

Broadview®
www.broadview.com.cn

行业畅销书升级版

← 告诉你职场图表背后的故事 →

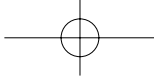
谁说菜鸟不会 数据分析

张文霖 刘夏璐 狄松 编著

入门篇



电子工业出版社
PUBLISHING HOUSE OF ELECTRONIC INDUSTRY
http://www.phei.com.cn



前言

经常有朋友询问：数据分析该怎么做？有什么分析技巧？这些数据怎么处理分析？

因为大量问题具有通用性，而且“懒”得挨个回答类似的问题，于是就结合大家关心的问题，编写这本通俗易懂的数据分析书。市面上数据分析的书大部分还停留在大雅的范畴，要么就是高深的统计学理论，要么就是专业的统计分析软件，给人感觉门槛非常高。而且，所讲解的案例大部分来自科研一线，让人看了摸不着北。这无形间在学习者与数据分析之间建起了鸿沟。

其实，通过多年的数据分析实践来看，数据分析还是件很有乐趣的事情。我们需要做的是：基于通用的工具Excel，加上必知必会的数据分析概念，采用通俗易懂的讲解方式。这样数据分析就不那么晦涩了，而且故事化的情境设计，让我们有一口气读下去的勇气，天堑也变通途了。

虽然积累了多年的数据分析实战经验，但是要上升到一本书还是花费了近1年的时间。她的第1章、第8章由狄松完成，第2章、第5章、第6章由张文霖完成，第3章、第4章、第7章由刘夏璐完成。这个创作过程是艰辛的，但也是很有成就感的。我们努力讲好数据分析的故事，同时把这个故事尽量展现得美丽动人。请允许我们以“她”来称呼这本与众不同的数据分析书籍，很多人翻开这本书的时候，可能会有大量疑惑，但，请耐着性子慢慢读下去，您将会有莫大的收获。

如果你觉得她看起来很轻松，千万别误以为她是一本小说，她其实是一本讲述数据分析的书籍

她抛开复杂的数学或者统计学原理，只和你讲必知必会的要点，关注解决实际问题；

她不去探究科班的学术问题，只和你耐心地分享职场中的实战案例；

她不板起脸和你讲大道理，只和你娓娓道来切身的趣味故事；

她天生丽质，图表漂亮绝伦；

她多姿多彩，还有卡通漫画风。

可能你会觉得她肤浅……

但是，当你揭开她华丽的外衣时，你会惊艳；

也会被她通俗而不庸俗，美丽而又深刻的本质所吸引。

把她珍藏起来吧，因为：

她会循循善诱地把你领进数据分析的大门；

她会让你简历更加具有吸引力；

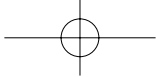
她会让你老板对你刮目相看；

她值得在你的书架上长期逗留，会让你的书架也增加色彩。

她讲述了职场三人行故事，她的故事还会让你偷着笑

牛董，关键词：私企董事、要求严格、为人苛刻。

小白，关键词：应届毕业生，刚入职场的伪白骨精（白领+骨干+精英）、牛董助手、单身



>> 谁说菜鸟不会数据分析（入门篇）

女白领、爱臆想。

Mr.林，关键词：小白的同事、数据分析达人、成熟男士、乐于助人、做事严谨。

哪些人会对她的故事有阅读兴趣呢

- ★ 需要提升自身竞争力的职场新人。
- ★ 在市场营销、金融、财务、人力资源、产品设计等工作中需要做数据分析的人士。
- ★ 经常阅读经营分析、市场研究报告的各级管理人员。
- ★ 从事咨询、研究、分析等工作的专业人士。

其实她还有效率助手

她的效率助手是一些常用的插件工具，此外她的配书数据文件可以到www.dozan.cn或<http://blog.sina.com.cn/xiaowenzi22>下载。拥有了这些实用的插件，可以让我们的数据分析如鱼得水。

致谢

感谢笔者的好朋友李治的鼓励和支持，让笔者下定决心写这本书。在此要衷心感谢成都道然科技有限责任公司 的姚新军先生，感谢他的提议和在写作过程中的支持。感谢参与本书优化的朋友：王斌、李伟、张强林、万雷、李平、王晓、景小燕、余松。非常感谢本书的插画师王馨的辛勤劳动，您的作品也让本书增色了不少。

感谢邓凯、黄成明、李双、刘晓霞、刘云锋、欧维平、石军、沈浩、张文彤、张立良、张志成、郑来轶、祝迎春、王雍、伍昊等书评作者，感谢他们在百忙之中抽空阅读书稿，撰写书评，并提出宝贵意见。

