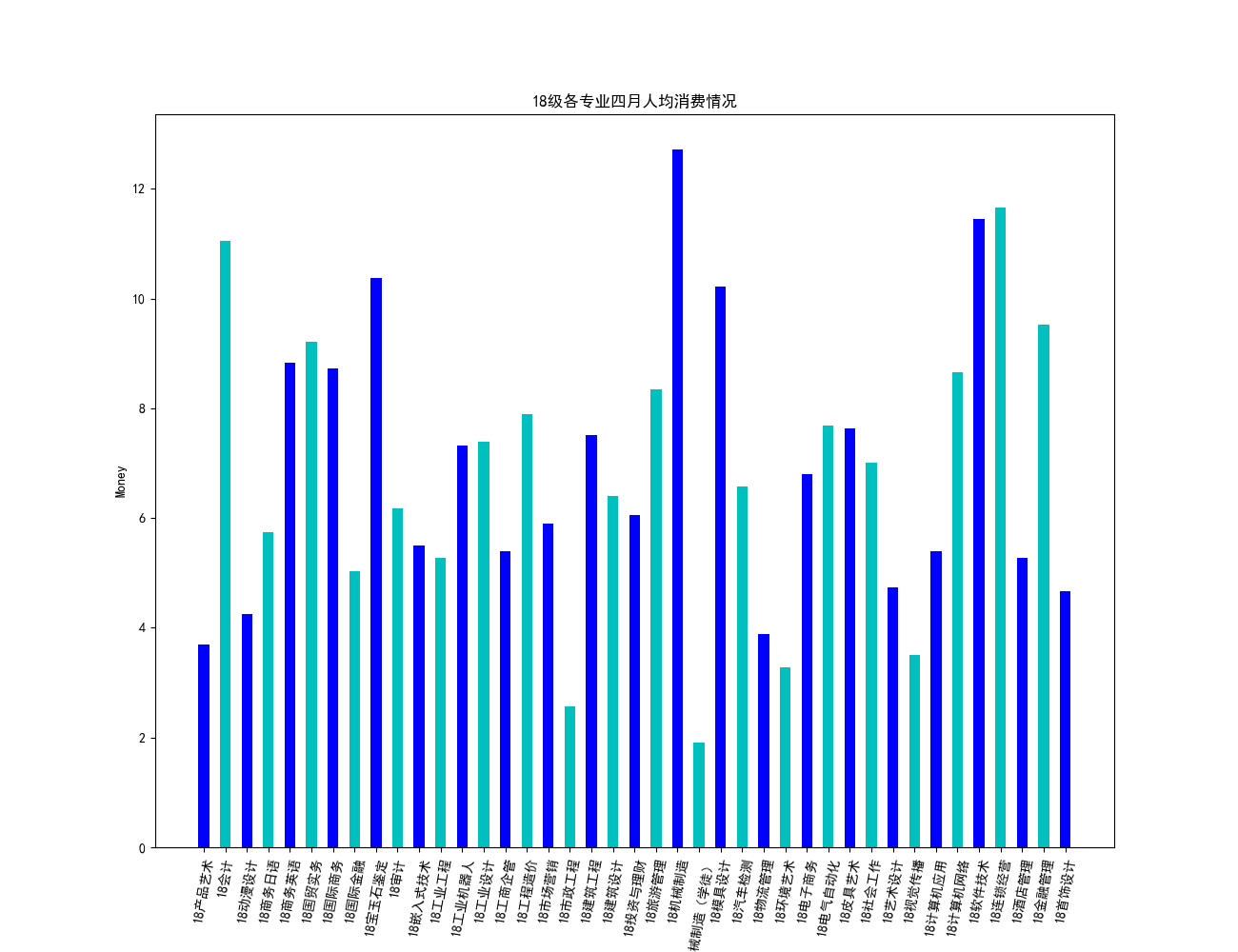


图5.2 18级各专业月人均消费情况

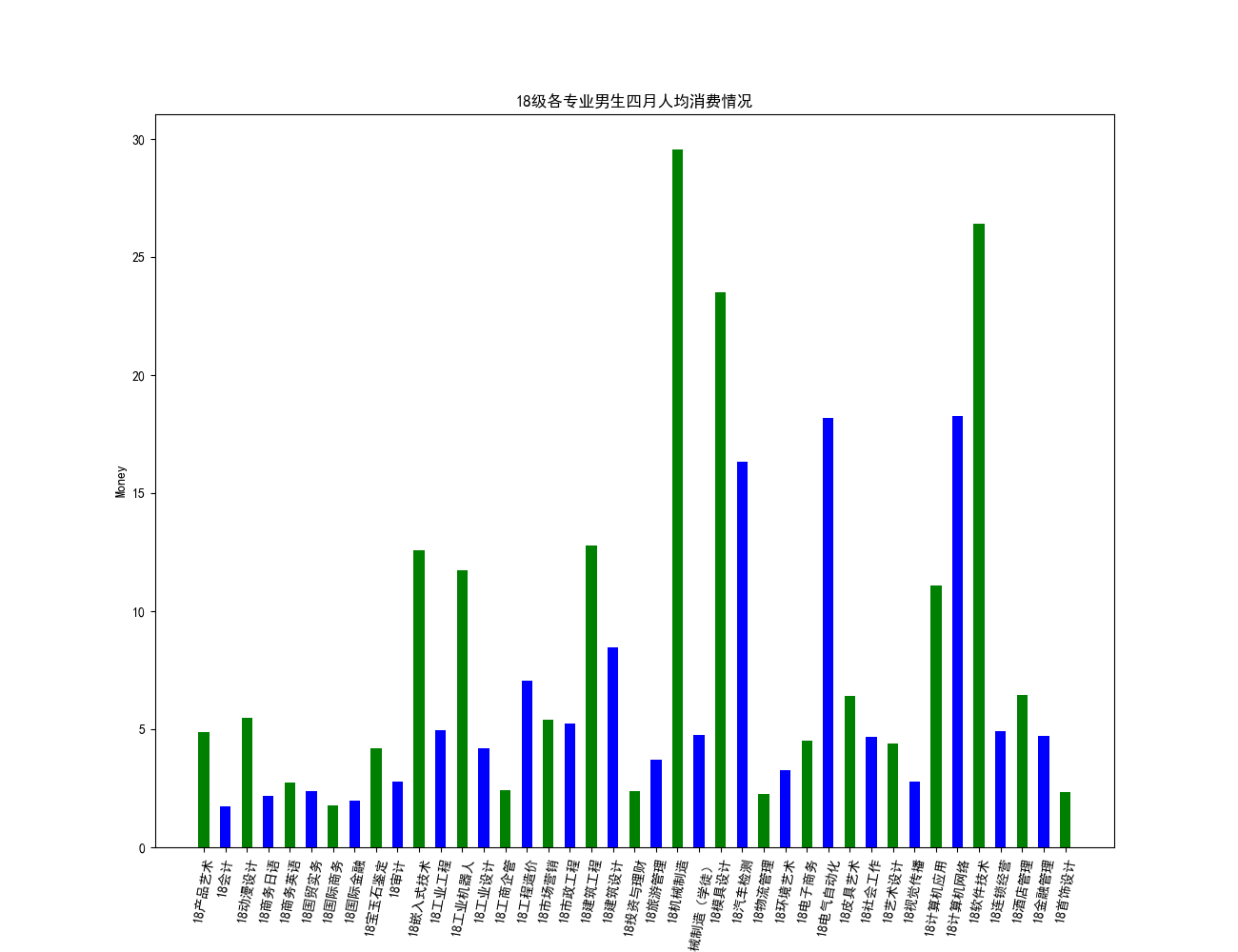


图5.3 18级各专业男生月人均消费情况

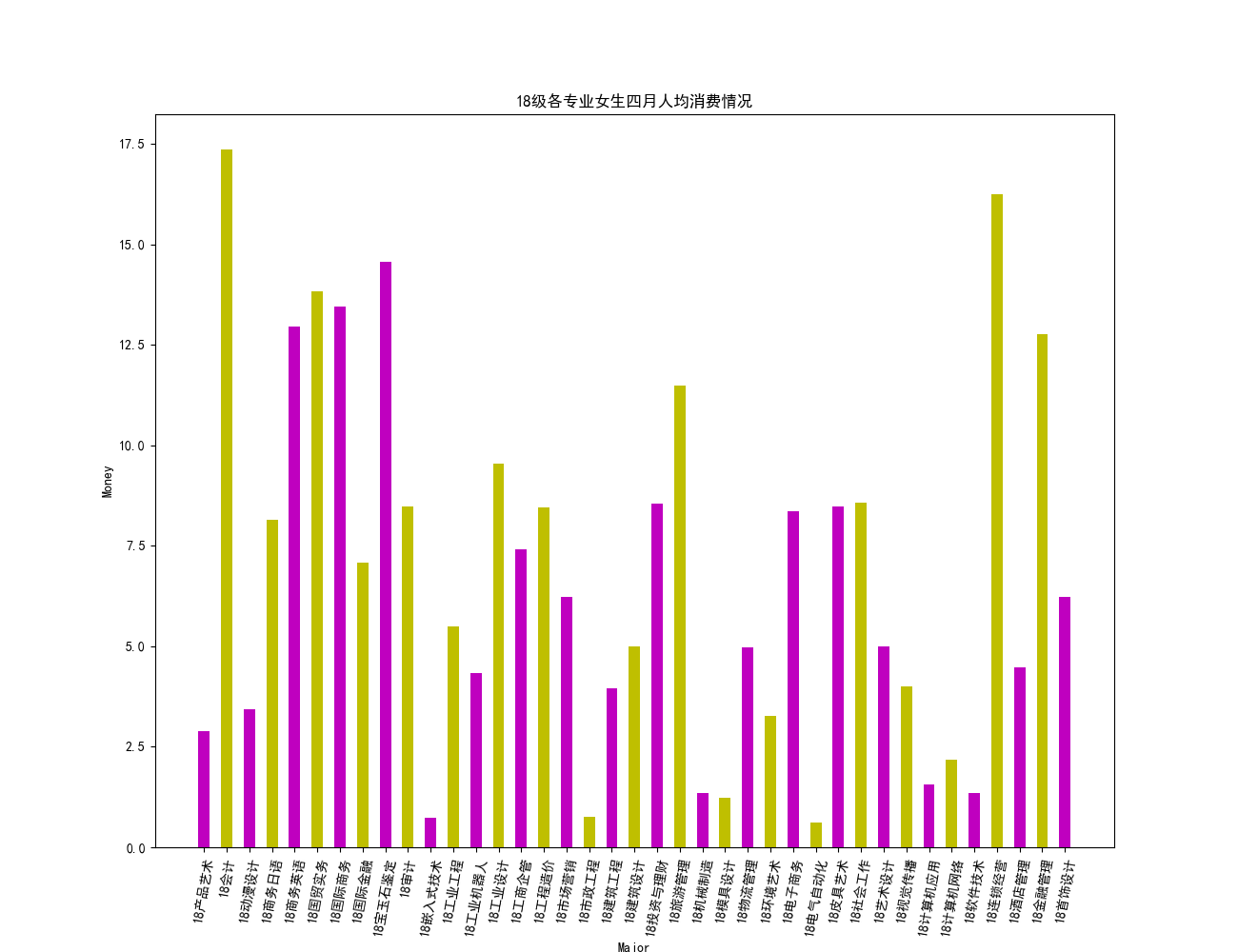


图5.4 18级各专业女生月人均消费情况

分析：运用Python计算得出18级人均刷卡频次为72次，人均消费为281元。由图5.2、图5.3和图5.4可知，各专业人均消费及各专业男女各消费情况， 其中女生消费整体要比男生高，且男生各专业消费情况差异很高。

1. 根据学生的整体校园消费行为，为学校判定学生的经济状况，我们可以认为生活饮食消费低且学习等其他消费高，存款低的学生经济状况较差。因此，选择生活饮食消费、其他消费和存款为特征构建聚类模型，分析学生群体的消费特征。

