

七周成为数据分析师

数据分析思维

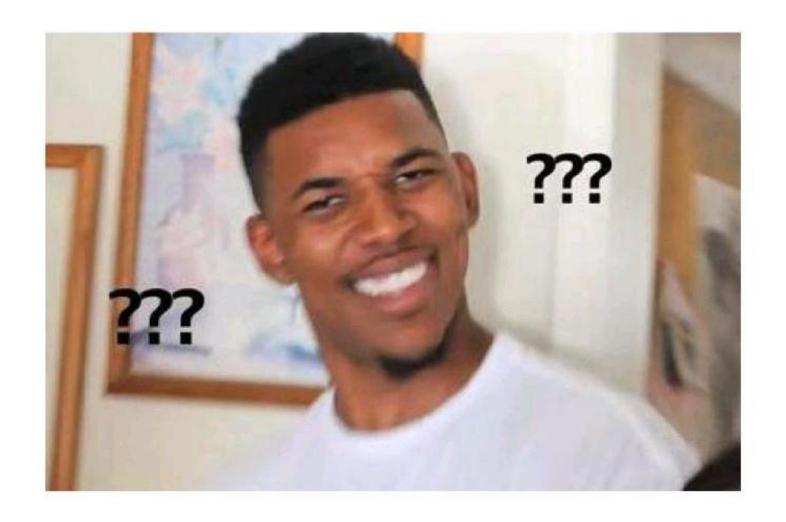
讲师:秦路



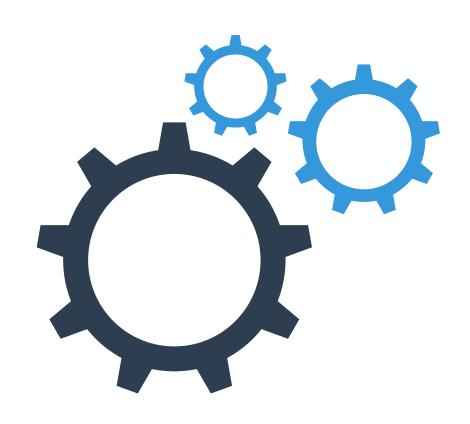
数据分析思维



为什么思维重要







What

三种核心思维

Why

数据分析的思维技巧

How

如何在业务时间锻炼分析能力





数据分析的三种核心思维











又改数据?!

个分析结论我们已经知道,多what?

活跃用户数又下降了!来我办公室一起

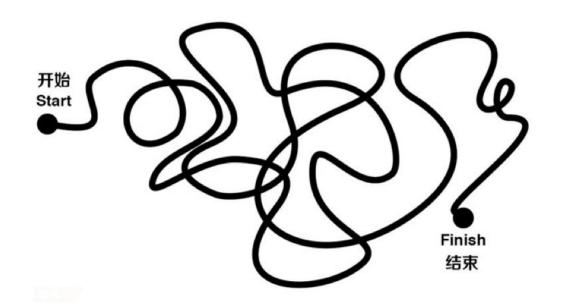
为什么分析的思路

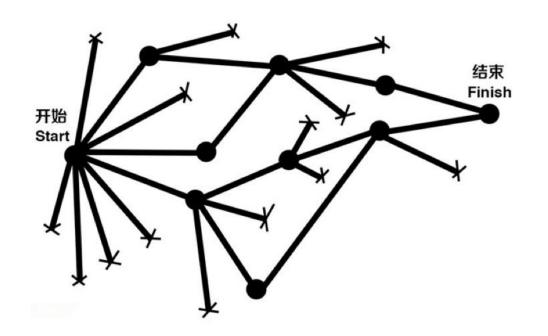
总是一团古し麻?

这个月KPI再不完成,只能去人力办么

为啥别人能拿20k的薪资

运营部的翠花,从来不满意我的分析





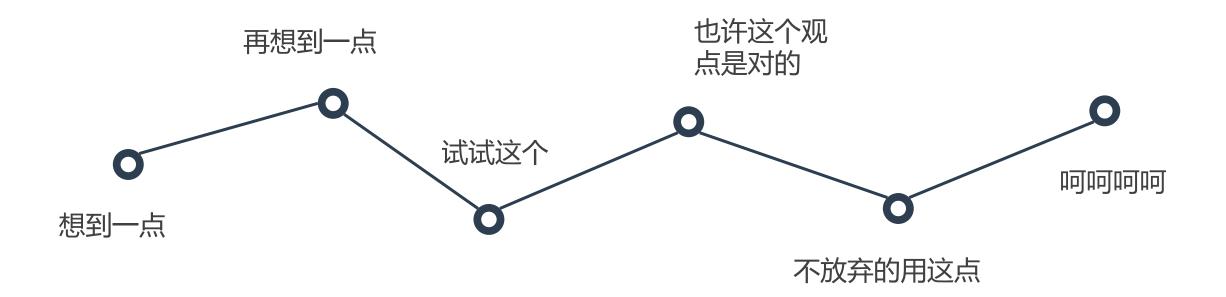




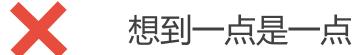
现在有一个线下销售的产品。 我们发现8月的销售额度下降,和去年同比下降了 20%。我想先观察时间趋势下的波动,看是突然暴 跌,还是逐渐下降。再按照不同地区的数据看一下差 异,有没有地区性的因素影响。我也准备问几个销售 员,看一下现在的市场环境怎么样,听说有几家竞争 对手也缩水了,看一下是不是这个原因。顾客访谈也 要做,但是往常一直找不出原因,这次我也不抱希 望,姑且试试吧。要是还找不出原因,那我也很绝望











