

6 个步骤，学会数据驱动产品的秘诀

作为前 LinkedIn 的首位做变现盈利的数据科学家，张溪梦负责的业务就是如何能用数据分析来增加销售，促进产品的研发效率、做更好的风险控制等。这个世界上第二大的 B2B 的云端软件企业，每年的销售额从七千万美元跃升至将近 30 亿美元，一个重要的推手就是数据的帮助。那到底怎么样才能用数据来推动产品优化呢？

以下为分享正文，由人人都是产品经理团队 @张婷 依据嘉宾现场分享内容整理，编辑有修改：



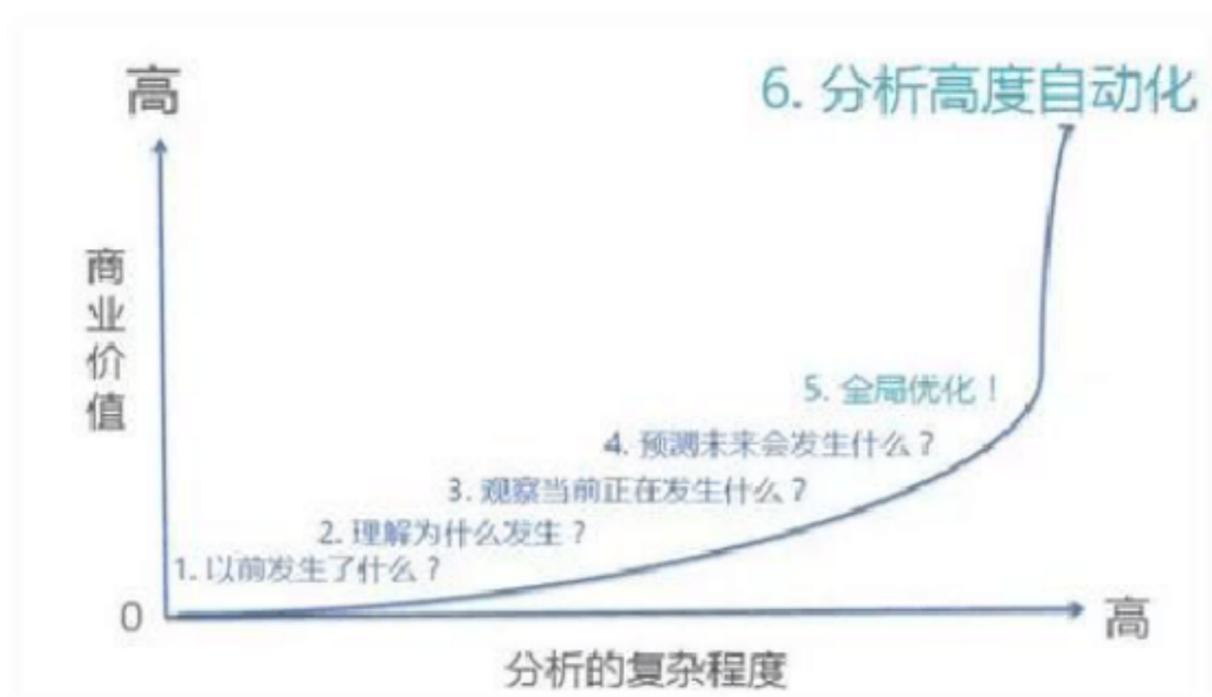
数据有四个属性：时间、地点、人物和事件；举个例子，在电子商务这个环境里，假设分析一个交易数据。这个交易数据在时间上，它必须有一个时间戳，这个数据必然会在某一个地点发生，还有对应的交易者，在事件的级别，消费者要购买某种产品，花某种价格，通过某种支付手段，进行交易。所以说，这是一个事件。真正的用户的行为，往往都是以事件为核心的数据点。这些数据点，会造成一个整体的事件流。

