

Analysys·数据说

用户留存分析白皮书



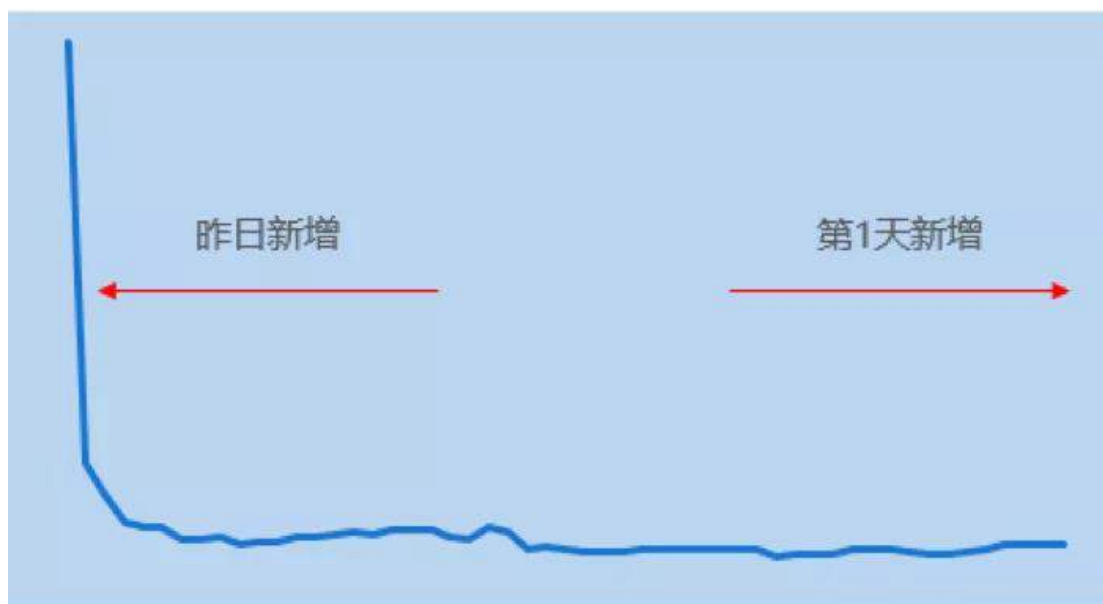
留存分析的重要性

在过往的互联网从业经历中，曾遇到一个特别棘手的问题：每天都有 2-3 万的注册用户，可是为什么每天的活跃用户总数却不见增长呢？

出于常识的解释是，用户规模不变，那一定是用户来了又走了，新增多少就流失多少。于是很自然的从流失着手，做了诸如流失人群画像、流失原因分析、甚至流失行为预测等调研。而在所有研究方向中，有两组数据在运营人员中得到广泛的关注，形式如下：

- 昨日活跃的 75 万用户中，都是哪天注册的新增用户？有多少是上周注册的？有多少是一年前注册的？
- 去年今日新增的 2.5 万用户中，还有多少在今天活跃？

通过前一个问题得出广泛引用的结论：每天的活跃用户都是由之前每一天的新增用户的存量用户数加和。并且里当前日期越近，占的比重越高。



每日活跃用户中的注册时间分布

而通过另一个问题则发现同一批新增的用户，随着日期增长，每日仍活跃的数量呈现快速衰减的走势（节假日或特殊活动时会出现反弹）。



这两个问题都指向一个关键词：**用户留存，而不是用户流失。**

留存分析即研究一个或一批用户，在首次使用产品之后，之后每天的活跃情况。活跃天数越多，说明用户质量很好、粘性高、生命周期长；反之说明用户质量低、粘性低和生命周期短。

通过留存分析的数据，得出的计算简便并且非常有效结论：

1、新用户普遍在前几天流失

数据显示新用户在第二日就流失 30~50%，一周后仅剩 10%也是非常普遍的，并且很多用户根本没有使用到产品的核心功能就流失了，十分可惜。

基于这个结论，近些年来，运营活动和产品设计方向上越来越针对新用户的前一周做关怀或新手任务，强化用户去体验到产品的基本功能。例如信用卡首刷礼，互联网理财产品赠送新用户加息券。