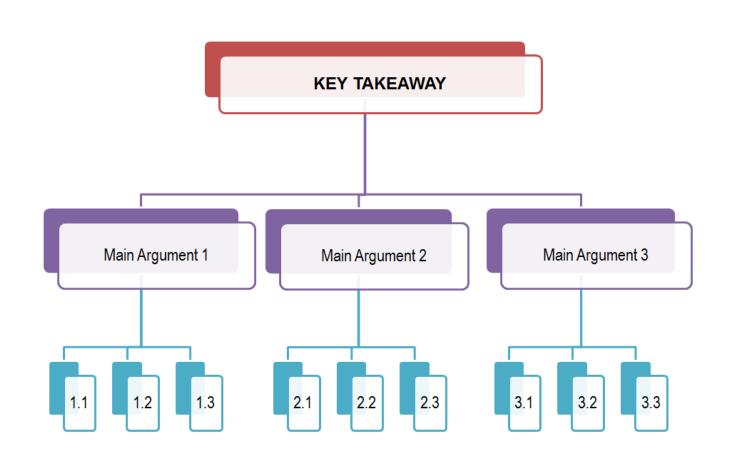
做假设,但乱打一枪

一次性分析,没有复用性

业务看了会流泪,老板看了想打人

还是没结果





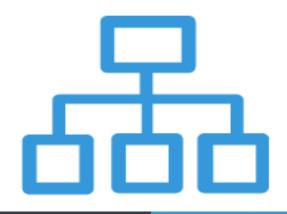
将分析思维结构化

将论点归纳和整理

将论点递进和拆解

将论点完善和补充





核心论点

寻找金字塔的塔顶,它可以是假设,是问题, 是预测,是原因

结构拆解

自上而下,将核心论点 层层拆解成分论点,上 下之间呈因果或依赖关 系

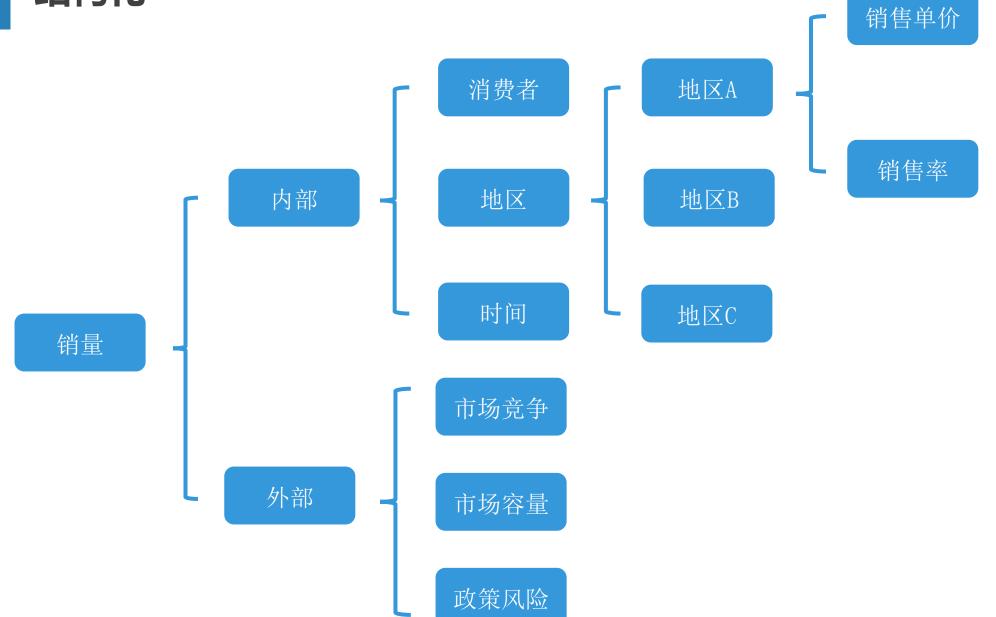
MECE

相互独立,完全穷尽。 论点之间避免交叉和重 复,分论点们要尽量完 善

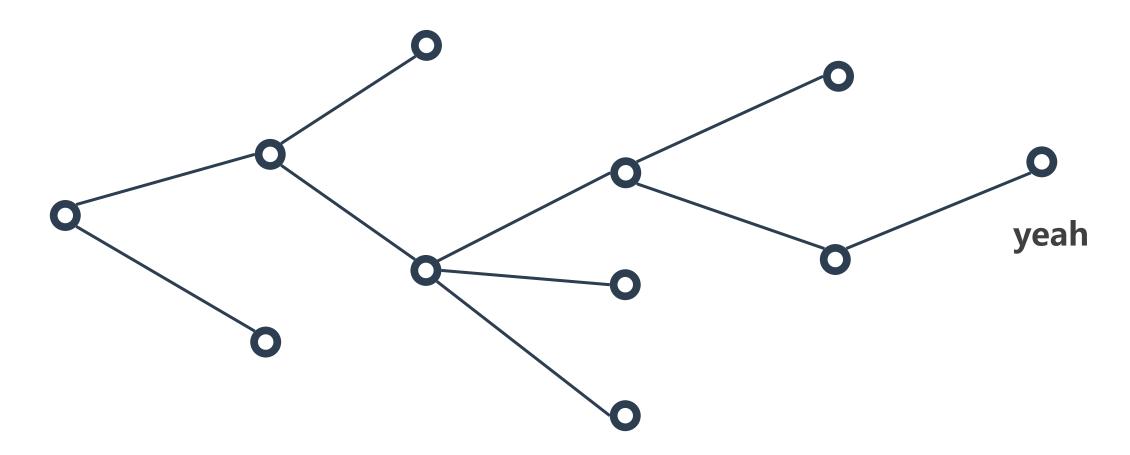
验证

不论核心论点还是分论 点,都应该是可量化的 用数据说话。它们必然 是可验证的

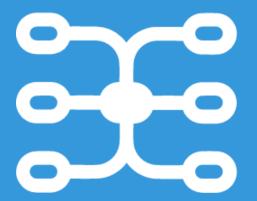












思维导图



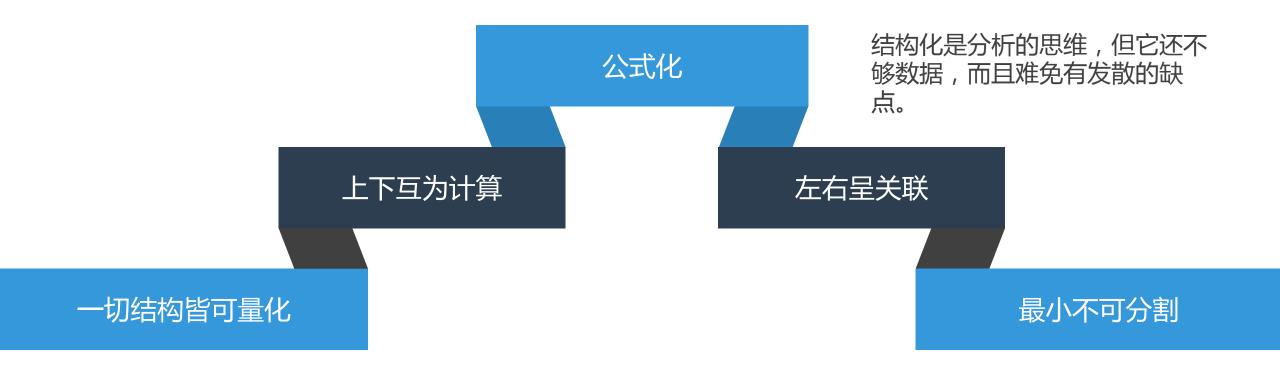


- •查看资料及背景,将结论列成一张表/卡片
- •把表上的结论,依据主题分类
- •将同一类型的结论,按顺序区分
- •讨论同一级别的共通结论,将其结论放在上
- 一段位置