二 数据的几何级数增长

举几个例子，像 LinkedIn 这家公司每天大约会产生 100 亿条事件；通过用户行为产生的事件，在不到 5 年的时间之内，咱们全球大约会有 40 亿的人每天会产生大量的数据，同时，它是一个 4 万亿级别的市场；在全球范围内，将近有 2500 万个 app 诞生，不但有针对普通用户的 app，还有对企业级服务的 app；另外会有将近 250 亿活跃用户的各种设备将会接入我们的网络；最后，每年将近产生 5 万亿 GB 的数据。

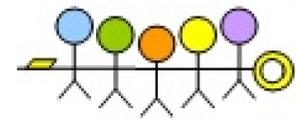
三 什么是数据分析

1、数据分析的基本概念



数据量这么大，呈爆发式增长，它的价值曲线，数据量，数据设备数，人的关联度不断增加，但这并不表示数据本身能够带来价值，真正的商业分析的核心，是通过 5 个步骤不断把低价值的信息或者数据变成高价值的产出。一般数据分析可以分成 5 个步骤。

- 了解以前发生了什么；
- 分析为什么会发生这件事；
- 观测目前正在发生什么事；
- 预测未来将会发生什么；



- 全局优化。

前四步是数据分析中的基础步骤，包括数据预测，它本身并不能产生太大的价值，只有通过全局优化，才能产生有价值的产出。

2、数据分析的发展历程

对于产品经理来讲，数据可以作为产品优化的依据，比如交互、留存。实际这一切都是基于对数据基础的理解，在过去的 5 年，数据框架在技术的推动下，有几次主要的迭代时期：

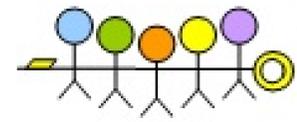
第一个时期

数据库和数据仓库的时期。传统的公司都使用数据库或者数据仓库来存储数据，他们应用一些数据库管理工具获取数据，再通过一些经典的 BI 软件，比如 Cognos 或者 Businessobject 进行数据可视化，形成报表发给管理层。一般对于这些数据，只有管理层决策层才使用。这个阶段发展了将近 15--20 年。

第二个阶段

第二个就是我们现在所处的阶段，至少在过去的五年，数据量呈爆发式增长。数据研究者开发出了一套分布式的技术，这套技术以 Google 和 Yahoo 为代表 ----Google table 和 Yahoo 的 Hadoop，它能处理大量的半结构化数据，然后通过分布式的方法，能并行地计算处理很多信息。

Google 和 Yahoo 的办法增加了数据处理的容量，大大降低了数据库的成本，分布式技术将数据库的技术提升了一个层次。以往数据分析只能支持几十个人的决策，现在能够支持几千甚至上万人的决策。它使得数据分析从一个非常贵的、缓慢的流程变成了一个有效的、价格相对较低的层次。



但这其中还有一个问题： 以往的软件都是比较结构化的， 很多人都可以使用。 但是 Hadoop 这种大数据结构， 很多人都很难用， 这就造成很多技术性瓶颈。 具体体现在：

- 数据收集来源增加，文件收集复杂性提高。
- 对工程能力操作感各个公司拆解很大。

3、数据分析驱动产品设计的意义

如果能够用好数据分析公司， 会产生一种战略性优势。 用不好数据分析的公司， 在发展上则会产生很多瓶颈。

在现在的美国和将来的中国，要想做好一个产品经理、设计师、产品运营人员。核心不再是

依靠分析师来分析数据， 而是能够熟练地应用各种企业级应用， 来达成数据驱动产品的手段。

如何利用数据驱动产品设计？

1. 必须要有数据驱动产品设计的意识，这个意识已经被美国的 Facebook 、 LinkedIn 证明；
2. 善于观察用户的行为， 用户的行为是不会说谎的， 用户的行为往往能够反映出用户对于产品的容忍度和熟练度；
3. 我们要很聪明的用技术做决策；
4. 今天的企业已经由流量型向用户体验和产品设计方向迅速进化和衍化。 这种进化和衍化的结果就是：每一个企业，不论大小，都需要用精益化运营的思维做好产品。

