# 数据介绍

该数据集是模拟 Starbucks rewards 移动 app 上用户行为的数据，包括顾客特征，推送消息的属性，以及用户触发的事件（如打开消息、购买等）。顾客收到的推送可能是不同的，一些顾客可能一连几周都收不到任何推送。每种推送都有有效期，可以认为顾客在有效期内都可能受到这条推送的影响。也有可能顾客购买了商品，但没有收到或者没有看推送。商家发送的推送类别有三种，第一，买一送一类优惠信息推送，标记为BOGO；第二，折扣类优惠信息推送，标记为discount；第三，无优惠纯信息推送，标记为informational。

该数据集包含三个数据文件

portfolio – 包括推送的 id 和每个推送的元数据（持续时间、种类等等）

profile– 每个顾客的人口统计数据

transcript – 交易、收到的推送、查看的推送和完成的推送的记录

1. **portfolio表**

|  |  |  |
| --- | --- | --- |
| 字段名称 | 字段含义 | 举例 |
| id | 推送的id |  |
| offer\_type | 推送的种类 | BOGO、discount、informational |
| difficulty | 满足推送的要求所需的最少花费（成本） |  |
| reward (int) | 满足推送的要求后给与的优惠（折扣力度） |  |
| duration | 推送持续的时间，单位是天（优惠持续天数） |  |
| channels | 信息发送的渠道 | 电子邮件，移动，社交，网络 |

1. **profile表**

|  |  |  |
| --- | --- | --- |
| 字段名称 | 字段含义 | 举例 |
| age | 客户年龄 |  |
| became\_member\_on | 客户注册该程序日期 |  |
| gender | 客户性别 | 男性M、女性的F、其他O |
| id | 客户ID |  |
| income | 客户收入 |  |

1. **transcript表**

|  |  |  |
| --- | --- | --- |
| 字段名称 | 字段含义 | 举例 |
| Event | 记录的描述 |  |
| Person | 客户id |  |
| gender | 客户性别 | 男性M、女性的F、其他O |
| Time | 时间（H） | 从推送发出开始计时 |
| value | 该笔交易数额 |  |