2、各渠道导入的新用户留存差异巨大，渠道策略有待提升

将所有渠道的留存比例排序后，可以非常容易发现某些渠道带来用户规模较大，却留存却很低，主要原因分就是渠道质量较低或者渠道作弊。停止向这些低效的渠道投入费用，将节省很多营销费用。留存分析就成为优化渠道策略的必要工具。

3、根据留存率计算用户生命周期

计算每个用户的生命周期时，通常采用用户最后一日的活跃日期减去首次访问日期，除以当日新增用户就得到所有用户平均生命周期。这个结果恰巧是每日的留存率之和。

4、需要通过同业对比确认是否正常

通常来讲，每个产品的新用户留存水平千差万别，但是在一个行业内留存水平往往存在趋同的现象，而不同行业的基准值也相差悬殊。例如游戏、视频等行业的留存就会高于证券、电商等行业。那么与同业的产品对比后，就会发现自身是否健康。

如何提升转化与留存

在互联网发展的初期，流量相对价格低廉获客成本低，大家的关注点在于如何快速的获客，实现跑马圈地式的增长，线上线下各种各样的方式投入大量资源购买流量，寻找最优质的获客渠道。

但随着移动互联网覆盖人群增速逐渐放缓，人口红利消失，流量成本激增，以前粗狂的获客方式不再管用，ROI 持续走低。获客逐渐开始从之前的“跑马圈地式的无序增长”，逐渐转入到“有质量的稳定增长”，留存超越引流，成为最关注的话题点。

本文的讨论，主要围绕以下内容：

- 留存分析为什么对产品是非常重要的？
- 做留存分析之前需要做哪些工作？
- 如何解读留存图和留存曲线？
- 如何找到提升留存的入手点？

1、为什么用户留存对产品至关重要

在这张图片里，我引用了一个小学时候很常见的一道数学题，就是一个蓄水池，一边放水，一边进水，然后问多长时间能够把水池灌满...

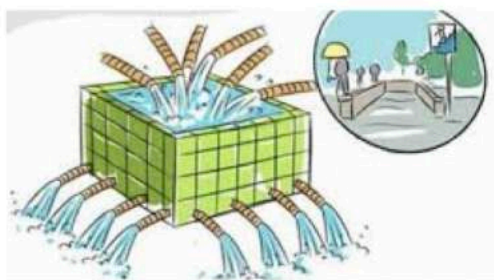
1、（头条号：方老师数学课堂）已知某水池有进水管与出水管一根，进水管工作 15 小时可以将空水池放满，出水管工作 24 小时可以将满池的水放完；对于空的水池，如果进水管先打开 2 小时，再同时打开两管，问注满水池还需要多少时间？

解：设注满水池还需 x 小时。

$$\frac{1}{15} \times 2 + (\frac{1}{15} - \frac{1}{24})x = 1$$

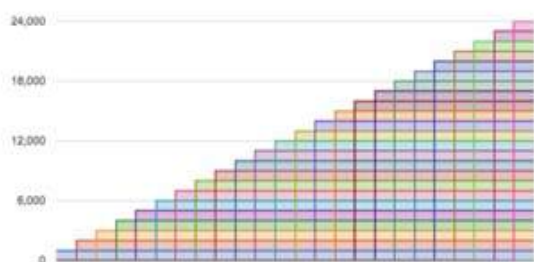
解得： $x = \frac{104}{3}$

答：注满该水池还需要 $\frac{104}{3}$ 小时。

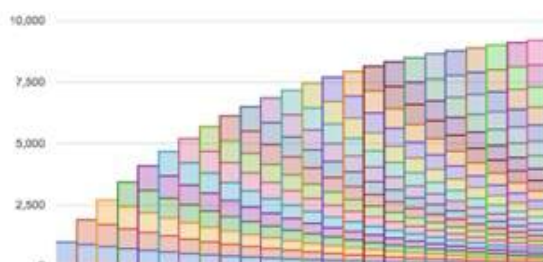


很多产品力都有类似的情况，每天都会新增用户，每天也会有流失的用户，池子里剩下的水就是我现在每天的活跃用户。其实做留存分析也是一样，分析多少用户进来了，多少用户流失了，他的情况是什么，对整个产品的影响究竟是什么至关重要。

