# 第1章 数据分析方法论

## 1.1 概述

### 1.1.1 为什么要学习数据分析方法？

对初级分析师来说，在数据分析工作中，经常是拿到数据不知如何下手，或者拿到问题没有清晰的思路。因此，初级分析师通常会有如下问题：

**（1）问题一：**

* 症状：**没有**数据分析**意识；**
* 症状表现：经常会说我觉得、我感觉、我认为；

→ 问题所在：分析数据靠感觉，而不是靠数据分析。

**（2）问题二：**

* 症状：**只有统计**的数据分析；
* 症状表现：做了很多图表，却发现不了业务中存在的问题；

→ 问题所在：只能从图表中得出已经知道的现象，不会深入分析现象背后的原因，无法给出有建设性的意见和建议。

**（3）问题三：**

* 症状：只会工具的数据分析；
* 症状表现：学了很多工具（excel、sql、python、spark），但是面对问题，还是不会分析（听过很多道理，却依然过不好这一生）；

→ 问题所在：工具用的很熟练，但是没有分析思路。

**（4）问题四：**

* 症状：分析结果无法得到认可；
* 症状表现：给出的分析结论得不到业务认可，或者不知如何去表达；

→ 问题所在：呈现数据分析结果的方式不对，不能给人以清晰的思路。

这样，我们就要思考：

* 当我们面对一个问题时，怎样才能有分析的思路？
* 当我们面对一堆数据，如何展开去分析？
* 当我们面对实际业务问题，如何去回答？
* 当我们面对分析结论汇报时，如何有效表达？

因此，这里就体现了数据分析方法的作用，分析方法可以为我们提供解决问题的思路和答案，是分析师乃至每个人都必须掌握的基本技能。

### 1.1.2 什么是分析方法？

简单来说，可以通过下图直观认识到，分析方法就是把我们零散的想法构建为有条理的分析思路，就是**从具体问题中抽象出来的普遍规律**。

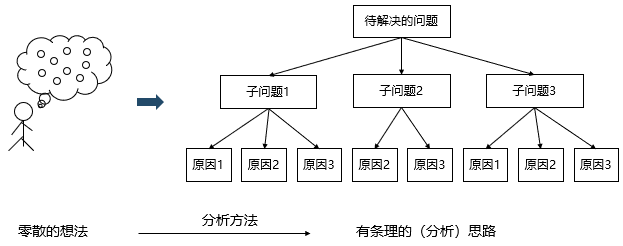


图1.1.1 零散想法构建有条理的分析思路

我们可以把数据分析工作类比盖房子的过程：

表1.1-1盖房子和数据分析有点相似

|  |  |  |
| --- | --- | --- |
|  | **盖房子** | **数据分析** |
| **要解决的问题** | 如何盖房子 | 某个业务问题 |
| **分析问题的思路** | 设计图 | 分析方法 |
| **技术工具** | 挖掘机等 | excel、sql、python、spark等 |

### 1.1.3 常用的分析方法有哪些？

总的来说，数据分析的方法非常多，不过万变不离其宗，不管有多少方法都可以纳入同一个体系下，**经过系统性的反思后，我们会发现这些分析方法都是普遍联系的，解决的是同一个问题**。

数据分析的一般路径就是这个体系，后面的各种方法都可以在这个体系下找到自己的位置。

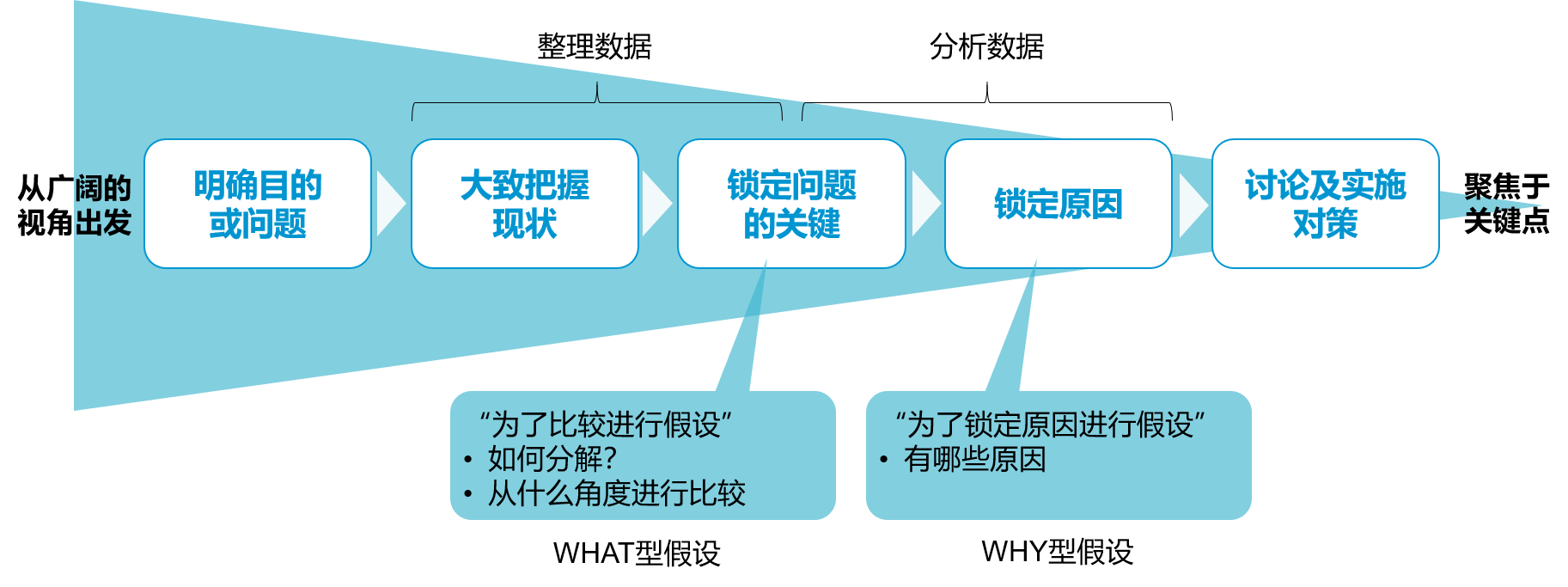


图1.1.2 数据分析应该这么玩儿

在我们遇到一个具体问题时，通常会用到以下方法的多种组合，只要我们以这个一般路径为中心，就能找到合适的分析方法，解决对应的问题。

表1.1-2 分析方法有很多

|  |  |  |
| --- | --- | --- |
| **分析目的** | **分析方法** | **实例** |
| 将复杂问题变得简单 | 逻辑树分析 | 宁德时代靠什么赚钱？ |
| 分析行业宏观环境 | PEST分析 | 新能源汽车的未来 |
| 多个角度思考 | 多维度拆解分析 | 如何降低新能源车的成本？ |
| 没有对比就没有伤害 | 对比分析方法 | 如何评价一家公司好不好？ |
| 如何分析原因 | 假设检验分析方法 |  |
| A和B有什么关系 | 相关分析 | 水泥的需求量的影响因素？ |
| 留存和流失分析 | 群组分析 | 账龄分析 |
| 用户价值分析 | RFM分析（聚类分析） | 航空价值分析 |
| 如何进行量化决策 | AHP层次分析法 | 专家评分卡 |
| 用户价值分类 | AARRR漏斗分析 | 福寿园客户的转化率 |
| 财务分析 | 杜邦分析 | 三类公司的财务分析 |
| 业务分析 | 波士顿矩阵 | 同一行业哪个公司更有投资价值？ |
| 行业分析 | 波特五力模型 | 茅台的盈利能力 |
| 战略分析 | SWOT模型 | 管理层的战略是否靠谱？ |

### 1.1.4 从哪些方面来介绍分析方法？

本专题首先介绍数据分析的一般路径，了解一个完整的体系。然后介绍具体的分析方法，分析方法主要通过以下四个方面展开：

（1）是什么？

（2）有什么用？

（3）如何用？

（4）注意事项

## 1.2 分析方法总论

### 1.2.1 数据分析的常见问题

1. **拿到数据就分析**：数据分析关注的不是这些代表结果的数据，而是造成这些结果的原因，分析的目的应该是针对“造成当下结果的原因是什么，需要采取哪些行动”等问题得出结论。拿到数据就分析，结果就是大多数工作都是徒劳的，效率很低。
2. **分析思路不清晰**：所有业务都出现同样问题的情况比较少见，大多数情况是只有其中的某些关键部分才是主要问题。在所有数据上做文章，由于多个要素掺杂在一起，会很难找到问题的关键；分析问题之前没有逻辑清晰地分解问题，造成后续进展不顺利。
3. **视野过于狭窄**：只关注眼前看到的内容，会使分析的范围局限在个人有限的认知维度内，这样可能会遗漏真正的问题。
4. **难以说服客户，让客户接收建议**：如果分析问题之前没有做好完整的规划，造成分析过程过于零散，无法构成完整的流程脉络，单纯给出几个结论很难说服客户接受。

### 1.2.2 数据分析的一般流程

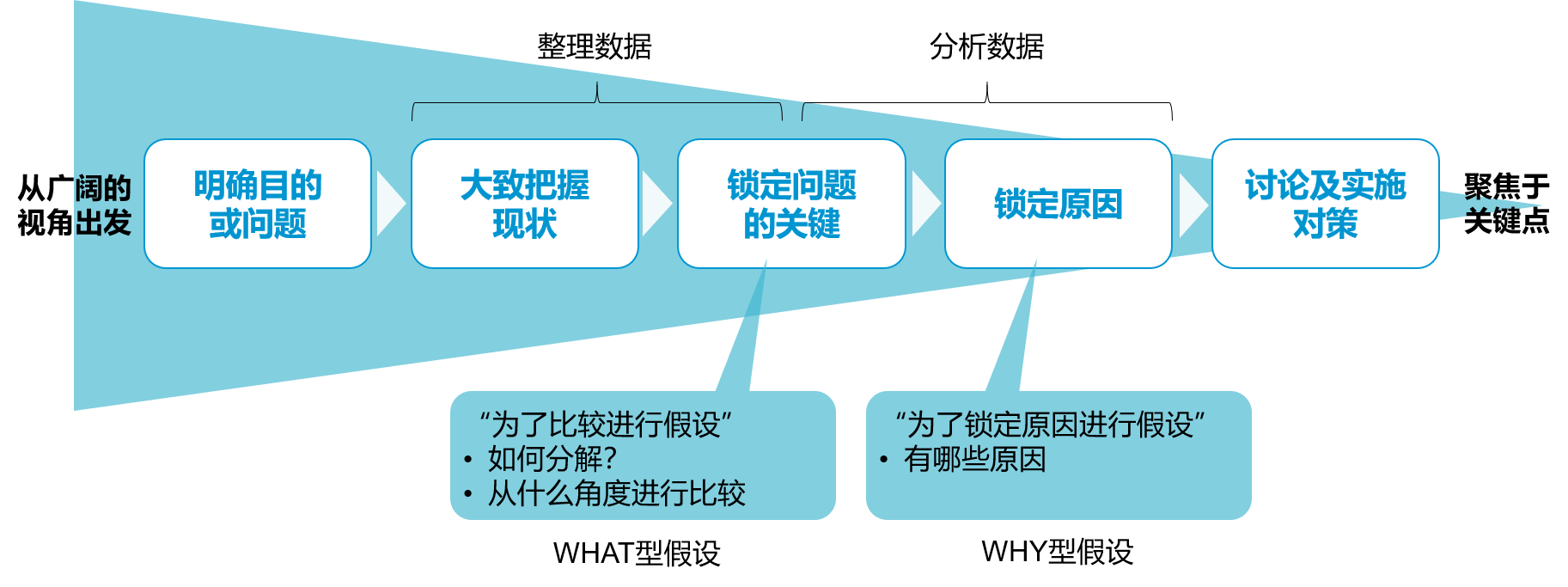


图3.15 数据分析的一般流程

为了保证数据分析的过程逻辑完整、思路清晰，解决数据分析中常见的问题，将数据分析的一般化流程归纳如图3.15所示，下面结合常见问题对整个流程作如下说明：

（1）**为了解决第一个问题，首先要做的是明确分析的目的和问题。**针对要解决的问题，有针对性地回到数据中找答案，而不是直接从数据中找答案（这样如同大海捞针，会使工作陷入停滞不前的状态）。确定目的或问题后，再决定假设、方法、所需数据等具体内容。

注：问题定义不要加入自己的设想，可以先考虑想要达到理想的状态是什么，在此基础上定义问题，即现状尚不满足这个理想状态。

（2）**为了解决第二个问题，进行数据分析时，“假设”可以发挥重要作用。**假设是数据分析的核心内容，或者说数据分析本质上就是在做假设检验。图中所示的数据分析流程中，有两个环节需要假设：在“锁定问题的关键”时，需要通过“WHAT型假设”对数据进行分解和比较；在“锁定原因”时，需要通过“WHY型假设”来列举出候补原因。

提出假设后，就可以确定检验各假设的方法及所需数据，然后才正式开始数据分析，根据数据来检验假设。采用“WHY型假设”来进行假设检验的一般路径可参考图3.16。

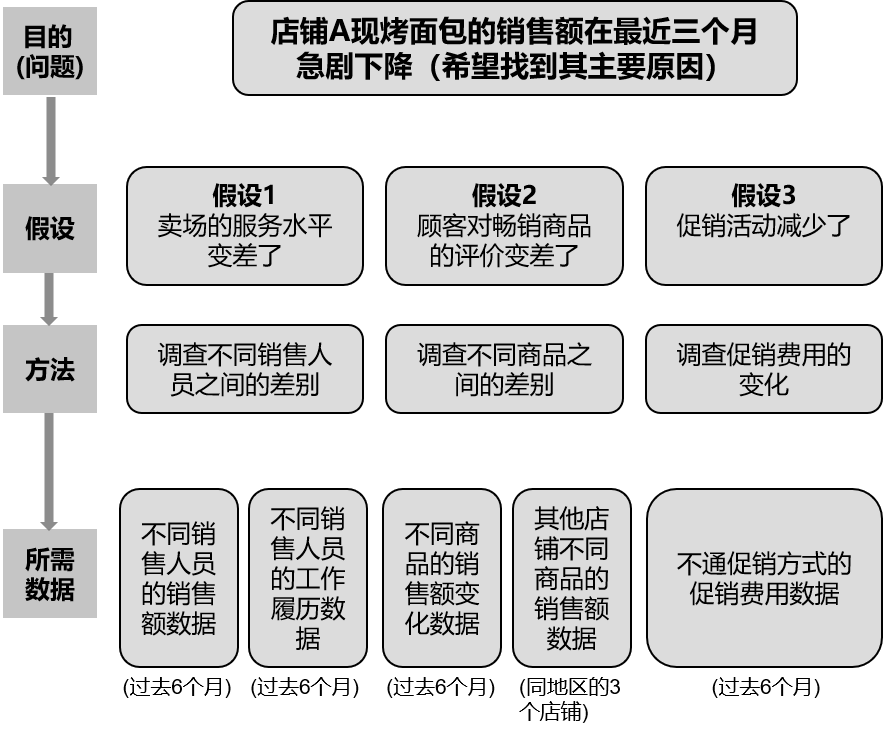


图3.16 通过假设得出所需的方法和数据

（3）**为了解决第三个问题，在实际动手处理数据之前，应该先缜密地思考。**将“问题”“假设”“分析结果”“结论”等以符合逻辑的方式展示给其他人的主要内容或流程确定下来，这样可以提高工作的效率，这一阶段的内容和质量基本上决定了后面分析的质量和整体故事的有效性。为了扩展分析视野，提出假设后，可以请业务部门对自己的假设进行确认，也可以运用业务自身的“框架”作为思考问题的线索，完善分析的架构，避免出现图3.17中的问题。

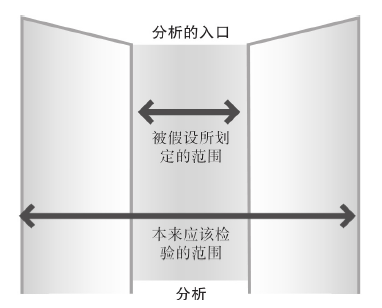


图3.17 分析问题的假设范围

（4）**为了解决第四个问题，这就要求我们能够逻辑清晰地呈现解决问题的思路。**从接受提案的一方来看，他们认同的并不是数据分析的内容和方法，而是以清晰明确的、能够说服别人的逻辑思路为基础提出的建议。解决问题的目标是准确地锁定主要原因，让相关人员接受自己的建议，在获得认同的基础上“推动其采取措施”。

### 1.2.3 分解数据，找到问题关键

对较大范围（包括多个要素）的数据，即使进行比较也很难找到差异。此时需要运用“四则运算”来分解数据，使其变得更为详细和具体。这就叫作“WHAT 型假设”， 我们可以采用这种方法分解问题。

分解问题后，再考虑从何种“维度”进行比较，才能找到问题的关键。这里要结合业务，从中优先挑选与问题联系最紧密，最能体现差异的维度进行比较，便可以找到问题的关键。如图3.18就是对一个待解决的问题采用“WHAT 型假设”分解后的结果[3]。

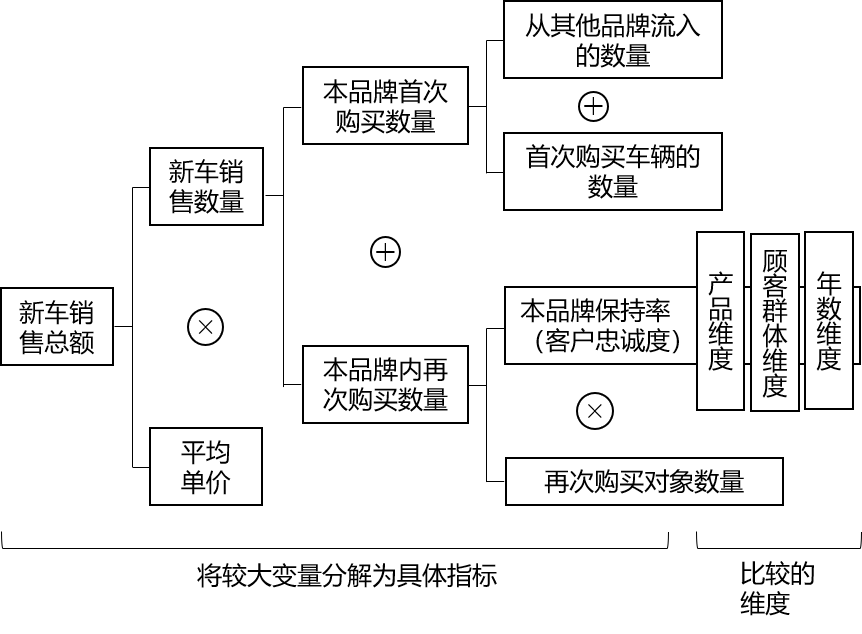


图3.18 采用“WHAT 型假设”分解问题

将问题分解为以上具体指标后，按照快照和趋势两个视点进行分析，得到如表3-4的矩阵图。

表3-4 不同视点分析的矩阵图

|  |  |  |
| --- | --- | --- |
|  | **“快照”视点** | **“趋势”视点** |
| **大小的维度** | 平均值、中位数 | 时间序列走势 |
| **波动的维度** | 标准差、变异系数 | 标准差的变化 |
| **比较方式** | 数值量 | 可视化 |

注：在通过不同的视点进行分析时需注意以下问题：

（1）大小的维度可以考虑平均值和中位数：平均值受离群点影响大，容易隐藏数据背后真相，要注意所分析数据构成的差异；中位数不易受离群点影响，可以结合直方图查看数据的分布。

（2）波动可以视为“风险”，波动的维度可以考虑标准差、变异系数或时间序列：标准差要比较相对大小才有意义；“变异系数”可以消除数据大小的差异，不必在意原始数据的大小，直接比较不同范围数据的波动程度；时间序列可以查看数据的走势。

### 1.2.4 采用交叉视点，锁定原因

与原因有关的假设叫作 WHY 型假设。采用WHY 型假设寻找接近结果的原因，即在关联更密切的数据之间考察相关关系，这里通过相关系数或散点图来找出对最终目标具有密切影响的原因。原则在于：确定与目标相关性较高的因素，并筛选有较大改进余地的因素。

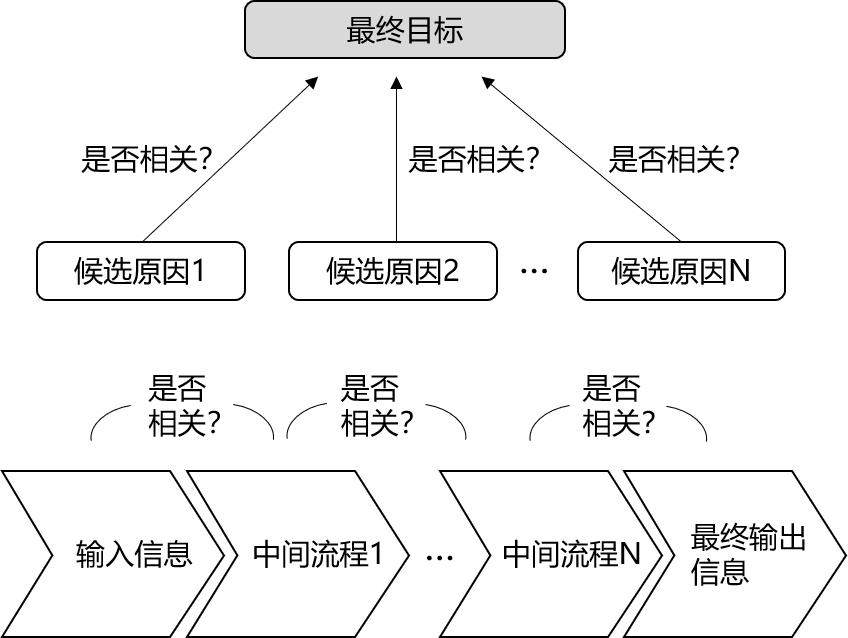


图3.19 相关分析的着眼点

运用相关分析的着眼点如图3.19所示，第1种模式通过相关分析，找出对最终目标具有密切影响的原因。例如，针对“销售额”这个最终目标，（通过假设）找出“顾客满意度”“降价”“产品魅力”等可能影响该目标的因素，观察这些因素与其相关程度的高低，可以得知按动哪个按钮（原因），会对销售额产生剧烈影响。

第2种模式是在某个业务流程中找到瓶颈。业务流程中包含金钱、信息等的流动，理想的状态是所有要素都畅通无阻地抵达最终阶段（输出）。但如果某个环节出现了停滞，预想的输出就无法实现，或者其中某个环节停滞时间较长，造成整体的处理效率偏低。

另外，在需要客观地决定不同影响因素的优先顺序时，可以通过 2 个维度展现其相对定位，如图3.20所示。

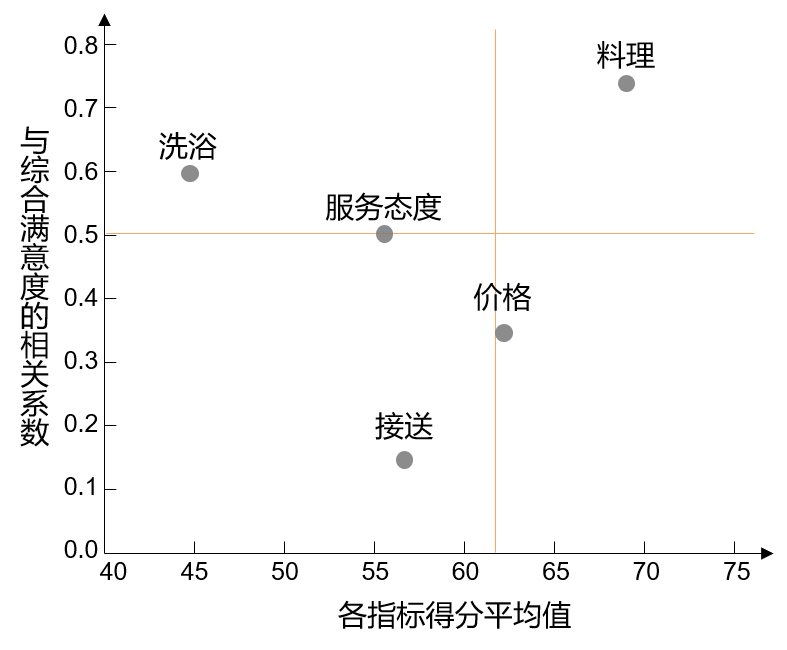


图3.20 提高某温泉度假酒店的综合满意度的因素

### 1.2.5 拟合数据，制定对策

相关分析可以锁定原因，但是无法得知原因会对目标产生多大影响，我们可以采用回归分析来确定影响程度，比如：用回归方程检验花费的成本能否带来足够的销售增长。比较成本与收益，确定对应的付出是否可以获得更大的回报，从而制定相应的策略。

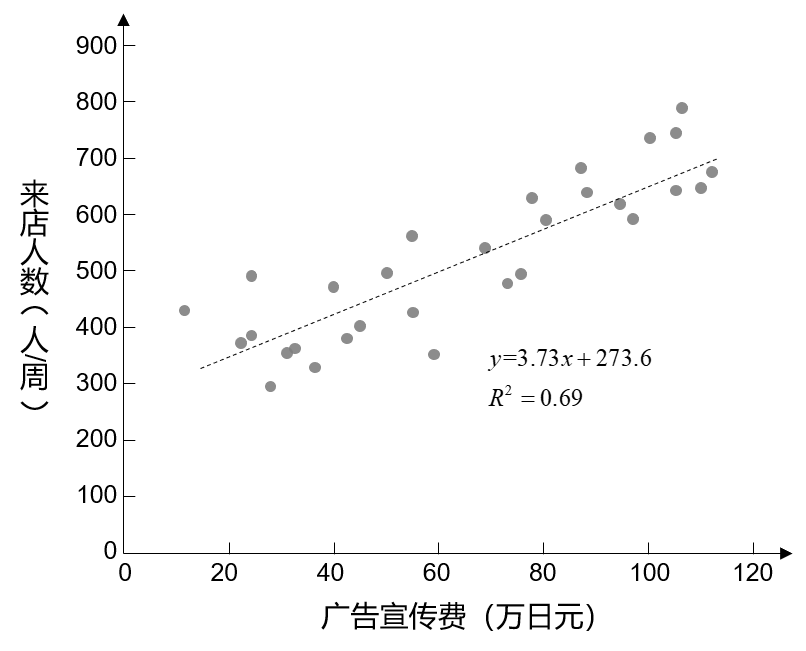


图3.21 利用回归分析制定决策

比如，图3.21这个问题，为了配合某月新商品上市，吸引更多的顾客来店，我们想知道投入多少广告宣传费可以达到预设的来店人数 3 200人。按照每月 4 个星期计算，则相当于每周 800 人。为了实现 800人的目标，需要的广告宣传费如下：

800（人）= 3.73 ×广告宣传费（万日元）+ 273.6

计算得到，每周所需的广告宣传费约为 141 万日元，所以这个月（4 周）需要的预算约为 560 万日元。

与缺乏可靠根据的预估金额相比，这样不仅能够提高预算金额的准确度，提交高层管理者或上司审批时也更具说服力。

### 1.2.6 向客户呈现解决问题的思路

按照我们分析问题的思路，即：“明确目的或问题”→“大致把握现状”→“锁定问题的关键”→“锁定原因”→“设定行动措施及 KPI（和所需资源）”，对于提升新车销售总额的问题，绘制分析过程的流程图，如图3.22和3.23所示。相比于直接给客户一个结论或改进措施，这样的展示方式自然可以得到客户的认可。