最后，要感谢三位作者的家人，感谢他们默默无闻的付出，没有他们的理解与支持，同样也没有本书。

尽管我们对书稿进行了多次修改，仍然不可避免地会有疏漏和不足之处，敬请广大读者批评指正，我们会在适当的时间进行修订，以满足更多人的需要。

升级版说明

《谁说菜鸟不会数据分析》自2011年7月出版以来得到广大读者朋友的大力支持，而且很荣幸获得中国书刊发行业协会颁发的“2011年度全行业优秀畅销书品种”称号。这个荣誉的取得与广大读者的大力支持是分不开的。为了让这本书精益求精，在征集了大量的读者反馈意见后，我们进行了本次的升级。

与故事作者的联系：

博客：<http://blog.sina.com.cn/xiaowenzi22>

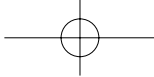
邮箱：xiaowenzi22@qq.com

新浪微博：[@小蚊子乐园](#)

与故事的策划者联系：

邮箱：info@dozan.cn

新浪微博：[@长颈鹿27](#)



业内人士的推荐（排名不分先后，以姓氏拼音排序）

本书将看似“浮云”的数据分析知识，蕴于商业化的场景之中，生动形象地让读者了解到“给力”的数据分析师是如何炼成的！本书引导非专业人士从数据的角度，认识、剖析、解决商业问题；对专业人士而言，亦能提供一次梳理知识和提高技能的机会。

邓凯

数据挖掘与数据分析博主，资深数据分析师

这是一本适合普通大众的“专业”数据分析图书，由浅入深，富有体系。让人有一口气读完的冲动，想马上找一台电脑试一试这些“新奇”的分析方法，更想拿一些数据来分析其中的规律。

读完本书，你会发现数据分析的乐趣，它并不是那么枯燥，数据背后的故事简直太有意思了。从此你将发现：无论新闻媒体，还是企业报表中的数字都有生命，因为它们在那里用自己的语言和你交流！

祝愿大家早日练就一颗数据分析的“心”！

黄成明

数据化管理顾问及培训师，零售及服装企业数据化管理咨询顾问

本书内容实用，语言简洁生动，通俗易懂。通过富有逻辑的路径式图示引导，对复杂步骤的图文分解，让读者快速掌握用Excel进行数据分析的各项实用技巧，令人耳目一新。不仅便于学习，也便于上机操作。

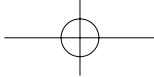
李双

数据分析与挖掘交流站，站长

数据分析的门槛有多高？可以很高，也可以不高。小蚊子的这本书可以给初学者一些实战性的指引。书中介绍的多个常用数据分析方法，对于初学者甚至是一定经验的数据分析师都有很好的启发。

刘晓霞

慧聰那白氏研究e-Eyes事业部，副总经理



>> 谁说菜鸟不会数据分析（入门篇）

本书是市面上少见的一本系统讲解数据分析的书籍。它没有针对高级分析方法和统计函数的介绍，而是针对职场人士日常工作中遇到的问题提出解决方案。书中通过小白跟师父学艺的整个过程，生动形象地描述和解释了什么是数据分析以及如何有效地进行数据分析，通俗易懂，趣味性非常强，是一本非常适合初学者的书籍。

刘云锋

北京简博市场咨询有限公司，高级经理

本书有四大亮点。其一，抛开了烦琐的统计公式与数理推论，完全以实践应用为导向，十分适合于没有统计背景的普通白领；其二，本书基本是小蚊子实际工作经验的总结，与大部分作者言之不尽的陋习不同的是，小蚊子在著作中对自己掌握的知识倾囊相授，这也是小蚊子一贯的品性；其三，作品除了教会你使用Excel简单处理日常工作涉及的数据分析，还在统计分析图表的展示方面为读者提供了重要的指导；其四，本书写作方式有新意，如小说一般的写作，使枯燥的数据分析过程兴趣盎然。

如果你正在为复杂的统计公式而头痛，如果你正在为学习统计软件而烦恼，如果你正在为如何将数据分析的结果进行专业的呈现而伤神，那么，选择这本书，也许你就找到了终南捷径。

欧雅平

广州市准诚市场咨询有限公司，研究总监

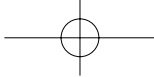
数据分析理论、公式和方法对部分初学者来说是枯燥、乏味的，或坠入云山雾海中不得其道。本书最大的特点是使用幽默风趣的语言，结合工作中典型案例加以分析、解读，是一本值得数据分析工作者一读的好书。

石军

安徽同微信息技术有限公司，总经理

当谈到用数据解决问题时，我经常用这样的语言去诠释：“如果你不能量化它，你就不能理解它，如果不理解就不能控制它，不能控制也就不能改变它”。数据无处不在，信息时代的最主要特征就是“数据处理”，数据分析正以我们从未想象过的方式影响着日常生活。