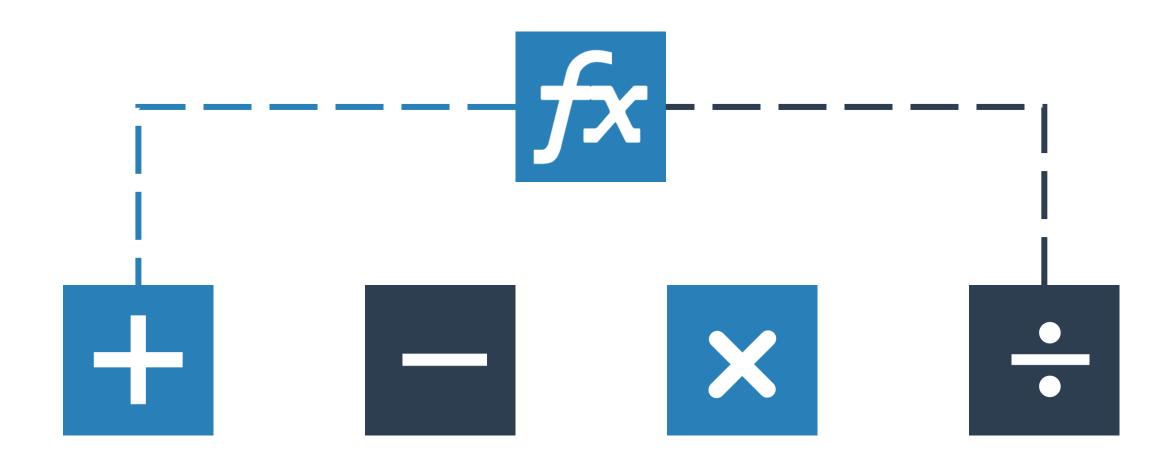


结构化不是完美的











销售额由什么组成的?销量和客单价相乘

利润由什么组成的?销售额收益和成本相减

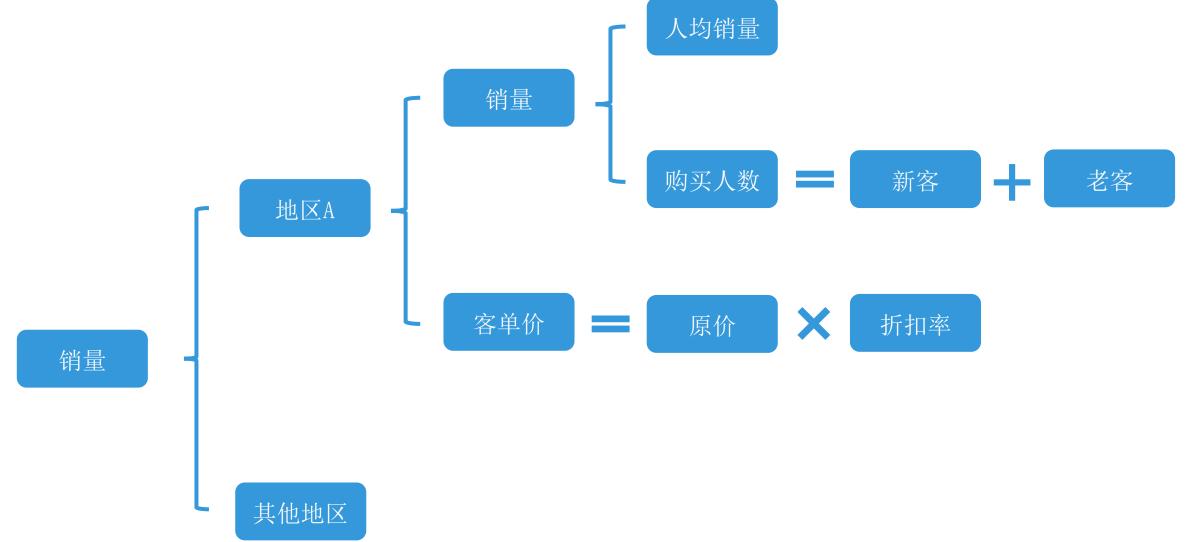
ショュロア

销售额是单一的维度么?不是。销售额是多个商品 / SKU的总和

地区的销量由什么组成?是不同线下渠道的累加

销量还能再细挖么?不妨想成人均销量和购买人数







公式化







不同类别的业务 叠加可以用加 法。



减法常用来计算 业务间的逻辑关 系。



乘法和除法是各 种比例或者比 率。







主动流量 营销、活动、推广



被动流量 邀请、应用商店搜索



















不同渠道

曝光量

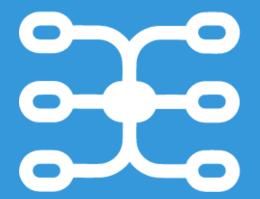
转化率

邀请人数

人均邀请量

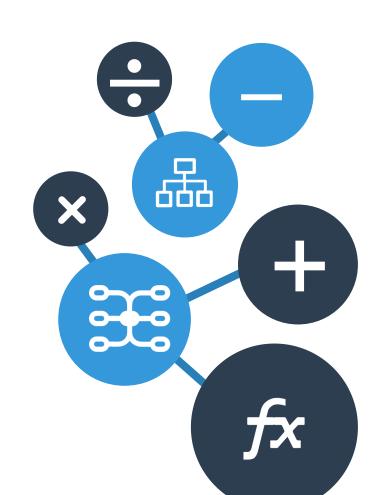
转化率





思维导图



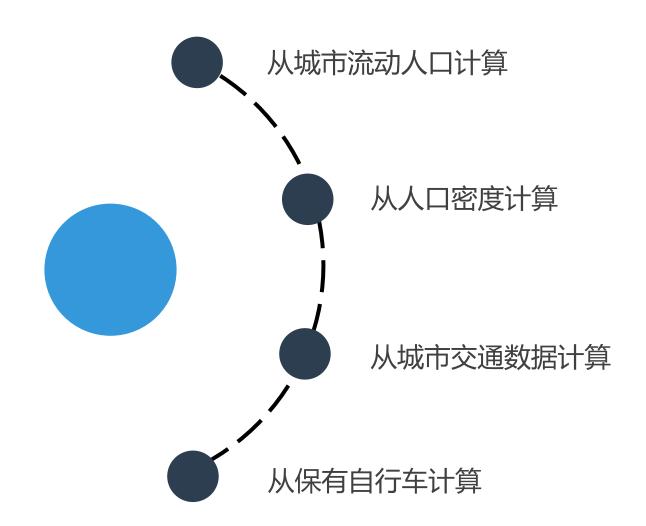


结构化和公式化,似乎已经能解决大部分的问题?



如何预估上海地区的共享单车投放量?