为什么说“留存”是产品能够活下去的一个非常关键点呢，虽然用户每天都会不断地补充，但是如果你不能做好留存，用户进一个走一个，其实你的产品永远不能做增长。这张图中两个图很客观地说明了留存对于一个产品用户增长的重要性。



每月新增1000个用户，24个月后，总数为24000



每月新增1000个用户，留存率90%，24个月后，
每月大概9000个，然后呢？，不再增长！

左侧这个图，假设每月新增 1000 个用户，每月新增没有任何人流失，那四个月之后，大概会有 24000 个用户，但在实际情况中，这种假设根本就不可能实现，每个月都会有人进来，也有人流失。

假设每个月新增只有 1000 个用户，且留存率做到 90%（这已经是一个非常高的值了），24 个月之后会发生什么呢？答案是 24 个月之后，产品的用户总数就不会再发生大的增长或者大的变化，因为每月 9000 加上 1000 是 10000，10000 流失 10% 就还是 9000。反正从 24 个月之后我的产品几乎不会有大的增长，这就是为什么说“留存”其实是一个产品活下去的关键。

做留存分析之前需要做哪些工作？

既然留存这么重要，怎么去做留存呢？那其实在做留存分析前有两个点需要提前明确。

- 产品内的关键事件是什么？关键事件是用户感受到产品核心价值所采取的行动，用户不断的重复关键事件，证明产品对于用户在不断的产生价值，行程良性循环；
- 用户平均多久会使用一次你的产品？使用周期决定了你观测留存的最终结果，比如一个电商 App，大部分用户不会每天都在买买买，因此观察电商类 App 时使用日留存毫无意义，因为电商的客户不可能每天都在买，一般都是看 30 日或者 60 日的留存。

首先，在我们确定产品关键事件时，我们可以通过两个点来思考：

- 在你的产品内，你希望用户每次使用时都会做什么？
- 对于你的业务，哪个指标是能够直接衡量发展状况的或者是你最想它保持持续增长的？用户的哪个行为和这个指标最相关？

| | | |
|--------|-----------|------|
| Airbnb | 预订房间的累计天数 | 预订 |
| 京东 | 销售额 | 下订单 |
| 方舟 | 续约 | 创建分析 |

在这里举几个例子：比如 Airbnb，他的业务衡量的核心指标之一是“预定房间的累计天数”，那跟这个房间天数相关的用户行为肯定是用户“预订”行为。京东的业务衡量的核心指标是“销售额”，那个销售额相关的用户行为可能就是“下订单”。而对“易观方舟”这种偏向 SaaS 类的产品，“续约”其实是最重要的一个业务指标，跟续约相关的就是用户创建分析的这件事情，只有用户不断的在产品内创建分析，并通过分析结果优化产品和业务，才会证明这个产品对他是有价值的，只有这样他才能不断地进行续约。

通过上面的分析，基本上可以确定产品内的关键事件是什么，下面再明确产品内的用户平均使用周期究竟是多长，就可以完成一次最基本的留存分析了。

其次，如何确定产品的使用周期？一般的计算方法是观察用户 60 天或 90 天的数据，将重复 2 次关键事件的用户分布情况按照不同的使用周期计算出来，当用户分布达到 80%以上

时所对应的周期，我们可以认为是产品在用户侧实际的使用间隔，进而将这个周期作为留存分析的周期。怎么去理解这件事情呢？

| | | | | | | | | |
|-----------------|------|-----|-----|-----|-----|-----|-----|----|
| 2 次关键事件发生的时间间隔 | 84 | 56 | 30 | 14 | 7 | 3 | 2 | 1 |
| 做过 2 此关键事件的用户分布 | 100% | 855 | 81% | 40% | 20% | 14% | 10% | 3% |

