

数据产品经理必备基础知识

最近几年随着内卷加剧，不少的研发和产品小伙伴转行当数据产品经理去了，目前相对来说竞争压力较小且行业需求越来越大；下方内容虽然是数据产品经理必须掌握的基本数据知识，但实际上作为非数据产品经理来说这些内容也是需要掌握学习的，只有掌握了这些数据基础知识才能够说入门数据产品经理。



一、基本术语

1. 埋点：一般意义上的埋点，是利用如友盟、Talkingdata 等第三方公司的 SDK，通过在 App 中嵌入一段 SDK 代码，设定触发条件，当满足条件时，SDK 会记录日志，并将日志发送到第三方服务器进行解析，并可视化地呈现给我们，过程就叫埋点。

2. 维度：维度就是指我们平时看事物的角度，也可理解成分析一个数据能从哪些方面去分析，这些“角度”都是有值且可被枚举的。

比如我们注册用户数有 10 万，那可分析的维度有：用户所在省份、用户性别、用户角色、用户来源等。不同维度来观察数据，可以得出不同结论，能否拓展观察维度，也是评估数据分析能力的一个关键。

3. 度量：度量和维度相辅相成，是指可量化的数值，用于考察不同维度观察的效果，也可理解成“数据指标”。观察度量值可总体查看，如 App 总用户数，也可配合“维度”分层查看，如不同省份的注册用户数、活跃用户数，不同来源的 App 启动次数、平均日使用时长等。

4. 渠道：指 App 的不同安装来源，如通过第三方应用市场安装，通过广告点击安装，通过地推二维码扫码安装，通过官网下载安装等。互联网公司的商务工作一般就是拓展渠道，观察不同渠道带来的数据表现，不断优化渠道质量。

二、基本指标

1. 网站来源分析：来源渠道、PV、UV、IP；跳出率、平均访问时长、平均访问页数；使用设备、操作系统等

2. 用户行为分析：年龄、性别、职业、地域、婚否、收入、兴趣、学历、关注点等。

3. 网站订单分析：转化率、成单量、平均订单金额、客单价；付费金额、付费毛利、付费利润；当前新增购买用户数、累计购买用户数、重复购买率等

4. 核心数据：拉新（新用户）、活跃、留存、转化等

5. 新增用户：安装 App 后，首次启动 App 的设备数，需要按“设备号”去重。新增用户主要为了衡量推广效果，以及当前产品在整个生命周期所处阶段。

6. 活跃用户：时间段内，启动过 App 的设备数，需要按“设备号”去重。活跃用户主要为了衡量运营效果，以及产品使用情况。

7. 启动次数：时间段内，启动 App 的次数，无需去重。启动次数主要为了衡量推送效果，以及 App 的内容是否足够吸引人。

8. 留存率：时间段内的新增用户，经过一段时间仍启动 App 的用户，占原新增用户的比例。“时间段”的划分方式有：按日、按周、按月，对应指标还可细分为“日留存率、周留存率、月留存率”。

9. 使用时长：时间段内，从启动到结束 App 使用的总时长。所谓“结束 App”，通常指杀掉进程，或者将 App 退到后台超过 30 秒。一般会按“人均使用时长、次均使用时长、单次使用时长”分析，衡量产品粘性和活跃情况。

10. 使用频率：用户上次启动 App 的时间，与再次启动的时间差。使用频数分布，可观察到 App 对用户的粘性，以及运营内容的深度。

三、基本技术

1. 数据采集—抓包

“抓包”，一般指观察 App 上传到服务器上的数据都有哪些。通过“抓包”观察，一方面可判断自家 App 是否正确上传了想要统计的数据，另一方面还可抓取到手机上安装的其他 App 的上传数据，用来分析竞品内容更新情况。

2. 数据提取—sql 语言

sql 语言一般用于从数据库中进行数据的增删改查，需要企业运维人员或 DBA 人员开启权限才可访问，大公司的产品经理基本没机会用到，如果是小公司的高级产品经理，且和技术商议仅开启只读权限，还是可以尝试使用的。

3. 数据处理—Excel、Python、JS

提取出来的数据，要深入分析，肯定得进行二次加工。按使用的难度高低，需要掌握工具如下：

Excel: 极其普通的 office 工具，有着极其强大的数据处理能力。常用数据分析功能有透视表和命令行。

Python、JS: Python、JS 是一种通用脚本语言，不止适用于数据分析，但由于其安装、使用方便，函数库丰富，适合有开发基础的同学尝试。

四、统计分析方法

1. 比较法: 多个数据提取进行比较。根据对比方法不同，分为“横向对比”和“纵向对比”。横向: 指空间维度的对比。相当于一个指标，在不同条件下的对比，但每个条件都属于一个层级。纵向: 指时间维度的对比。包括“同比、环比”。同比: 指本期数据与上年同期数据对比。环比: 本期统计数据与上期比较。

2. 分解: 某个维度对比后发现问题后需要找原因时，需要进一步分解指标。如: 某日的销售额只有昨日的 50%，就需要对销售额指标拆分为: 成交用户数 \times 客单价，而成交用户数 = 访客数 \times 转化率。

接下来就可分别针对：访客数、转化率、客单价，观察今日和昨日相比的数据变化，找出原因。

3. 降维：当维度太多时，我们不可能全部分析，这时就要筛选出有代表性的关键维度，去除掉那些无关数据，这就是“降维”。如：“成交用户数 = 访客数 x 转化率”，当同时存在这三个指标时，其实我们只要三选二就能得出结论了。

4. 增维：增维和降维是相对的，如果当前观察的维度无法解释当前问题，就需要对数据进行运算，多增加一个指标。在可视化分析领域，也可将不同类型的图表嵌套使用，能达到增加信息展现维度，扩展分析广度的目的，

5. 分组：也可叫聚类，合适的分组能更好地理解业务和场景。如用户画像创建，就是一个按不同维度对数据分组的过程。通过用户画像，可以很清晰地知道产品的用户地区、用户兴趣、用户年龄、用户性别等属性占比，产品经理可通过画像进一步了解用户需求。

6. 漏斗分析：主要用于分析产品使用的关键路径，通过设定一系列操作步骤，统计每一步中的操作用户数，并将用户数以柱状图纵向排列，就可形成用户流失漏斗，我们可分析漏斗每个环节的流失率，并观察改进环节交互体验后，流失用户的变化情况，以此来验证改动效果。

7. AARRR 模型：该模型不仅仅适用于游戏数据分析，是 Acquisition、Activation、Retention、Revenue、Refer，这五个单词的缩写，分别对应一款移动应用生命周期中的 5 个重要环节。AARRR

本身是一个循环，使用者需观察每个环节的数据情况，以此来分析产品是否在执行一个正循环过程。这其中的任一环节除了问题，都会导致产品数据的异常。