互联网已经从 web2.0 进化到 web3.0 ，意味着流量为王的时代已经结束；这样的形态下，

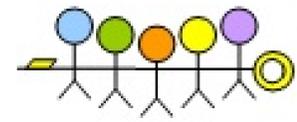
企业往往不注重流量的转化， 而在乎流量获取的速度。 而在今天这个时代， 一个企业如何能

够在一定时间内快速地对用户进行转化， 增加用户的体验， 增加用户的留存， 这是当代互联

网产品的核心。

四 如何通过数据分析驱动产品设计

精益化运营



精益化运营不得不提的三个指标 --- 转化率、用户的活跃度、用户留存率。这三点组成了一个产品最基本的一个骨架。

以往我们谈论的都是流量：你有多少用户，多少活跃度，仅此而已。但今天来说，那些粗浅的指标已经不足以衡量产品的优秀。一个好的产品，我们首先要关注的是用户的留存率和用户粘度。在传统互联网企业，用户的留存率和粘度一般是后期才会关注的；但是在今天，在产品设计的早期，我们就需要关注这个指标。这个指标可以细化成 5 个不同的点：

1. 能够通过非常简单的方法让用户知道这个产品的价值；
2. 产品使用的流畅度。今天产品设计的思路已经演变为流式的设计，一步步引导用户如何使用你的产品，降低用户的使用门槛很重要；
3. 产品的区隔。这也是精益化运营的思维。不知道大家在平时做产品的时候有没有考虑到用户是可以分成不同的属性，同样的产品，针对不同的用户应该有不同的功能点，不同的体现，这就要求产品经理和运营人员花足够多的时间去了解这些不同的用户对产品交互的不同的需求，喜欢什么，不喜欢什么；
4. 要分析用户的行为。用户的行为是产品经理了解用户最好的方法，用户行为的分析往往能代表最真的用户需求；
5. 结合前边五点的分析，对产品进行改造。

Facebook 提倡一句话：Move fast and break things。它讲的是一种黑客的文化——快速迭代。可以犯错，但要快速迭代；持续迭代，小范围改进；最后会产生几何级数的变化。传统的方法，则是建立一个较长的项目周期。这样来做的话会有几点风险。