如何预估上海地区的共享单车投放量?

单车是有损耗的,计算公式中应该考虑单车的消耗因素



结构化+公式化

道理懂了很多,但离分析水平大成,总还差了那么一些。不知道原因在哪里?

业务化

为分析而分析,却没有深入理解业务,俗称不接地气。好的数据分析思维,本身也是具备业务思维。



你的分析贴不贴合业务?



有没有从业务方的角度思考



真的分析出原因了吗?



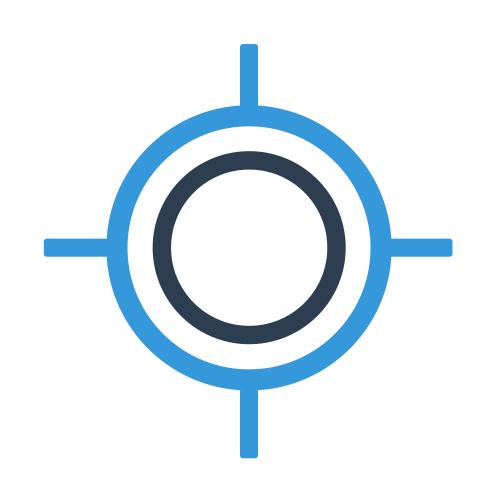
能不能将分析结果落地



一家销售公司业绩没有起色,对它进行了分析

- 销售人员的效率降低,因为士气低落
- 产品质量不佳,和同期竞争对手比没有优势
- 价格平平,顾客并不喜欢





用结构化思考+公式化拆解,获得的最终分析论点。很多时候,是现象。数据是某个结果的体现,但不代表原因。



我是数据分析师, 我会设立哪些指标



换位,如果我是参与其中的人, 我会怎么考虑/哪些行为



结构化思维 —— 结构化数据 —— 结构化业务数据

捋顺思路

将其可数据化 落地,贴合业务



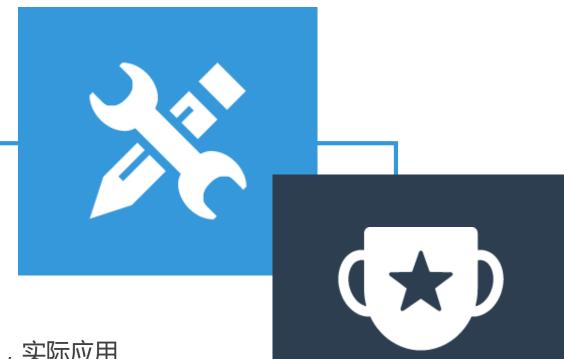
某一天,天善学院的课程学习人数下跌了,梁老板急得头发都要掉光了,现在,你能不能通过三种思维方式,做出一个假设型的分析案例?