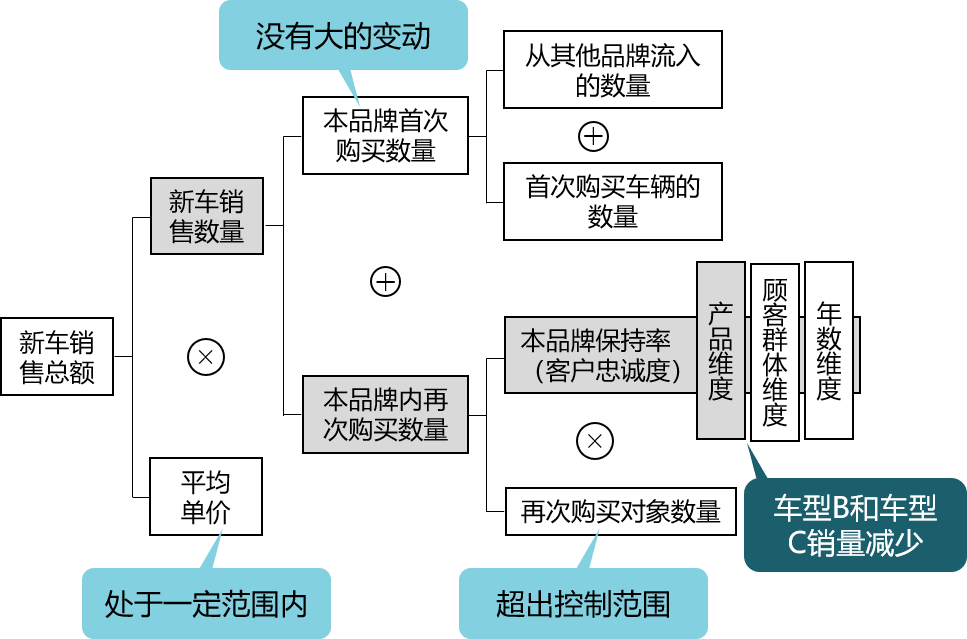


图3.22 分解问题，锁定问题关键

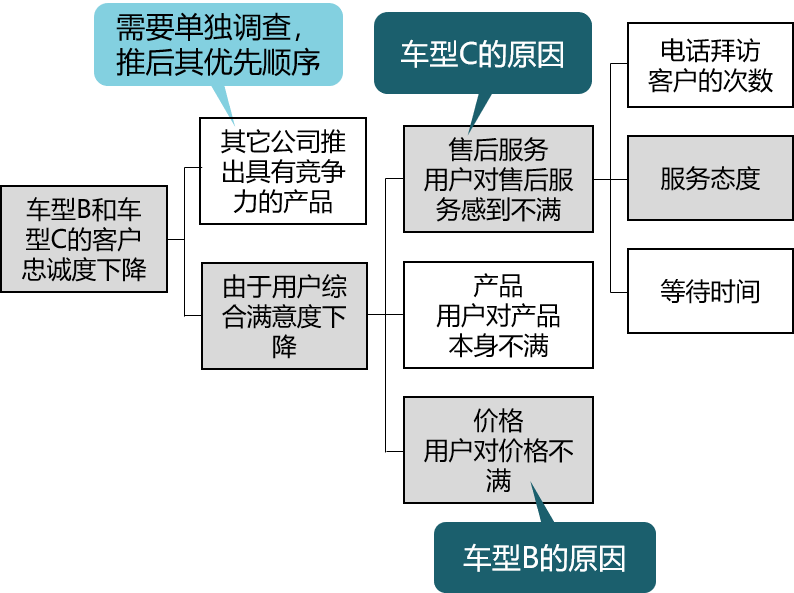
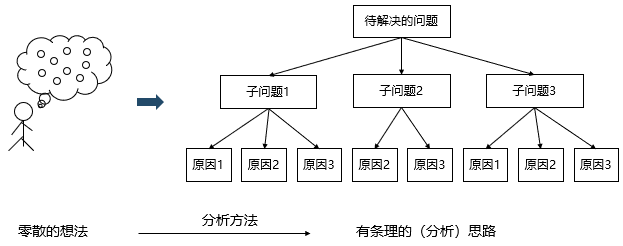


图3.23 分析问题产生的原因

## 1.2 逻辑树分析

**1.2.1 是什么？**



逻辑树分析方法是将问题的所有子问题分层罗列，从最高层开始，并逐步向下扩展，可以帮助我们理清思路，将零散的想法转化为有条理的分析思路。对照1.2中数据分析的一般思路，逻辑树分析对应其中的“WHAT型假设”，即分解问题。

**1.2.2 有什么用？**

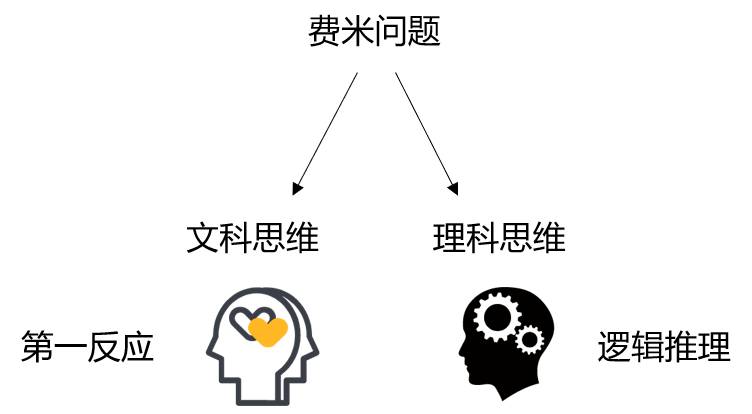
简单来说，使用逻辑树分析方法有下面几个用处：

* 形成有条理的思路，不进行重复和无关的思考；
* 按照层次分解，将复杂问题变得简单；
* 能保证解决问题的过程的完整性；

**1.2.3 如何用？**

下面，通过案例来看一下如何使用：

**案例一：费米问题**



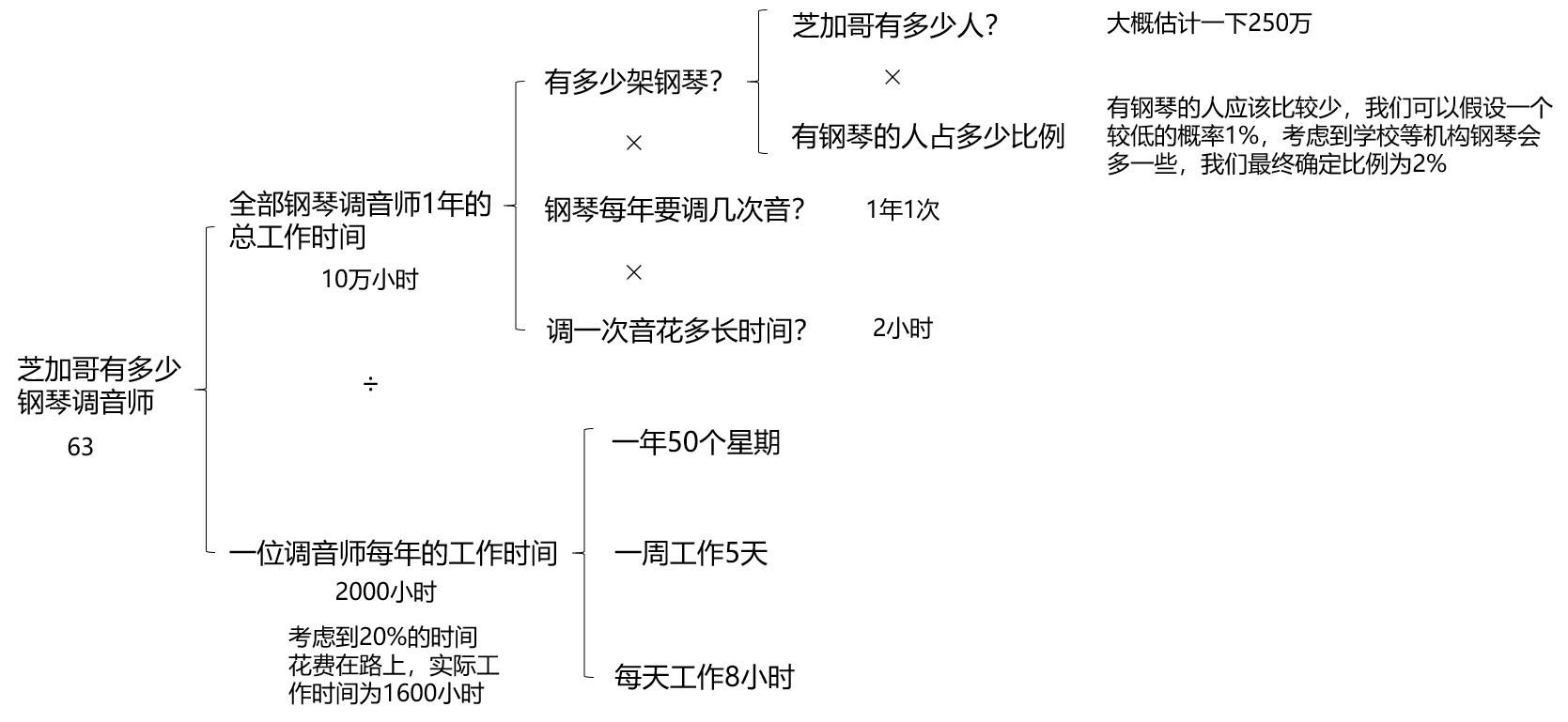
通常在企业招聘的时候，经常会出现费米问题，考察费米问题并不是为了要给出确切的答案，而是想要验证求职者是文科思维，还是理科思维。

一般来说，文科思维体现的是第一反应，比较感性；理科思维是经过逻辑推理的思维方式，比较理性。通常来说，没有经过刻意训练的人，很难通过逻辑推理的方式来解决问题。我们一般说一个人思路清晰，其实就是说这个人具有逻辑推理的能力。

像面试中的费米问题，比如：北京有多少辆公共汽车？一个正常成年人有多少根头发？都是这类问题。下面我们举个例子来看下如何解决费米问题：

比如：芝加哥有多少钢琴调音师？

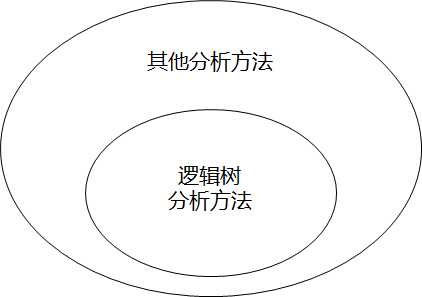
对于这个问题，我们可以按照下面的逻辑树分析方法来分解问题：



注：这里其实就是数据分析的一般路径中所说的运用“四则运算”来分解数据的方法。

从上面的例子我们可以看到：逻辑树分析方法不仅可以用在数据分析工作中，还能提高逻辑推理的能力，帮助我们形成有条理的思路，更好的解决问题。

**1.2.4 注意事项**



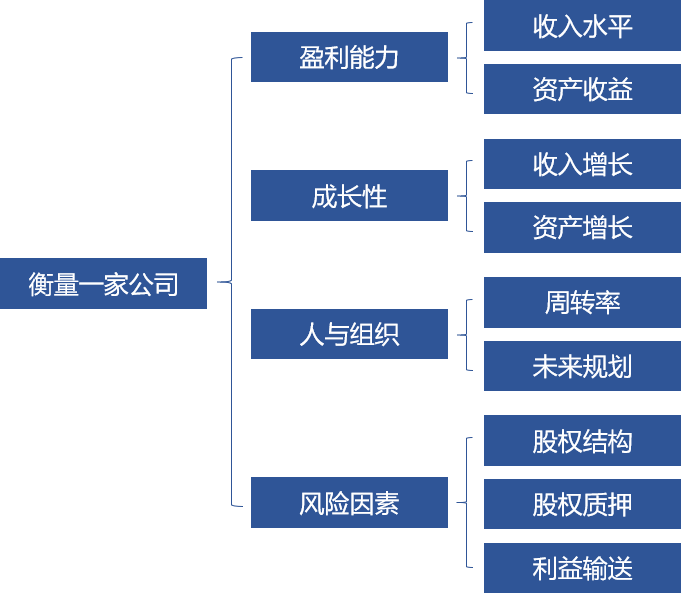
值得注意的是，逻辑树分析作为一种最基础的分析思想，通常会出现在其他分析方法中。所以，在数据分析任务中，我们要善于利用这种方法去分解问题，简化问题。

## 1.4 多维度拆解分析

**1.4.1 是什么？**

对于一个问题，我们想要从不同的角度去分析，就要将此问题拆分为不同的维度去衡量，最终得到我们的结论。

比如，对于一家公司，从投资的角度，你想要衡量这家公司好不好？我们就可以从不同的维度去拆解，这里我们简化一下这个问题，简单地从四个角度去拆解，即盈利能力、成长性、人与组织和风险因素。



如果经过拆解，加上后续的分析，我们得到的结论是：这家公司盈利能力较好，未来具有很好的成长性，管理层的运营能力和对企业未来的规划较好，并且没有明显的风险存在，那么，基本上我们就可以说这是家好公司。

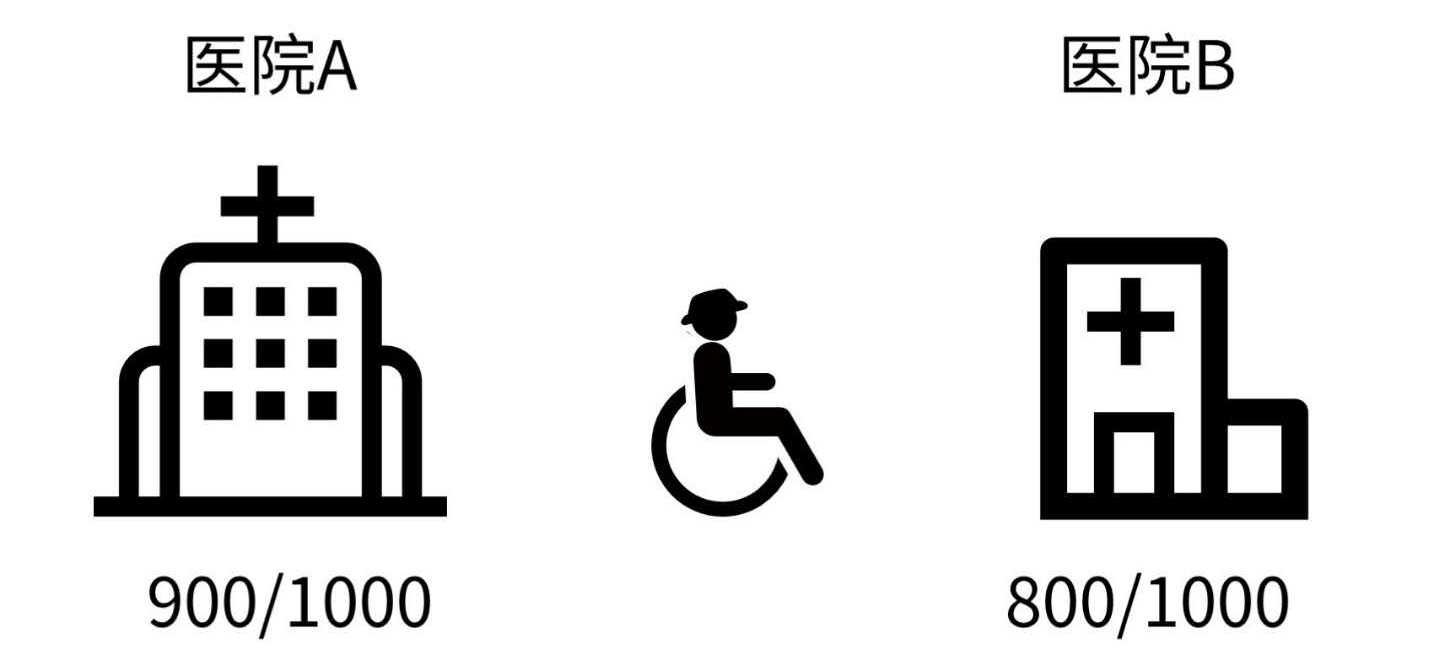
当然，只是分析这四个维度还不足够，比如我们还要了解这家公司的商业模式，还要结合PEST分析一下外部宏观环境，还要采用波特五力模型分析一下行业内的竞争格局等等。

不过，当我们遇到一个问题的时候，不可能完全考虑周全，我们可以根据经验进行适当地拆解，先从几个维度出发，在分析过程中我们又会发现新的问题，逐步对我们拆解的维度进行补充，最终会得到一个全面的思考。

**1.4.2 有什么用？**

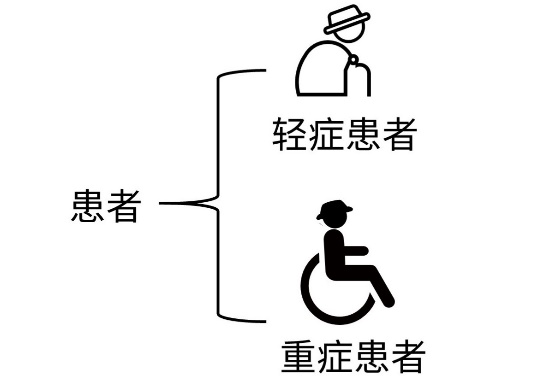
下面我们从一个更加实用的例子出发，来看一下多维度拆解分析方法有什么用？

比如说本地区有两个医院，一名患者想要选择医疗条件更好的医院，然后我们统计了这两个医院的存活率。

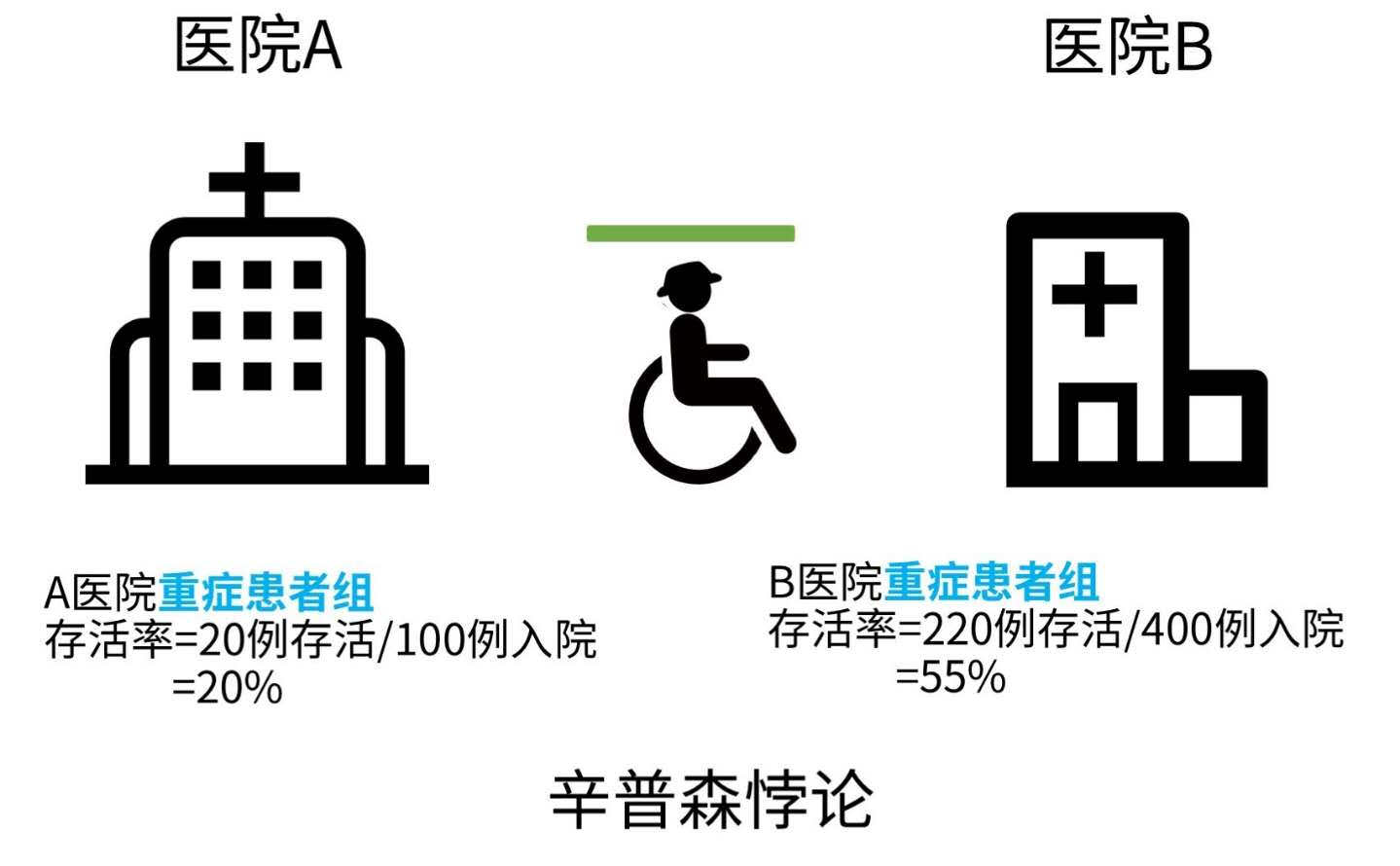


统计结果如下，A医院存活率为90%，B医院存活率为80%。从数据上来看好像是A医院的医疗水平更好，可是实际是这样么？这里就可以使用多维度拆解分析：

只看整体的话，我们可能看不到数据内部构成的差异，下面我们按照患者患病的程度将病人划分为两类：轻症和重症。



经过拆解之后，我们再去统计两个人群的存活率，结果如下：



比较两个医院的统计结果，对于重症患者组，我们发现A医院的存活率为20%，而B医院的存活率为55%。这样，我们会发现应该是B医院更好。

通过多维度拆解数据，我们发现，有时候考察数据整体和考察数据的不同部分会得到完全相反的结论，这种现象叫做**辛普森悖论**。

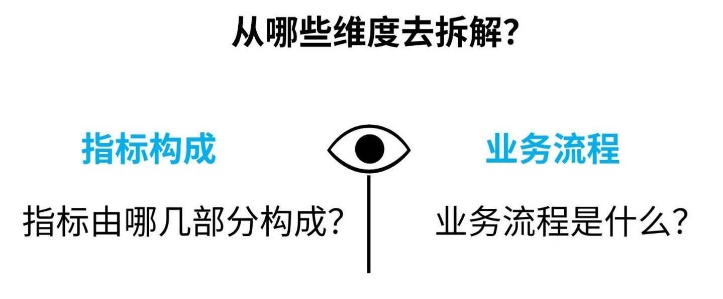
为什么会出现这种现象呢？因为只比较数据整体，我们会注意不到数据各个部分构成的差异，如果忽略这种差异，我们就可能忽略这种差异造成的印象。比如上面那个例子，可能A医院只是个普通医院，病人大多都是轻症患者，而B医院是更好的医院，重症患者较多。在这种情况下，我们单纯从整体上比较存活率显然是有问题的，因为重症患者的死亡率本来就高，所以得出的结论是不靠谱的。

因此，在分析过程中，我们要从多维度观察数据，才能得到相对比较可靠的结论。

**1.4.3 如何用？**

前面说了拆解问题，但是关键在于我们应该从哪些维度去拆解呢？

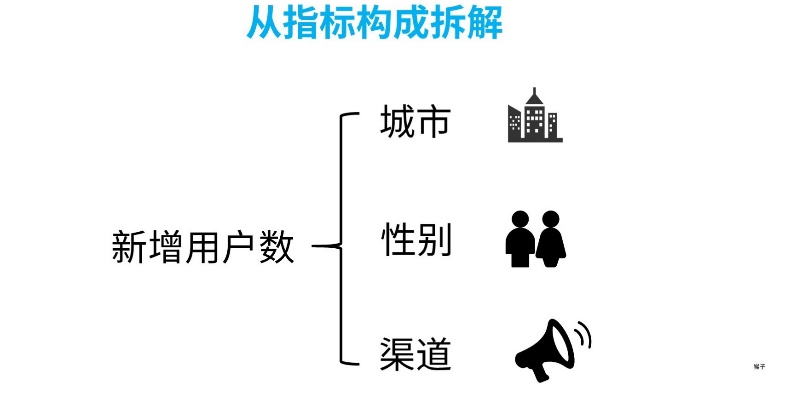
一般我们有两个角度，从**指标构成**或者**业务流程**的角度去拆解。



下面通过两个案例来看下如何拆解：

**案例1：**一家线上店铺做了一波推广，老板想看一下推广效果如何？

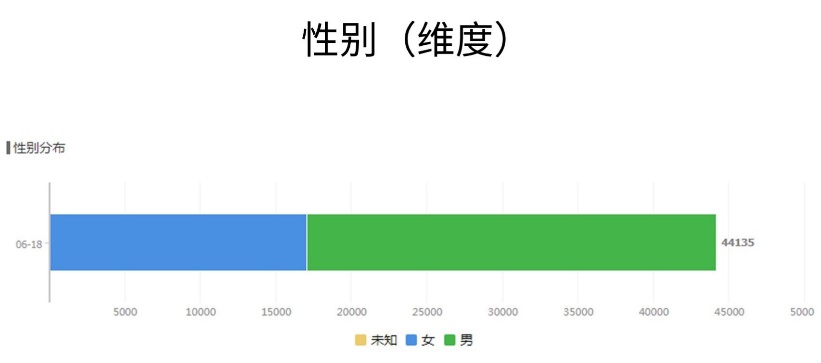
一般衡量推广效果，就是看一下推广前后用户增加了多少，也就是新增用户数。下面按指标构成来拆解新增用户数，比如按照城市、性别、渠道等维度去细分：



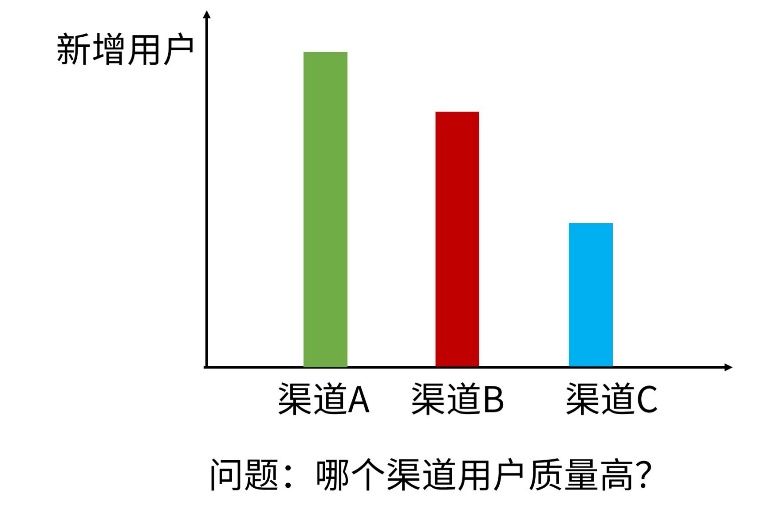
（1）从城市来看，我们发现：一线城市的客户对产品更感兴趣。



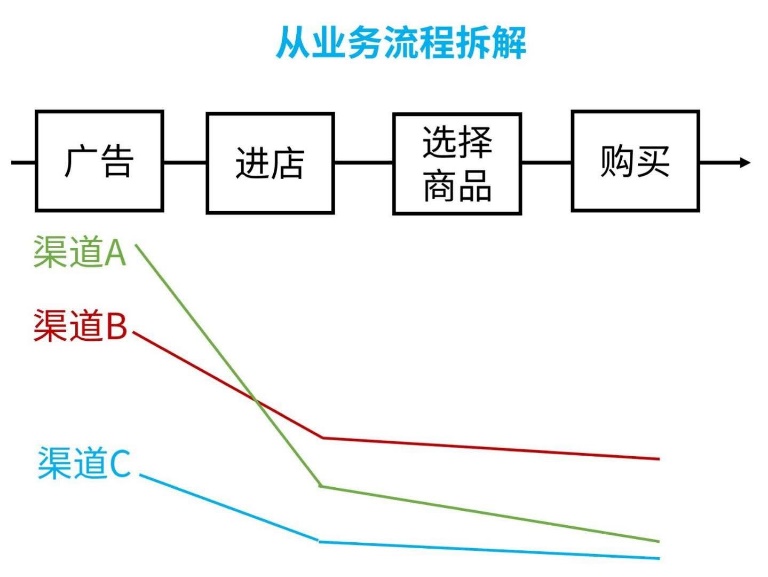
（2）从性别来看，我们发现：男性用户多于女性用户。



（3）从渠道来看，我们发现：渠道A、B、C的用户数依次下降。



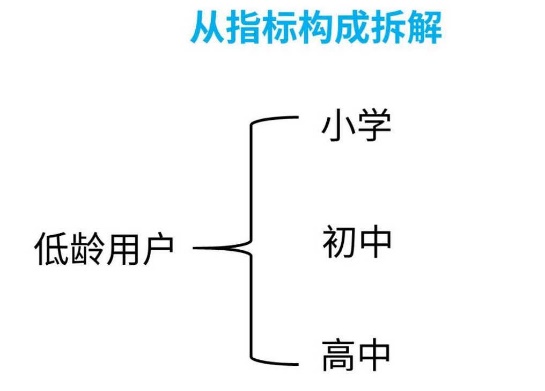
这里我们就要思考，哪个渠道的用户质量最高呢？这里我们就可以继续从业务流程拆解：



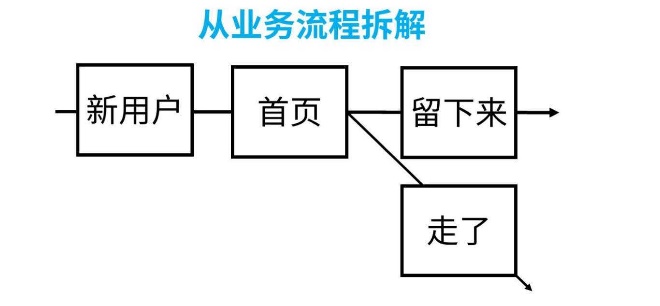
通过比较，我们可以看出：由于渠道B最终购买的用户是最多的，或者说渠道B的用户转化率是最高的，所以，渠道B的用户质量最高。