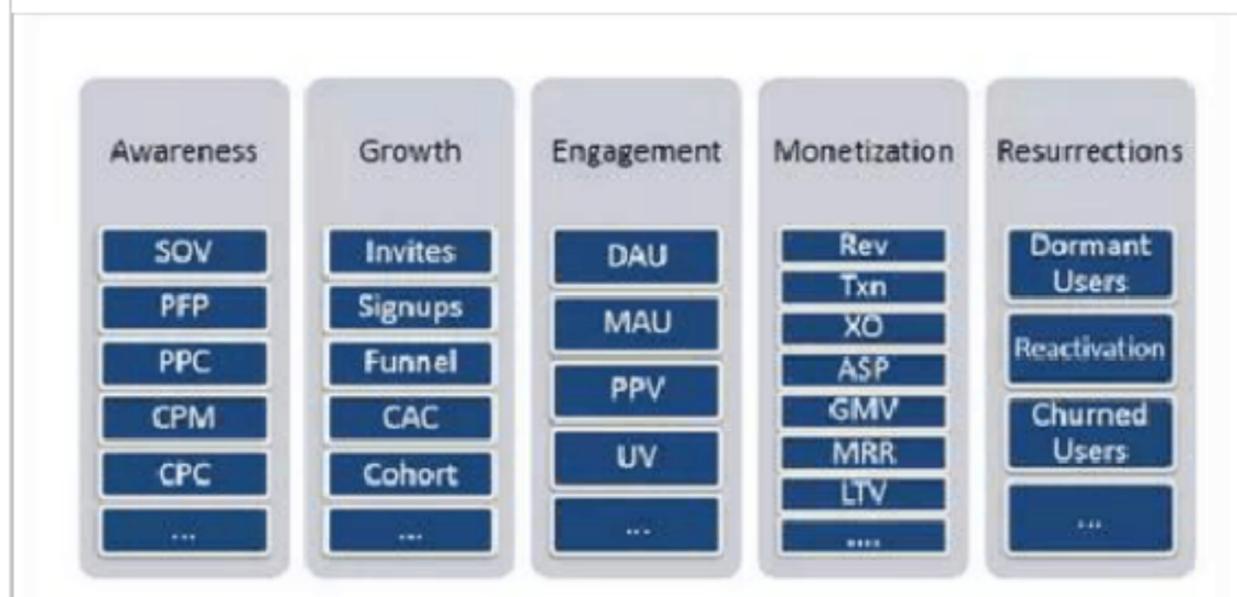
- 实施周期长；
- 迭代速度慢；
- 成本高；

- 不能在短期之内理解用户需求。

所以在过去几年，硅谷提倡 MVP（这个概念在《精益创业》中有提到）。一个核心的概念是：产品设计者如何能够产生一个最小的产品，迅速把它推向市场或用户；测试完之后衡量产品的数据，通过数据来验证假设，进一步完善产品；再进入下一个迭代的循环。

精益创业概念的核心，并不只适用于创业公司，其实在很多公司都有体现。比如 Facebook，就应用了这种概念，对产品进行小范围体验，数据分析，然后迅速迭代。

产品各阶段关注指标



在普通的传统的公司，当我们做数据分析时，很多人用 PV、转化率来衡量产品的好坏，而用数据驱动则会用更全面的数据来衡量产品。

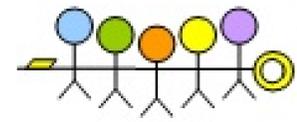
如何用数据衡量一个产品呢：

首先，它应该有一定的认知度，否则就不会有用户知道。SOV、PPC、CPM 这些词都是显示广告投放的；

其次，需要衡量用户的增长。这种增长的指标在美国已发展成为一种学科，有多少新增用户，你的漏斗模型是怎样的，你的霍克模型是怎么样，这个用户的存留区间是怎么样；

第三，用户活跃度。衡量每月，每天的活跃度，以及有多少独立访客

第四，衡量如何赚钱。比如电子商务网站，衡量的是 GMV--- 整个交易额，ASP--- 每个商品的平均价格，LTV--- 生命周期价值；这些 KPI 对产品分析是一个整体性的东西



一个产品，我们必须全面衡量。产品经理需要有这种直觉——不能用一个指标来衡量产品，而是要用多个指标来衡量。

比如，9年前，我刚加入 eBay，要衡量 150 个指标，每天发送一份报告，这份报告大概有 20-30 个指标，来衡量业务的发展。这些指标的增长或者降低都是由业务战略的决策造成的。

如果每分钟见证指标 0.1 的变化，对当时的 eBay 来讲，将是一个 5000 万金额的变化。

好的产品不能只看产品的 KPI，需要通过市场、运营、产品设计的角度以及最后销售的角度来衡量。

比如 LinkedIn，每天关注的是 50-60 个核心指标，公司 CEO 早晨起来的第一件事，就是要看这些 Report。在线营业额增加了 65%，是由于哪条产品线上的哪个功能引起的，为什么会带来这些变化，这些为什么会分解到每个产品线？产品经理会对此进行细化，找出原因，进而优化产品。

数据分析需关注指标以及指标拆解

首先，我们核心要关注的指标往往不是 150 个指标，或者 500 个指标，我们重点关注的指标往往不超过三个指标，这三个指标根据不同的业务线会有很大的差别。这三个指标有：

用户留存（用户活跃度）---- 最核心的指标；LinkedIn 会衡量月活跃用户和日活跃用户。

一般会把 MAU 分成三个档次 ----- 明天或下星期或下个月或今年我对用户的成长预期是什么，这是一个中位数。然后，还会做一个高点和低限，就是按照我们的预期，这个增长最高