数据分析的思维技巧





数据分析技巧

三种核心思想是框架型的指引,实际应用中也应该借助思维的技巧工具,达到四两拨千金的效果。并且,它们应该足够简单和有效



数据分析的思维技巧

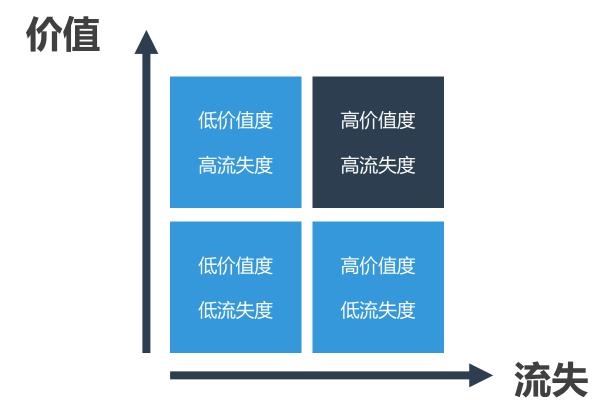




象限法

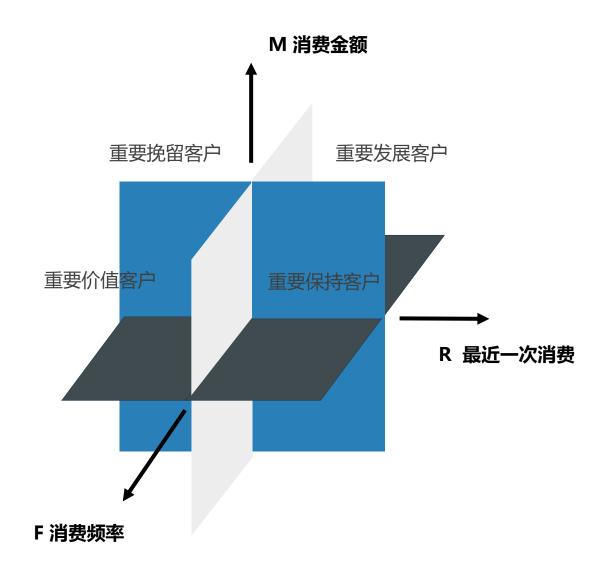








象限法







象限法是一种策略驱动的思维



适用范围广、战略分析、产品 分析、市场分析、客户管理、 用户管理、商品管理等



须知

象限划分可以按中位数,也可 以按平均数,或者是经验

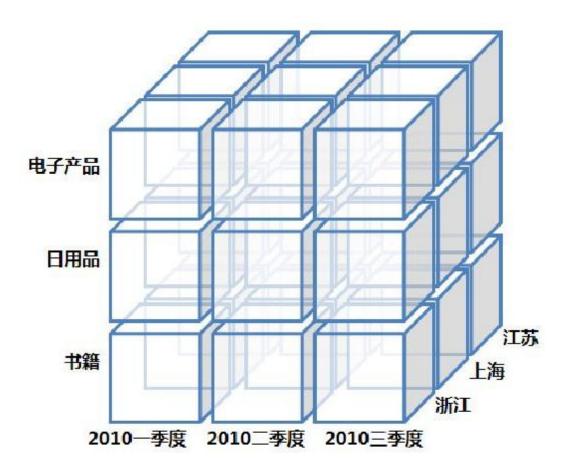


直观,清晰,对数据进行人工 的划分。划分结果可以直接应 用于策略











用户统计维度:性别、年龄...

用户行为维度:注册用户、用户偏好、用户兴趣、用户流失...

消费维度:消费金额、消费频率、消费水平...

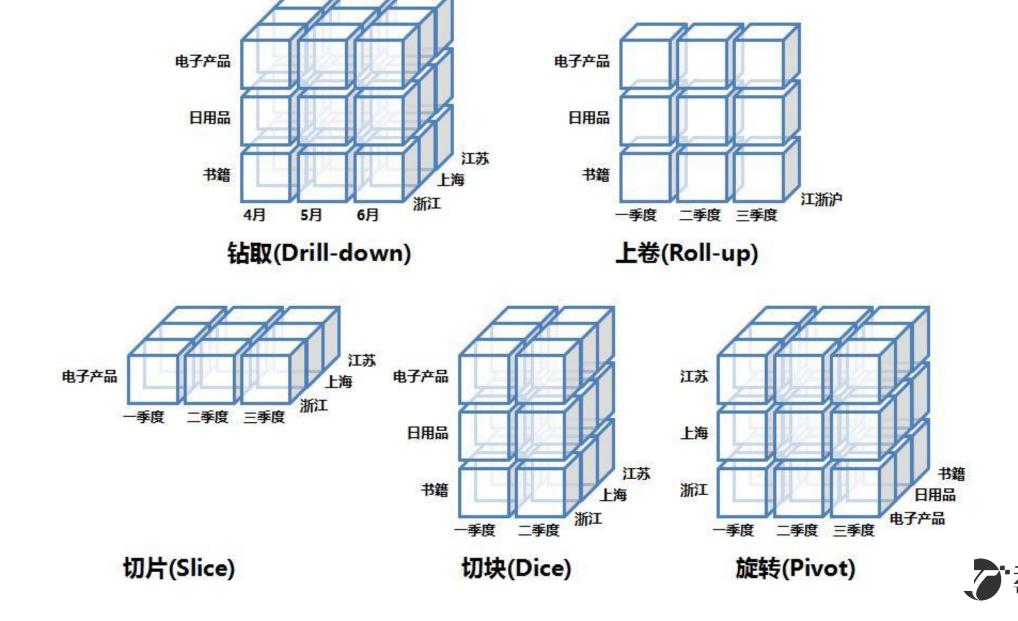
商品维度:商品品类、商品品牌、商品属性...