**案例2：**对于一款APP，我们想要分析为什么低龄用户的留存比较差？

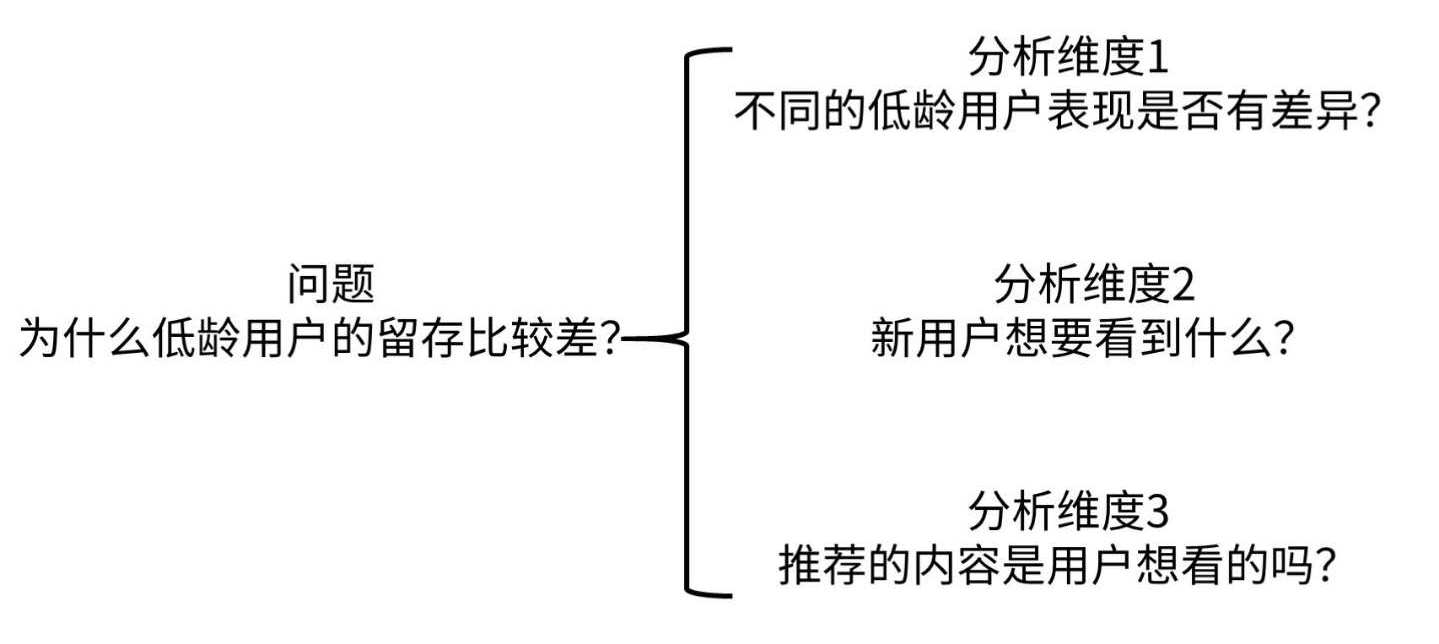
首先，对低龄用户可以先做拆解，得到分析维度1：不同的低龄用户表现是否存在差异？



然后，按照业务流程进行拆解，得到分析维度2：新用户在APP想要看到什么？以及分析维度3：推荐的内容是用户想看的么？



所以，我们将这个问题拆解为三个维度：



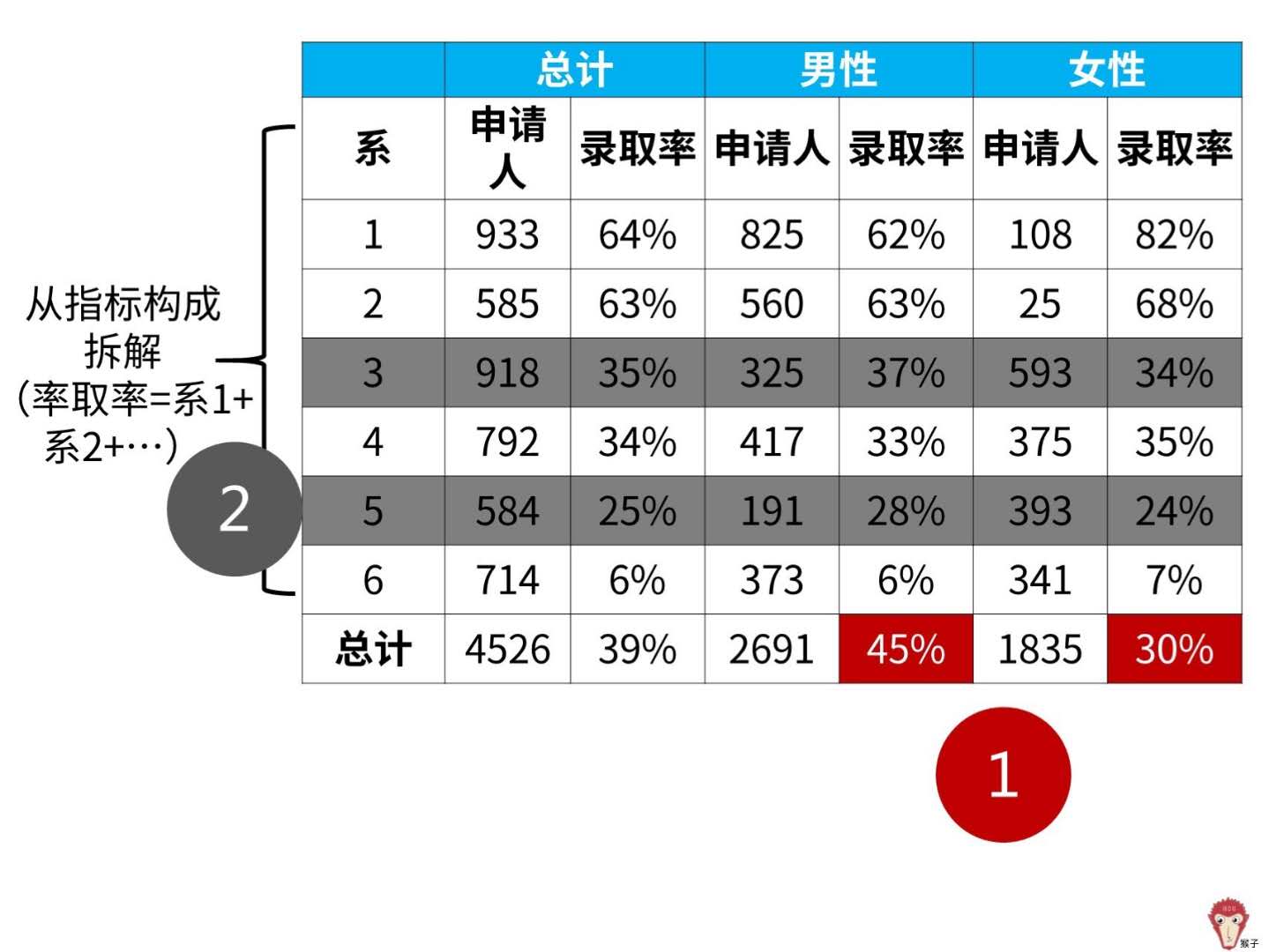
**1.4.4 注意事项：辛普森悖论**

下面再看两个辛普森悖论的案例：

悖论1：上世纪70年代有人指出加州大学伯克利学院研究生院歧视女性申请人，证据如下：

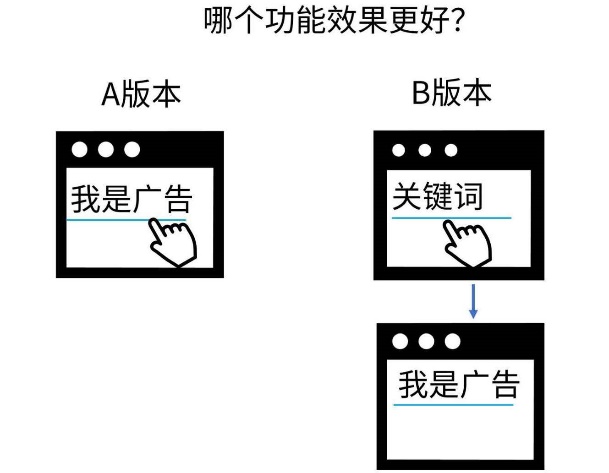


法院为了搞清楚这个问题，启动了如下调查：

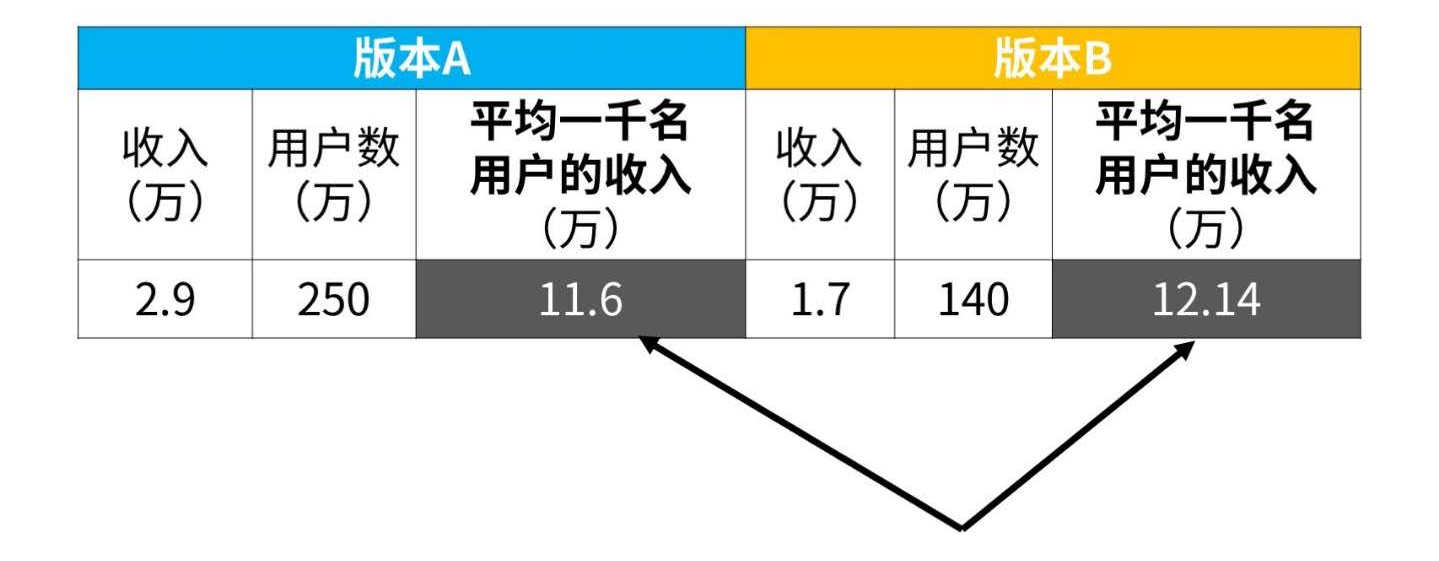


按照系进行划分后，我们发现并不存在男性申请人的录取率高于女性申请人的情况，表中可以看出大多系都是女性录取率大于男性，仅有两个系男性录取率略微大于女性，这是由于女性申请人较多，使得总体的录取率被拉低。

**案例2**：一家互联网公司想要验证哪种广告的效果更好。



如果从整体看，我们会选择版本B的广告。



但是，采用多维度分析法拆解后，我们发现：不管是国内还是国外，都是版本A的收入更高，应该选择版本A。



当然对于多维度拆解分析，我们还可以参考1.3节PEST分析中的实例，关于如何降低新能源车的成本？这个例子拆解了新能源车的结构，找到了降低新能源车成本的方法。

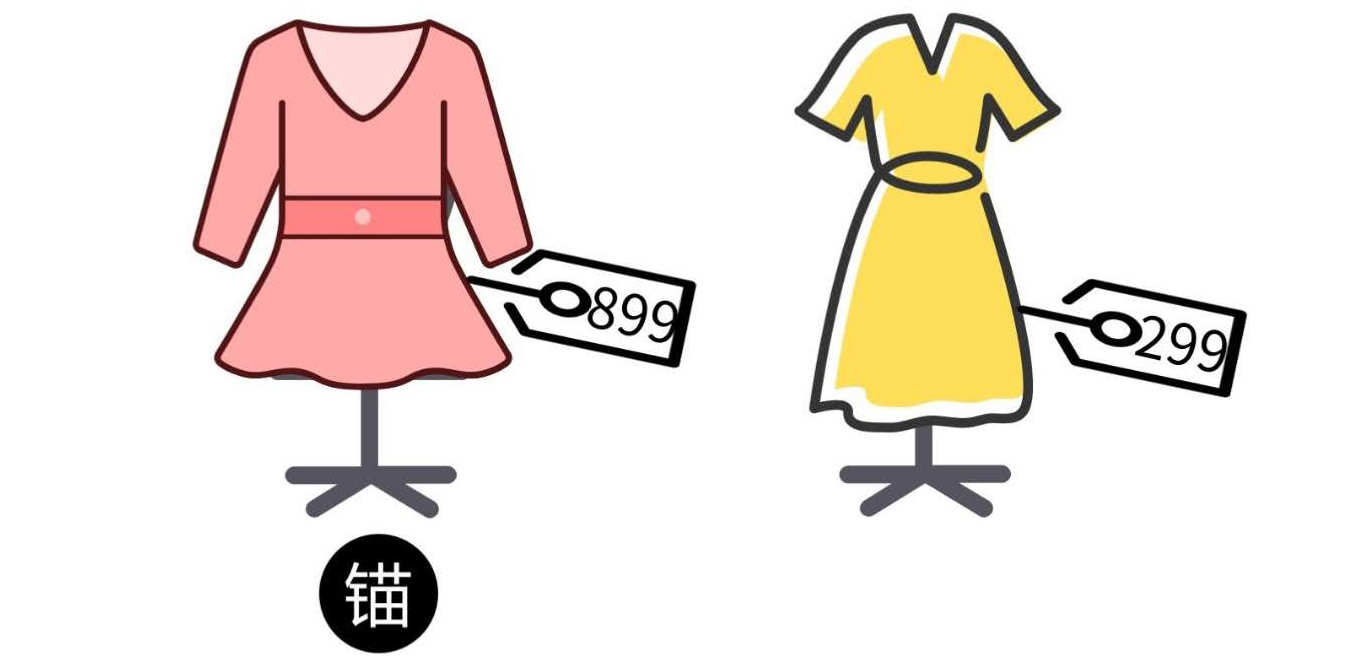
## 1.5 对比分析

**1.5.1 是什么？**

为什么要做对比，因为没有对比，就没有好坏。

**1.5.2 有什么用？**

（1）价格锚定



锚定效应，就是指人们在做判断的时候，易受到对事物第一印象的影响。这个第一印象，会像一个沉锚一样，把人们的思想固定在某处，继而影响人们的后续行为。

人们在做决策的时候，会不自觉地给予锚定信息过多的重视。

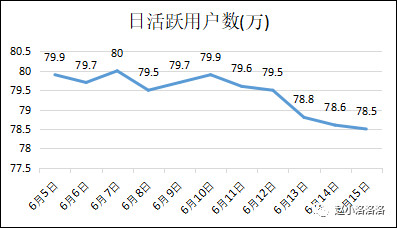
比如，服装店里标价非常高的衣服，就可以起到价格锚定的作用，因为跟这件衣服比起来，消费者会觉得其他衣服性价比更高，哪怕它并不值这个价钱。

再比如，星巴克除了卖30块一杯的咖啡外，他们还出售20块一瓶的矿泉水，而且会放在非常显眼的位置。

这么贵的矿泉水基本上卖不出去。但是星巴克的主营产品是咖啡，矿泉水卖的好不好并不重要，重要的是让咖啡卖得好。20块钱高价矿泉水的存在， 就是一种锚定，可以使30块钱的咖啡显得非常划算。

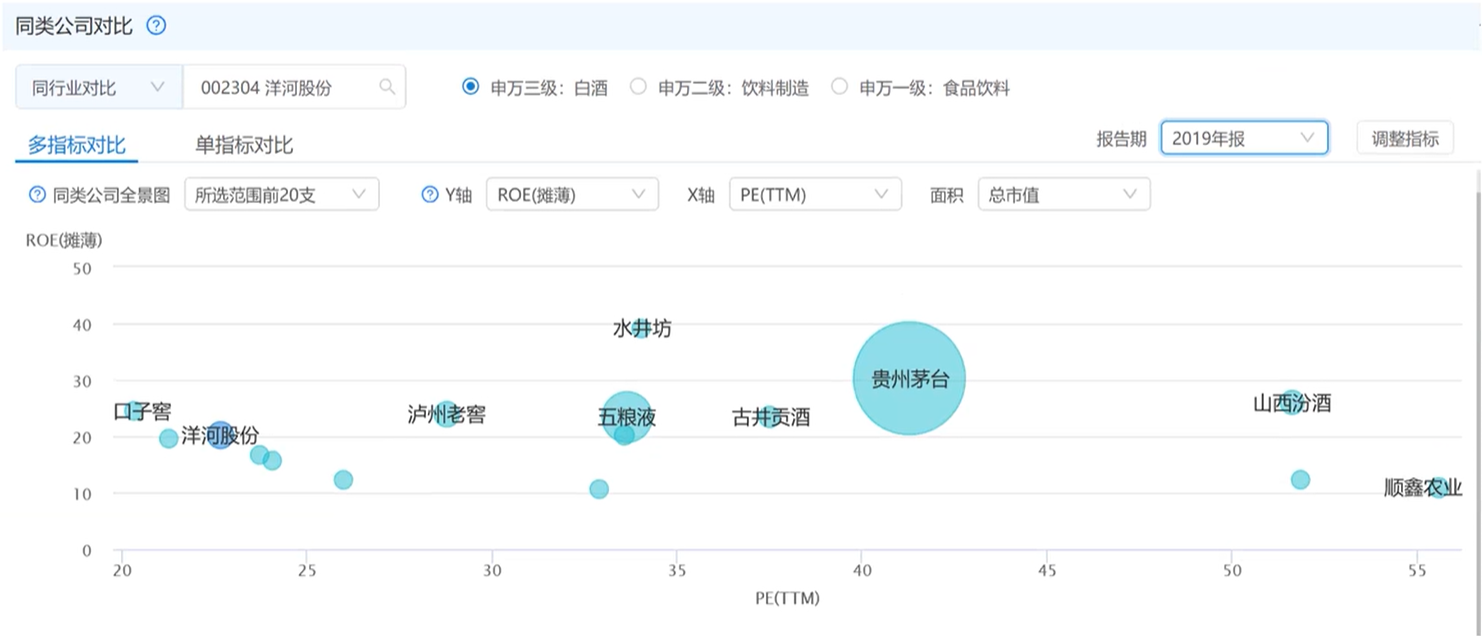
（2）追踪业务是否存在问题

比如，互联网公司关注的一个指标——DAU(Daily Active User)日活跃用户数量，毕竟互联网公司本质上就是要**占用用户时间**。想要知道日活是否正常，通常就是和历史数据作比较，如果某段时间发生数据大幅下降，说明肯定是某个环节出现了问题，接下来就应该去定位问题所在。



（3）筛选具有投资价值的公司

通常，在做股票投资的时候，很多时候都要作对比分析去发现更有投资价值的企业。对于一个公司，通常我们要与其历史业绩作对比，要与行业内同类公司作对比。比如，下图是2019年白酒行业市值较高的一些公司的PE和ROE，可以从资产收益率和估值两个角度去比较同行业的公司，作为投资过程中的参考。



此外，川建国同志也深谙此道，建国同志交易的艺术就是漫天要价（价格锚定），然后不断推进，获得一个不错的价格（Great Deal!）。

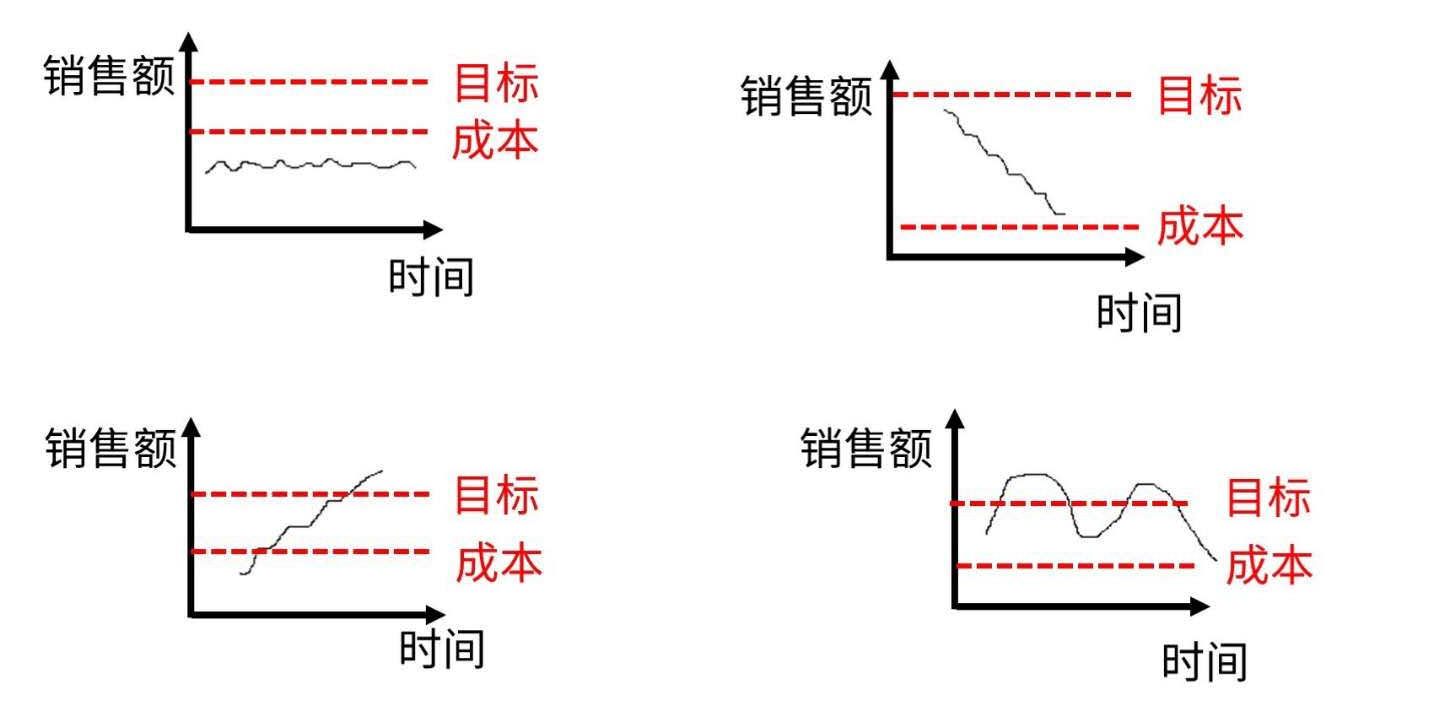
**1.5.3 如何用？**

（1）和谁比

首先要解决的一个问题就是和谁比，比如，下面的四个销售额，单独拿出来是没有意义的，因为不知道跟谁比，没有比较就没有高低。

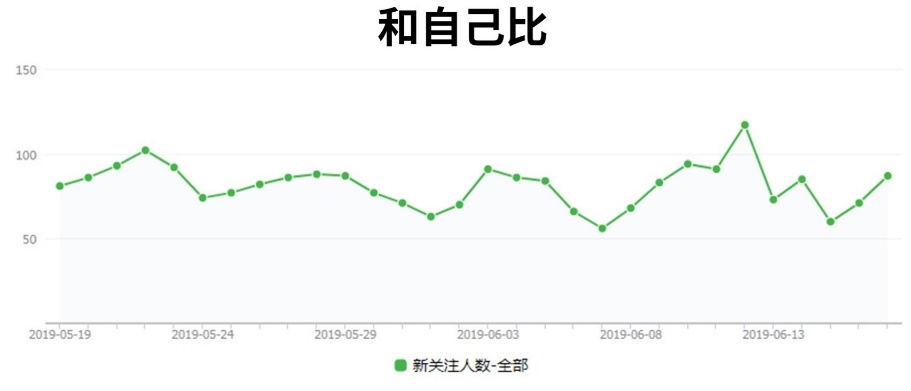


下面，我们将目标和成本画到图中，这样我们就知道跟谁比了。

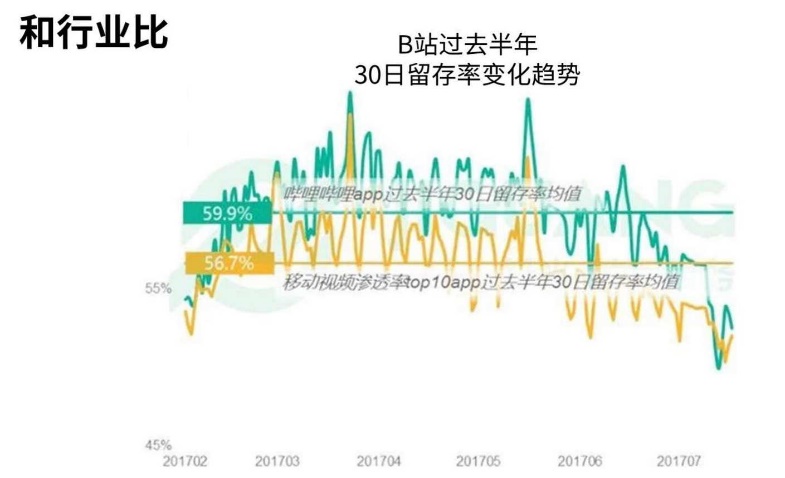


所以，和谁比的问题，我们可以从两个方向来思考：和自己比、和行业比。

* 和自己比



* 和行业比



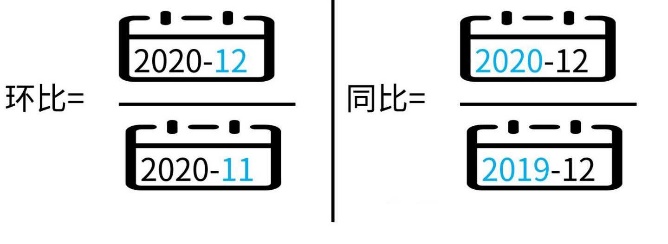
通过与自己和行业作对比，我们可以判断业务指标的变化是自身原因还是行业趋势，从而采取对应的措施。

（2）如何比？

知道了跟谁比，接下来的问题就是如何比较？在实际应用对比方法时，为了防止遗漏，我们可以用下表来记录，防止遗漏重要的信息。

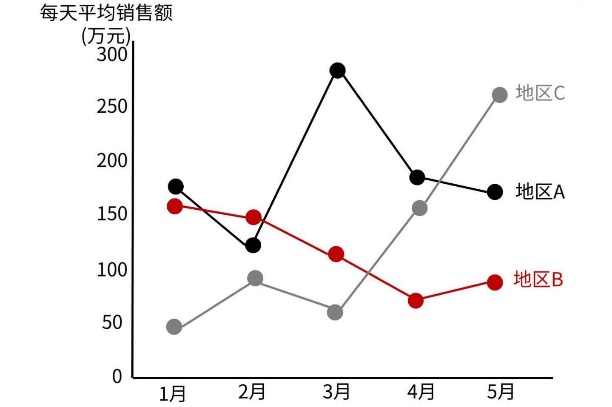
|  |  |  |
| --- | --- | --- |
|  | **“快照”视点** | **“趋势”视点** |
| **大小的维度** | 平均值、中位数 | 时间序列走势 |
| **波动的维度** | 标准差、变异系数 | 标准差的变化 |
| **比较方式** | 数值量 | 可视化 |

在比较的过程中，我们可以从两个角度去对比：环比和同比。其中，环比是和上一个时间段对比，用于观察**短期**的数据；同比是和去年同一个时间段对比，用于观察**长期**的数据；