在知识与信息技术时代，每个人都面临着如何有效地吸收、理解和利用信息的挑战。那些能够有效利用工具从数据中提炼信息、发现知识的人，最终往往成为各行各业的强者！



这本书向我们清晰又友好地介绍了数据分析方法、技巧与工具，强烈推荐读一读这本书，它或许会给你带来意想不到的收获！

沈浩

中国传媒大学电视与新闻学院，教授

调查统计研究所，副所长

数据挖掘研发中心，主任

IPSOS公司，首席技术顾问

对我们财务人员来说，每个月都要写经济活动分析，但总是列出一堆干巴巴的数字，领导不爱看，自己也脸上无光，而这本书却能改变这一切。不懂数据库？不会Excel？不知如何做图表？没关系，这本书充分考虑了初学者的知识背景，让你从入门到精通。更难能可贵的是，本书设计的场景风趣幽默，让人仿佛是在看小说，把枯燥无味的学习变得生动有趣。

张立良

Excel必备工具箱，开发者

统计学是一门很难，但是很有趣，更是很有用的工具学科。懂得如何使用它的人总是乐在其中，而尚未入门的人则畏之如虎。国内讲述统计学理论，以及讲述统计软件操作的书籍可谓汗牛充栋，但是多数流于理论，疏于应用和实践指导。很大一部分读者需求未被满足。

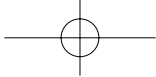
近年来随着信息技术的普及，各行各业的业务数据自动化趋势愈来愈明显，使得数据分析的需求开始从统计专业人士向各行业人员全面扩展。在此背景之下，出版一本能够深入浅出，从实际应用的角度介绍统计分析基础知识的书就变得很有必要。

本书在理论和实践的平衡方面做了很有价值的尝试，基于最为普及的Excel、5W2H、PEST等数据分析方法论，深入浅出地介绍了如何满足具体工作中的常见统计分析需求，对于需要应用统计分析，但是又未接受过这方面系统培训的读者来说，本书应当是一本非常合适的数据分析入门教材。

张文彤博士

英德知联恒市场咨询（上海）有限公司 新事业开发部，全国技术总监

这是一本真正介绍数据分析而不是介绍数据分析工具的书，全书内容就是按照数据分析流程而组织的，每一章节均通过案例来讲解，语言生动有趣。更加可贵的是，案例的“剧情”大多都



>> 谁说菜鸟不会数据分析（入门篇）

是作者多年来在现实中遇到的，因此实用性较高！在宏观结构上采用的经典结构能够带领读者按照正确的顺序稳步前进，在微观上采用的叙述方式极大增强了“渴”读性，使得应用技术更加直观。本书还是一本经典案例大全，内容涵盖人力资源等各方面的应用。因此，本书适合所有工作中需要做数据分析的朋友阅读！

张志成

<http://blog.soufun.com/site>，[远址分析师](#)

数据分析是一种能力，更是一种思想。本书结构有层次、内容全面、通俗易懂，一步步带你走进数据分析的世界，让数据分析变得既简单又有趣。

郑来秋

中国统计网站长，资深数据分析师

这是一本简单易懂，但又不缺乏深度的数据分析图书。该书作者常年从事数据分析工作，熟悉日常数据分析工作常见的问题和解决方案，所以该书积累了大量数据分析的实用性方法与技巧，需要细致的深入其境的学习，最好直接跟着书中内容实际操作，边做边学边领悟，这样可以达到事半功倍的效果。数据分析需要不断在工作中实践，这是一本入门性的书籍，最终还是需要读者靠自己的意志力克服畏难情绪去学习。付出才会有收获，学习任何东西都是如此。

祝迎春

高等学校教材《SPSS统计分析高级教程》，合作者

数据分析圈的朋友应该都知道小蚊子或小蚊子乐园这个博客，本书是小蚊子同学多年数据分析工作的积累总结，是一本简单实用的书，是数据分析技巧与工作实例充分结合的一本书。全书通过幽默的对话勾勒出数据分析的全景，彻底打开了以往数据分析相关专业书籍单一枯燥的局面。

王雍

数据元素博主，资深数据分析师

你们想想，带着这本书出了城，吃着火锅还唱着歌，突然就学会数据分析了……

伍昊

五号咨询，Excel首席培训师

目 录

第1章 数据分析那些事儿

- 1.1 数据分析是“神马” /14
 - 1.1.1 何谓数据分析 /15
 - 1.1.2 数据分析的作用 /16
- 1.2 数据分析六步曲 /17
 - 1.2.1 明确分析目的和思路 /18
 - 1.2.2 数据收集 /20
 - 1.2.3 数据处理 /21
 - 1.2.4 数据分析 /21
 - 1.2.5 数据展现 /22
 - 1.2.6 报告撰写 /23
- 1.3 数据分析的三大误区 /24
- 1.4 数据分析师的职业发展 /25
 - 1.4.1 数据分析的广阔前景 /25
 - 1.4.2 数据分析师的职业要求 /27
 - 1.4.3 数据分析师的基本素质 /28
- 1.5 几个常用指标和术语 /32
- 1.6 本章小结 /36