例如，在过去的 60 天内，有 100 个用户可能做过两次产品内的关键事件，然后我把这 100 个人挑出来去计算他们做第一次和做第二次关键事件之间的时间间隔。统计发现，3% 的人一天之内做了两次，有 20% 的人在 7 天之内做了两次，在第 30 天的时候，已经有 81% 的用户都做过两次以上。

我们基本上就可以认定 30 天是用户侧对于产品的一个实际反馈的平均使用周期，它就可以作为我们留存分析里面所观测到的留存的一个周期状态，也就是月留存率，看这个指标可以得到的一个相对有意义的分析结果。

确定产品的关键事件和留存周期之后，就可以通过工具计算用户的留存情况了。

如何解读留存分析曲线？

常见的是这种数据留存图表，但很多人其实只会看留存表，不太了解背后的计算逻辑是什么，简单解释一下：

| | 第 0 天 | 第 1 天 | 第 2 天 | 第 3 天 | 第 4 天 | 第 5 天 | 第 6 天 | 第 7 天 |
|-------------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| 所有用户 9,701 位用户 | 9,701 | 290 | 83 | 60 | 32 | 23 | 0 | 0 |
| 2018年9月12日 1,752 位用户 | 1,752 | 73 | 27 | 14 | 9 | 23 | 0 | 0 |
| 2018年9月13日 1,852 位用户 | 1,852 | 54 | 21 | 12 | 23 | 0 | 0 | |
| 2018年9月14日 1,717 位用户 | 1,717 | 60 | 14 | 34 | 0 | 0 | | |
| 2018年9月15日 1,254 位用户 | 1,254 | 49 | 21 | 0 | 0 | | | |
| 2018年9月16日 1,307 位用户 | 1,307 | 54 | 0 | 0 | | | | |
| 2018年9月17日 1,819 位用户 | 1,819 | 0 | 0 | | | | | |
| 2018年9月18日 0 位用户 | 0 | 0 | | | | | | |

以这张表为例，查看 9 月 12 日到 9 月 18 日这 7 天用户的留存的情况，可以计算出来 9 月 12 日一共新增了 1752 个用户，而在第 3 天的时候只有 14 个用户又来进行了访问，相当于 9 月 12 日对应的第 3 日留存就是 $14/1752=0.8\%$ 。以此类推，可以算出来所有的日期对应的第 3 日留存是什么，进而可以得到产品内的所有用户的第 3 日留存。



有了上面那张表以后，就可以画出产品内对应本次分析周期的一条用户留存曲线。我们希望留存曲线在一段时间之后尽量保持平缓，而不是处于下降的态势。因为只有当留存相对平缓，才能够证明已经在市场上找到了与产品相匹配的用户群体，也才能证明产品在市场上具有一定的价值。



当然，仅仅满足于此还不够留住用户，通过持续的用户运营活动提升使用粘性、产品改版提升用户体验等，也是提升用户留存的有效方式。当做了一些运营活动，或者完成一次产品改版之后，我们希望整个留存曲线是抬升的，意味着后续的这个产品的用户留存的基础和数量，相对之前都是有提升的。

如何找到提升留存的入手点？

当我们发现留存曲线并不是平缓或者整个留存率并不是特别理想的时候，我们可以分析下用户流失的原因，一般可能从五个角度来做一些留存分析和对比。

- 第一要对比不同用户群之间的留存情况。
- 第二要对比产品内不同功能之间的留存情况。
- 第三要对比用户不同行为之间的留存情况。
- 第四要对比不同维度下的留存情况。
- 第五要对比用户重复关键事件不同频率的留存情况。