不会超过这个水平面，最低也不会低于这个水平面。这样，我们把 KPI 分解成了四个不同的

KPI 值 ---- 平均值、高值、低限、还有实际值。这样说，实际值如果高于高点或者低于低

限，对于业务运营来说，都是一个很大的问号；

再举个例子，比如我把月度活跃用户进一步进行分解，分解成：新注册用户、既有用户今天

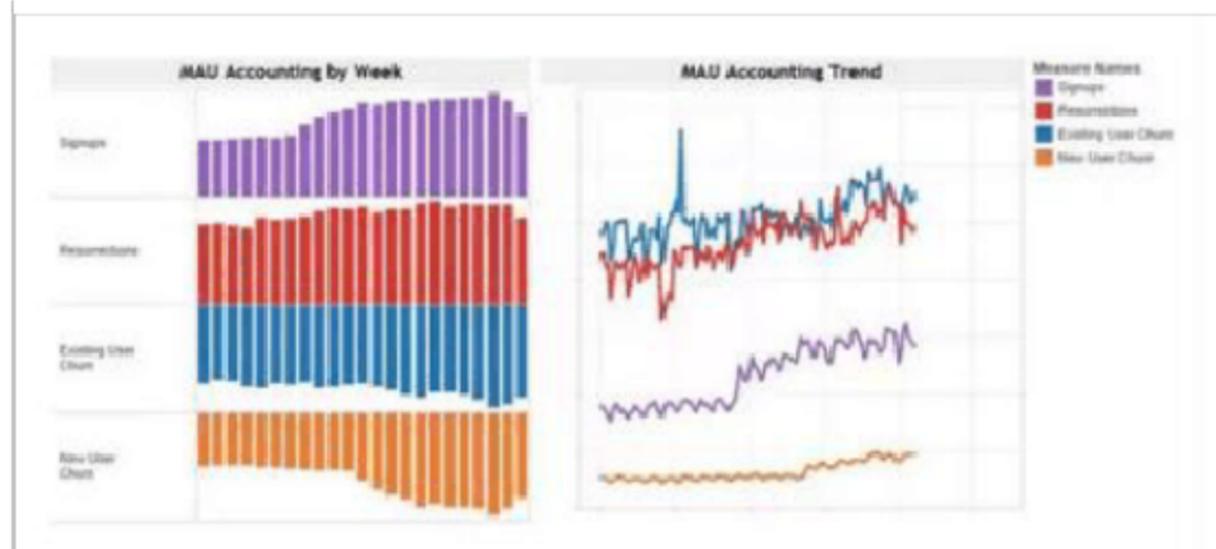
重新回来的用户、以往用户今天开始不使用的用户、上周新来用户这周不使用的用户。把月

活跃用户分解成了新接入用户，持续活跃用户，以往活跃今天不活跃，刚刚新增用户就开始不活跃用户。为什么将这四个指标分解成这四个指标，因为这四个指标可以和产品关联起来。

比如拉新，如何有拉新的手段；第二，如何促进用户的活跃。第三，已经流失的用户，如何减低它的流失，第四，新增的用户如何减少流失。这个数字相加，就是每月的活跃用户。这四个指标是可以采取行动的指标。

再比如，LinkedIn 新增的用户，来源的渠道是不一样的，活跃的用户，流失的用户，是我们需要持续关注的，因为它不断地把我们新增的用户活跃度减低，因此我们要尽量减低流失率。新用户的流失，在一个产品里面，非常忠诚的用户流失的可能性比新增用户流失的可能性低很多。但是如果我们的忠诚用户在迅速流失，那么说明我们这个平台上的用户的粘性有很大的问题。但是新增用户的流失，可以用培训的方法，引导用户的方法，促进它的活跃度。