第2章 结构为王——确定分析思路

- 2.1 数据分析方法论 /38
 - 2.1.1 数据分析方法论与数据分析法的区别 /38
 - 2.1.2 数据分析方法论的重要性 /39
- 2.2 常用的数据分析方法论 /40
 - 2.2.1 PEST分析法 /40
 - 2.2.2 5W2H分析法 /43
 - 2.2.3 逻辑树分析法 /44
 - 2.2.4 4P营销理论 /45
 - 2.2.5 用户行为理论 /47
- 2.3 本章小结 /48

第3章 无米难为巧妇——数据准备

- 3.1 理解数据 /50

>> 谁说菜鸟不会数据分析（入门篇）

3.1.1 字段与记录 /51

3.1.2 数据类型 /52

3.1.3 数据表 /53

3.2 数据来源 /57

3.2.1 导入数据 /57

3.2.2 问卷录入要求 /63

3.3 本章小结 /65

第4章 三心二意——数据处理

4.1 何为数据处理 /67

4.1.1 “三心二意”处理数据 /67

4.1.2 数据处理的内容 /69

4.2 数据清洗 /70

4.2.1 重复数据的处理 /71

4.2.2 缺失数据处理 /76

4.2.3 检查数据逻辑错误 /80

4.3 数据加工 /82

4.3.1 数据抽取 /82

4.3.2 数据计算 /87

4.3.3 数据分组 /91

4.3.4 数据转换 /92

4.4 数据抽样 /97

4.5 本章小结 /98

第5章 工欲善其事必先利其器——数据分析

5.1 数据分析方法 /101

5.1.1 对比分析法 /102

5.1.2 分组分析法 /105

5.1.3 结构分析法 /106

5.1.4 平均分析法 /107

5.1.5 交叉分析法 /108

5.1.6 综合评价分析法 /109

5.1.7 杜邦分析法 /113

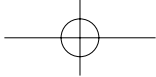
5.1.8 漏斗图分析法 /114

5.1.9 矩阵关联分析法 /115

5.1.10 高级数据分析方法 /120

5.2 数据分析工具 /121

5.2.1 初识数据透视表 /121



目 录

- 5.2.2 创建数据透视表的三步法 /122
- 5.2.3 数据透视表分析实践 /124
- 5.2.4 数据透视表小技巧 /130
- 5.2.5 多选题分析 /132

5.3 本章小结 /135

第6章 给数据量体裁衣——数据展现

6.1 揭开图表的真面目 /138

- 6.1.1 图表的作用 /138
- 6.1.2 经济适用图表有哪些 /139
- 6.1.3 通过关系选择图表 /140
- 6.1.4 图表制作五步法 /145

6.2 表格也疯狂 /146

- 6.2.1 突出显示单元格 /146
- 6.2.2 项目选取 /147
- 6.2.3 数据条 /148
- 6.2.4 图标集 /149
- 6.2.5 迷你图 /150

6.3 给图表换装 /151

- 6.3.1 平均线图 /152
- 6.3.2 双坐标图 /153
- 6.3.3 竖形折线图 /156
- 6.3.4 瀑布图 /159
- 6.3.5 帕累托图 /160
- 6.3.6 旋风图 /165
- 6.3.7 人口金字塔图 /169
- 6.3.8 漏斗图 /171
- 6.3.9 矩阵图（散点图） /173
- 6.3.10 发展矩阵图 /176
- 6.3.11 改进难易矩阵（气泡图） /178

6.4 本章小结 /180

第7章 专业化生存——图表可以更美的

7.1 别让图表犯错 /183

- 7.1.1 让图表“五脏俱全” /183
- 7.1.2 要注意的条条框框 /185
- 7.1.3 图表会说谎 /196

7.2 浓妆淡抹总相宜——图表美化 /200

>> 谁说菜鸟不会数据分析（入门篇）

- 7.2.1 图表美化的三原则 /200
- 7.2.2 略施粉黛——美化图表的技巧 /203
- 7.2.3 图表也好“色” /209

7.3 如虎添翼的招儿 /213

- 7.3.1 我的图表模板 /214
- 7.3.2 快速制图 /215
- 7.3.3 添加标签小工具 /216
- 7.3.4 修剪超大值 /218

7.4 本章小结 /220

第8章 专业的报告——体现你的职场价值

8.1 初识数据分析报告 /222

- 8.1.1 数据分析报告是什么 /222
- 8.1.2 数据分析报告的写作原则 /222
- 8.1.3 数据分析报告的作用 /224
- 8.1.4 数据分析报告的种类 /225

8.2 数据分析报告的结构 /228

- 8.2.1 标题页 /228
- 8.2.2 目录 /230
- 8.2.3 前言 /231
- 8.2.4 正文 /233
- 8.2.5 结论与建议 /234
- 8.2.6 附录 /234