学院		女生 申请	女生 录取	女生 录取率	男生 申请	男生 录取	男生 录取率	合计 申请	合计 录取	合计 录取率
总计	•	120	50	42%	120	25	21%	240	75	31.3%

学院	女生申请	女生 录取	女生 录取率	男生申请	男生录取	男生录取率	合计申请	合计 录取	合计 录取率
商学院	100	49	49%	20	15	75%	120	64	53.3%
法学院	20	1	5%	100	10	10%	120	11	9.2%
总计	120	50	42%	120	25	21%	240	75	31.3%





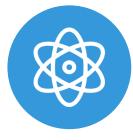


应用

只要数据齐全且丰富,均可以 应用

核心

象限法是一种精细驱动的思维



优点

处理大数据量,维度丰富且复杂的数据有较好的效果。但是 维度过多,会消耗不少时间



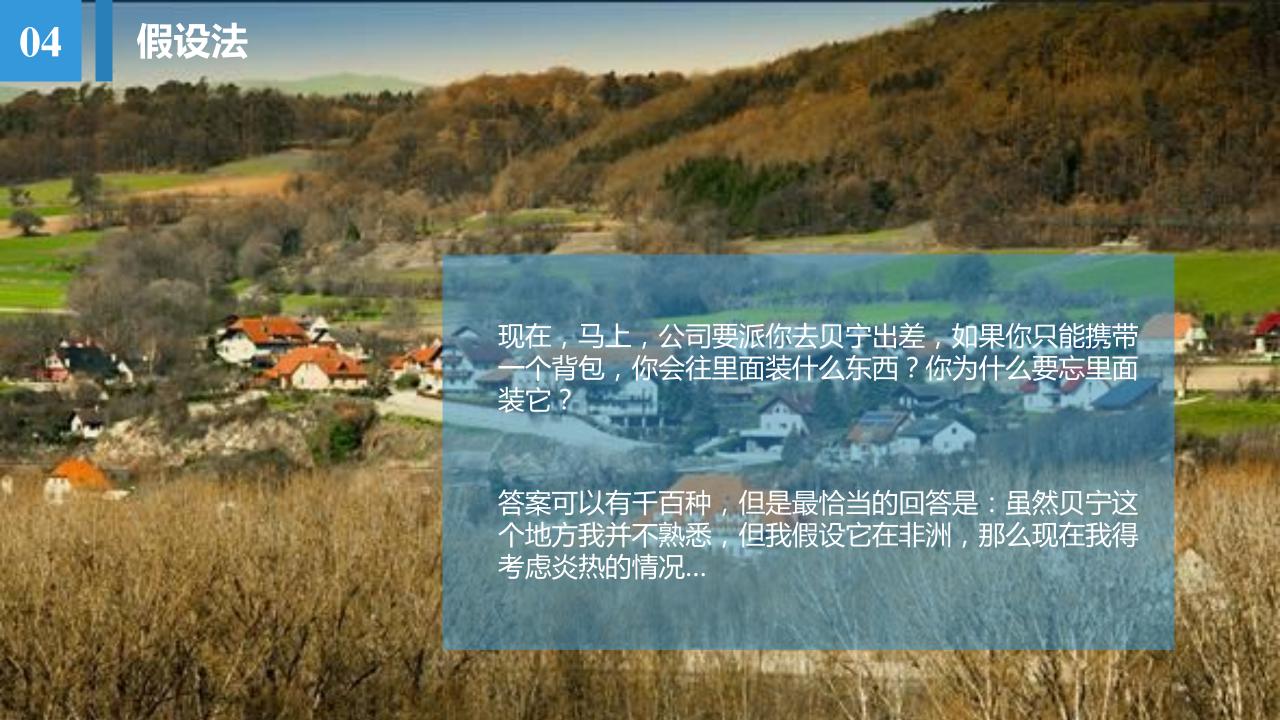


须知

对不同维度进行交叉分析时, 需要注意辛普森悖论









很多时候,数据分析是没有数据可明确参考的: 比如新进入一个市场,公司开拓某样产品。老板 让你预测一年后的销量,或者产品的数据基础非 常糟糕,你拿不到数据。



公司在节日进行了一次营销活动,APP上的销量数据整体比上周上升了20%。因为统计失误问题,拿不到明细数据,也就是说,活动效果是一个黑盒。现在的问题是,销量本身就有可能因为节日而提高,那么怎么证明活动是有效或者无效的呢?



假设活动是有效的

- 思考一下,活动有效的话,会发生什么事情?
- 会有一定数量的用户购买,如果能证明这条,那么我们有理由相信活动是有效的。
- 用户通过活动购买商品,会发生什么可观测的行为呢?假设有一些用户会评论留言,那么可以统计提及活动的字眼。
- 当用户提及了这次营销活动,接下来的问题是,有效了多少?10%,20%?
- 假设参与活动的用户行为没有变化,那么通过历史数据的用户评论占比,反推购买人数。



你是自营电商的数据分析师,现在想商品提价后,收入会不会有变化?,你会怎么做?

- 假设商品提价后,销量一定会下跌,问题是销量下跌多少?
- 首先假设流量不会有变化,流量和渠道营销正相关,商品价格影响转化率,那么现在确定转化率的波动。
- 找出平时的转化率(譬如为20%), 预估提价后的转化率变化。假设各类型用户对价格敏感度不同, 那么将用户划分忠诚XX、普通XX、羊毛XX...
- 不同用户层次数量不同,反应不同。忠诚用户转化率变化极低,羊毛几乎不会转化...这些数据可以凭借经验做出假设。最后汇总。



核心

假设是一种启发思考驱动的思 维



应用

它更多是一种思考方式,假设—验证—判断。



当没有直观数据或者线索能分析时,以假设先行的方式进行推断,这是一个论证的过程。





须知

不止可以假设前提,也能假设 概率或者比例,一切都能假设, 只要自圆其说。







中国今年的经济指标如何?

美国NBA最佳球星是谁?