举一个例子，帮大家去理解为什么从这五个角度来切入更容易得到留存分析的结论：

常见的用户类型

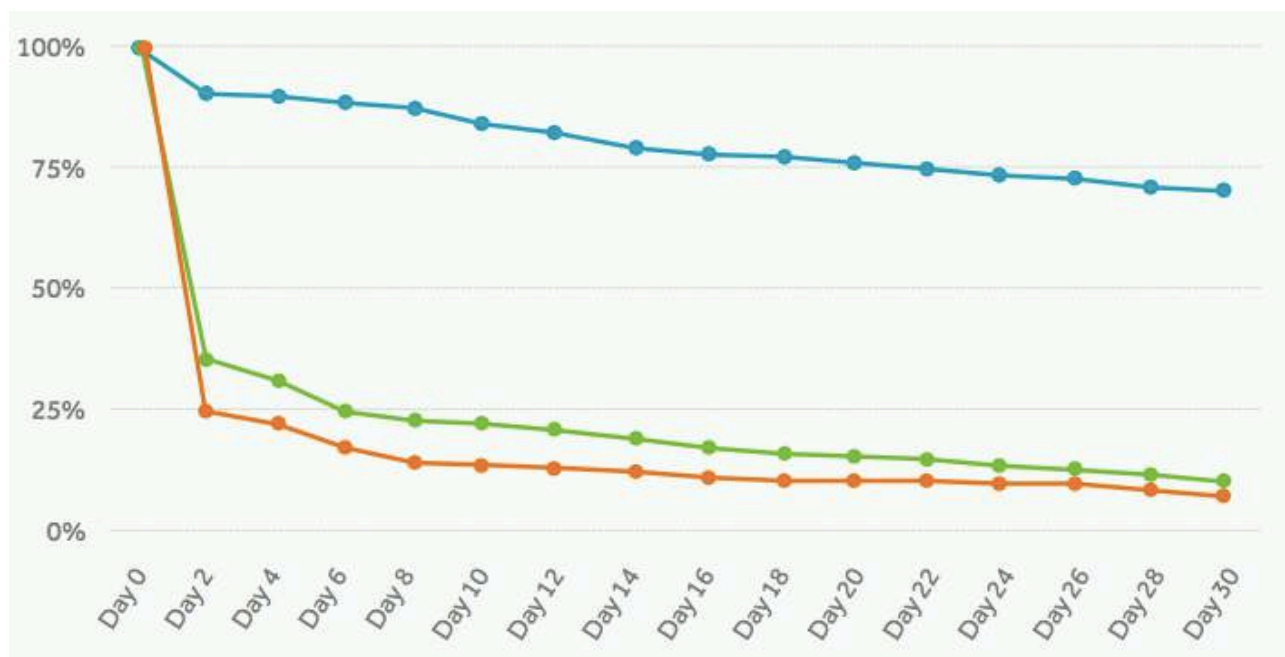


假设已经知道典型用户的使用周期，要针对不同人群去做一次留存分析，比如常见的用户类型可以分为活跃用户、新用户和召回用户三类，那我们怎么去定义这三类用户呢？可以将：

活跃用户：在这个周期内活跃了，前一个周期也活跃了的用户定义为活跃用户；

新用户：指的是第一次在这个周期内产生这种访问行为，定义为新用户；

召回用户：指的是在这个周期内活跃了但在前一个周期内并不活跃，甚至更早周期内它活跃过，那就把这一类客户定义为召回用户；



图例（蓝色：活跃用户，绿色：召回用户，橙色：新用户）

下面我们就要对这三类人群，去看他们分别的留存情况做对比。通过这张图我们可以发现，“活跃用户”留存率是最高的，其次是“召回用户”，表现最差的其实是“新用户”，整体留存率介于 70%到 30%之间。通过人群对比发现，其实不同特征的人群表现是不一样的，如果要提升整体留存率，就需要提升新用户和召回用户的留存率，同时还要将活跃用户保持住，也就意味着我们要保持蓝线的稳定，同时提升橙线和绿线。

下面再做更进一步的分析：

- 为什么活跃用户能够保持这么高的留存率？
- 为什么召回用户和新用户的留存率会这么低？
- 之前已经召回的用户，保持较高留存率的原因是什么？
- 是否能够挖掘更多的沉默的户，让他变成可召回用户？

通过提升召回用户的总量，并将新用户转化为活跃用户，最终提升整体的留存率。

以易观方舟产品的“活跃用户”为例，我们认为用户的 aha moment 是“创建分析”。其实在方舟里面创建分析其实包含了很多的功能，有可能创建的是事件分析，有可能是漏斗分析也可能是留存分析，我把这些行为分别做下对比，然后去观测做过这些分析的活跃用户和做过这些分析之后沉默/流失了的用户，他们的留存情况，那这时候我就能构建一组留存曲线对比。