这样，就把一个宏观的指标，分解成了更细化的指标。另外一点，分解指标的好处是，我们分解能产生增长或者衰竭的原因。比如下边这张图

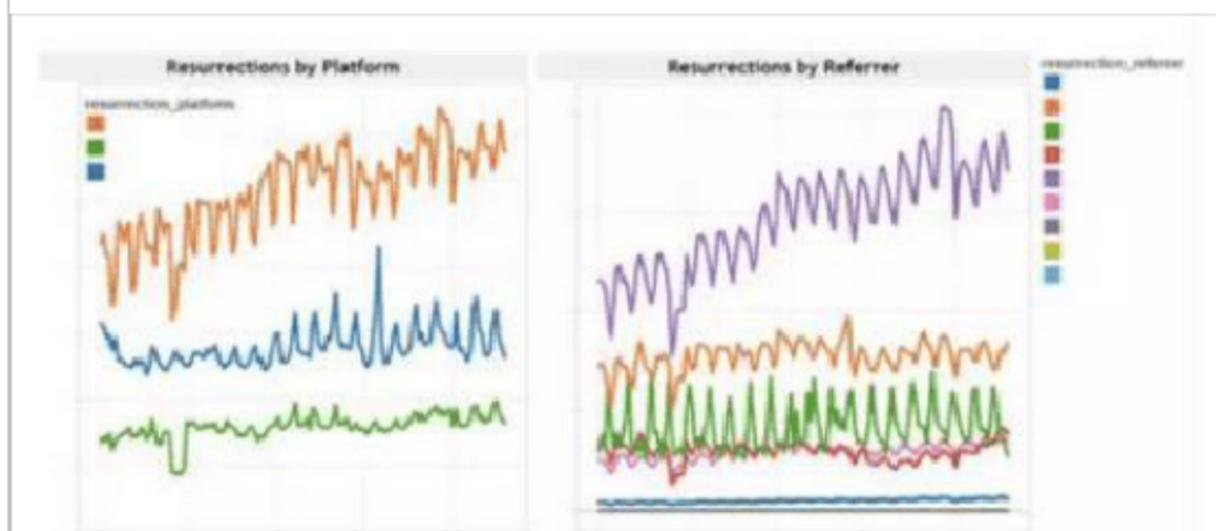


我们可以看到每条线，变化是不一样的。最后的增长就知道它的成因。

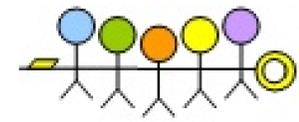


一个活跃用户，或者流失用户，可以进一步对它进行分解。这种方法就是通过维度的分解。

举个最简单的例子，比如说我把用户的流失率，用男性和女性的方法拆解开来，比如上图中，绿色是男性，红色是女性，大家可以看到，绿色的活跃度是高的，说明平台比较适合男性用户，女性用户偏少。如果我们想刺激活跃的话，可以针对这两种不同的人群采取不同的方法。



再讲一个分解核心 KPI 指标的方法，咱们讲一个促活的例子。比如说，用户很少用一个平台的产品，我们发现，当把这个指标进行拆解的时候，假设上图中橘黄色的这条线是 IOS 用户，蓝色的线是安卓用户，绿色线是 PC 用户，大家来看，我们把一个重新激活的用户分成了三个组，可以看见，IOS 用户在不断增长，它的活跃度是在不断增加的，安卓用户表示的线比较平缓，PC 线相对来说是衰减趋势的，这样说，在哪个平台上，你的用户是十分活跃的。



AARRR 以及产品指标拆解案例

在美国有一个概念 ---AARRR 。 A--- 获取用户 ， A--- 激活用户 ， R--- 用户的留存 ， R--- 用户产生多少价值。 R---- 用户与用户之间的推荐。 用户之间的推荐， 是最好的获取客户的方法。

对这个指标进一步进行分解，比如说在 LinkedIn 上的推荐，它是通过下面几种方法来实现的：

- 通过你可能认识的人来把 LinkedIn 推荐给一个用户；
- 通过上传地址录来推荐给别人；
- 通过电子邮件来做产品的推荐；

当我们做了这样的一个渠道的分类之后，我们发现，在产品上的模块 --- 您可能认识的人，产生了最多的推荐用户数量。这对于产品运营人员来说，可以进行各种优化。

给大家举个实例，在最早的时候， LinkedIn 研究如何使用户爆发式的增长，他们发现从首页上登录的用户活跃度和忠诚度比从 EMAIL 地址邀请来的用户的活跃度高的多。当时的产品经理做了一种尝试： 大范围优化从网站上来的用户。 怎么做呢，就是对每个人的简历进行优化，优化之后，使它更方便 SEO 爬虫的搜索，这样一来，从网站上获取的用户就远远大于从邮件获取的用户。它的核心的方法论是这样：