竞争对手产品表现的如何?

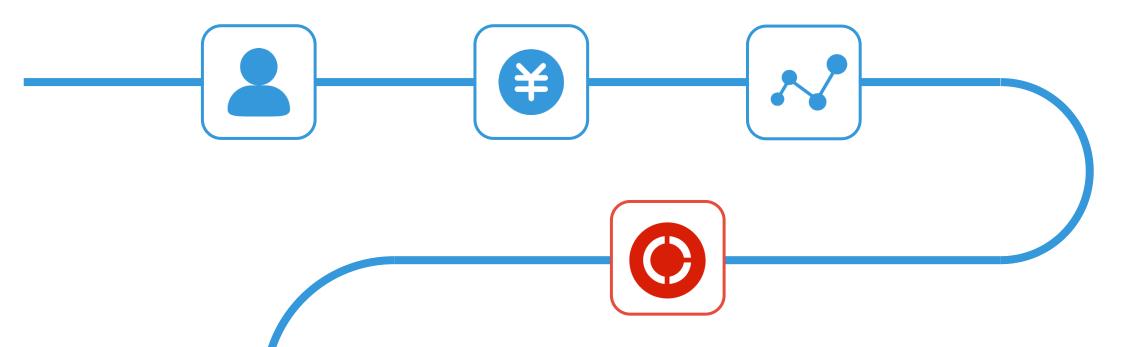
哪位是天善学院最帅的男人?



NBA比赛数据贡献值:

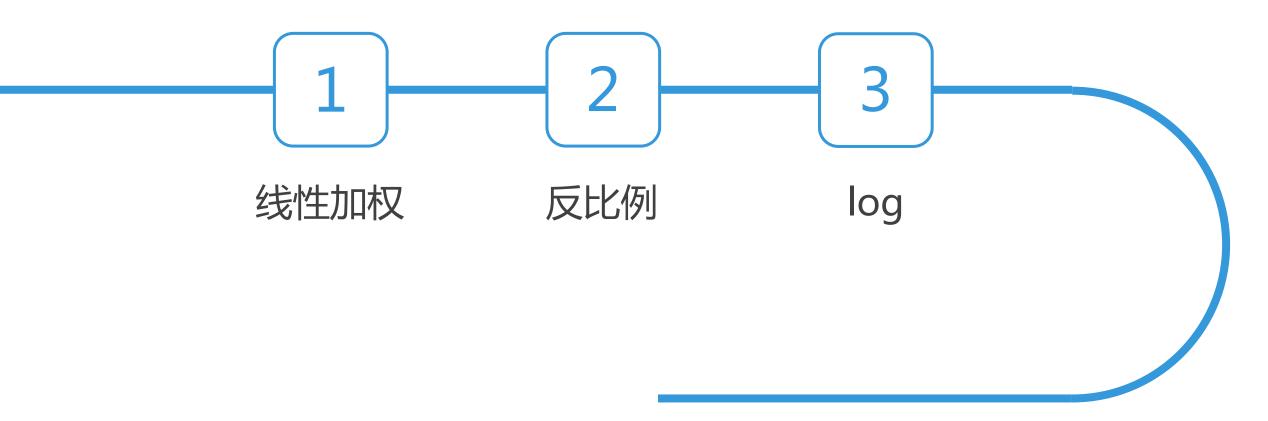
(得分+篮板+助攻+抢断+封盖)-(出手次数-命中次数)-(罚球次数-罚球命中次数)-(失误次数/球员上场比赛的场次)