这张图就是活跃用户和沉默用户之间做过事件分析的留存差异。围绕创建分析可能存在创建事件分析、创建漏斗分析、创建留存分析等等，将这些行为作为起始行为，来分别比较活跃用户和沉默用户之间的留存情况，我们找到留存差异最大的行为是——事件分析。然后紧接着，分析重复这个行为达到多少次以上时，对留存的影响最大。而后在运营活动、产品迭代上，让用户不断的去重复这个行为，把整体留存率做一个提升。

通过活跃用户的分析，我们找到了活跃用户一直保持这么高留存率的一个原因，就是他一直在去重复使用“事件分析”这个功能。“事件分析”的使用深度是对用户留存是有正向影响的，因此针对新用户我们可以做进一步的细化分析。

所有的产品，对于新用户的时候都应该有一个 Onboarding（上手引导）过程，用户可以通过 Onboarding，来熟悉整个产品基本的使用技巧，那么我们可以对比新用户里完成和未完成 Onboarding 的留存差异。我们发现没有完成 Onboarding 和完成了的用户差了三倍的留存，那我们基本就可以确定这个行为对新用户的留存有正向的影响。

我们要做的就是不断的通过产品的迭代或者体验的优化，能够让用户更顺利地完成 Onboarding 的过程，这样就可以有助于做好新用户的留存。当然 Onboarding 的优化可能只是其中一个因素，还有更多的因素需要我们在实际分析中发掘并不断进行优化。



最后，再分析下召回用户的留存情况，还是以易观方舟为例。当用户沉默一段时间之后又突然开始使用我们的产品，有可能是搞了一次运营活动，或者收到了我们的短信、EDM 邮件。

这个时候就可以通过留存分析找到：

- 究竟是哪种触达用户的形式导致了召回用户的留存差异？
- 召回用户回访来源之间的差异是什么？然后找到留存最好的回访来源；

通过以上几个角度分析，基本上可以找到影响召回用户的留存最重要的因素是什么，下一步要做的就是将他做的这些事情复制到其他沉默用户中，并且刺激这些沉默用户，然后让他们能够成为我们的召回用户。其实很多时候增长团队要做的事情就是不断刺激沉默用户，然后把他们变成我们的活跃用户，这其实是巨大的潜在增长的流量池。

小结：

本文中我们介绍了，留存分析为什么对产品是非常重要的？去做留存分析之前，如何去选择关键事件？如何确定产品的平均使用周期？以及，如何解读留存曲线，通过进一步不同角度的对比分析，找到背后的原因是什么，从而制定后续整个产品的增长策略，或者用户运营策略。

从“跑马圈地式”的无序增长，逐渐转入到“有质量的稳定增长”，产品的经营不仅仅需要持续得获客，还需要通过“精益运营”，提高留存的质量和效果，带来有效增长。

关于易观方舟：

易观方舟 (ark.analysys.cn) 是由易观 (www.analysys.cn) 推出的一款面向用户分析、产品运营和营销闭环场景的企业级数据智能平台，以数据驱动企业营销闭环效果评估、用户运营留存转化、产品精益迭代优化，赋能企业市场、运营、产品和数据分析人员。支持采集全端数据采集及 API 导入，支持代码埋点和可视化埋点，提供完整的 SaaS 及私有部署方案。还可通过易观自有数据增补多达 8000 多个细分数字标签，构建更加立体的全景用户画像。目前已经覆盖了银行、证券、新零售、电商、教育等行业，服务了首汽约车、宁波银行、上品商业等知名客户。如果希望了解更多，请咨询 4006 - 010 - 230，将有专业的人员为您解答。



试用产品



了解易观

易观方舟·数据驱动精益成长

地址：北京市朝阳区酒仙桥路10号恒通商务园B12C座三层
易观官网：<http://www.analysys.cn/>
易观方舟：<http://ark.analysys.cn>
客服电话：4006-515-715