- 我要增长核心用户；
- 通过分析用户的来源， 找到了那些不断增长的活跃用户大部分来源于网页， 而不是从电子邮件来的；
- 迅速做一个产品尝试 ---- 优化每个人的简历，增加 SEO 的曝光率。

这次改变以后，为当时的 LinkedIn 增加了 50% 的扩张速度。这次成功尝试背后的一套核心观点是：如何用数据验证它的这种假设， 或者迅速验证这种业务的判断。 就是那么一个简单

的例子，带着一个假设去做一个产品的改动， 改动之后再数据证明自己的改动是有效率的。

如果这个改动是没有效率的，那么需要立刻停止这个改动。

再举个例子，也是今天我认为产品分析驱动产品设计最核心的指标 ---- 用户的留存度。一个

好的产品，是能够让很多用户在你的平台上产生价值的， 用户必然会花更多的时间会有更高

的粘度。这是今天衡量一个产品最核心的一个指标。



这张图，是一个衡量隔日留存率的图， 就是隔天有多少用户在使用。 这张图为什么很重要呢？

因为在很多产品里边有一些魔术数字， 很多人都尝试着在产品中找到这个魔术数字， 这个魔

术数字往往是从这张图里迭代出来的。比如说我的产品利用某个功能的隔日留存，还不到

10% ,利用另外一个功能的隔日留存变成了 30% ,不用某个功能的用户的隔日留存是 10% ,

用这个功能的用户的隔日留存提到了 30%。这个告诉了我们什么？

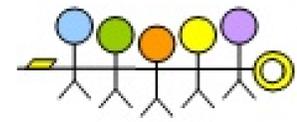
我们的产品里边，用户是有区隔的，不同群组的用户在用不同的功能

用某种功能的用户，对我们的平台或产品有很高的粘度和依赖度。 这个提醒我们，

要增加这种有高粘度的产品功能点， 让大量的用户能够用这个功能点， 潜在的或

者间接的就会增加用户的留存度。

总结一下：



通过留存图找到不同用户的群组在留存率上的区别。比如用功能 A 和用功能 B 的用户在留存上是有区别的，那么我们就要专注打那个功能点，让更多的用户来使用这个产品 FEATURE。

听起来很直接很傻的方法，但是被很多公司屡试不爽。LinkedIn 找到了这个魔术数字是：

在第一周连续增加 5 个社交网络链接的用户群，这类用户的留存度和他们在未来为平台贡献的价值，是第一周贡献少于 5 个社交连接的用户群的 5 倍以上。

发现这个魔术数字后，当时的产品经理立刻将这个增加链接的功能点做到了很多页面上去：

- 在首页的入口大范围增加了一个功能栏，让用户上传它的地址簿（因为已上传地址簿，自然会增加用户连接）；
- 给每一个新登录用户推荐和他在一个公司工作的他可能认识的人；

数据分析发现，当用户要邀请他的好友的时候，需要打字才能邀请。LinkedIn 想能不能不

要让用户打字，直接推荐给他可能认识的人，这个功能点再一次增加了第一周之内登录用户增加社交连接的可能性。

大家看，通过那一个点找到了 magic number，然后产品经理通过各种不同的产品设计来

促进这个魔术数字的发生，就产生了未来 LinkedIn 病毒式的增长。这种病毒式增长，就是

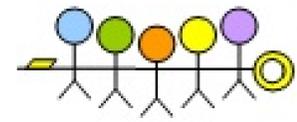
成长黑客最核心的一个概念：如何用最快的最低的成本驱动最小的产品改动，通过数据驱动的方法来大规模获取客户。

然后再讲讲如何用数据驱动用户体验：

首先，我们必须能分析今天的数据，这是所有做最小化产品，精益运营的最核心的一个概念。至少我们能够看到我们的产品有多少人在使用，如果没有数据的话，是无法证明自己脑子里的假设。