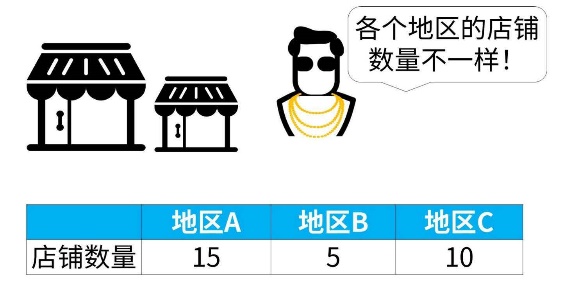


**1.5.4 注意事项**

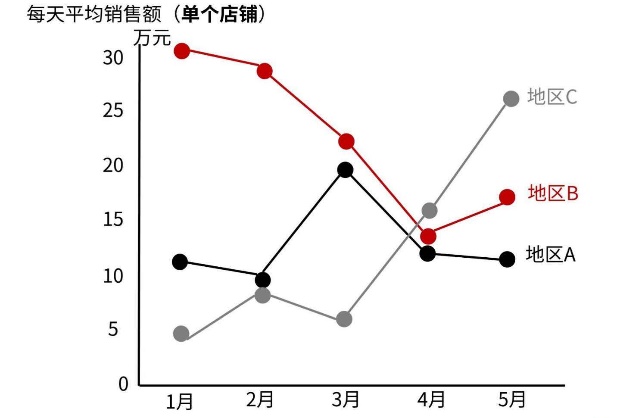
**注意比较对象的规模要一致。**比如，在我们比较不同地区商品的销售额时，得到如下结果：



这样，我们可以得出结论：地区B的销售情况最差？其实不然，因为各个地区的店铺数量不一样，根本没有比较的意义。



那么如何解决呢？可以按照店铺取均值，比较单个店铺的每天平均销售额，这样我们可以看出地区B的销售额并不是最差的。



**案例一：A/B测试**



A/B测试是一种广泛使用的对比方法，我们设置实验组和对照组，在满足其他条件基本相同的情况下，改变一个条件，对比两个分组是否存在差异。广为人知的一个例子就是airbnb做过的一个实验：测试房屋拍摄更专业的照片，预定量是否会上升，比较结果表明：实验组的房源预定量高于平均预定量的2-3倍，这里说明了用户更倾向于预定具有专业照片的房子。

**实例二：比较一家公司的盈利状况**

和自己的历史比、和同行业比

## 1.6 假设检验分析

**1.6.1 是什么？**

统计学中的假设检验听起来很高端，其实只不过是披了层数学的皮，底层的思想非常简单，就是基本的**逻辑推理**的套路。

用福尔摩斯的话说，就是：一旦排除所有的不可能，剩下的不管多么难以置信，一定就是真相。

所以，从这个思路出发，假设检验的推理过程可以分成四步：

* 要解决的问题是什么？
* 证据是什么？
* 判断标准是什么？
* 作出结论

**1.6.2 有什么用？**

总的来说，假设检验有两个作用：

（1）提高逻辑思维能力；

（2）分析问题发生的原因（归因分析）；

**1.6.3 如何用？**

我们知道假设检验的推理过程可以分成四步，下面从一个简单的例子出发，来看看这个推理过程：

我们假设福尔摩斯接了一个案子，就是老李头觉得自己妻子可能出轨了（或许是隔壁老王），但是又不确定，想让福尔摩斯帮忙调查一下，于是福尔摩斯就通过假设检验，来证明老李的妻子有没有出轨。

**（1）要解决的问题是什么？**

问题：判断老李的妻子有没有出轨

根据这个问题，我们可以提出两个对立的假设，即：

* 零假设H0：老李的妻子出轨了。
* 备择假设H1：老李的妻子没有出轨。

这里为什么要提出两个假设呢？从逻辑学上来说，如果我们能够证明零假设不成立，那么其对立假设必然成立，用前面福尔摩斯的话就是，排除所有不可能，剩下的就一定是真相。

那么，接下来我们就要想办法去找证据，看看零假设到底成不成立。

**（2）证据是什么？**

通过中心极限定理，我们知道，合理且足够多的样本可以代表总体，于是福尔摩斯便出去寻找合理的样本数据来作为证据。

福尔摩斯去随机调查了老李妻子最近一段时间的样本数据，包括通话、出行、财务等等。

根据这些数据显示出来的证据，老李的妻子经常同隔壁老王通电话，手机中发现很多暧昧信息，并且在酒店发现了开房记录，通过统计学的知识，计算出，在零假设成立的前提下，这些证据出现的可能性只有0.01%。

这个概率值在统计上有个专门的术语叫做 p 值。也就是，在零假设成立的前提下，得到样本观测结果出现的概率。

这里 p 值就是在零假设成立的前提下（即老李妻子没有出轨），样本证据发生的概率 p=0.01%。

**（3）判断标准是什么？**

虽然说有了一定证据，但是也不能就直接判断老李妻子出轨了，这还得看老李头的心理承受能力，说不定我们认为老李妻子出轨了，但是老李头觉得还好，这个程度可以接受也说不定呢。

所以，福尔摩斯要跟老李头约定一个定罪标准，符合这个标准才判断老李妻子出轨了。

因为我们的零假设是：老李妻子没有出轨。在这个前提下，这里定的标准是：如果老李妻子没有出轨的概率 ≤ 5%（也可以说出轨的概率 ≥ 95%），那么就直接否定零假设，接收备择假设，认定老李妻子出轨了。

这里做出决策的标准5%，在假设检验里叫做“显著性水平”，用符号α表示，是一个概率值。

**（4）作出结论**

那么，老李妻子到底有没有出轨呢？

我们将样本证据计算出的 p 值与判断标准 α比较下就可以了：

* 如果p≤α，那么拒绝零假设，接收备择假设；
* 如果p>α，那么零假设成立；

这里，在零假设成立的前提下，样本证据发生的概率（p=0.01%），远远小于我们定义的标准5%。

也就是说，本想零假设证明老李妻子没有出轨，但是样本数据给出的证据证明没有出轨的概率只有0.01%，远远小于老李头心理可接受的标准5%，那么，没办法，根据样本证据只能说老李被带绿帽子了。

总结一下假设检验的思路：

* 要解决的问题是什么？

提出零假设和备择假设；

* 证据是什么？

假设零假设成立的前提下，得到样本均值的概率 p-值；

* 判断标准是什么？

显著性水平α（1%, 5%）；

* 做出结论

p≤α，拒绝零假设，接收备择假设；

p>α，接收零假设；

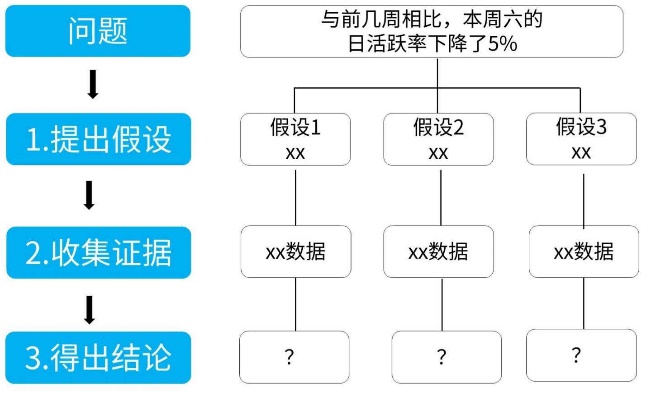
**1.6.4 注意事项**

（1） 第三步得出的结论不能是主观猜想出来的，要从第二步收集证据得出结论。

（2）假设检验的三个步骤是不断重复的过程，完成之后，要多问自己几个为什么，然后用数据验证，直到找到问题原因（合理假设，小心求证）。

（3）假设检验过程中还要用到其他分析方法。

（4）在假设分析之前，可以先通过下面这样的图理清思路。



到此为止，我们已经知道了假设检验的过程，但是这个过程中会遇到一个问题，就是如何提出假设？实际上，我们可以从以下几个角度考虑：

（1）从用户、产品、竞品三个维度

（2）4P营销理论

（3）业务流程

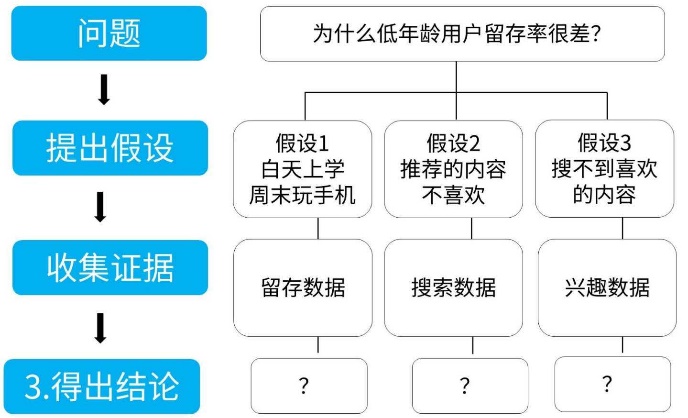
**案例一：对于一款APP，分析为什么低年龄用户的留存率很差？**

首先，从日常生活的角度可以提出一个假设：是不是低年龄用户白天上学，只能周末玩手机？

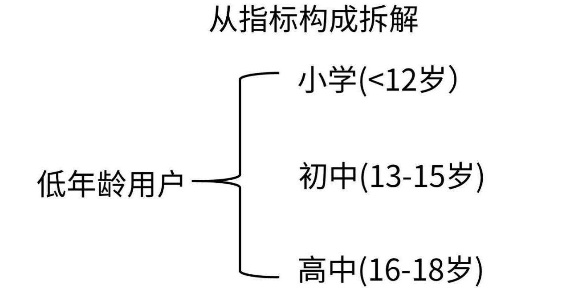
接下来可以从业务流程提出假设，比如APP的使用流程如下，我们可以根据流程的环节提出两个合理假设：一是假设推荐的内容低年龄用户不喜欢，二是假设低年龄用户搜不到自己想看的内容。



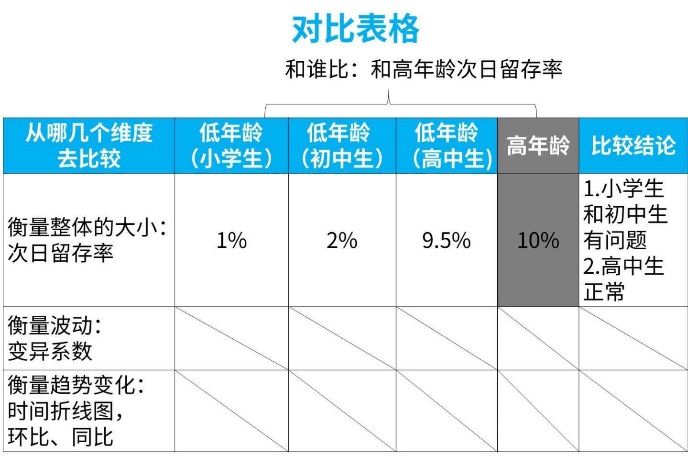
于是，我们便可以得到如下思路图：



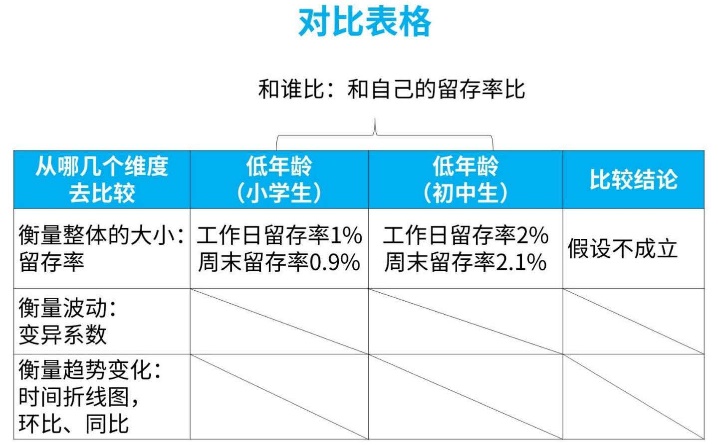
此外，为了避免用户群体自身的差异性，我们将低年龄用户进行拆解，锁定存在真正存在问题的目标客群：



拆解之后，我们发现高中生的次日留存率和高年龄用户相差不大，主要是小学生和初中生的留存率存在问题。



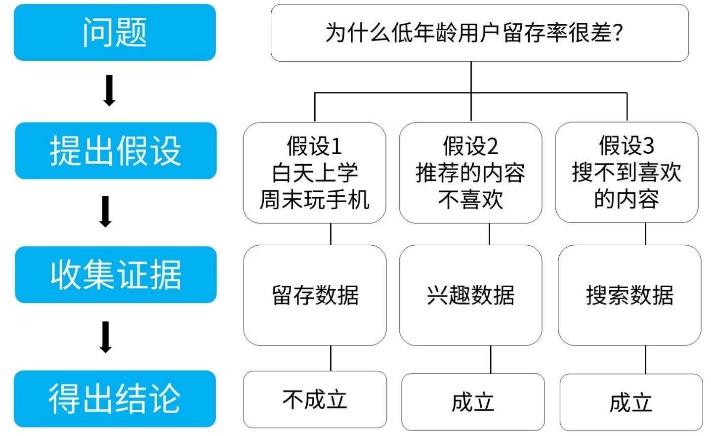
接下来，针对小学生和初中生，我们分析第一个假设，是否是白天上课，周末才能玩手机，统计结果如下，我们发现工作日和周末的留存率相差不大，说明假设一不成立。



然后，我们考虑假设二：是否是推荐的内容低年龄用户不喜欢。我们将小学生、初中生的兴趣与APP推荐的内容作比较，发现两者确实存在较大差异，说明推荐的内容确实不是用户喜欢的。



接着，我们就可以考虑假设三：推荐的内容用户不喜欢，那低年龄用户能否搜到自己想看的内容？答案是搜不到，该APP主要涉及的是电影、读书类的内容，与低年龄用户的兴趣不一致。因此，我们可以得到如下分析结果：



我们发现留存率差的低年龄用户主要是小学生和初中生，原因在于app的内容与小学生的兴趣相悖。这样，我们就可以采取相应措施：如果低年龄客群存在推广价值，那么app可以增加相应内容；如果低年龄客群带来的收益低于增加内容的成本，那么可以考虑是否放弃这部分客户，把主要资源用于维护高年龄用户的留存率。

**案例二：同为高端白酒，为何贵州茅台的股价远远高于五粮液？**

## 1.7 相关分析

**1.7.1 是什么？**

研究两种或两种以上数据之间有什么关系。

**1.7.2 有什么用？**

（1）相关分析考察数据之间的相关关系，找出对最终目标具有密切影响的原因。

（2）扩大分析思路；

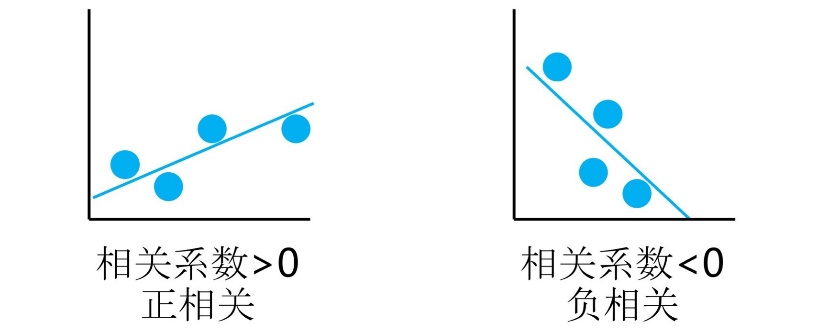
（3）通俗易懂；

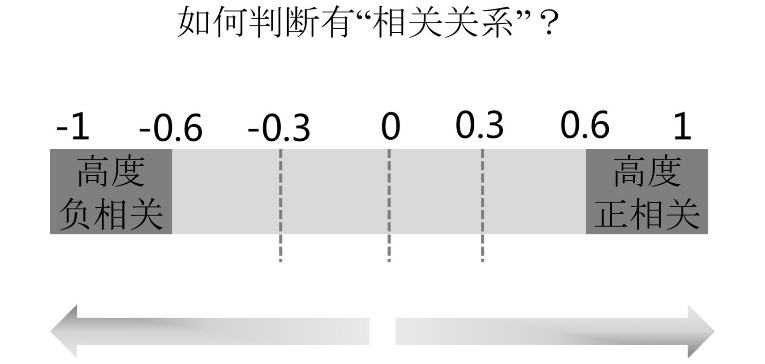
（4）可以结合其他分析方法，进行更深入的分析；

**1.7.3 如何用？**

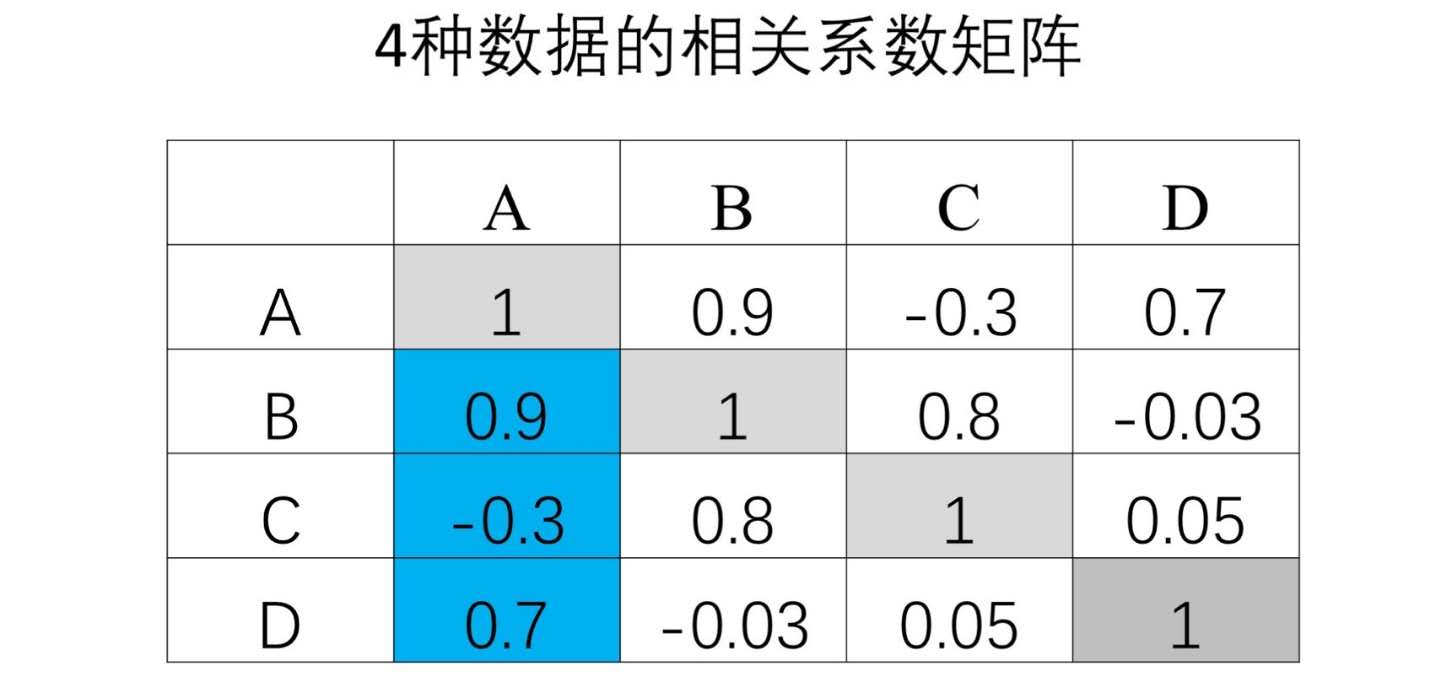
**（1）相关系数**

相关性包括正相关和负相关，相关系数在-1~1之间，一般相关系数大于0.7可认为高度正相关，相关系数小于-0.7认为高度负相关。





对于多维数据，我们可以采用相关系数矩阵观察不同数据之间的相关性。

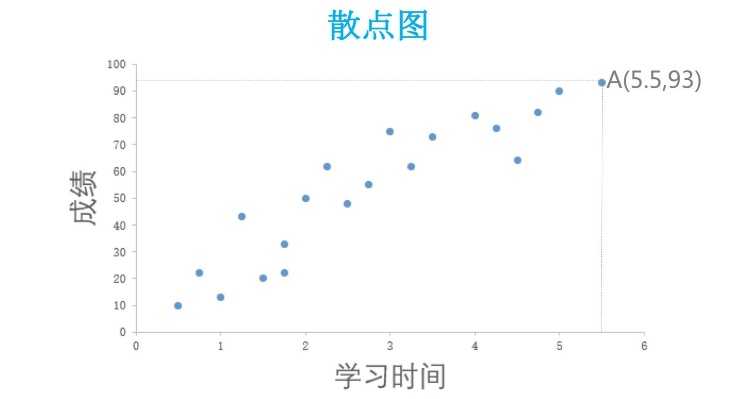


**（2）散点图**

散点图能够可视化可视化，相对于相关系数有如下好处：

（1）即使不懂相关分析也能理解；

（2）可以查看异常值；



**（3）着眼点**

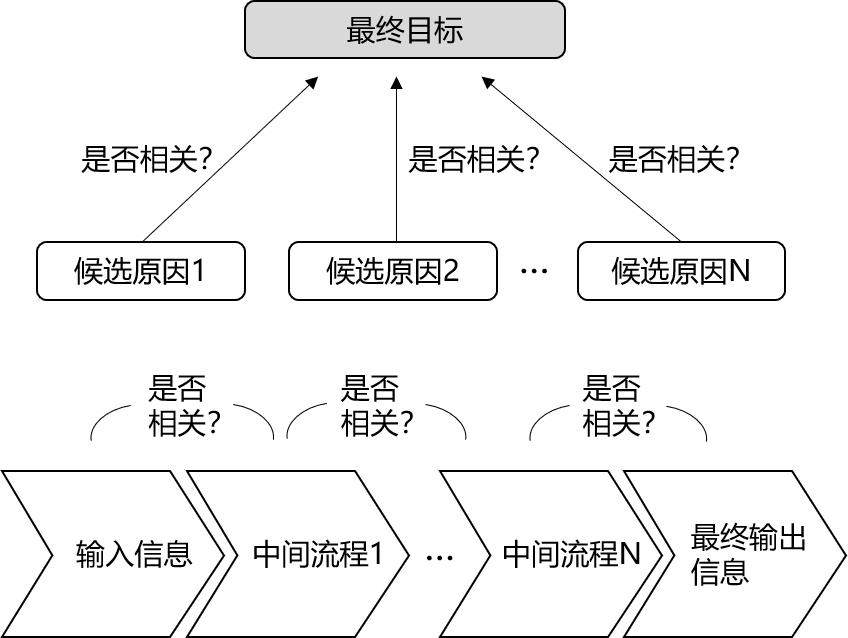
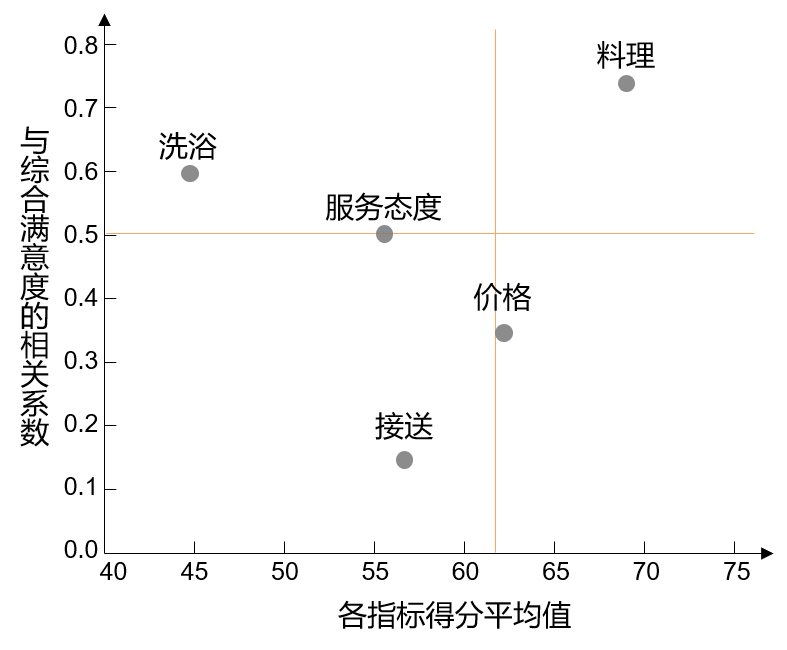


图3.19 相关分析的着眼点

运用相关分析的着眼点如图3.19所示，第1种模式通过相关分析，找出对最终目标具有密切影响的原因。例如，针对“销售额”这个最终目标，（通过假设）找出“顾客满意度”“降价”“产品魅力”等可能影响该目标的因素，观察这些因素与其相关程度的高低，可以得知按动哪个按钮（原因），会对销售额产生剧烈影响。

第2种模式是在某个业务流程中找到瓶颈。业务流程中包含金钱、信息等的流动，理想的状态是所有要素都畅通无阻地抵达最终阶段（输出）。但如果某个环节出现了停滞，预想的输出就无法实现，或者其中某个环节停滞时间较长，造成整体的处理效率偏低。

另外，在资源有限的情况下，通常优先解决影响大的因素。在需要客观地决定不同影响因素的优先顺序时，可以通过 2 个维度展现其相对定位，选择影响最大的指标。如图，可选择用户评分低，且与综合满意度相关性较高的指标，即洗浴和服务态度。



**1.7.4 注意事项**

相关关系不等于因果关系。因此，在判断是相关关系还是因果关系时需要进行单变量控制。

有些问题我们只要找出相关关系就行了，比如在推荐系统中。但有些问题需要在相关的基础上找出因果关系。比如有一个任务，需求是分析防晒霜的销量，看看如何进行品牌营销？

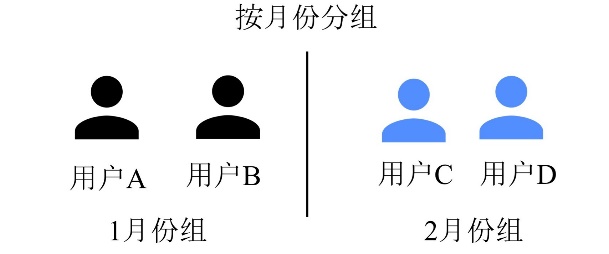
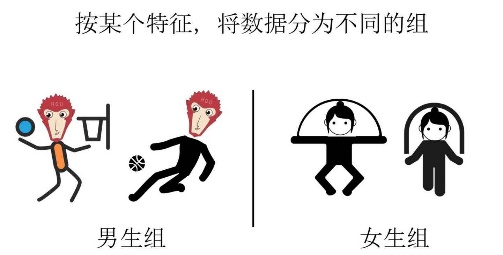
通过分析发现：蜜蜂的数量与防晒霜的销量存在相关关系。通过进一步分析发现：蜜蜂多的日子里通常天气也会比较好，大家也会适当增加室外活动，出外运动则会涂抹防晒霜，所以户外活动和防晒霜才是因果关系。

