

S U P E R

后浪出版公司

统计学魔鬼

Ian Ayres

[美] 伊恩·艾瑞斯 / 著

刘清山 / 译

Why Thinking-by-Numbers
Is the New Way to
Be Smart

你以为不能算的
其实可以算

九州出版社
JIUZHOU PRESS

魔鬼统计学

[美]伊恩·艾瑞斯
刘清山 译

版权信息

- 丨 书名：魔鬼统计学
- 丨 作者：[美]伊恩·艾瑞斯
- 丨 译者：刘清山
- 丨 版权：后浪出版咨询（北京）有限责任公司

目录

前 言 超级数据分析师的崛起

第一章 谁在替你思考？

回归方程帮你找到灵魂伴侣

赌徒的痛点可以量化

当公司比客户更理解客户

消费者的反击

让人信服的回归方程

世界是个数据矿

“坏蛋”是可以分析出来的

寻找神奇数字

第二章 反直觉：随机化与最优解

第一资本掷骰子

你看到的可能是随机测试网页

随机测试是对直觉的检验

随机测试——应用广泛的利器

随机测试为什么没有大范围普及？

第三章 概率治国

为省钱而花钱

州级思想实验室

随机应用的随机测试

概率的世界

第四章 医生该如何看待循证医学？

拯救十万人的生命

错误模型导致错误思维

随查随用——统计结果的快速传输

未来已来

第五章 传统专家 vs 算法方程

“令人不安的小书”

为什么人类不善于预测？

数据分析与经验知识是否可以兼容？

我们还能做什么？

第六章 世界正在迅速数字化

从90到300万

数据的商品化

数据混搭

技术还是方法？

计算机可以像人类一样思考吗？

哪一部电影会赚钱？

超级数据分析革命来临

惊心动魄的未来图景

第七章 我们找到乐趣了吗？

老师是什么？花瓶吗？

帝国的反击

地位之争

你会购买超级数据分析师的二手车吗？

好莱坞的破坏者

提防超级数据分析者的赠礼

隐性歧视

我知道你明年夏天会做什么

约翰·洛特是谁？

假如数据分析师犯错

第八章 未来需要的人才

未来属于统计学家

一个数字中蕴含的海量信息

概率领跑者

反向推导

波拉克（Polak）的怀孕问题

结 语

后 记 关于革命的进一步解释

致 谢

出版后记

[返回总目录](#)

前言 超级数据分析师的崛起

奥利·阿森费尔特（Orley Ashenfelter）非常喜爱葡萄酒。“你只要把优质红酒放上一段时间，它就会发生一些非常神奇的事情。”阿森费尔特不仅对于葡萄酒的品鉴很着迷，而且希望知道优质葡萄酒和普通葡萄酒背后的推动力量。

“当你购买优质红酒时，”他说，“你是在进行一项投资，因为它未来很可能会变得愈发诱人。你希望知道的不是它现在的价值，而是它未来的价值，尽管你可能不会卖掉它或者喝掉它。你可以通过推迟品尝获得多少快乐呢？这是一个非常有趣的话题。”过去25年，这个话题占据了他的很大一部分时间。

阿森费尔特的日常工作是分析数字。他用统计量提取隐藏在大型数据集中的信息。作为普林斯顿的经济学家，他曾考察同卵双胞胎的工资，以估计多上一年学所带来的影响。他曾考察限速差异，以估计各州对于统计寿命的重视情况。他曾在多年时间里担任美国顶级经济学期刊《美国经济评论》的编辑。

阿森费尔特个子很高，留着浓密的白色络腮胡子，他那洪亮而友好的嗓音往往会成为会场上的主旋律。他并不胆小。你可能认为数字分析师是一些懦弱而腼腆的人，但阿森费尔特会使你迅速消除这种成见。我曾见过阿森费尔特在教室里踱步，用和蔼而热情的语气阐述一篇研讨会论文背后的原理。当他以高度赞扬的语气开始他的评论时，你就要当心了。

阿森费尔特通过分析数字评价波尔多葡萄酒品质的做法使他遇上了很大的麻烦。他不是使用罗伯特·帕克（Robert Parker）等葡萄酒大师“痛饮和回味”的方法，而是用统计量寻找与拍卖价格高低相关的年份酒特征。

“这是显而易见的，”他说，“葡萄酒是农业产品，它会受到不同年份天气的极大影响。”根据法国波尔多地区数十年的天气数据，阿森费尔特发现，较少的收获期降水量和较高的平均夏季气温可以给人们带来品质最佳的葡萄酒。正如彼得·帕塞尔（Peter Passell）在《纽约时报》中所说，阿森费尔特的统计方程与数据吻合得相当完美。

当葡萄成熟、汁液浓缩时，波尔多酒是最好的。在夏季特别热的年份里，葡萄会充分成熟，其酸度会降低。在降水量低于平均水平的年份里，水果汁液会浓缩。所以，你往往会在炎热干燥的年份得到具有传奇色彩的年份酒。成熟的葡萄可以酿造出口感柔和的（酸度低的）葡萄酒。汁液浓缩的葡萄可以酿造出风味浓郁的葡萄酒。