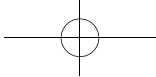
8.3 撰写报告时的注意事项 /235

8.4 报告范例 /236

8.5 本章小结 /244

写在后面的Q/A

附录A 网络学习资源



第 1 章

数据分析那些事儿

数据分析是“神马”

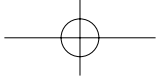
数据分析六步曲

数据分析的三大误区

数据分析师的职业发展

几个常用指标和术语





>> 谁说菜鸟不会数据分析（入门篇）

小白过五关斩六将，通过严格的面试，最终从众多优秀毕业生中脱颖而出，成为公司的一员。在报到的第一天，公司HR向小白介绍了她的职位——公司牛董的助理，负责文秘工作，可能需要做一些数据分析之类的活儿。小白一听到数据分析这个词，就感觉头皮发麻，这时，她的脑子里幻想出一些穿白大褂的科研人员在实验室的计算机前不断忙碌的场景。虽然在上大学时也使用过Excel，但是如果要做数据分析工作，她还真的有些不知如何下手。无数个问号涌到她的脑海中：数据分析到底要做什么呢？我要怎么做数据分析？老板想要看什么样的结果呢？……唉，只好边走边干了。

HR看出小白的心事，说道：关于数据分析你不用太担心，如果遇到难题，你可以请教我们公司的Mr.林，他在这方面可是专家喔！

小白一听有救星，立马兴奋起来，好像抓到救命稻草一样，想赶紧找到这位大师级人物，然后一股脑把疑问全倒出来。

HR：小白，你跟我来吧，我给你引荐下Mr.林。

小白：好的。

说着小白跟HR来到了Mr.林的办公桌旁，HR说道：Mr.林，这是我们公司新来的同事，叫小白，现担任牛董的助理，其中部分工作涉及数据分析，到时候有问题还要麻烦您多多指点。

小白紧跟着说道：Mr.林，您好，请多多指教。

Mr.林：别客气，有问题直接来找我就可以了。

小白趁机说道：我现在就有问题，您现在是否有时间帮我解答下？数据分析是干什么的，具体要怎么做？

Mr.林听完笑了起来：你还真是不客气呀。好吧，你刚进公司，我就先给你做个简单的培训，带你入个门吧，以后的修行可就靠你自己努力了。

小白用力点着头，HR见这“师徒”二人颇有一见如故之感，大致交待一番后就离开了。

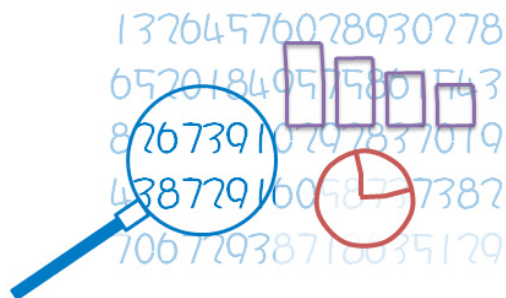
1.1 数据分析是“神马”

Mr.林接着说道：小白，既然要学习数据分析，借用现在网络流行语，你能说说数据分析是“神马”么？

小白：呵呵，数据分析肯定不是“浮云”。从字面上理解，数据分析就是对数据进行分析。Mr.林，我只能理解到这一层面，专业的解释就需要您来指教了。

1.1.1 何谓数据分析

Mr.林: 简单来说, 数据分析就是你刚才讲的——对数据进行分析。较为专业的说法, 数据分析是指用适当的统计分析方法对收集来的大量数据进行分析, 将它们加以汇总和理解并消化, 以求最大化地开发数据的功能, 发挥数据的作用。数据分析是为了提取有用信息和形成结论而对数据加以详细研究和概括总结的过程。这里的数据也称观测值, 是通过实验、测量、观察、调查等方式获取的结果, 常常以数量的形式展现出来。



数据分析

数据分析的目的是把隐藏在一大批看似杂乱无章的数据背后的信息集中和提炼出来, 总结出所研究对象的内在规律。在实际工作中, 数据分析能够帮助管理者进行判断和决策, 以便采取适当策略与行动。例如, 企业的高层希望通过市场分析和研究, 把握当前产品的市场动向, 从而制订合理的产品研发和销售计划, 这就必须依赖数据分析才能完成。

在统计学领域, 有些学者将数据分析划分为描述性数据分析、探索性数据分析以及验证性数据分析 (参见图1-1)。其中, 探索性数据分析侧重于在数据之中发现新的特征, 而验证性数据分析则侧重于验证已有假设的真伪证明。



图1-1 数据分析类别

从另一个角度看, 描述性数据分析属于初级数据分析, 常见的分析方法有对比分析法、平均分析法、交叉分析法等。而探索性数据分析以及验证性数据分析属于高级数据分析, 常见的分析