**实例：水泥的需求量的影响因素？**

## 1.8 群组分析

**1.8.1 是什么？**

按照某个特征，将数据分为不同的组，观察指标随时间的变化。比如，我们可以按性别或月份分组：



**1.8.2 有什么用？**

群组分析常用来分析用户留存率随时间发生了哪些变化？找出用户留下或离开的原因。比如：按月份分组比较。

**1.8.3 如何用？**

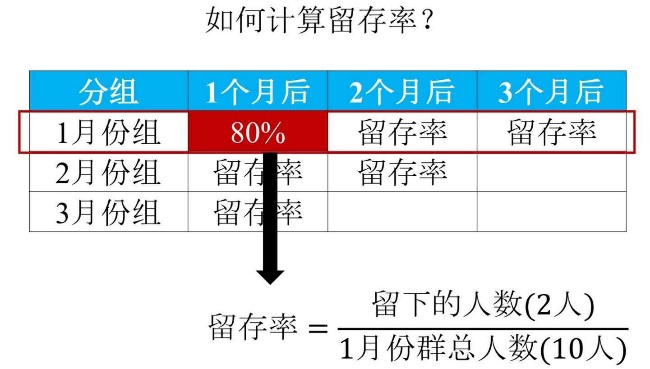
按照月份分组，观察留存率随时间是如何变化的，其过程如下：



（1）首先，构建群组分析表格



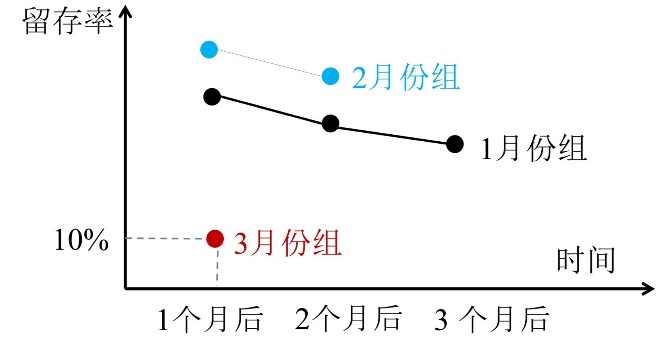
（2）计算用户留存率



（3）比较留存率随时间的变化



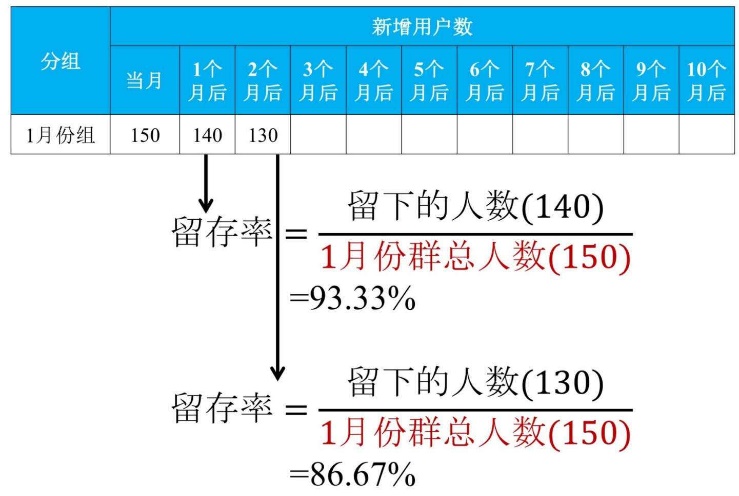
（4）通过可视化比较更加直观



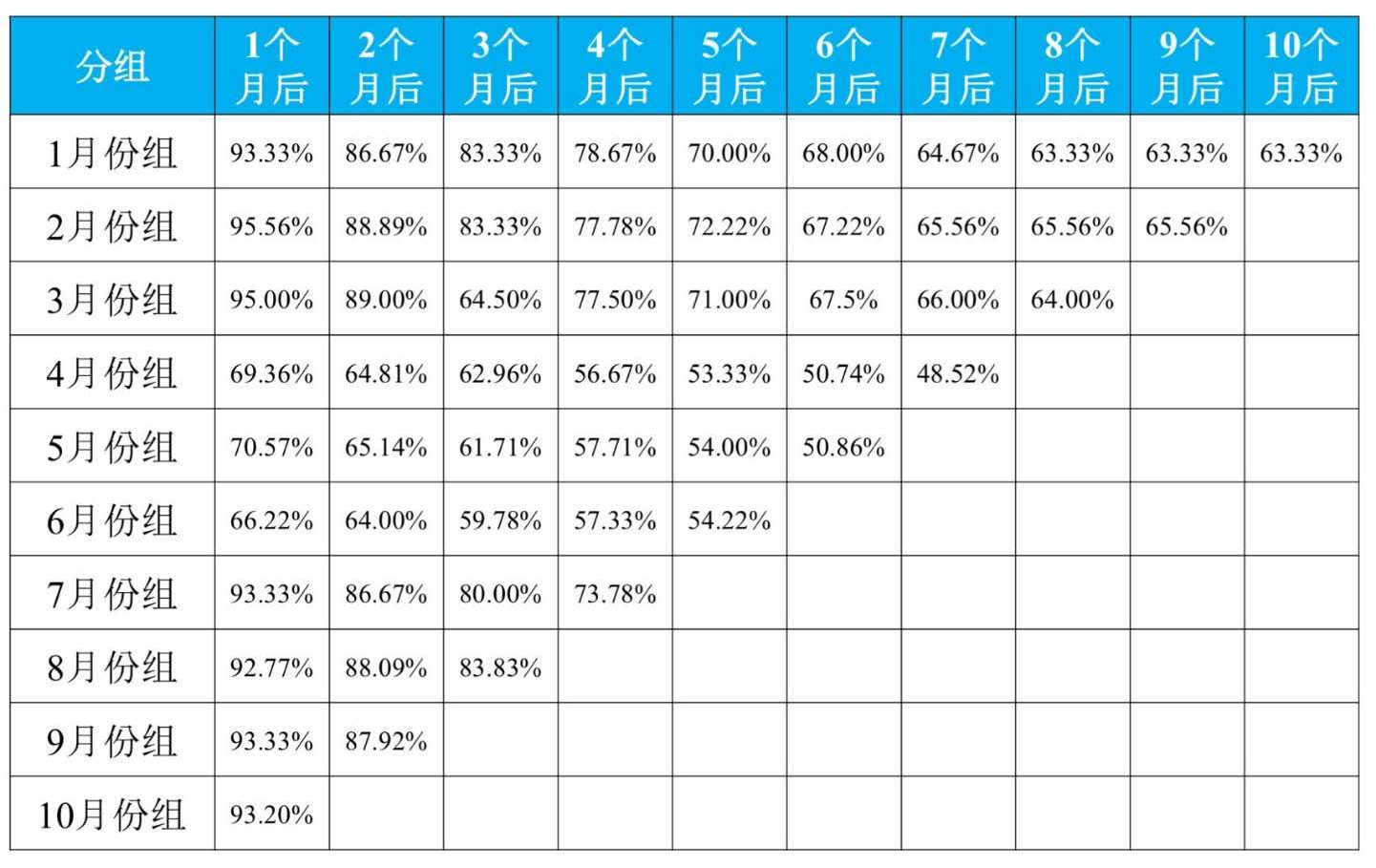
**案例1：对视频网站的用户流失进行分析**

（1）首先，按照注册月份对用户分组，统计之后每个月留存的客户数量。

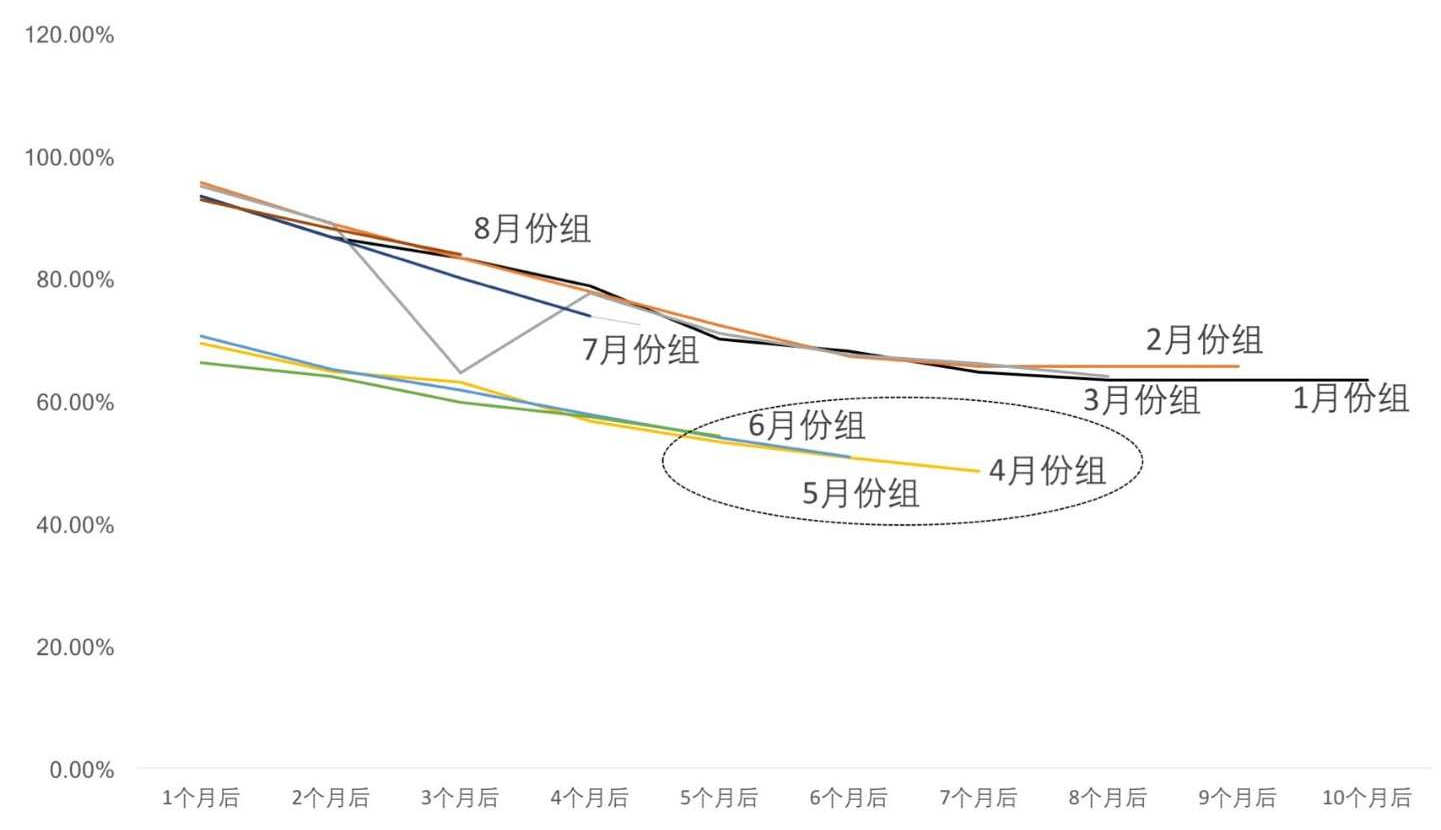
（2）计算用户留存率



（3）生成用户留存率统计表



（4）可视化不同分组的留存率



（5）分组较多，简化结果，提取1月份和4月份的留存率进行比较



图中明显可以看出4月份组客户的留存率较低。

（6）锁定问题后，可以通过假设检验或相关分析找到留存率低的原因



**案例2：用户留存分析**

为什么推特为新注册用户推荐30个最可能感兴趣的账号可以使用户成为长期活跃用户？

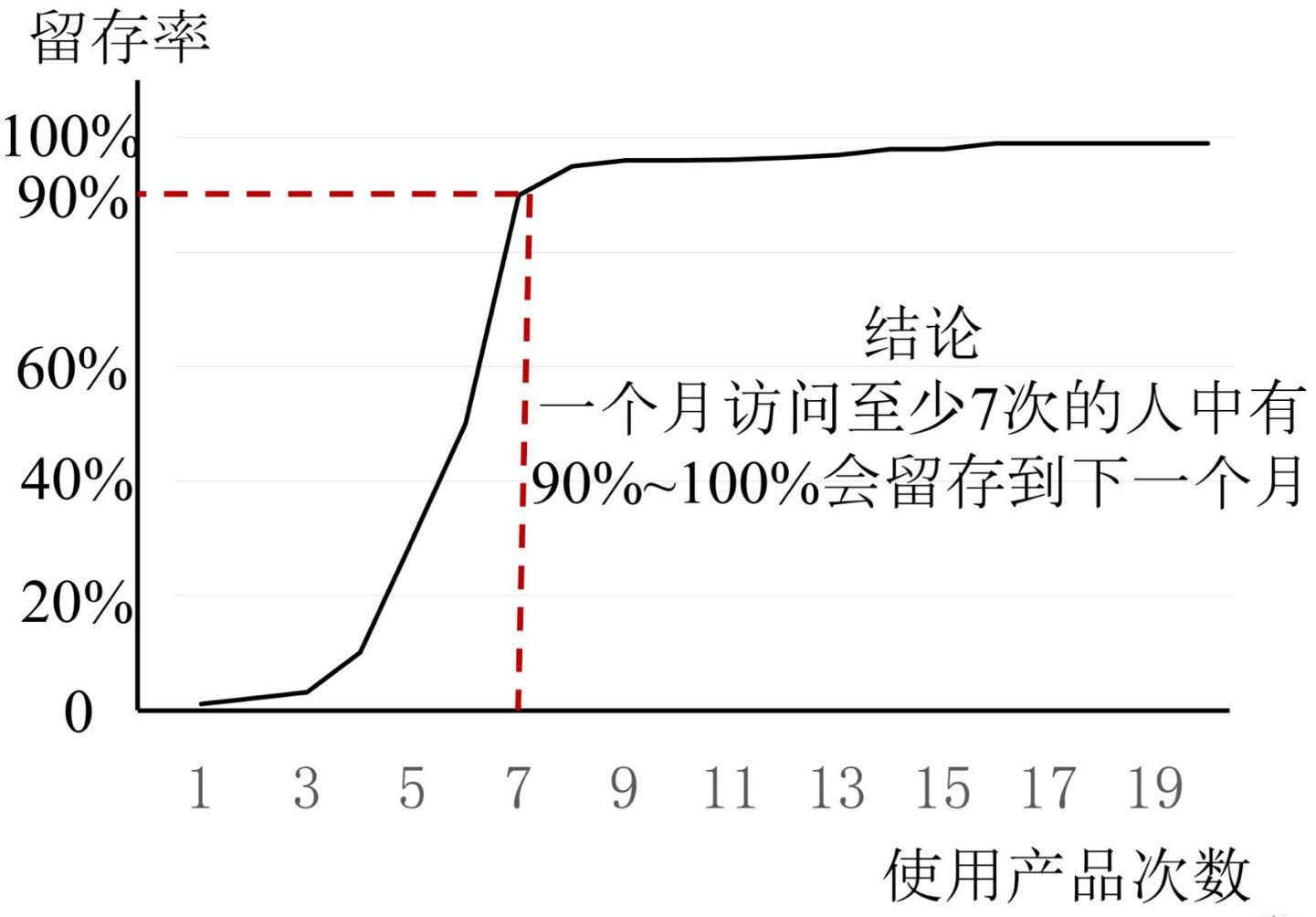
（1）按照某月使用产品的天数对客户分组，统计当月用户数，以及1个月后留存的用户。



（2）计算每个分组客户的留存率：



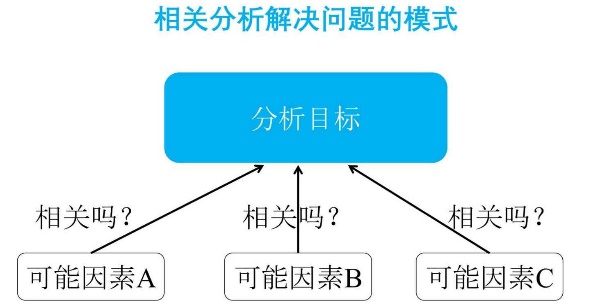
（3）可视化比较不同使用次数的客户的留存率，发现使用产品一个月访问至少7次的人中有90%~100%会留存到下一个月。

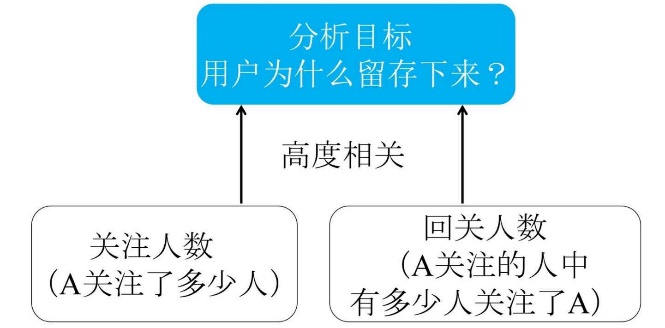


（4）按照使用天数将客户划分为核心客户、一般客户和冷漠客户。



（5）通过相关分析寻找可能的原因

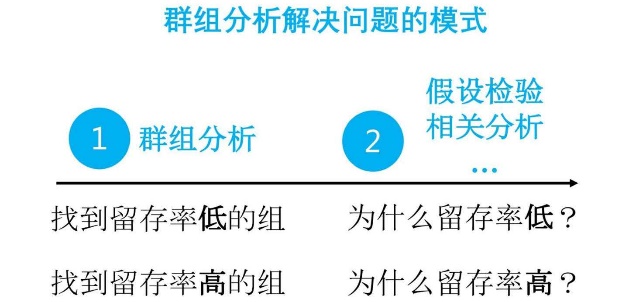




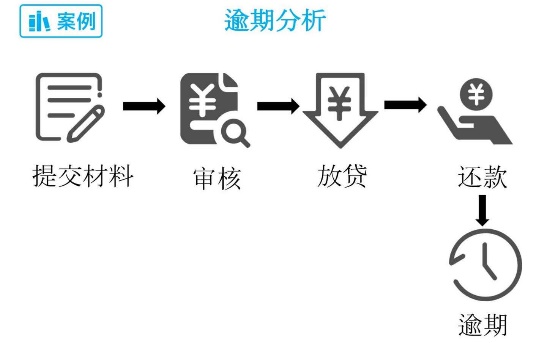
通过相关分析发现：关注的人数和回关人数与客户留存高度相关。

（6）通过电话回访，寻找背后的原因，最终发现：回关人数为1/3是最好的。

所以，群组分析解决问题的模式为：



**案例3：逾期分析**



某产品提供的是贷款款期限为7天的短期小贷服务，为了分析用户逾期率的变化，下图是这款产品的“群组分析表”（也叫做账龄分析报表）。



表格中，第一列是按每周（7天贷款期限）的还款日期分组，比如W1是“第1周的贷款”，W2是“第2周的贷款”。每一行是这一组用户贷款到期后的第几天，比如“1”表示的是贷款到期后的第一天，以此类推。

单元格里的数据是逾期率，也就是当前天数未还款的用户的占比。例如W1（第1周）贷款到期后的第1天对应的数据是31%，表示逾期率是31%，也就是第1周放款的用户当中，在逾期1天时仍未归还的用户占是31%。



这张表的制作时间是第8周贷款到期后的第2天，而此时我们还无法获取第2天及后续的数据，因此W8这一行只有一个数据。同理，此时第7周也只有8天的数据，剩余数据我们暂时还无法获取。

那么，我们从这个表里可以分析出什么呢？



首先，我们看第一列中，方框中的W2和W3对应的逾期率分别是36.3%和39.5%，而其余这一列中的逾期率几乎都是31%左右。

通过比较不同组的数据，我们发现相对于W1而言，W2和W3的第1天逾期率明显地升高了。随后产品调整了风控策略，W4的首逾率下降为31.2%，恢复到了正常水平。



其次，我们观察表的右半部分。方框中的数据是从W2到W7的第8天逾期率。从W2到W7，它是逐步下降的，从22.3%下降到了16.4%。这表示从第二周开始，用户逾期8天的占比有明显的下降。

不仅仅是第8天，我们对第8天以后每周的逾期率进行分析，也可以得出这一结论。这说明在调整了风控策略以后，不仅仅在第1天逾期率上的表现有所提升，后续的表现同样也有所提高。

最后，我们可以得出结论，通过调整风控策略，我们的用户质量有了显著的提升，这证明了风控策略的调整是有效的。

**1.8.4 注意事项**

分组方式不固定，可以根据业务灵活分组。

## 1.9 RFM分析

参考聚类分析方法论

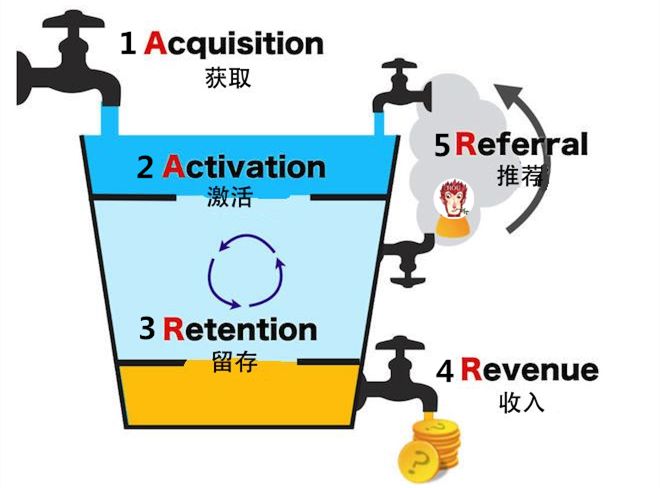
## 1.10 AARRR分析

AARRR模型是一个用于产品运营和用户行为分析的数据模型，可以应用在面向大众消费者的各个行业中，帮助业务成长。

AARRR模型对应产品运营的 5 个重要环节，分别是：

* 获取用户（Acquisition ）：用户如何找到我们？
* 激活用户（Activation）：用户的首次体验如何？
* 提高留存（Retention）：用户会回来吗？
* 增加收入（Revenue ）：如何赚到更多钱？
* 自传播（Refer）： 用户会告诉其他人吗？

如果我们把产品看做一个蓄水池的话，这5个环节可以画成下面这样：



在这个漏斗过程中，用户可能会一步一步慢慢流失，剩下的部分用户实现最终转化。我们分别看下每一个环节：

**一、获取用户（Acquisition ）：用户如何找到我们？**

在互联网行业中，很多创业公司死掉并不是因为他们的业务或产品不行，而是因为他们的获客成本很高，并且没有办法降下来。肖恩最初是在硅谷的Dropbox公司工作，这是一家做云存储业务的创业公司。这家公司开始时每获得一名用户的成本高达400美元，而它提供的付费服务每年是99美元。

肖恩对低成本获客，提供了两个建议：一是语言——市场匹配，二是渠道——产品匹配。

**1、语言——市场匹配**

语言——市场匹配，就是你怎么说才能打动用户的心。实际上考验的是抓住用户注意力的能力。现在人们关注一条网络信息的时间通常是8秒，如果不能在8秒内告诉用户你的产品对他有什么用，你就失去了一次获客的机会。

2001年iPod问世时，乔布斯完全可以在他传达的内容中解释为何他的播放器与众不同，性能更加优越，但他没有这样做。他明智地决定不使用任何形容MP3播放器功能的表述。他只用了一个简单而迷人的句子——“将1000首歌放在你的口袋里”，就彻底重塑了人们对便携式播放器魅力的认知。

**2、渠道——产品匹配**

渠道——产品匹配是说产品投放在什么渠道才能直达用户的视线范围。

渠道可以分为三类，一种是口碑渠道，适合病毒营销；一种是有机渠道，适合搜索引擎优化、内容营销；还有一种是付费渠道，有哪些付费渠道呢？

**1）展示位广告**

在网站或手机 App 的顶部、App 的开屏广告位等。开屏广告，就是当用户打开手机 App 时，会有几秒的广告时间，比如微博、知乎打开会先给你展示一个开屏广告。

这种类型的广告通常是按展示次数付费（CPM，Cost Per Mille），也就是有多少人看到了该广告。

**2）搜索广告**

比如搜索引擎（谷歌、百度等）的关键字搜索，电商搜索广告比如淘宝直通车等。广告主为某一个搜索关键词出价，用户看到的搜索结果是按广告主出价的高低来排名的。

这种类型的广告是按点击次数付费（CPC，Cost Per Click），也就是有多少人点击了该广告。

**3）信息流广告**

比如微博、今日头条、知乎、朋友圈（信息流）里的广告。这种广告是根据用户的兴趣爱好来推荐广告的。这种类型的广告是按点击次数付费（CPC）或者按投放的实际效果付费（CPA，Cost Per Action）。

按投放的实际效果付费（CPA，Cost Per Action）包括：

* CPD（Cost Per Download）：按 App 的下载数计费；
* CPI（Cost Per Install）：按激活 App 的数量计费；
* CPS（Cost Per Sales）：按完成购买的用户数或者销售额来计费。

在将产品投放去哪个渠道的时候，要清楚你的目标用户是谁，目标用户在哪。如果你是一款为企业服务的软件，在娱乐网站打广告就非常不合适。目前主流广告平台都支持这三种方式的付费：按展示次数付费（CPM）、按点击次数付费（CPC）和按投放的实际效果付费（CPA）。广告主可以按自己的产品需求来灵活选择。

**3、经典案例**

下面我们通过几个案例，来看如何获取用户：

**案例1：美国的流媒体视频巨头网飞公司（Netflix）**

美国的流媒体视频巨头网飞公司（Netflix），2017年，Netflix实现了一个似乎不可能实现的目标，一个视频网站的用户数量超过了美国全部有线电视用户的总和，成为全球最大的娱乐供应商。

为了寻找更多的用户，创业初期的Netflix选择与DVD播放机厂家合作——在DVD包装盒里放入自己租赁影碟的优惠券。在购买DVD光盘不太方便的当时，用户打开DVD就能看到租赁光盘的优惠券，非常方便。不仅给Netflix带来了用户，也促使更多用户购买DVD播放机。如此循环，Netflix的用户越来越多。

**案例2：健身app Keep**

创立四年的健身软件Keep，2018年用户数已经超过了1.3亿。Keep在从0到1的阶段是如何获取用户的呢？

创始人王宁发现很多社交媒体都是运动爱好者、减肥者的聚集地，于是他和团队进到这些社群里，以优质、原创内容作为种子。

王宁认为“内容即产品、内容即流量、内容即用户，要把内容做得足够扎实、足够好、足够广”，让用户从社交媒体的角度了解Keep。

等产品上线后，再告诉用户，这些好内容都在一款app上聚合了，大家可以来看也能跟着这些内容训练。Keep的第一波用户完全是靠内容沉淀下来的，此后这些用户又去通过口碑传播给更多人。

王宁给这个行为起了一个名字叫“埋雷计划”，就是先要付出一些东西，再慢慢收获一些东西。不过，使用“埋雷计划”时需要搞清楚自己的用户在哪里，再去深耕他们真正的需求，然后让这些用户开始喜欢你，愿意去感受你。

Keep从上线到用户破百万用了105天，在第921天的时候，用户数破亿。王宁说，第一个100万用的可能时间会很长，但是第一个一个亿就会很快的速度，不到三年的时间。

**案例3：社交软件WhatsApp（类似国内的微信）**

用户拉新一定要降低用户参与的门槛。有的app好不容易把用户吸引过来，但是整个注册过程超级反锁，用户一看太麻烦就跑了。

社交软件WhatsApp（类似国内的微信）被Facebook以190亿美元收购。WhatsApp将自己所有的精力都用在了如何让自己的产品更加简单易用，甚至不需要你创建用户名或密码就能使用。正因为它的简单易用，WhatsApp的口碑传播力非常强。

**案例4：付费渠道**

在2017年和2018年微信公众号“吴晓波频道”采购了外部增粉服务，分为三种形式：在其他公众号发送的图文消息中嵌入；通过智能秤屏幕广告投放，扫码后关注微信公众号方可获得体重信息推送；通过wifi平台，用户关注微信公众号后方可实现wifi联网。

其中，2017年通过外部渠道累计新增公众号关注用户人数8467人，占2017 年全年累计新增公众号关注用户的 0.66%，2017 年采购增粉服务所产生费用账面合计约1万元；

2018 年通过外部渠道累计新增公众号关注用户41.74万人，占 2018 年全年累计新增公众号关注用户的36.45%，占345.88万总关注用户数的12.07%，2018年采购增粉服务所产生费用账面合计约 40.02 万元。

**4、需要关注哪些指标？**

* **渠道曝光量**：有多少人看到产品推广的线索
* **渠道转换率**：有多少用户因为曝光转换成用户
* **日新增用户数**：每天新增用户是多少
* **日应用下载量**：每天有多少用户下载了产品
* **获客成本（CAC）**：获取一个客户所花费的成本

**二、激活用户（Activation）：用户的首次体验如何？**

很多手机应用，注册用户不少，但是打开率不高。你要做的是激活他们，让他们真正地使用产品。

想唤醒休眠用户，就得先摸清楚产品的“啊哈时刻”（Aha moment）。什么叫“啊哈时刻”呢？

说白了就是让用户感受到产品亮点情不自禁地发出赞叹的时刻，它对应的就是打动用户的产品亮点。比如网易云音乐，在众多音乐软件中突围的亮点就是评论，用户打开软件，可能第一件事不是听歌，而是看评论，心里面会有一些赞叹。

要想激活用户，得绘制一幅通往“啊哈时刻”的路线图。比如你负责的产品是个购物软件，在新用户体验到啊哈时刻之前，必须要完成下面这些步骤：下载app，找到所需商品，放入购物车，创建账户，输入姓名，加入信用卡和配送信息，然后点击购买。

在这一系列动作中，到底用户停留在了哪一步？是搜不到要的东西，还是创建账户太麻烦，或者是页面设置不合理？你要计算每个节点用户的流失率，来相应地提高产品性能，改善用户体验。

比如下面的某淘宝店铺，用户购物过程中每个节点的流失率。



很多人不知道自己app或者网站因为响应时长，因为不人性化的按钮设计，流失了多少人。

**2、经典案例：拼多多如何激活用户？**

电商app在反复冗余的跳转之后，用户最终放弃了支付。而拼多多为了减少购物环节和复杂性，从进入首页到支付仅有4个环节。要知道，淘宝搜索式购物流程有6个环节，浏览式购物流程有7个环节。拼多多是如何做到的呢？