阿森费尔特大胆地将他的理论简化成一个公式：

$$\text{葡萄酒品质} = 12.145 + 0.00117 \times \text{冬季降水量} + 0.0614 \times \text{生长季平均气温} - 0.00386 \times \text{收获期降水量}$$

没错。通过将当年的天气数据代入这个公式，阿森费尔特可以预测任何年份酒的总体品质。通过另一个更加复杂的公式，他可以更加准确地预测100多个庄园的葡萄酒品质。“这看上去可能有一点数学成分，”阿森费尔特说，“但这正是法国人在著名的1855年分类中为葡萄酒厂排名时使用的方法。”

传统葡萄酒评论家并不接受阿森费尔特基于数据的预测。英国《葡萄酒》杂志说：“这个公式显然很可笑，不值得尊重。”纽约葡萄酒商威廉·索克林（William Sokolin）说，在波尔多葡萄酒从业者眼里，阿森费尔特的工作“介于野蛮和歇斯底里之间”。阿森费尔特时常遭到葡萄酒交易者的轻视，当他在佳士得葡萄酒部门发表关于葡萄酒的演讲时，会场后排的交易商公然对他的演讲发出嘘声。

罗伯特·帕克也许是世界上最有影响力的葡萄酒评论家，他也是《葡萄酒先锋》的出版人。他生动地将阿森费尔特形容为“彻头彻尾的骗子”。虽然阿森费尔特是世界上最受尊重的计量经济学家之一，但是在帕克看来，他的方法“是尼安德特人看待葡萄酒的方式。它太荒谬了，简直不值得嘲笑”。帕克否认了数学公式有助于鉴别优质葡萄酒的可能性，“我不想去他家做客，喝他家的葡萄酒”。

帕克说，阿森费尔特“就像一个从不看电影、仅仅根据演员和导演判断电影好坏的影评人一样”。

帕克的观点有一定的道理。既然通过观影得到的判断更加准确，通过品尝葡萄酒得到的判断不也应该更加准确吗？但是，这里有一个问题：波尔多和勃艮第葡萄酒需要在橡木桶里存放18到24个月，然后封装在酒瓶里。像帕克这样的专家在葡萄酒装进酒桶4个月以后才能开封品尝。即便此时，他们喝到的仍然只是正在发酵的非常难闻的混合物。我不知道品尝这种无法饮用的早期葡萄酒能否为品酒师带来关于葡萄酒未来品质的非常准确的信息。例如，巴特菲尔德拍卖行葡萄酒部门前主管布鲁斯·凯泽（Bruce Kaiser）说过：“早期葡萄酒变化很快，在其存放至少10年甚至更长时间以前，没有一个人可以对葡萄酒做出准确评价。”

与之形成鲜明对比的是，阿森费尔特通过分析历史数据找到了天气和价格之间的关系。他通过这种方式发现，冬季降水量每增加1厘米，预期价格往往会增加0.00117美元。当然，这只是一种趋势。不过，通过分析数字，阿森费尔特可以在葡萄收获时立即预测出未来的年份酒品质——比首次品尝提前几个月，比首次销售提前几年。由于葡萄酒期货交易频繁，因此阿森费尔特的预测为葡萄酒收藏者带来了巨大的竞争优势。

20世纪80年代后期，阿森费尔特开始在半年度简报《流动资产》上发布他的预测。他首先在《葡萄酒观察者》上用小篇幅广告宣传他的简报，逐渐积累了大约600名订阅者。订阅者来自世界各地，包括许多百万富翁和品酒专家——其中大多数人来自接受计量经济学方法的葡萄酒收藏者小圈子。罗伯特·帕克的简报《葡萄酒先锋》有3万订阅者，年费为30美元。相比之下，阿森费尔特的订阅群体显得不值一提。

1990年初，《纽约时报》头版发布了一篇关于阿森费尔特新型预测机器的文章，使更多的人接触到了他的思想。他公开批评帕克对于1986年波尔多葡萄酒的评估。帕克认为1986年葡萄酒“非常好，简直称得上出类拔萃”。阿森费尔特不同意这种观点。他认为这种酒的生长季平均气温低于平均水平，收获期降水量高于平均水平，因此该年份酒一定很平庸。

不过，这篇文章中真正的重磅炸弹与阿森费尔特对1989年波尔多葡萄酒的预测有关。这些葡萄酒在酒桶里只存放了3个月，还没有机会得到评论家品尝，但阿森费尔特认为它们将成为“这个世纪的葡萄酒之选”

欢迎访问：电子书学习和下载网站 (<https://www.shgis.com>)

文档名称：《魔鬼统计学》[美]伊恩·艾瑞斯.pdf

请登录 <https://shgis.com/post/2246.html> 下载完整文档。

手机端请扫码查看：