>> 谁说菜鸟不会数据分析（入门篇）

方法有相关分析、因子分析、回归分析等。我们日常学习和工作中涉及的数据分析主要是描述性数据分析，也就是大家常用的初级数据分析。

1.1.2 数据分析的作用

Mr.林：了解了数据分析是“神马”后，那么小白，你能想到数据分析在对企业日常经营状况的分析工作中具体有哪些作用，体现在哪几方面吗？

小白：您刚才说过，数据分析就是把隐藏在一大批看似杂乱无章的数据背后的信息集中和提炼出来，总结出所研究对象的内在规律，帮助管理者进行有效的判断和决策。我觉得它就好比是从矿山中挖掘出金子。

Mr.林：没错，但刚才说的是数据分析的最终目的，要达到这些目的，我们在日常工作中该做些什么呢？比如日常通报、专题分析等，这些都是数据分析具体工作的体现。而什么时候做通报，什么时候该开展专题分析，则需要我们根据实际情况做出选择。很多人经常做这些工作，但不知为何而做，只是为做而做，没有想清楚做的目的，所以常常不得要领被数据所湮没。只有当你对数据分析目的及工作有了足够清晰的认识，开展数据分析时才会如鱼得水，游刃有余。

数据分析在企业的日常经营分析中主要有三大作用，如图1-2所示。



图1-2 数据分析三大作用

🎯 现状分析

简单来说就是告诉你过去发生了什么。具体体现在：

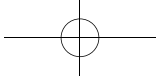
第一，告诉你企业现阶段的整体运营情况，通过各个经营指标的完成情况来衡量企业的运营状态，以说明企业整体运营是好了还是坏了，好的程度如何，坏的程度又到哪里。

第二，告诉你企业各项业务的构成，让你了解企业各项业务的发展及变动情况，对企业运营状况有更深入的了解。

现状分析一般通过日常通报来完成，如日报、周报、月报等形式。

🎯 原因分析

简单来说就是告诉你某一现状为什么发生。



第1章 数据分析那些事儿

经过第一阶段的现状分析，我们对企业的运营情况有了基本了解，但不知道运营情况具体好在哪里，差在哪里，是什么原因引起的。这时就需要开展原因分析，以进一步确定业务变动的具体原因。例如2012年2月运营收入环比下降5%，是什么原因导致的呢？是各项业务收入都出现下降，还是个别业务收入下降引起的？是各个地区业务收入都出现下降，还是个别地区业务收入下降引起的？这就需要我们开展原因分析，进一步确定收入下降的具体原因，对运营策略做出调整与优化。

原因分析一般通过专题分析来完成，根据企业运营情况选择针对某一现状进行原因分析。

🎯 预测分析

简单来说就是告诉你将来会发生什么。

在了解企业运营现状后，有时还需要对企业未来发展趋势作出预测，为制订企业运营目标及策略提供有效的参考与决策依据，以保证企业的可持续健康发展。

预测分析一般通过专题分析来完成，通常在制订企业季度、年度等计划时进行，其开展的频率没有现状分析及原因分析高。

Mr.林：小白，现在对数据分析有了进一步认识了吧。什么时候开展什么样的数据分析，需要根据你的需求及目的来确定。

小白：是的，Mr.林，那么请问数据分析具体如何开展呢？有哪些准备工作要做呢？

1.2 数据分析六步曲

Mr.林：小白，你的问题问得好，现在我们就来看看如何开展数据分析工作。

数据分析过程主要包括6个既相对独立又互有联系的阶段。它们是：明确分析目的和思路、数据收集、数据分析、数据处理、数据展现、报告撰写等6步（参见图1-3）。

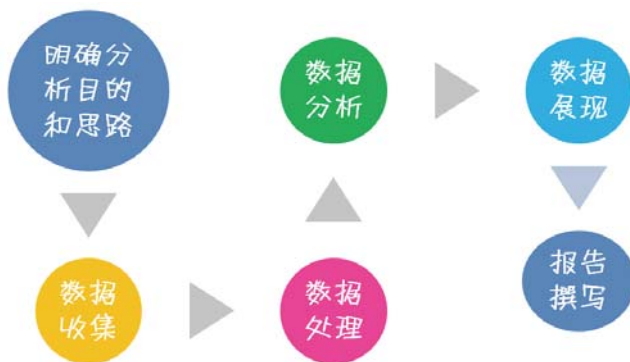


图1-3 数据分析六步曲

1.2.1 明确分析目的和思路

明确分析目的

Mr.林：做任何事都要有个目标，数据分析也不例外。小白，我们先来看看菜鸟与数据分析师之间的差别。刚好我这里有一张菜鸟与数据分析师想法对比图（见图1-4），你可以对比一下，看看自己在哪些方面想法还属于菜鸟级别，在哪些方面想法已经达到了数据分析师的水平。从他们思考问题的方式，你就能发现为什么一个是菜鸟，而另外一个和数据分析师了。

小白：好的，我来看一下菜鸟与数据分析师各是什么样的。

菜鸟与数据分析师的区别	
菜鸟会想	分析师会想
这张曲线图真好看，怎么作的？	数据变化的背后真相是什么？
这些数据可以做什么样的分析？	从哪些角度分析数据才系统？
高级的分析方法在这里能用吗？	用什么分析方法最有效？
要作多少张图表？	图表是否表达出有效的观点？
除了为数据添加文字说明还需说什么？	数据分析的目的达到了吗？
数据分析报告要写多少页？	数据分析报告有说服力吗？
.....

