的属性或者关键指标。确认事件与变量可以通过AARRR(Acquisition Activation Retention Revenue Referral)海盗模型或者UJM(User Journey Map,用户旅程图)模型进行逐步拆解,理清用户生命周期和行为路径,抽象出每一个步骤的关键指标。

AARRR模型和UJM模型会在第6章详细讲解。

2. 明确事件的触发时机

不同的触发时机代表着不同的事件计算口径,因此触发时机是影响数据准确性的重要因素。以用户付款为例,是以用户点击付款界面作为触发条件,还是以付款成功作为触发条件进行埋点呢?二者口径不同,数据肯定会有一定差异,因此明确事件的触发条件非常重要。

在用户付款这个例子中,笔者建议使用两个字段记录用户付款行为:一个字段记录点击付款界面这个行为,另一个字段记录是否付款成功。

3. 明确事件的上报机制

上报机制也是数据准确性的重要影响因素之一。客户端上报数据可能会由于网络连接原因出现丢包的情况,前面章节已经详细介绍过上报机制之间的异同,这里就不再赘述。数据分析师在完成埋点工作的时候也需要确定数据是实时上报还是异步上报,以确定埋点是否合理,并及时调整数据埋点方案。

4. 统一表结构

统一数据表结构,可方便团队内部进行数据的管理和数据复用,建议在团队内部形成一套统一的数据结构规范。例如,将表分为不同的层级,第一层记录用户的基础信息,包括用户ID、地区、昵称等;第二层记录用户行为信息。

5. 统一字段名规范

有了统一的数据表结构规范还不够,统一数据命名规范也是数据埋点工作的重要一环。如果有条件的话,可以建立数据字典,以统一数据命名规范。例如,确保同一变量在所有的数据表中都用统一的字段名。