拼多多通过四大做法减少网购的复杂性：

1）拼多多首页直接推荐产品，不放无效信息，不需点击多次才到达产品页面；

2）拼多多没有设置购物车，一件产品直接购买支付，减少了消费者犹豫时间；

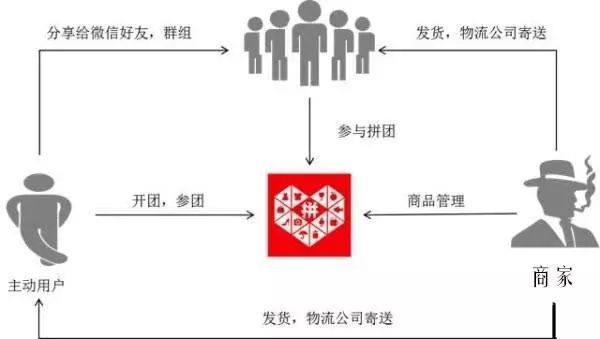
3）所有产品都包邮，减少与客服沟通的需要；

4）先付款后拼团，拼多多把拼团环节放在支付环节之后，让消费者提前用团购价完成支付，随后再找人拼团，拼成功了便发货，不成功便退款，将支付环节提前，尽快锁定消费者。

你还可以通过游戏化的做法，给用户奖励来唤醒用户，比如打卡、积分、发优惠券等。

在看拼多多是如何利用玩法激发活跃之前，先说说一个心理学上的“厌恶损失”现象。大家都应该听过一个故事，你丢了100块钱，后来又捡到100，但是你依旧感觉是痛苦的。失去100块的痛苦要大于得到100块的快乐。损失带来的痛苦远大于收益给你的满足。

拼多多很好的利用了这个心理，为什么大家乐于转发拼团？他们真的很在乎这几块钱吗？他们在乎的是不能亏损，不拼团就意味着损失。这样的心理下，你就能看到各种群里充斥着拼团，砍价的链接。



**3、需要关注哪些指标？**

* **日活跃用户数**(简称日活)：一天之内，登录或使用了某个产品的用户数。比如一天内打开我微信公众号的人数。类似的还有周活跃用户数，月活跃用户数。
* **活跃率**（活跃用户占比）：某一时间段内活跃用户在总用户量的占比。根据时间可分为日活跃率（DAU）、周活跃率（WAU）、月活跃率（MAU）等。

下面是三大电商2018年3月活跃用户数。我们可以看出拼多多基于微信生态的拼团游戏玩法，小程序月活跃数远高于淘宝和京东。



* 产品使用每个节点用户的流失率。用户流失率和留存率恰好相反，如果某产品新用户的次日留存为20%，那么反过来说明有80%的用户流失了。

如果是网页端的产品，还有两个指标来衡量网页端活跃。

* PV（PageView）是页面浏览量，用户每打开一个网页可以看作一个PV，用户看了十个网页，则PV为10。
* UV（Unique Visitor，独立访客数）是一定时间内访问网页的人数。在同一天内，不管用户访问了多少网页，他都只算一个独立访客。怎么确认用户是不是同一个人呢？技术上通过网页缓存cookie或者IP判断。如果这两者改变了，则用户算作全新的访客。

**三、留存：用户会回来吗？**

在第二个环节用户终于被激活唤醒了，这时候你的任务就变成了如何让用户变成回头客。留存的核心目标是让用户养成使用习惯。在这个阶段，你要明白习惯是如何形成的。

**1、经典案例**

**案例1：亚马逊会员服务**

亚马逊的会员服务在刚推出时，许多分析人士都说这个计划必定会失败，因为美国的配送成本是很高的，而99美元的会员服务给会员免费配送一年，亚马逊肯定会亏。但亚马逊的真正目标是改变人们的习惯，让他们在购物中习惯会员优惠价格，习惯免运费，从而不再去其他家买东西，从而提高用户留存。

**案例2：蚂蚁森林游戏**

蚂蚁森林是阿里巴巴推出的是一款养成类的小游戏，在支付宝里就可以打开看到。截至到2018年5月份，蚂蚁森林已经累计了3.5亿用户，成功拉升了支付宝一直渴望的用户活跃度。

用户激活蚂蚁森林的账号后，她的一部分日常行为，例如“乘地铁”、“走路”、“在线购买电影票”等都会被视为减排行为，然后，这些行为会被换算成一定的能量，显示在个人的主页面上，用户只要点击屏幕上的能量，就可以收取。

用户可以到其他用户的页面上去偷取对方的能量。当能量累计到一定程度，就可以申请种一棵树。蚂蚁金服会为用户发放一个带有特定编号的证书，并且真的在现实中种一棵树。同时，政府还会给你发放一个植树的证书。

蚂蚁森林累计种植真树超过1000万棵。如果我们按照种一棵树，需要用户在线上持续使用蚂蚁森林2个月（事实上实际需要的时间会超过2个月），这很好的增加了用户的粘性，提高了留存。

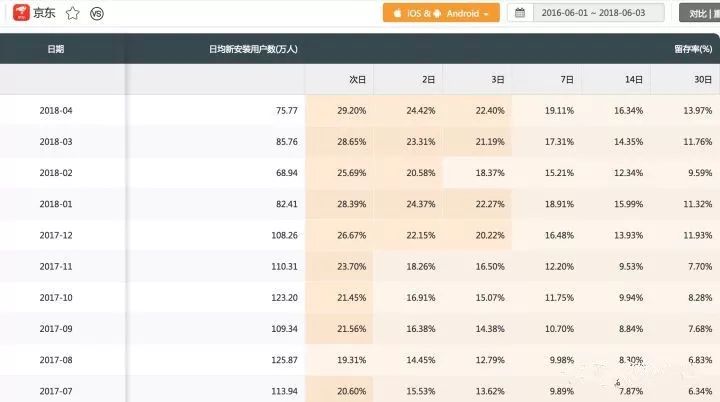
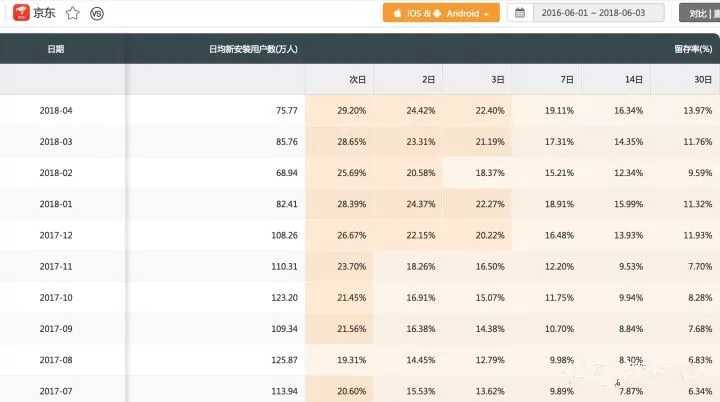
**2、需要关注哪些指标？**

* **次日留存率**：（当天新增的用户中，在第2天使用过产品的用户数）/第一天新增总用户数
* **第3日留存率**：（第一天新增用户中，在第3天使用过产品的用户数）/第一天新增总用户数
* **第7日留存率**：（第一天新增的用户中，在第7天使用过产品的用户数）/第一天新增总用户数
* **第30日留存率**：（第一天新增的用户中，在第30天使用过产品的用户数）/第一天新增总用户数

比如某个app第一天新增用户100个，第二天这100个人里有30个人打开过app，那么次日留存率=30/100=30%。如果第七天这100个人人里还有10个人打开过app，那么称七日留存率=10/100=10%

Facebook有一个著名的40-20-10法则，即新用户次日留存率为40%，7日留存率为20%，30日留存率为10%，有这个表现的产品属于数据比较好的。

下面是手机淘宝和拼多多的日活和留存率。（数据来自Questmobile）。我们可以看到，手机淘宝和拼多多的次日留存率都在 30% 左右。京东留存相对差些。



**四、增加收入：如何赚到更多钱？**

完成了上面所有的步骤，终于来到最有挑战性的一关，那就是如何从用户身上获得真正的收益。比如让用户续订、创造更多的广告位等等。

虽然商业模式不同，变现的方式不同，但是有一个概念需要重视，那就是夹点（pinch point），它指的是损失潜在收益的地方。

对于电商企业来说，用户从挑选商品到完成购买之间是一个危险区，很多人中途会放弃购买。要评估这些常见夹点，分析用户在这些点放弃的原因。

**1.经典案例：Facebook**

有段时间Facebook发现自己用户流失得很厉害，之前注销账号的确认页面有一段提示：你是不是要注销啊，你确认下，你再多确认一下，OK你去意已决那就给你注销。

后来Facebook改变了下确认页面，在注销结束之前增加了一个页面。这个页面它会把跟你互动比较亲密的五个朋友头像列出来，然后配上文字说：你确定你要注销嘛？

你如果要注销的话，这些人再也看不到你了，你也再也关注不了这些朋友的新的情况了，你是不是要再考虑一下？

就是这个东西增加了3％的留存可能，相当于Facebook一年有三百万个流失就这样被避免掉了。

**2、需要关注哪些指标？**

这里我们要注意“**用户**”和“**付费用户**”的区分，这也将影响收入的计算。（下面提到的“用户”，如果没有特别说明是“付费用户”，那么就是指全部用户（包括付费用户和非付费用户））。

* **客单价**：每位**用户**平均购买商品的金额。客单价=销售总额/顾客总数
* **PUR(Pay User Rate)**：付费用户占比
* **ARPPU**(Average Revenue Per Paying User)：某段时间内，**付费用户**的平均收入（不包括未付费的用户）。ARPPU=总收入/付费用户数
* **生命周期价值**(LTV，Lift Time Value)：，即平均一个用户在首次登录游戏到最后一次登录游戏内，为该游戏创造的收入总计。常用于游戏行业。
* **复购率**：一定时间内，消费两次以上的用户数 / 总购买用户数。比如有10个用户购买了产品，5个用户产生了重复购买，则重复购买率=5个重复购买用户数 / 10（总购买用户数=50%
* **销售额**=用户数\*转化率\*客单价\*购买频率（购买频率是指消费者或用户在一定时期内购买某种或某类商品的次数。）

我们来看一个例子，截止 2018 年 3 月 30 日，拼多多单个买家过去 12 个月的消费额是 673.9 元（客单价），同时期阿里巴巴用户的年消费额是 8696 元，京东用户的年消费额则是 4426 元是拼多多的数倍乃至十几倍。



**五、自传播： 用户会告诉其他人吗？**

前面的4个环节做完，就到了第5个环节自传播，也叫病毒式营销。

美国作家马尔科姆·格拉德威尔的书《引爆点》(The Tipping Point)中用流行病类比流行，提出了这样一个观点，引爆一种流行病不只有一种途径，流行病爆发需要三个条件：

1）传染物本身

2）传染物发挥作用所需的环境

3）人们传播传染物的行为

* **传染物本身**

传染物本身是说要对自己的产品有足够的了解。试着问自己一个问题：我的产品是否真正解决了用户的痛点？

如果你是写文章的，就要考虑你的文章能为用户带来哪些真正的价值。如果你是做餐饮的，你的菜品是否真的为用户所喜欢。

* **传染物发挥作用所需的环境**

也就是你的用户在哪里？对应我们前面讲的AARRR模型第一关环节（获取用户）。要去思考使用你产品的刚需用户经常在哪些环境（社区，大学等等渠道）中出现。

* **人们传播传染物的行为**

在对自己的产品有了深刻洞察，同时找到你的目标人群后，还要考虑到人们会因为什么目的去分享你的产品，让更多的人看到你的产品。这个条件对应的就是AARRR模型第五个环节（推荐，病毒式营销）。

**1、经典案例**

比如滴滴通过红包这样的超级营销工具一样，实现了产品的自传播。

还有民宿平台Airbnb有一个“旅行基金”的"功能，当你每推荐一位好友注册Airbnb，你和好友都可以获得礼金券。

如果没有这种自传播，你的信息流就中断，从而无法实现自增长。好比，你在朋友圈转发了一篇自己写的文章，与此同时，有10个朋友看到这篇文章，但是都没有进一步去转发到他们各自的朋友圈去分享，那么你的这篇文章的信息流到此已经中断，不会进行二次传播了。

需要注意的是，坚持把每一次营销当做一种产品体验的理念，热点不能只是噱头，而是希望同时兼顾一些普惠型的营销活动，真正能够为消费者带来好处。

**2.第五环节推荐：需要关注哪些指标？**

* **转发率**：某功能中，转发用户数 / 看到该功能的用户数。比如我的微信公众号（猴子聊人物）推送一篇文章给3万用户，转发这篇文章的用户数是3000，那么转发率=3000（转发这篇文章的用户数） / 3万（看到这篇文章的用户数）=10 %。
* **转化率**：计算方法与具体业务场景有关。我们举几个栗子。

比如，淘宝转化率=某段时间内产生购买行为的客户人数/所有到达店铺的访客人数，假如双11当天，有100个用户看到了你店铺的推广信息，被吸引进入店铺，最后有10个人购买了店铺里的东西，那么转化率=10（产生购买行为的客户人数）/100（所有到达店铺的访客人数）=10%。

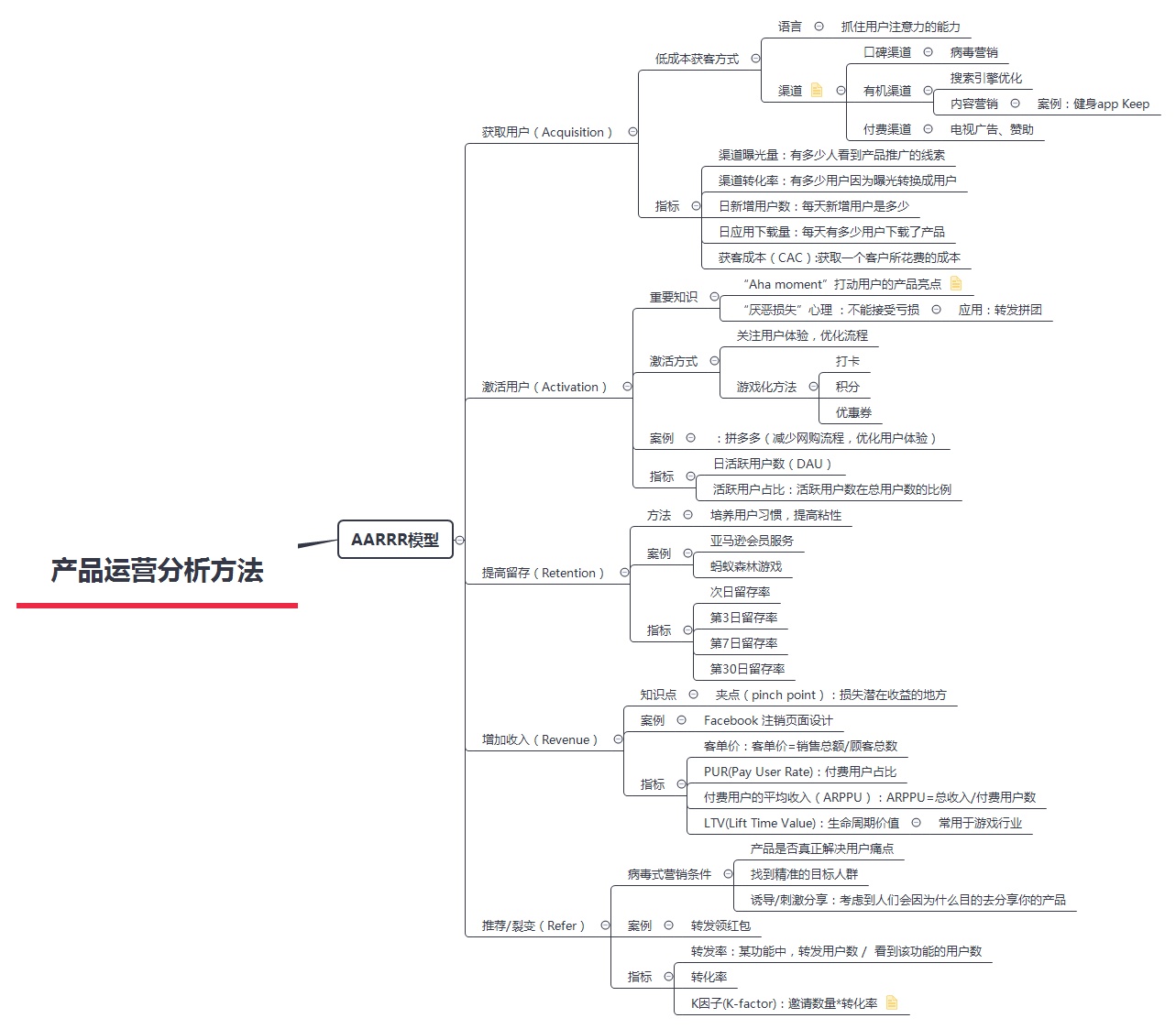
广告转化率=点击广告进入推广网站的人数 / 看到广告的人数，比如我们经常使用百度，搜索结果里会有广告，如果有100个人看到了广告，其中有10个人点击广告进入推广网站，那么转化率=10（点击广告进入推广网站的人数） / 100（看到广告的人数）=10%。

* **K因子**(K-factor)：用来衡量推荐的效果，即一个发起推荐的用户可以带来多少新用户。

K 因子= (每个用户向他的朋友们发出的邀请的数量) \* (接收到邀请的人转化为新用户的转化率)。

假设平均每个用户会向20个朋友发出邀请，而平均的转化率为10%的话，K =20\*10%=2。这个结果还算是不错的效果。

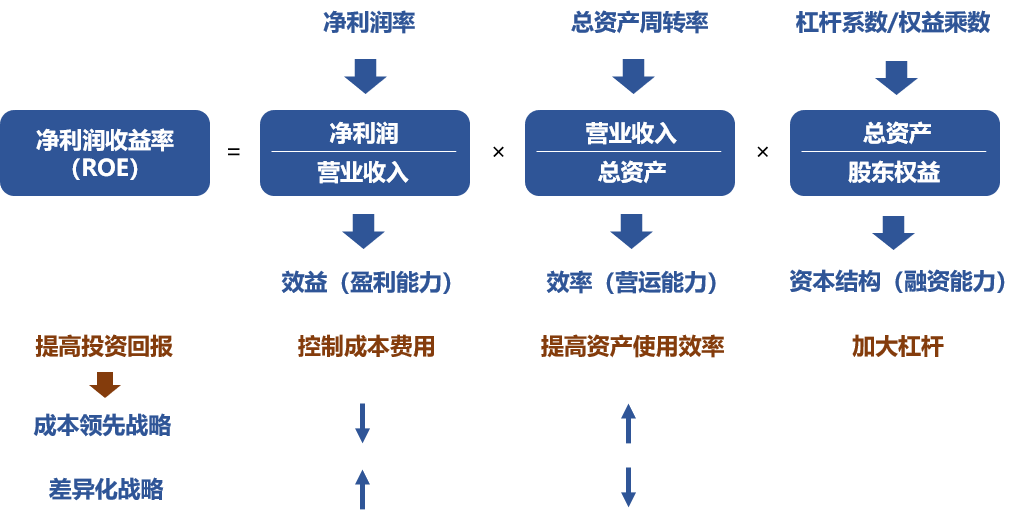
当K>1时，用户群就会象滚雪球一样增大。如果K<1的话，那么用户群到某个规模时就会停止通过自传播增长。



## 1.11 杜邦分析

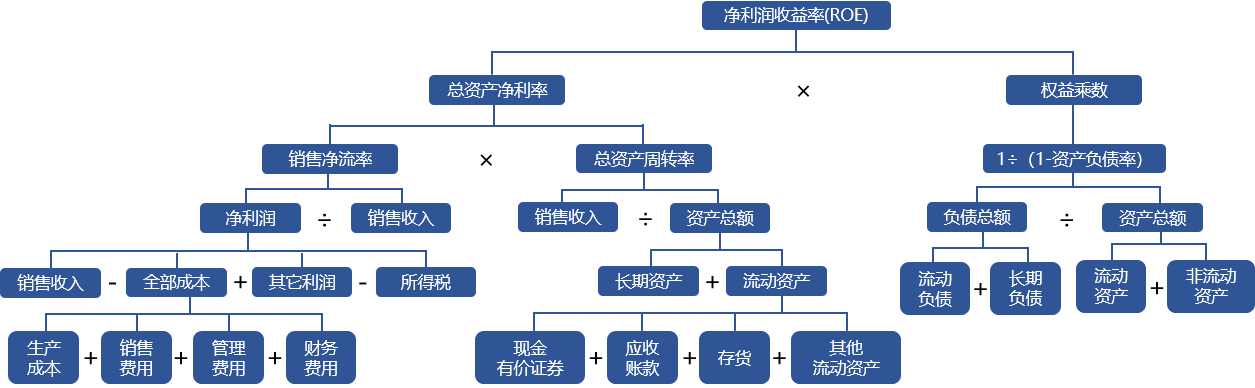
**1.11.1 是什么？**

杜邦分析法是将企业**净资产收益率**逐级分解为多项财务比率的乘积，这样有助于深入分析比较企业的经营业绩。



**1.11.2 有什么用？**

将净资产收益率逐步分解，形成一个完整的指标体系，如下图。图中显示了销售净利率、总资产周转率和债务比率之间的相互关联关系，为分析人员全面仔细地了解企业的经营和盈利状况提供方便。



杜邦分析法很好地揭示了决定企业获利能力的三个因素：

（1）**成本费用控制能力**：因为销售净利润率=净利润/销售收入=1-(生产经营成本费用+财务费用本+所得税)/销售收入，而成本费用控制能力影响“(生产经营成本费用+财务费用+所得税)/销售收入”，从而影响销售净利润率。

（2）**资产的使用效率**：用资产周转率(销售收入/总资产)反映。它表示融资活动获得的资金(权益+负债)，通过投资形成的总资产,其每一单位资产能产生的销售收入。虽然不同行业资产周转率差异很大，但对同一个公司，资产周转率越大，表明该公司的资产使用效率越高。

（3）**财务上的融资能力**：用权益乘数(总资产/股东权益)反映。若权益乘数为4，表示股东每投入1个单位资金，公司就能借到3个单位资金，即股东每投入1个单位资金，公司就能用到4个单位的资金。权益乘数越大，即资产负债率越高，说明公司过去的债务融资能力越强。

一个企业，若想赚到更多的钱，就要提高净资产收益率。想要提高净资产收益率，就要从三个方面着手：

* 提高销售净利率
* 提高总资产周转率
* 提高权益乘数

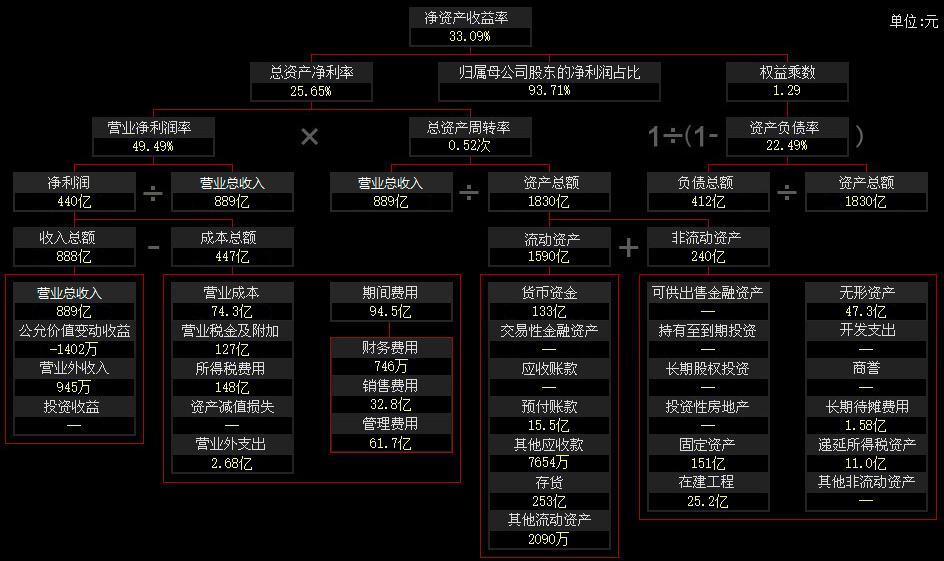
**1.11.3 如何用？**

如上所述，大体来说，企业最终利润的高低由三个因素决定，即毛利率高低、周转率快慢和经营杠杆大小。按照不同的盈利方式可以分成三种模式：“**茅台模式”、“沃尔玛模式”和“银行模式”**，这三种情况分别代表产品竞争力强弱、管理层运行能力的高低、企业承担的风险大小。对于不同的企业，都可以考虑一下他们靠什么赚钱，然后就能发现这些公司需要自己追踪的要点。

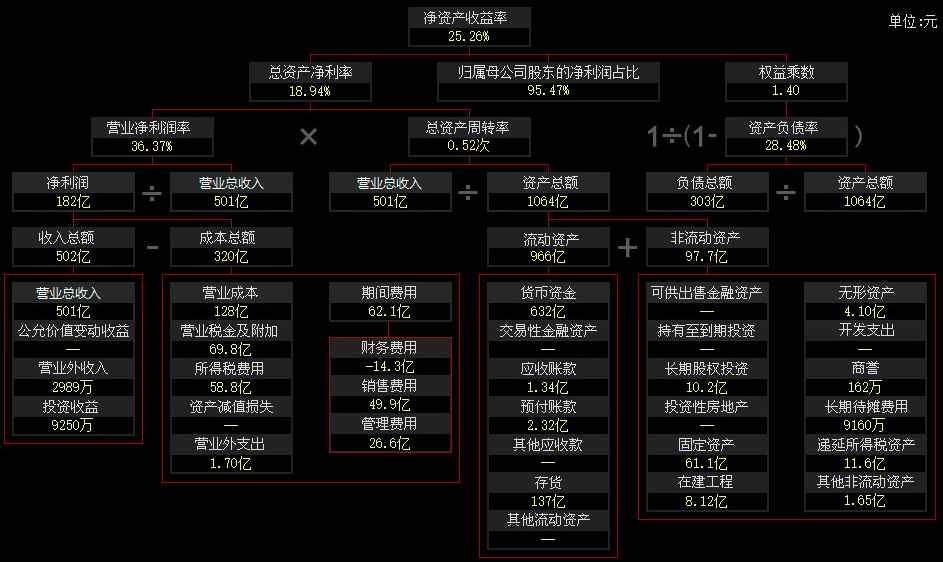
* 茅台模式企业：随时注意产品口碑，竞品替代的可能，例如白酒、游戏行业。
* 沃尔玛模式企业：管理层是否出现较大变化，例如制造业。
* 银行模式企业：宏观上资金的松紧，例如金融、房地产行业。

**案例：以大消费板块的四个白酒企业为例**

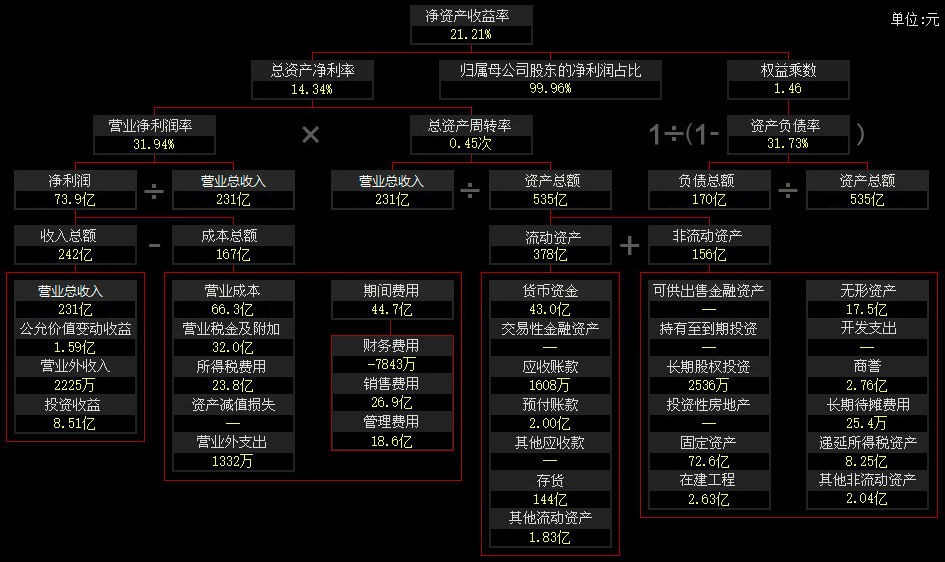
贵州茅台杜邦分析结构图（2019年年报）



五粮液杜邦分析结构图（2019年年报）



洋河股份杜邦分析结构图（2019年年报）



古井贡酒杜邦分析结构图（2019年年报）



我们就杜邦分析法的四要素（净资产收益率、销售净利率、总资产周转率、杠杆倍数）对上述四个公司进行对比分析，列表如下：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | ROE | 销售净利润 | 总资产周转率 | 杠杆倍数 | 市值 | PE | PS |
| 贵州茅台 | 33.09% | 49.49% | 0.52次 | 1.29 | 2.3万亿 | 51.58 | 24.22 |
| 五粮液 | 25.26% | 36.37% | 0.52次 | 1.4 | 1.07万亿 | 55.08 | 19.25 |
| 洋河股份 | 21.21% | 31.94% | 0.45 | 1.46 | 3296亿 | 44.4 | 15.74 |
| 古井贡酒 | 25.55 | 20.71 | 0.79 | 1.47 | 1371亿 | 72.4 | 13.33 |