图1-4 菜鸟与数据分析师的区别

Mr.林：小白，从菜鸟与数据分析师想法对比图中，你发现菜鸟与数据分析师的区别了吗？

小白：我发现，菜鸟很迷茫，目的不明确，而且一味追求高级的分析方法。而数据分析师的目的就很明确，一切都是以解决问题为中心。

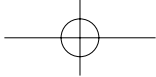
Mr.林：说得好！其实他们最主要的区别就在于目的是否明确，如果目的明确，所有问题就自然迎刃而解了。例如，分析师是不会考虑“需要多少张图表”这样的问题的，而是思考这张图表是否有效表达了观点？如果没有，需要怎样调整？如果目的不明确，菜鸟自然会在用什么样的分析方法、作多少张图表、需要多少文字说明、报告要写多少页等这些问题上纠结。

小白：我明白了，我不要做菜鸟，我要朝着数据分析师的方向努力。

Mr.林：菜鸟与数据分析师的区别就在于菜鸟做分析时目的不明确，从而导致分析过程非常盲目。所以在开展数据分析之前，要想想：为什么要开展数据分析？通过这次数据分析我要解决什么问题？只有明确数据分析的目标，数据分析才不会偏离方向，否则得出的数据分析结果不仅没有指导意义，甚至可能将决策者引入歧途，后果严重。

确定分析思路

Mr.林：当分析目的明确后，我们就要梳理分析思路，并搭建分析框架，把分析目的分解成



第1章 数据分析那些事儿

若干个不同的分析要点，即如何具体开展数据分析，需要从哪几个角度进行分析，采用哪些分析指标。

只有明确了分析目的，分析框架才能跟着确定下来，最后还要确保分析框架的体系化，使分析结果具有说服力。

小白：Mr.林，前面说的角度与指标我都理解，但分析体系化应该如何理解呢？

Mr.林：体系化也就是逻辑化，简单来说就是先分析什么，后分析什么，使得各个分析点之间具有逻辑联系。这也是很多人常常感到困扰的问题，比如经常不知从哪方面入手，分析的内容和指标常常被质疑是否合理、完整，而自己也说不出个所以然来，所以体系化就是为了让你的分析框架具有说服力。

小白：如何使分析框架体系化呢？

Mr.林：问得好！根据我的经验，就是以营销、管理等理论为指导，结合实际业务情况，搭建分析框架，这样才能确保数据分析维度的完整性，分析结果的有效性及其正确性。

营销方面的理论模型有4P、用户使用行为、STP理论、SWOT等，而管理方面的理论模型有PEST、5W2H、时间管理、生命周期、逻辑树、金字塔、SMART原则等。这些都是经典的营销、管理方面的理论，需要在工作中不断实践应用，你才能体会其强大的作用。如图1-5所示，就是以PEST分析理论为指导，搭建的互联网行业PEST分析框架，故而使数据分析变得有血有肉有脉络，真正做到理论指导实践。其他具体应用我在后面会为你介绍。

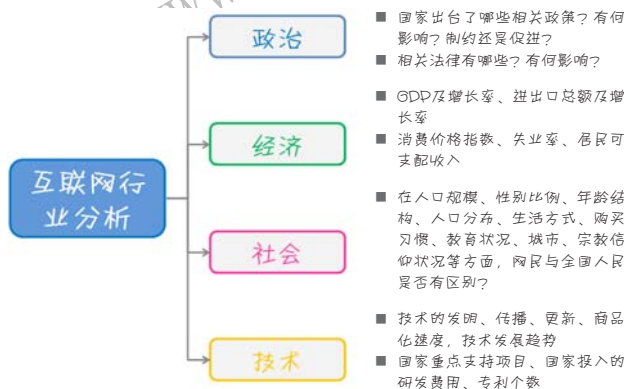
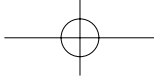


图1-5 互联网行业PEST分析框架

明确数据分析目的以及确定分析思路，是确保数据分析过程有效进行的先决条件，它可以为数据的收集、处理及分析提供清晰的指引方向。

小白：嗯，在做数据分析前，先想想到底为什么要做数据分析，明确数据分析方向，才能保证不迷失在数据的丛林中。



>> 谁说菜鸟不会数据分析（入门篇）

1.2.2 数据收集

Mr.林：数据收集是按照确定的数据分析框架，收集相关数据的过程，它为数据分析提供了素材和依据。这里所说的数据包括第一手数据与第二手数据，第一手数据主要指可直接获取的数据，第二手数据主要指经过加工整理后得到的数据。一般数据来源主要有以下几种方式，如图1-6所示。



