【参考答案】

(1)该数据没有缺失值，离群点考虑为真实情况，不做异常值处理。*（缺失值、异常值各2分，共4分）*





R含义：在2023年12月1日到2023年12月31日期间，顾客最近一次交易的时间间隔*（3分）*

F含义：在2023年12月1日到2023年12月31日期间，顾客交易订单号的数量，即一段时间内的消费频次*（3分）*

M含义：在2023年12月1日到2023年12月31日期间，顾客的总消费金额（或一段时间内的平均单次消费金额，平均需按照订单号计算）*（3分，两种说法均可给分）*

（2）*（每个顾客的R、F、M各2分，共18分）*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 顾客编号 | 时间间隔 | 购买频次 | 总消费金额 | 平均消费金额 |
| 12748 | 26 | 6 | 15.08 | 2.51 |
| 16218 | 30 | 1 | 15 | 15 |
| 17850 | 29 | 2 | 26.7 | 13.35 |

(3)

|  |  |  |  |
| --- | --- | --- | --- |
| R中位数 | F中位数 | 总消费金额中位数 | 平均消费金额中位数 |
| 22 | 1 | 15.08 | 15 |

*（中位数每个1分，共3分）*

以R的中位数作区隔进行划分，高于中位数记作“1”，代表该用户的最近一次消费时间间隔较长，在R维度上价值较低；反之记为“2”，代表该用户的最近一次消费时间间隔较短，在R维度上价值较高。

以F的中位数作区隔进行划分，高于中位数记作“2”，代表该用户的消费频次较高，在F维度上价值较高；反之记为“1”，代表该用户的消费频次较低，在F维度上价值较低。

以M的中位数作区隔进行划分，高于中位数记作“2”，代表该用户的平均单次消费金额较高，在M维度上价值较高；反之记为“1”，代表该用户的平均单次消费金额较低，在M维度上价值较低。*（RFM综合标签规则每个1分，共3分）*

确定RFM综合标签，划分客户类别，共8类。

不同类别客户的数量统计表*（根据上文等号不会和规则不同，合理即可5分）*

|  |  |
| --- | --- |
| 客户类别 | 数量 |
| 111 | 50 |
| 112 | 33 |
| 121 | 4 |
| 122 | 12 |
| 211 | 50 |
| 212 | 52 |
| 222 | 6 |

(4)222类客户为近期消费、消费频繁且额度高的重要价值用户，客户类型如下表所示*（描述不同类别客户的特征，合理即可，4分）*超市的重要保持和重要价值用户较少，建议组织相关促销活动，回访客户，提高用户忠诚度。*（给出相应的营销建议，合理即可，4分）*

|  |  |  |  |
| --- | --- | --- | --- |
| **客户类型** | **R** | **F** | **M** |
| 最近一次交易时间差 | 一段时间内的消费频次 | 一段时间内的消费金额 |
| 重要价值用户 | 1 | 2 | 2 |
| 重要保持用户 | 2 | 2 | 2 |
| 重要发展用户 | 1 | 1 | 2 |
| 重要挽留用户 | 2 | 1 | 2 |
| 一般价值用户 | 1 | 2 | 1 |
| 一般发展用户 | 1 | 1 | 1 |
| 一般保持用户 | 2 | 2 | 1 |
| 一般挽留用户 | 2 | 1 | 1 |