A.生活饮食消费：认为'在第一食堂','第二食堂','第三食堂','第四食堂','第五食堂','好利来食品店','红太阳超市','教师食堂','水电缴费处'的消费为生活饮食消费。

B.其他消费：非生活饮食消费的为其他消费。

C.存款：为存入校园卡的金额。

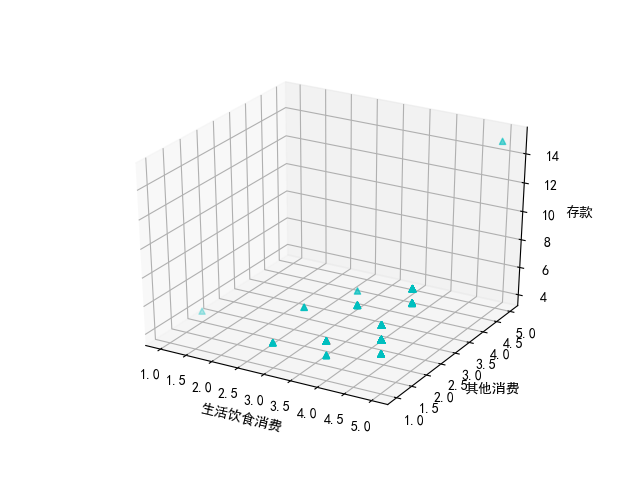


图5.5 18级学生得分情况分布

分别计算每个学生的生活饮食消费、其他消费和存款的总额，并设定其值得分为1,2,3,4,5等五个等级计算每位同学的得分情况并保存为task3\_X5，我们可以认为得分较高的前500名学生为“准援助对象”。由图5.5可看出18级学生的校园卡消费情况都较为集中在几个点，因此可以根据点的颜色深浅及各点间的距离判断出18级学生整体消费得分情况，四月份消费均高，存款较低，其中生活饮食消费较其他消费高。

**六、小结**

通过以上分析，我们可以对18级的校园消费情况得出以下结论：

1.根据学生的消费行为判断其消费喜好，学校各食堂运营可根据各个时间段的消费人次配备合理的食材数量避免食材浪费或亏损，并可根据人流情况在高峰期开设多个窗口，避免因人流影响就餐。

2.由各专业的消费图可知，不同专业的人均消费额相差很大，尤其男生明显，学校可对部分专业人均消费较低可加多援助名额。

3.以生活饮食消费、其他消费和存款三个指标构建学生消费细分模型，各指标数值可分为五个等级，计算每位学生的综合得分，将得分较高的前500名学生确定为“准援助对象”。