通过杜邦分析法的对比分析，五个标的之间，资产结构的优劣、盈利能力的高下、资产利用效率的高低如下：

（1）净资产收益率：**贵州茅台**>古井贡酒>五粮液>洋河股份

（2）销售净利率：**贵州茅台**>五粮液>洋河股份>古井贡酒

（3）总资产周转率：古井贡酒>**贵州茅台**=五粮液>洋河股份

（4）财务杠杆：古井贡酒>洋河股份>五粮液>**贵州茅台**

可见茅台科技吊打其他白酒企业，古井贡酒总资产周转率较高，若是同样的净资产规模，古井贡酒的使用效率为1.16，而贵州茅台为0.67，说明古井贡酒的资金使用效率明显更高，管理团队更好。

## 1.12 波士顿矩阵

**1.12.1 是什么？**

波士顿矩阵（BCG Matrix），又称市场增长率-相对市场份额矩阵，是通过“市场引力”以及“企业实力”两个基本因素来分析企业产品结构。

进行波士顿矩阵分析的目的是为企业通过此种方法分析，来确定业务发展方向以保证企业收益。核心在于要解决如何使企业的产品品种及其结构适合市场需求的变化。关键是如何协助企业分析与评估其现有产品线，利用企业现有资金进行产品的有效配置与开发。

**1.12.2 有什么用？**

波士顿矩阵认为，一般决定产品结构的基本因素有两个：市场引力与企业实力。

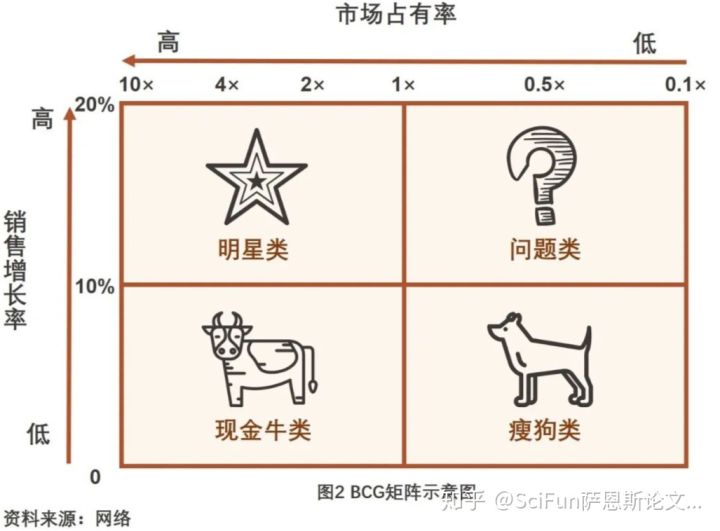
* **市场引力：包括“整个市场的销售量（额）增长率”、“竞争对手强弱”及“利润高低”等。其中“销售增长率”是决定企业产品结构是否合理的“外在因素”，它是反映市场引力的综合指标。**
* **企业实力：包括“市场占有率”、“技术、设备、资金利用能力”等，其中“市场占有率”是决定企业产品结构的“内在要素”，它直接显示出企业竞争实力。**

通过“销售增长率”与“市场占有率”两个因素相互作用，会出现四种不同性质的产品类型即“问号类”、“新星类”、“现金母牛类”、“瘦狗类”，并形成不同的产品发展前景：

* 明星类产品：销售增长率和市场占有率“双高”的产品群 ；
* 瘦狗类产品：销售增长率和市场占有率“双低”的产品群；
* 问号类产品：销售增长率高、市场占有率低的产品群；
* 现金牛类产品：销售增长率低、市场占有率高的产品群。

应用到波士顿矩形坐标图中时，企业销售增长率做纵轴，市场占有率做横轴，各以10%和20%作为区分高、低的中点，将坐标图划分为四个象限，依次为“问号”、“明星”、“现金牛”、“瘦狗”。

企业可将产品按各自的“销售增长率”和“市场占有率”归入不同象限，通过产品所处不同象限的划分，使企业采取不同决策，以保证其不断地淘汰无发展前景的产品，保持“问号”、“明星”、“现金牛”、“瘦狗”产品的合理组合，实现产品及资源分配结构的良性循环。



按照波士顿矩阵分析，最佳的产品组合是：

* **迅速淘汰“问题产品” ；**
* **让“明星产品”转化成“现金牛” ；**
* **保证“现金牛”不受冲击；**
* **暂时保留“瘦狗产品” 。**

波士顿矩阵分析对于企业产品所处的四个象限具有不同的定义和相应的战略对策：

**（1）明星型产品（Stars）——新增长盈利的产品**

这类产品可能成为企业的现金母牛产品，需要加大投资以支持其迅速发展。采用的发展战略是：增长战略，积极扩大经济规模和市场机会，以长远利益为目标，提高市场占有率，加强竞争地位。但只有具备规模效应综合实力较强的企业才宜实行 “明星”战略，因为将“新星”转化为“明星”不但有一较长时间过程、还要一个资金大量投入的过程，同时对企业相关条件要求过高，如有系统的新产品战略， 市场预测，新技术储备、人力资源，研发资金支持、以及有能较快的将“新星”产品转化为“明星”产品的企划思想和营销能力等。

**（2）问题型产品（Question Marks ）——衰退的产品**

对此类型产品，企业将面临选择加强(市场渗透、市场开发或产品开发战略)还是进行出售、放弃的决定。只有那些符合企业发展长远目标、企业具有资源优势、能够增强企业核心竞争力的业务才得到肯定的回答。得到肯定回答的问题型业务适合于采用增长战略，目的是扩大SBUs的市场份额，甚至不惜放弃近期收入来达到这一目标，因为问题型要发展成为明星型业务，其市场份额必须有较大的增长。得到否定回答的问题型业务则适合采用收缩战略。

**（3）现金牛型产品（Cash Cows）——现金流量大，保证公司运转的产品**

“母牛”产品作用是提供稳定销量，由于市场已经成熟，企业不必大量投资来扩展市场规模，同时作为市场中的领导者，该业务享有规模经济和高边际利润的优势，因而给企业带来大量现金流。可采用收获战略：即所投入资源以达到短期收益最大化为限。①把设备投资和其它投资尽量压缩；②采用榨油式方法，争取在短时间内获取更多利润，为其它产品提供资金。对于这一象限内的销售增长率仍有所增长的产品， 应进一步进行市场细分，维持现存市场增长率或延缓其下降速度。对于现金牛产品，适合于用事业部制进行管理，其经营者最好是市场营销型人物。

**（4）瘦狗型产品（Dogs)——已过成熟期的产品**

一般情况下，这类产品常常是处在低增长率、低市场占有率象限内的产品群。其财务特点是利润率低、 处于保本或亏损状态，负债比率高，无法为企业带来收益。对这类产品应采用撤退战略：首先应减少批量，逐渐撤退，对那些销售增长率和市场占有率均极低的产品应立即淘汰。其次是将剩余资源向其它产品转移。 最后是整顿产品系列，最好将瘦狗产品与其它事业部合并，统一管理。

**1.12.3 如何用？**

**（1）核算企业各种产品的销售增长率和市场占有率。**

* **销售增长率**：这一指标用以表示发展前景，可以用本企业的产品销售额或销售量增长率，时间可以是一年或是三年以至更长时间。
* **市场占有率**：这一指标用以表示竞争力。可以用相对市场占有率或绝对市场占有率，但是需要用最新资料。



**（2）确定纵坐标“销售增长率”的一个标准线，从而将“市场增长率”划分为高、低两个区域。**

比较科学的方法有两种：

* **把该行业市场的平均增长率作为界分点**
* **把多种产品的市场增长率(加权)平均值作为界分点**

需要说明的是，高市场增长定义为销售额至少达到10%的年增长率（扣除通货膨胀因素后）。

**（3）确定横坐标“市场占有率”的一个标准线，从而将“相对市场份额”划分为高、低两个区域。**

确定标准线时要考虑的情况较多，一种比较简单的方法是，高市场份额意味着该项业务是所在行业的领导者的市场份额。需说明的是，当本企业是市场领导者时，这里的“最大的竞争对手”就是行业内排行老二的企业。

**（4）标明各项业务在波士顿矩阵图上的位置。**

画图方法是以业务在二维坐标上的坐标点为圆心画一个圆圈，以圆圈的大小来表示企业每项业务的销售额。

到了这一步公司就可以诊断自己的业务组合健康与否。一个失衡的业务组合有太多的劣狗类或问题类业务，或太少的明星类和现金牛类业务。

例如有三项的问题业务，不可能全部投资发展，只能选择其中的一项或两项，集中投资发展；只有一个现金牛业务，说明财务状况是很脆弱的，有两项瘦狗业务，这是沉重的负担。

**案例应用：**

在充分了解波士顿矩阵的概念和分析应用法则之后，我们再来通过一个例题将理论运用到实际案例中去。

上海和达汽车零部件有限公司是某国内上市公司和外商合资经营的生产汽车零部件的企业。和达公司的主要产品分成五类，一是挤塑和复合挤塑类（密封条、车顶饰条等）；二是滚压折弯类（车门导槽、滑轨等）；三是普通金属焊接类；四是激光焊接类(铝镁合金横梁模块)；五是排档杆类(手动排档总成系列)。



针对上图所示情况，根据波士顿矩阵原理，进行如下分析：

滚压折弯类产品由于技术含量不高进入门褴低，未来市场竞争程度必然加剧。所以对于这类产品最好的选择就是舍弃。由于目前还能带来利润，不必迅速退出，只要目前依然保持必要的市场份额，公司不必再增加投入。当竞争对手大举进入时，可以舍弃。确认C、D品牌为新星品牌，虽然目前不是公司的主要利润来源，但发展潜力很大，决定加大资金投放力度，加快发展步伐，扩大与竞争对手的差距，力争成为公司新的利润增长点。

铝横梁的真空电子束焊接系统是国内第一家。具有技术上的领先优势。因此企业应该加大对这一产品的投入。以继续保持技术上的领先地位。对于排档杆类产品由于国内在这个领域的竞争程度还不太激烈，因此可以考虑进入。和达公司应该把这类产品作为公司的明星业务来培养。要加大对这方面的资金支持。在技术上应充分利用和寻找国外已具有同等类似产品的厂商进行合作，尽快回笼资金。

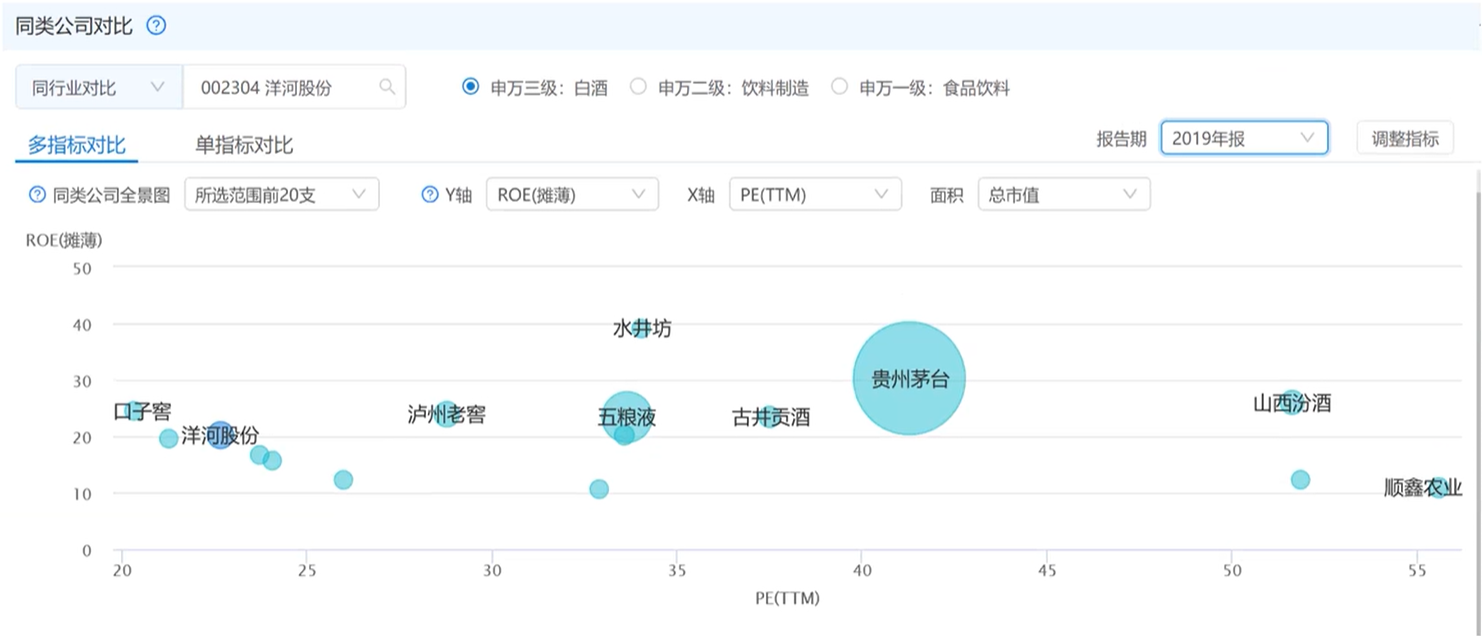
普通金属焊接类产品即是现金牛类产品。由于进入市场的时机较早，产品价格不错，每年能够给企业带来相当的利润。因此对于和达公司来说，对于金属焊接类产品，应该保持住目前的市场份额。把从这个产品中获取的利润投入到铝横梁和排档杆的产品中去。

普通塑料异型挤出和异型体复合挤出类产品因设备陈旧等原因。在国内已落后于主要竞争对手。从公司战略的角度出发，应该不断对这一块进行收缩。不必再投入更大的精力和财力，逐渐把注意力集中在激光焊接和排档杆的业务上去。

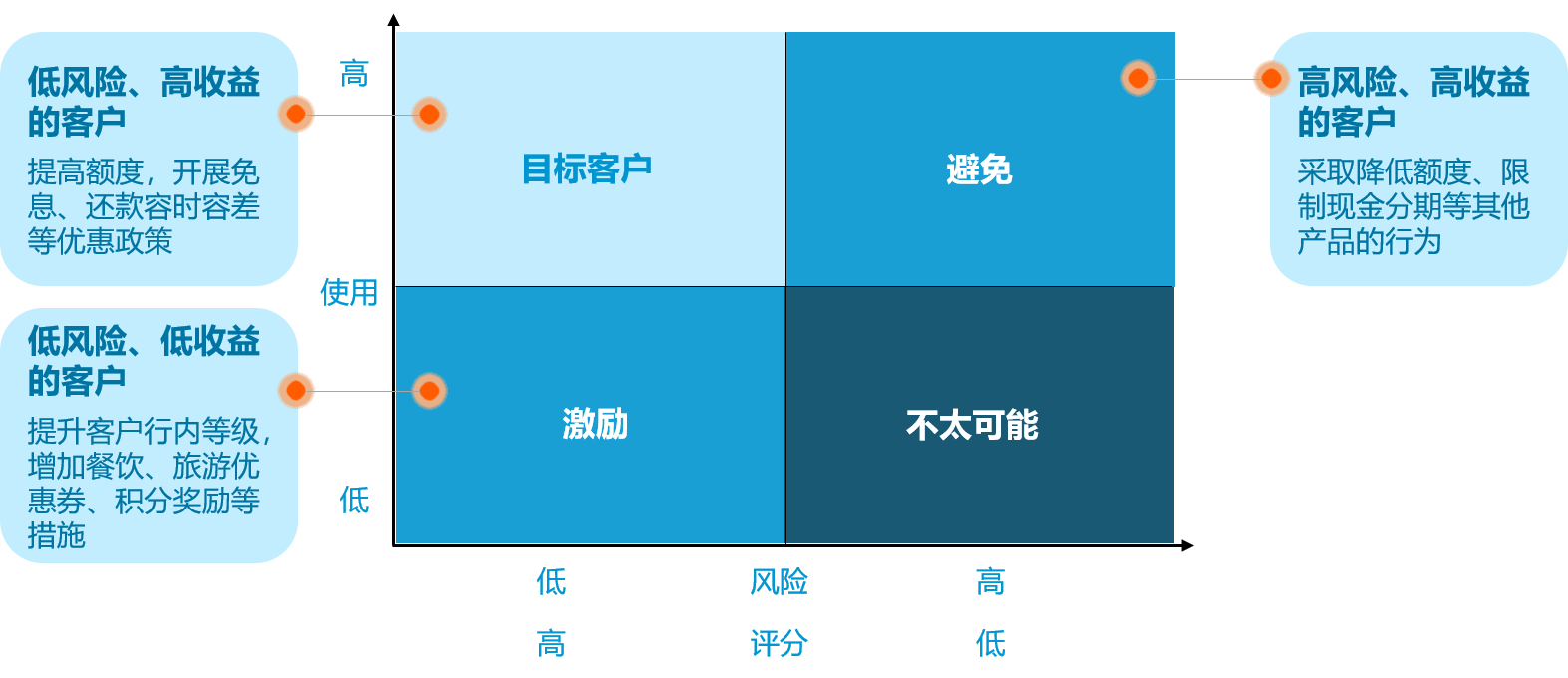
通过运用波士顿矩阵分析，可以使公司明确产品定位以及发展方向，对于企业投资的选择起到了举足轻重的作用。但波士顿矩阵在具体的应用还有很多困难，在使用波士顿矩阵法过程中要尽量查阅更多资料，谨慎分析，避免因方法上的缺陷而造成决策上的失误。

**1.12.4 注意事项**

波士顿矩阵分析方法可以作为一种分析框架，不能教条化的只从“市场引力”以及“企业实力”两个维度去分析，比较维度可以根据需求自行定义，比如从盈利能力和估值角度，我们可以定义ROE-PE分析矩阵，从盈利能力和成长性角度，我们可以定义净利润-利润增长率分析矩阵。



此外，在信贷领域进行授信时，我们可以考虑从风险和客户需求两个角度，定义信用评分-客户需求授信矩阵，也可从风险和收益两个角度，定义信用评分-客户使用率授信模型。

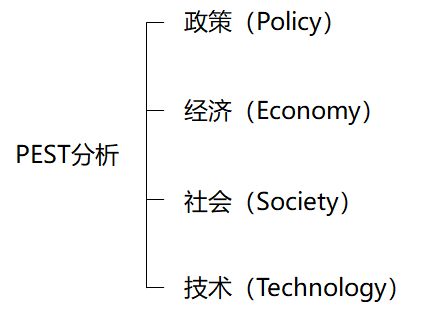


**应用：同一行业哪个公司更有投资价值？信贷领域的授信矩阵**

## 1.3 PEST分析

**1.3.1 是什么？**

PEST分析是战略咨询顾问用来帮助企业检阅其外部宏观环境的一种方法，从政策、经济、社会、技术四个方面来分析宏观环境。



**1.3.2 有什么用？**

* 外部因素主要包括p、e、s、t四个方面，作为战略决策依据，PEST可以从宏观角度全面地分析外部环境。
* 利用不同的角度，从变动的因素上探求某个行业可能的发展潜能，对企业的发展前景有一个大的整体把握。
* 对于各方面的变动可以及时地作出反应，制定出对应的改变策略。

**1.3.3 如何用？**

接下来，通过一个案例来看一下，如何使用PEST分析研究宏观环境。

**案例：新能源汽车行业的宏观分析**

**（1）政策**

双积分政策的托底以及新能源汽车产业规划（2020-2035）的引导，推动新能源汽车产业链的发展。

**（2）经济**

人民的生活水平和消费水平越来越高，国民经济稳中有进增长态势为新能源汽车产业发展提供了保障。

我国家庭汽车保有量逐年增多，新能源汽车也呈现了较高的增长态势，随着新能源汽车价格的下降，以及续航里程和充电基础建设的逐步提升，新能源汽车的产销量会大幅增长。

**（3）社会**

我国的能源消费量逐年增加，根据预测，2030年我国石油消耗将超过8亿吨，石油进口依存度将达到75%左右，在全球政治环境日趋紧张之下，能源安全正在面临着严峻挑战（想想当前芯片之痛，能源被卡脖子必将带来更大威胁），新能源汽车的发展能够降低石油的对外依存度，保证能源安全，并且有利于缓解环境污染问题。

**（4）技术**

从以上三个层面，可以确信整个大环境对新能源汽车的推动是确定的，新能源作为一个长期赛道，完全毋庸置疑。

如果说前两年还有人怀疑新能源汽车是不是未来，那么现在问题就变成了新能源汽车能走多远？能走多远一方面取决于别人愿不愿意买，另一方面取决于买不买得起，而想要解决这两方面就要取决于技术。

首先，第一点是买不买得起，想要更多人买得起新能源汽车，就要通过技术手段降低成本。接下来，以特斯拉Model3为例，来看一下如何降低特斯拉的成本（用逻辑树来分解）：



通过逻辑树分解Model3的工业成本，我们清楚地知道降低新能源汽车的成本主要靠降低电池的成本。电池成本的降低可以从电池技术的改进和电池工艺的改进着手，显然，工艺改进比技术改进容易多了，而这也是新能源汽车制造商正在做的。

然后，第二点是如何让更多的人愿意买，这就涉及到新能源汽车一直以来被质疑的最大问题——**续航能力**。续航能力当然要取决于电池中蕴含了多少电，电池里面蕴含多少电的核心是电芯，电芯的含电能力用能量密度来表示，单位kw/kg，也就是一公斤电芯能发多少电。如果说传统燃油车加满油可以跑600公里，目前新能源汽车的实际里程也就400-500公里。对于液态电解液电池，想要通过提高电芯的能量密度达到和传统燃油车相同的里程十分困难，当然也有解决办法，也就是固态电池，不过目前技术上还没有实现，什么时候实现，技术上的事，这谁说得准，不过总有一天能实现。

以上采用PEST分析方法对新能源汽车的外部宏观环境做了分析，容易看出，新能源汽车确实是未来，国家也在大力扶持，不管从环保上还是从能源安全上都必将不遗余力的支持。不过，我们也发现新能源车当前的关键问题出在电池技术上，技术的问题不是一朝一夕能够解决的，不过在国家的大力扶持下，能够砸出光伏行业的世界级龙头隆基股份，也必将能砸出电池行业的世界级龙头（目前看宁德时代已有王者风范）。不久之后，当大量新能源汽车奔跑于大街小巷之时，新能源汽车还只是汽车么？就好像苹果手机只是手机么？苹果手机只是终端，其核心应该是苹果OS，而新能源汽车的核心呢？当然应该是汽车OS。

**1.3.4 注意事项**

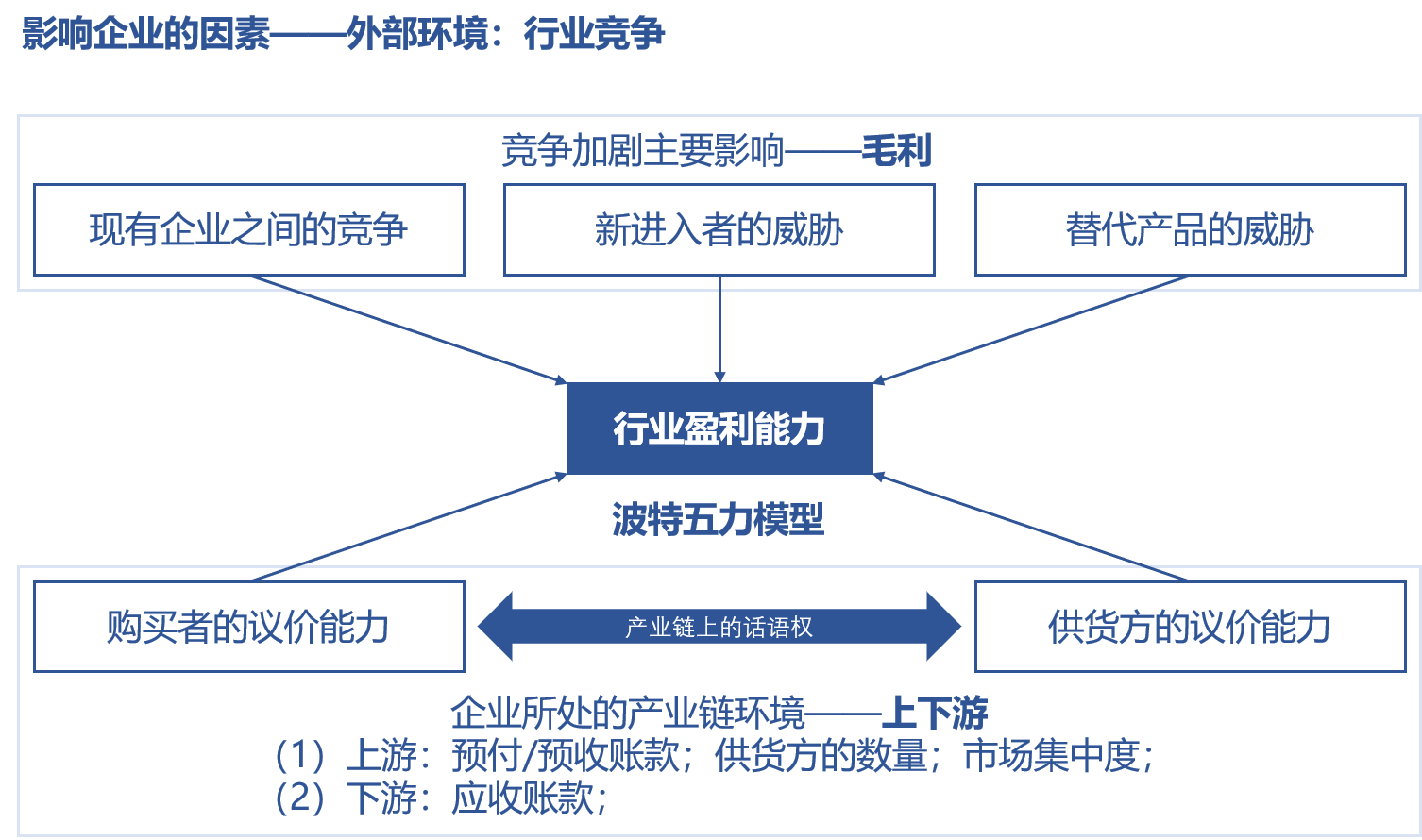
* 变化因素大
* 企业决策需要考虑各种因素，PEST作为参考
* 分析过程可以参考行业报告

## 1.13 波特五力模型

**1.13.1 是什么？**

迈克尔·波特(Michael Porter)于80年代初提出了波特五力分析模型，他认为行业中存在着决定竞争规模和程度的“五种力量”，即“供应商的议价能力、购买者的议价能力、潜在竞争者进入的能力、替代品的替代能力、行业内竞争者现有的竞争能力”。这五种力量综合起来，影响着产业的吸引力和现有企业的竞争战略决策。五种力量的不同组合变化最终影响行业利润潜力变化。

该种竞争战略分析方法，将大量不同的因素汇集在一个简便的模型中，帮助企业确定在某行业中的竞争优势和可能达到的资本回报率，并以此来分析该行业的基本竞争态势。从一定意义上来说，波特五力分析隶属于外部环境分析方法中的微观分析，通过它可以让企业预计自身的投资期望，同时找到企业自身的战略方向。



**1.13.2 有什么用？**

**（1）供应商的议价能力（Bargaining Power of Suppliers)**

供方主要通过**提高投入要素价格**与**降低单位价值质量**的能力，来影响行业中现有企业的盈利能力与产品竞争力。供方力量的强弱主要取决于他们所提供给买主的是什么投入要素，当供方所提供的投入要素其价值构成了买主产品总成本的较大比例、对买主产品生产过程非常重要、或者严重影响买主产品的质量时，供方对于买主的潜在讨价还价力量就大大增强。

一般来说，满足如下条件的供方集团会有较强大的讨价还价力量：

* 供方行业为一些具有比较稳固市场地位而不受市场剧烈竞争困挠的企业所控制，其产品的买主很多，以至于每一单个买主都不可能成为供方的重要客户。
* 供方各企业的产品各具有一定特色，以至于买主难以转换或转换成本太高，或者很难找到可与供方企业产品相竞争的替代品。
* 供方能够方便地实行前向联合或一体化（产业链纵向延伸的能力），而买主难以进行后向联合或一体化。