第二，最重要的一点，我们要找到最关键的问题进行思考，比如 LinkedIn 对产品做的这些

改进，通过数据分析的方法找到了问题：



- 从电子邮件邀请来的用户活跃程度低；
- 我们发现第一周登录增加了五个社交关系的人活跃度高，不增加的人活跃度低。找到了这个差异点，就要思考如何强力的推动某种改变。

在产品的早期，没有公司有能做 A/B 测试，往往是做很小范围的改动，随时看数据的变化，这就足够了。实验完了之后，又回到了数据上面，得到了分析结果，要记得看数据，最后，其实数据驱动是很简单的事，它只要求我们：

- 有信念；
- 要有一套非常基础的数据分析系统；
- 需要创新的精神，能够很迅速的进行产品迭代。

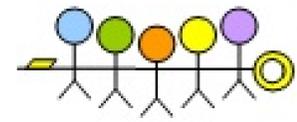
以下是小伙伴们现场提问的一些问题：

问：小公司是否应该使用 A/B 测试？

答：A/B 测试是产品经理一定会用的一个功能，但是在产品早期没必要做一套产品测试的 A/B 系统。

为什么呢？因为 A/B 测试对初创公司，是非常高的成本。需要同时上线两个版本，同时对这个网站以及 APP 的流量进行准确切割，而且还需要搭建一套非常实时的数据分析系统，或统计学检验的系统，特别是产品的版本控制需要做到非常精确。我觉得在早期，进行试错是非常重要的，但是并不是非得用 A/B 测试。比如新产品上线，我可以用 5 天时间来做产品 feature 的测试，是否能够被很多人喜欢，当你看到转化率有一个很核心的提升，那么我也能够知道这个用户的基本的反馈了。

但是，当产品达到越来越高的时候，尤其是当你的流量足够多的时候，就非常有必要做 A/B 测试了。还有，当一个公司里有大量的产品经理，有普通产品上线的时候，有一套 A/B 测



试就是非常重要的，因为其中一套非常重要核心的统计模型，比如如何切割流量，如何能分配 KPI，比如得到一个指标，是否有意义，还是需要这套系统的。

问：统计学分析学，数据挖掘的书看了不少，如何能做数据分析

答：这个取决于大家核心关注的东西。

如果你是一个产品经理或者产品运营，我觉得没有太多的必要大肆的读这方面的书。我建议

看一些入门级的书，比如说美国比较流行的《精益分析》，用简短的篇幅描述了不同形态的产品该关注什么指标；

如果你是专业的分析师，那该学的东西就太多了：统计模型、可视化工具、数据可视化等；

另外，对逻辑回归也要有一定的理解。还有，需要有商业直觉，并且必须用自己的产品。不用自己的产品的分析师，一定不是一个好的分析师。

问：如何收集自己需要的数据？如何保证数据的准确性？

答：这个是大多数 pm 的痛处。对于一个小型的公司来说，我强烈建议不要建太大的数据分析系统，要学会应用很多数据分析工具；比如：项目管理工具、电子营销工具。能熟练使用各种分析工具是一个产品能够灵活进行工作的基础。

对于如何保障数据准确这个问题，我认为数据不需要百分百准确，但是需要在趋势上一致。

因为好的产品经理需要利用趋势来判断用户趋势，银行需要百分之百的准确性，但是用户行为的数据只要趋势对就足够了

问：针对工具的 APP 有什么好的数据分析方法？

答：任何 app 它的核心的数据方法论没有本质上的区别。

AARRR 模型基本可以涵盖用户的存留激活转化营业额以及用户的推荐。只不过对每个指标

需要进行细化，最重要的一个指标是用户的留存率，这个留存率是如何体现出来的呢？就是

从用户如何使用 APP 中体现出来的。好的 APP 用户使用频次高，差点的 APP 用户使用频

次比较低。当然这个也和业务场景有关，比如一个旅游的 app，没有人天天去旅游，所以这个产品的粘度就和时间有很多关系。但是一个高频次的 app，比如订餐软件，使用频次就会高一些。你还可以将用户分成高中低三个层次，找到不同用户对产品不同的需求点。