很多时候,我们有数据,但不知道怎么应用。就是因为缺乏了一个有效性的方向。这个方向可以成为目标指数。通过将数据加工成指数,达到聚焦的目的。









指数法是一种目标驱动的思维



应用

和假设法不同,假设法是缺乏 有效的数据,指数法是无法利 用数据而将其加工成可利用 的。



须知

指数法没统一的标准,很多指 数更依赖经验的加工。



目标驱动力强,直观、简洁、 有效。对业务有一定的指导作 用。一旦设立指数,不易频繁 变动



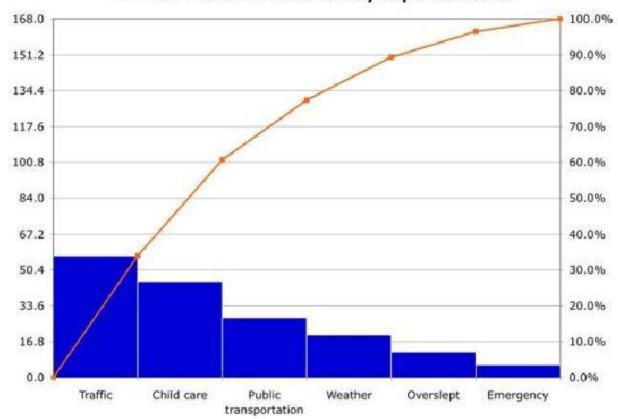






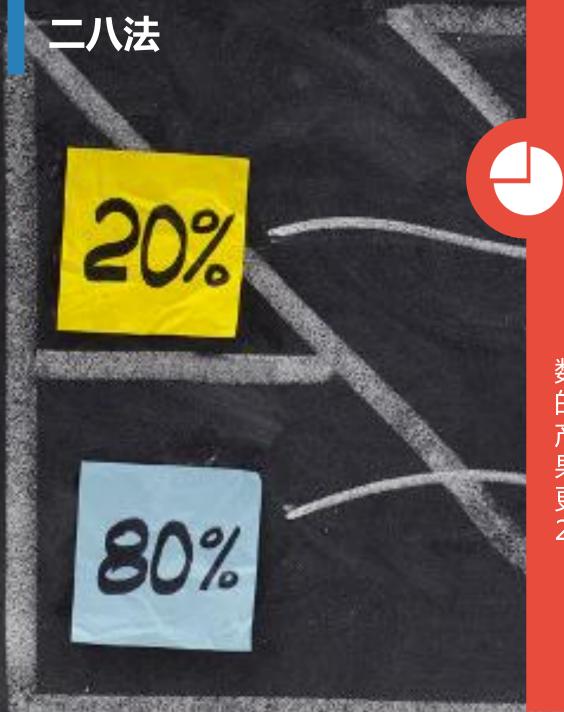






帕累托图





TopN

#

数据中,20%的变量将直接产生80%的效果,数据分析更应该围绕这20%作文章

持续关注TopN的数据,是一个非常好的习惯,尤其在部分行业

虽然指标很多,但往往某些指标更有价值,二八法则不仅能分析数据,也能管理数据

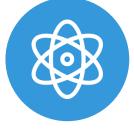
核心

二八法是一种只抓重点的思维



应用

二八法则存在于几乎所有的领域,所以这种分析思维没有局限。



优点

和业务紧密相关,和KPI更紧密相关。几乎花费最少的精力就能达到不错的效果,性价比很优





须知

在条件允许的情况下,数据分析依旧不能放弃全局,否则会让思路变得狭隘





好的数据指标,一定是比例或者比率

好的数据分析,一定会用到对比

——不是我说的



对比法

老王卖水果,今天卖了1000元水果,这个数据有分析价值么?

老王卖水果,今天卖了1000元,昨天卖了800源,这个数据有分析价值么?

老王卖水果,今天卖了1000元。隔壁的老马卖了2000元,这个数据分析有价值么?

老王卖水果,今天卖了1000元,昨天卖了800元,隔壁的老马昨天卖了3000元,今 天卖了2000元,这个数据分析有价值么?

老王花了3000元的营销成本卖了1000元水果,这个数据分析价值么?

老王其实是王健林,他有几百亿资产,他卖水果卖了1000元,这个数据分析有价值么?



节日大促,女生消费占比从60%变为70%,女生节日爱消费

- 这个结论是有问题的
- 占比提高了,不代表绝对值提高了。某商品平时销售额100万女生占60万,节日销售额80万女生占56万,女生真的消费变高了?
- 谁说节日销售额会提高?别忘了竞争对手
- 孤数不证



竞争对手对比

类别对比

特征和属性对比

时间同比环比

转化对比

前后变化对比



核心

对比法是一种挖掘数据规律的 思考方式



应月

对比更多是一种习惯,是数据分析的牛角尖,一次合格的分析,一定要用到n次对比



对比法可以发现很多数据间的 规律,它可以与任何思维技巧 结合,比如多维对比、象限对 比、假设对比等





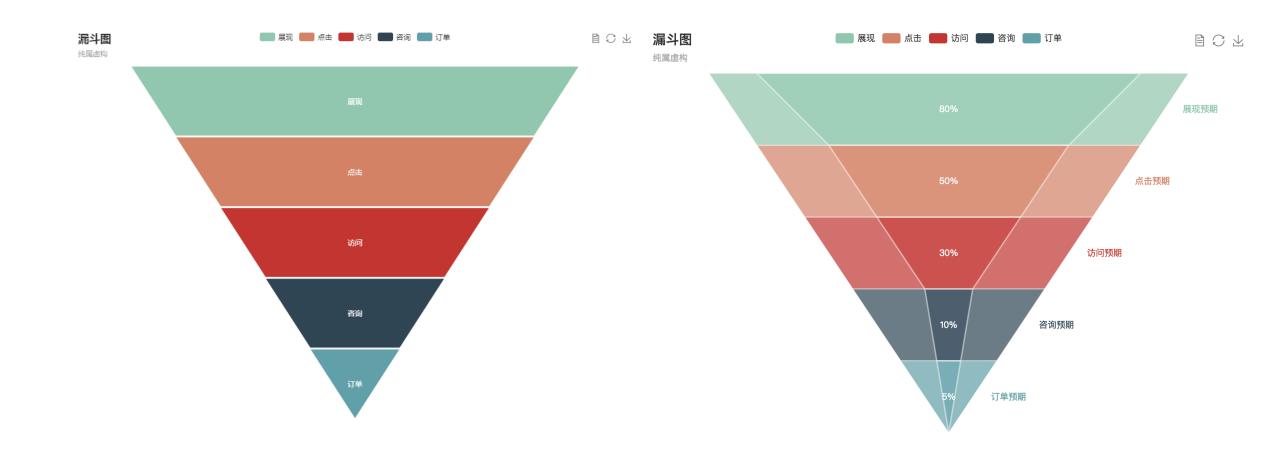
须知

在条件允许的情况下,数据分析依旧不能放弃全局,否则会让思路变得狭隘





漏斗法



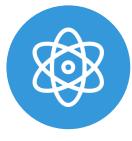


並用

涉及到变化和流程的都能用

核心

漏斗法是一种流程化思考方式







须知

单一的转化率没有,单一的转化率没有,单一的转化率没有,单一的转化率没有。

优点

单一的漏斗分析没有用,转化率20%,但是能说明什么呢? 它要和其他分析思维结合,比如多维,比如对比







如何在业务时间锻炼数据分析思维





好奇省己为什么?







买了尿布的人会买啤酒,但是买啤酒的人会买尿布么?

尿布旁边应该摆放其他东西么?啤酒是否是最好选择?

怎么摆放啤酒?部分还是全部品类?

场景型的摆放是否比品类摆放更好?

数据呢?







www.hellobi.com



—无二的





讲师. 唐宇迪 深度学习领域多年一线实践研究专家

Python R语言 机器学习

数据分析

BI、商业智能

数据挖掘 大数据

Tableau

QlikView

Hadoop Hive Cognos

BIEE ETL

数据库 MySql

SQL Server