**（2）购买者的议价能力** **(Bargaining Power of Buyers)**

购买者主要通过**压价**与**要求提供较高的产品或服务质量**的能力，来影响行业中现有企业的盈利能力。

一般来说，满足如下条件的购买者可有较强的讨价还价力量

* 购买者的总数较少，而每个购买者的购买量较大，占了卖方销售量的很大比例；
* 卖方行业由大量的规模相对较小的企业所组成；
* 购买者所购买的基本上是一种标准化产品，同时向多个卖主购买产品在经济上也完全可行；
* 购买者有能力实现后向一体化，而卖主不可能前向一体化。（注：简单来说即“客大欺主”）；

**（3）新进入者的威胁（Threat of New Entrants）**

新进入者在给行业带来新生产能力、新资源的同时，会希望在现有已被企业瓜分完毕的市场中赢得一席之地，这将可能导致与现有企业发生原材料与市场份额的竞争，最终使行业中现有企业盈利水平降低，严重的话还有可能危及现有企业生存。竞争性进入威胁的严重程度取决于两方面的因素：进入新领域的障碍大小与预期现有企业对于进入者的反应情况。

* 进入障碍主要包括规模经济、产品差异、资本需要、转换成本、销售渠道开拓、政府行为与政策、不受规模支配的成本劣势、自然资源、地理环境等方面，其中有些障碍很难借助复制或仿造的方式来突破。
* 预期现有企业对进入者的反应情况，主要是采取报复行动的可能性大小，取决于有关厂商的财力情况、报复记录、固定资产规模、行业增长速度等。

总之，新企业进入一个行业的可能性大小，取决于“进入者主观估计进入所能带来的潜在利益”、“所需花费的代价”与“所要承担的风险”这三者的相对大小情况。

**（4）替代品的威胁（Substitutes）**

两个处于同行业或不同行业中的企业，可能会由于所生产的产品是互为替代品，从而在它们之间产生相互竞争行为，这种源自于替代品的竞争会以各种形式影响行业中现有企业的竞争战略。

* 现有企业产品售价以及获利潜力的提高，将由于存在着能被用户方便接受的替代品而受到限制；
* 由于替代品生产者的侵入，使得现有企业必须提高产品质量、或者通过降低成本来降低售价、或者使其产品具有特色，否则其销量与利润增长的目标就有可能受挫；
* 源自替代品生产者的竞争强度会受到产品买主转换成本高低的影响。

总之，替代品价格越低、质量越好、用户转换成本越低，其所能产生的竞争压力就强。而这种来自替代品生产者的竞争压力的强度，可以具体通过考察替代品销售增长率、替代品厂家生产能力与盈利扩张情况来加以描述。

**（5）同行业竞争者的竞争能力(Rivalry)**

大部分行业中，企业间利益都是紧密联系在一起的。作为企业整体战略一部分的各企业竞争战略，其目标都在于使得自己的企业获得相对于竞争对手的优势。所以，在实施中就必然会产生冲突与对抗现象，这些冲突与对抗就构成了现有企业之间的竞争。现有企业之间的竞争常常表现在价格、广告、产品介绍、售后服务等方面，其竞争强度与许多因素有关。

一般来说，出现下述情况将意味着行业中现有企业之间竞争的加剧：

* 行业进入障碍较低，势均力敌竞争对手较多，竞争参与者范围广泛；
* 市场趋于成熟，产品需求增长缓慢；
* 竞争者企图采用降价等手段促销；
* 竞争者提供几乎相同的产品或服务，用户转换成本很低；
* 一个战略行动如果取得成功，其收入相当可观；
* 行业外部实力强大的公司在接收了行业中实力薄弱企业后，发起进攻性行动，结果使得刚被接收的企业成为市场的主要竞争者；
* 退出障碍较高，即退出竞争要比继续参与竞争代价更高。在这里，退出障碍主要受经济、战略、感情以及社会政治关系等方面考虑的影响，具体包括：资产的专用性、退出的固定费用、战略上的相互牵制、情绪上的难以接受、政府和社会的各种限制等。

行业中的每一个企业，或多或少都必须应付以上各种力量所构成的威胁，而且也需要注意行业中的每一个竞争者的举动。除非认为正面交锋有必要且有益处，例如要求得到很大的市场份额，否则可以通过设置进入壁垒，包括差异化和转换成本来保护自己。当确定了其优势和劣势时(参见SWOT分析)，必须进行定位，以便因势利导，而不是被预料到的环境因素变化所损害，如产品生命周期、行业增长速度等，然后保护自己并做好准备，以有效地对其它企业的举动做出反应。

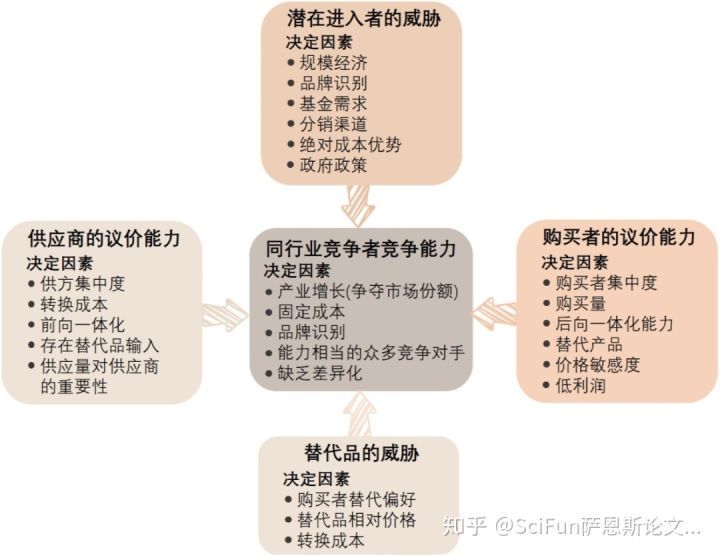
**1.13.3 如何用？**

利用波特五力模型进行分析的基本思路：



具体来说，运用波特五力模型进行产业结构分析，你需要：

* 通过分析产品范围和地理范围来剖析有关产业；
* 逐一分析五种力量所代表的有关角色，并分门别类；
* 估计每种力量背后的驱动因素。哪些强？哪些弱？为什么？



* 进一步分析整个产业的结构。哪些力量控制着营利性？并非所有的力量重要性都是平等的，要更加深入地挖掘公司所在产业最重要的力量。你需要思考：你的分析结果与当前以及今后的产业盈利水平是否一致？
* 分析每种力量最近的变化以及未来可能发生的变化。你需要思考：它们的发展趋势是什么？放眼未来，竞争主体或新进企业如何影响产业结构？
* 如何根据五种力量定位该企业。能否将企业定位在五种力量最薄弱的领域？能利用产业变动为自身服务吗？能按照符合自身利益的方式塑造产业结构吗？

**案例一：分析滴滴出行公司**

**替代品的威胁**：企业的利润率首先是由替代品的威胁决定的。移动出行的替代品指出租车、公交车和私家车等传统出行工具。这些传统的出行工具虽然可以代替网约车出行，但它们的便捷性远不及网约车。从经济实惠的角度出发，网约车更符合消费者偏好，因此替代品的威胁较小说明网约车持续盈利的能力强。替代品的相似性和差异性还依赖于不同客户划分的细分市场，网约车行业中划分为网约 “出租车”、网约 “专车”和 “顺风车，滴滴出行核心业务包含这些领域，因此替代品的威胁对滴滴出行来说较小。

**新进入者的威胁**：滴滴出行公司的生存和发展也一直处于瞬息万变的行业竞争环境中，也不断受到其他企业的进入威胁。如何与新进入者竞争，决定行业的进入壁垒，进入壁垒的来源有:

* 高额的资本成本；
* 规模经济和绝对成本优势可形成另一种进入壁垒；
* 顾客和司机对品牌的忠诚可以形成进入壁垒；
* 产品差异也会构成进入壁垒，若顾客认为产品差异实际上是难以模仿的，这时产品差异就会形成一个有效的壁垒。

**买方的还价能力**：网约车行业的利润率部分决定于买方的议价能力。买主可能是高度集中的，也可能是极度分散的，相比高度集中的买主来说分散的买主没有什么议价能力。滴滴出行的买方就是有打车需求的客户，其对价格十分敏感，对质量的敏感度中等，但随着安全事件的发生，顾客对质量的要求也越来越高。但是由于滴滴出行的覆盖面广，并且替代品较少，对于买方来说对网约车的选择余地较少，而且网约车的消费者是分散的，因此形成了买方还价能力低的局面。滴滴出行来说，其用户规模巨大，甚至出现供不应求的状况，因此滴滴出行在与买方讨价还价方面有优势。

**卖方还价能力**：滴滴出行的卖方就是广大的出租车、专车等交通工具的司机，他们在各个平台注册的成本都较低，因此转换平台的可能性就更大。并且市场价格对于司机是可掌握的，在比较工资后，也极易产生更换平台的念头。对于卖方来说，以往是由于滴滴出行补贴力度大，才使得卖方有动力使用该平台。但目前滴滴出行的竞争对手给卖方提供更多的优惠政策和更多的补贴，卖方成为了网约车行业争抢的对象，因此卖方的还价能力较高。正是因为供应商拥有较好的议价能力，才使网约车企业的利润率下降。

**现存竞争者之间的竞争**：在 2016 年以前，滴滴出行的最大竞争者是快的和优步中国，滴滴出行在与快的、优步中国合并后俨然占据市场最大份额，在同等阶梯上已经没有竞争者，而在行业的第二阶梯竞争十分激烈，不少网约车企业通过自身的特点向外延伸产业，不断伸向滴滴出行涉及较少的领域，较为代表性的神州专车和嘀嗒拼车，前者以高端用车为主打，后者以简单便捷的拼车为主打。除滴滴出行以外，当前其他企业市场份额均在 1%-3%。滴滴出行与现存竞争者之间的竞争是有优势的。但滴滴出行并购这两大强敌后并未取得稳定的垄断地位，其不断扩张业务范围使其碰到新的竞争对手。摩拜单车、嘀嗒专车、首汽约车、美团打车等等，看似单个弱小的敌人威胁有限，可如今抱团而成反滴滴联盟，正从滴滴的快车、出租车、专车等核心业务入手，将战火烧到分时租赁、无人驾驶等滴滴薄弱或者暂未取得领先优势的领域。目前，滴滴出行的主要竞争对手是美团点评和摩拜单车，因为这两个企业在不同领域拥有大规模用户，并且都在滴滴出行的核心领域有所行动。

**总结**：通过分析发现，滴滴出行在行业竞争中的最大的优势是其市场占有率高，最大的劣势是滴滴出行的使用者的忠诚度低。一旦有资金力量雄厚的新进入者参与网约车的竞争，就及其容易以优惠的政策抢夺滴滴出行的客户和司机，从而动摇滴滴出行在网约车市场的地位。

**实例二：分析贵州茅台**

**1.12.4 注意事项**

（1）研究、制定战略的理论假设（不管是隐含的还是明确的）应该与企业所处的实际环境和企业的战略目标保持动态一致性，这是获得良好战略的必要条件。

（2）全面研究企业内外环境条件而不有所偏废是制定竞争战略的基本出发点。既重视对环境的分析也重视对企业自身条件的研究，这应该成为战略理论家和企业界的共识。

（3）没有普遍适用的环境分析模型。由于企业内部及其环境的复杂性和异质性，不存在“标准的” 五力分析模型，每个企业都应该根据自己的实际情况来分析所受的作用力。

（4）不存在通用的竞争战略。不可否认，波特的所谓通用竞争战略具有广泛的适用性，但更一般的是，在不同的情境下应该适用不同的战略，也就是说，要不断进行战略创新。适应和创新才是战略的本质和灵魂，不跳出通用战略的藩篱就不可能获得持续的竞争优势。

（5）企业自身的资源和能力对于企业的持续成功非常关键。这意味着在企业内部寻求竞争优势的来源是更为可取的战略观。

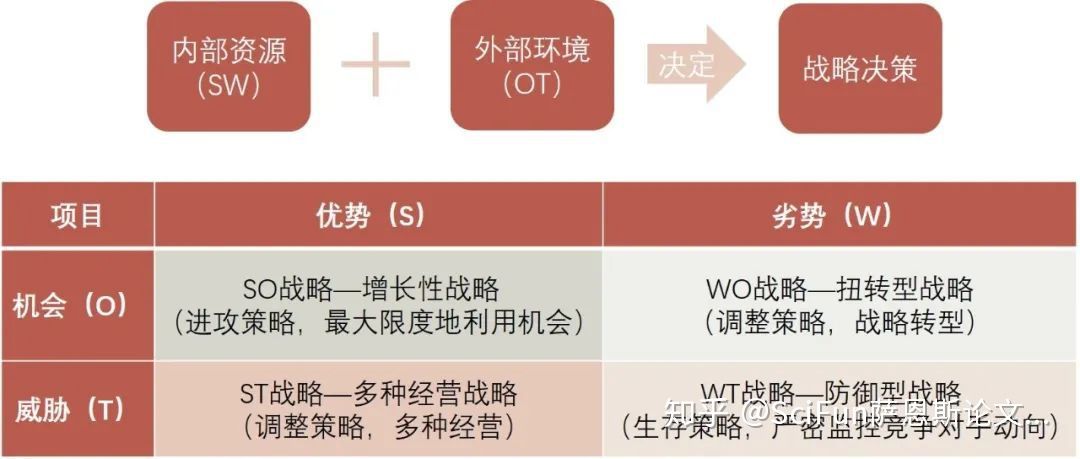
## 1.14 SWOT模型

**1.14.1 是什么？**

SWOT分析法，即态势分析法，在现在的战略规划报告里，SWOT分析应该算是一个众所周知的工具。来自于麦肯锡咨询公司的SWOT分析，其中S代表strength（优势），W代表weakness（劣势），O代表opportunity（机会），T代表threat（威胁），S、W为内部因素，O、T为外部因素。



SWOT分析就是将与研究对象密切相关的各种主要的内部优势、劣势以及外部的机会、威胁等，通过实际调查列举出来，并依照矩阵形式排列。然后利用系统分析的思想把各种因素相互匹配，加以分析，进而从中得到一系列相应的结论，而结论往往带有一定的决策性。



运用这种方法，可以对研究对象所处的情景进行全面、系统、准确的研究，从而根据研究结果制定相应的发展战略、计划以及对策等。SWOT分析法作为一种**企业战略分析方法**，其实也可以运用于区域、产业的发展战略研究，甚至还可以运用于个人职业规划的设计里。运用在什么地方，运用的是否恰当，关键在于对四个要素的含义要正确理解和把握。

**1.14.2 有什么用？**

**（1）机会与威胁分析（environmental opportunities and threats）**

外环境发展趋势分为两大类：一类表示环境机会，另一类表示环境威胁。

* **环境机会**就是对公司行为富有吸引力的领域，在这一领域中，该公司将拥有竞争优势。
* **环境威胁**指的是环境中一种不利的发展趋势所形成的挑战，如果不采取果断的战略行为，这种不利趋势将导致公司的竞争地位受到削弱。

对环境的分析也可以有不同的角度。比如，一种简明扼要的方法就是**PEST分析**，另外一种比较常见的方法就是**波特五力分析**。



**（2）优势与劣势分析（Strengths and weaknesses）**

识别环境中有吸引力的机会是一回事，拥有在机会中成功所必需的竞争能力是另一回事。**所谓竞争优势是指一个企业超过其竞争对手的能力，这种能力有助于实现企业的主要目标——赢利。**但值得注意的是，**竞争优势并不一定完全体现在较高的赢利率上，因为有时企业更希望增加市场份额，或者多奖励管理人员或雇员。**

竞争优势可以指**消费者眼中一个企业或它的产品有别于其竞争对手的任何优越的东西**，它可以是产品线的宽度、产品的大小、质量、可靠性、适用性、风格和形象以及服务的及时、态度的热情等。



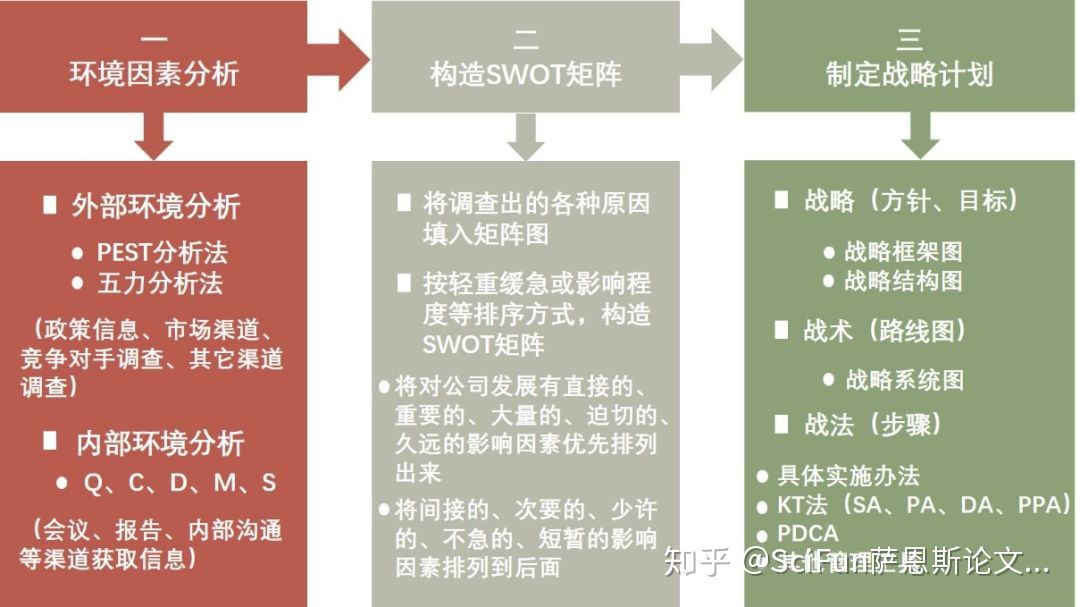
由于企业是一个整体，而且竞争性优势来源十分广泛，所以，**在做优劣势分析时必须从整个价值链的每个环节上，将企业与竞争对手做详细的对比**。**如产品是否新颖，制造工艺是否复杂，销售渠道是否畅通，以及价格是否具有竞争性等**。如果一个企业在某一方面或几个方面的优势正是该行业企业应具备的关键成功要素，那么，该企业的综合竞争优势也许就强一些。**需要指出的是，衡量一个企业及其产品是否具有竞争优势，只能站在现有潜在用户角度上，而不是站在企业的角度上**。

企业在维持竞争优势过程中，必须深刻认识自身的资源和能力，采取适当的措施。因为一个企业一旦在某一方面具有了竞争优势，势必会吸引到竞争对手的注意。一般来说，企业经过一段时期的努力建立起某种竞争优势，然后就处于维持这种竞争优势的态势，竞争对手开始逐渐做出反应。而后，如果竞争对手直接进攻企业的优势所在，或采取其它更为有力的策略，就会使这种优势受到削弱。

**影响企业竞争优势的持续时间，主要的是三个关键因素：**

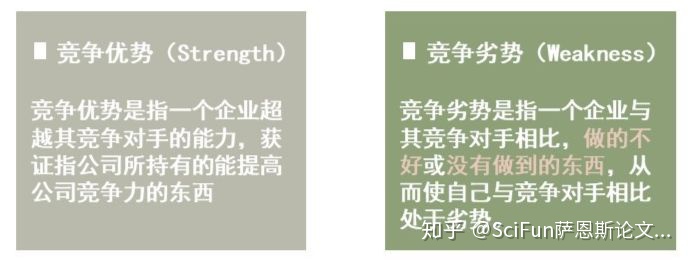
* **建立这种优势要多长时间?**
* **能够获得的优势有多大？**
* **竞争对手做出有力反应需要多长时间？**

**1.14.3 如何用？**



**（一）环境因素分析**

**（1）SW优势与劣势分析（内部环境分析）**



**SW分析主要从Q、C、D、M、S几个领域进行分析：**

* **Q-品质**

产品质量的安全性、稳定性、可靠性、美观性、适用性、耐久性、经济性等。

* **C-成本（价格）**

同样等级产品的生产成本、销售成本、服务成本等和销售价格（产品赢利能力）。

* **D/D-产量、效率、交付能力**

产品总量、生产能力（CT），综合效率、人均产量、人均附加量、交付按量准时。

* **D/L-产品研发**

新产品设计开发能力，开发周期，专利技术，专有技术，技术创新能力等。

* **M-人才/设备/物料/方法/测试**

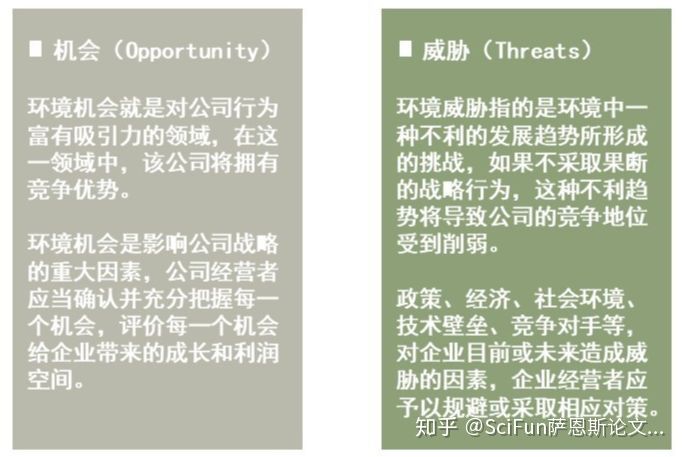
人才：经验丰富的优秀管理、技术人才，优秀的管理、技术团队，年轻/激情。  
设备：先进高效的生产线，现代化高精度的生产设备、检验设备。  
物料：优秀的供应商团队，一流的供应链，高品质、低价格的物料稳定的供应。  
方法：先进的管理方法、体系，畅通的信息（比其他对手更能优先获得信息）。测量：先进的测量仪器，科学的测量方法，完整的品质检验体系。

* **S-销售/服务**

销售：强大的销售网络，优秀的销售团队，丰富的销售经验和技巧，灵活的市场应对能力，优秀的品牌形象，婆娘拍的价值及市场认可度，良好的客户关系，忠诚的消费者。

服务：完善的售后服务体系，优质的服务，满意的客户群。

**（2）OT机会与威胁分析（外部环境分析）**



**OT分析主要运用PEST分析法、波特五力模型分析法：**

* **PEST分析法**

PEST是从政治（Politics）、经济（Economic）、社会（Society）、技术（Technology）四个方面，基于公司战略的眼光来分析企业外部宏观环境的一种方法。公司战略的制定离不开宏观环境，而PEST分析法能从各个方面比较好的把握宏观环境的现状及变化的趋势，有利于企业对生存发展的机会加以利用，对环境可能带来的威胁及早发现避开。



* **波特五力模型分析法**

五力分析模型是迈克尔·波特(Michael Porter)于80年代初提出，对企业战略制定产生全球性的深远影响。用于竞争战略的分析，可以有效的分析客户的竞争环境。五力分别是： **供应商的议价能力、购买者的议价能力、潜在竞争者进入的能力、替代品的替代能力、行业内竞争者现在的竞争能力。**五种力量的不同组合变化最终影响行业利润潜力变化。

**（二）构造SWOT矩阵**

在构造SWOT过程中，要将分析出来的内容按**轻重缓急**及**影响程度**，做出优先顺序，那些对公司发展有**直接的、重要的、大量的、迫切的、久远的**影响因素优先排列出来，而将那些**间接的、次要的、少许的、不急的、短暂的**影响因素排列在后面。



【优先顺序评价说明】



**（三）制定战略计划**

制定战略计划的基本思路是：

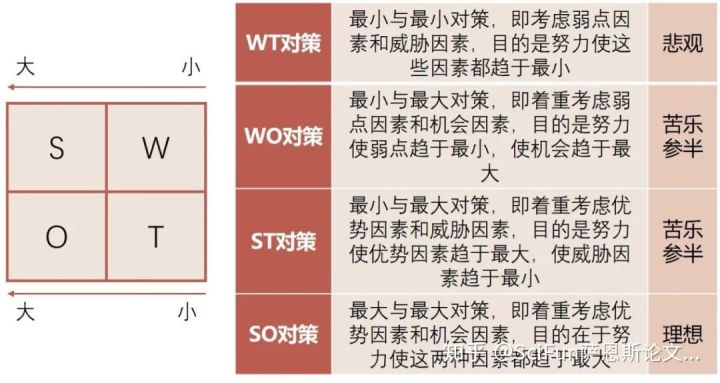
* **发挥优势因素，分析劣势因素，并克服劣势因素。**
* **利用机会因素，识别威胁因素，并规避或化解威胁因素。**
* **考虑过去，立足当前，着眼未来。**

运用系统分析的综合分析方法，将排列与考虑的各种环境因素相互匹配起来加以组合，得出一系列公司未来发展的可选择对策。

【威胁和劣势的原因分析及对策措施】

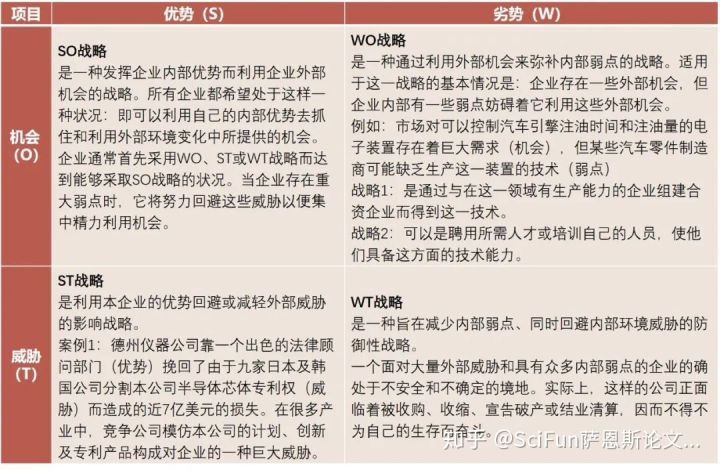
——目的：将威胁降到最低，将劣势减到最小！





SWOT分析只是战略发展的第一步，企业需要进一步找到内部要素与外部环境的结合点，有效调整整合内部各要素，以吻合或超越外部环境的变化，获取竞争优势。

此矩阵就是将内部要素与外部环境结合分析的工具。通过将强弱势与机会威胁进行分割，可得出企业应对环境变化的4个主要战略：



【例：对某炼油厂进行SWOT分析结果】



**1.14.4 注意事项**

SWOT分析范围很广，如果缺乏事实和数据，分析就会变得很笼统，制订的战略就会缺乏依据，成为没有价值战略方案，它带有时代的局限性。学生在运用SWOT分析的过程中存在的主要问题：

一是往往把“发展”及其“成绩”放入优势，比如已经做的“工作”、取得的“进展”等等，把“问题”及其“原因”纳入劣势，比如“观念保守”、“体制落后”、“市场单一”、“发展方式粗放”，等等，甚至把“成绩”与“机会”、“问题”与“威胁”混为一谈，导致**六个因素搅在一起，前后矛盾，难解难分，一笔糊涂账**；

二是在做完SWOT分析后，不与后面的“对策建议”结合，忽略了**做分析的目的是在于寻找到更有高度的、针对性的化解问题的“对策”**，使SWOT分析部分成了与其他部分毫无关联的纯粹凑字数的部分。

进行SWOT分析时需注意的是：

* 进行SWOT分析的时候必须对公司的优势与劣势有客观的认识；
* 进行SWOT分析的时候必须区分公司的现状与前景；
* 进行SWOT分析的时候必须考虑全面；
* 进行SWOT分析的时候必须与竞争对手进行比较，比如优于或是劣于竞争对手；
* 保持SWOT分析法的简洁化，避免复杂化与过度分析；
* SWOT分析法因人而异；

实例：通过管理层讨论，看管理层对公司的发展战略、计划和对策是否合理。

## 1.15 5W2H分析

市场调研方法

（1） WHAT——产品的背景是什么？

（2） WHERE——产品的市场空间有多大？

（3） WHO——目标用户是谁？有多大量级？

（4） WHY——产品定位，我们要解决用户什么痛点？

（5）WHEN——什么时间完成？什么时机最适宜？

（6） HOW ——怎么做？如何实施？方法怎样？

（7） HOW MUCH——做到什么程度？费用产出如何？

## 1.16 3W2H分析

商业模式

（1） WHERE——目标区域市场在哪里？

（2） WHAT——服务定位有哪些？

（3） WHO——目标用户是谁？有多大量级？

（4） how to make money——如何盈利？盈利方式有哪些？

（5） how to achieve——如何去做？怎么获取？