图1-6 数据的来源

数据库

每个公司都有自己的业务数据库，存放从公司成立以来产生的相关业务数据。这个业务数据库就是一个庞大的数据资源，需要有效地利用起来。

公开出版物

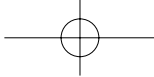
可以用于收集数据的公开出版物包括《中国统计年鉴》《中国社会统计年鉴》《中国人口统计年鉴》《世界经济年鉴》《世界发展报告》等统计年鉴或报告。

互联网

随着互联网的发展，网络上发布的数据越来越多，特别是搜索引擎可以帮助我们快速找到所需要的数据，例如国家及地方统计局网站、行业组织网站、政府机构网站、传播媒体网站、大型综合门户网站等上面都可能我们有需要的数据。

市场调查

进行数据分析时，需要了解用户的想法与需求，但是通过以上三种方式获得此类数据会比较困难，因此可以尝试使用市场调查的方法收集用户的想法和需求数据。市场调查就是指运用科学的方法，有目的、有系统地收集、记录、整理有关市场营销的信息和资料，分析市场情况，了解



第1章 数据分析那些事儿

市场现状及其发展趋势，为市场预测和营销决策提供客观、正确的数据资料。市场调查可以弥补其他数据收集方式的不足，但进行市场调查所需的费用较高，而且会存在一定的误差，故仅作参考之用。

小白：看来我以后做数据分析时，不用担心没有数据来源啦。

1.2.3 数据处理

Mr.林：数据处理是指对收集到的数据进行加工整理，形成适合数据分析的样式，它是数据分析前必不可少的阶段。数据处理的基本目的是从大量的、杂乱无章、难以理解的数据中，抽取并推导出对解决问题有价值、有意义的数据。

数据处理主要包括数据清洗、数据转化、数据提取、数据计算等处理方法（参见图1-7）。一般拿到手的数据都需要进行一定的处理才能用于后续的数据分析工作，即使再“干净”的原始数据也需要先进行一定的处理才能使用。



图1-7 数据处理

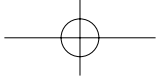
小白：哦，数据处理是数据分析的前提，对有效数据的分析才是有意义的，明白啦！

1.2.4 数据分析

Mr.林：数据分析是指用适当的分析方法及工具，对处理过的数据进行分析，提取有价值的信息，形成有效结论的过程。



数据分析



>> 谁说菜鸟不会数据分析（入门篇）

在确定数据分析思路的阶段，数据分析师应当为需要分析的内容确定适合的数据分析方法，等到真正进入进行数据分析阶段时，就能够驾驭数据，从容地进行分析和研究了。

由于数据分析多是通过软件来完成的，这就要求数据分析师不仅要掌握各种数据分析方法，还要熟悉主流数据分析软件的操作。一般的数据分析我们可以通过Excel完成，后面也将重点为你介绍，而高级的数据分析就要采用专业的分析软件进行了，如数据分析工具SPSS Statistics等。

小白：好的，期待您后续精彩的数据分析讲解。

📍 数据分析与数据处理的区别

小白：Mr.林，我有个问题。我经常听到数据处理和数据分析这两个词，它们之间有何关系呢？

Mr.林：从前面介绍数据处理和分析的过程不难看出，数据处理是数据分析的基础。通过数据处理，将收集到的原始数据转换为可以分析的形式，并且保证数据的一致性和有效性。如果数据本身存在错误，那么即使采用最先进的数据分析方法，得到的结果也是错误的，不具备任何参考价值，甚至还会误导决策。

📍 数据分析与数据挖掘的关系

小白：还有个问题，我常听到数据挖掘，那么数据挖掘是什么呢？它与数据分析有什么关系呢？

Mr.林：这个问题问得好！数据挖掘其实是一种高级的数据分析方法。

数据挖掘就是从大量的数据中挖掘出有用的信息，它是根据用户的特定要求，从浩如烟海的数据中找出所需的信息，以满足用户的特定需求。数据挖掘技术是人们长期对数据库技术进行研究和开发的结果。一般来说，数据挖掘侧重解决四类数据分析问题：分类、聚类、关联和预测，重点在寻找模式与规律。

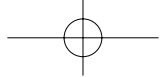
数据分析与数据挖掘的本质是一样的，都是从数据里面发现关于业务的知识。

数据挖掘现在还不是你学习的重点，如果你有兴趣，可在掌握一定的数据分析知识后，再进一步查阅相关资料学习。

1.2.5 数据展现

小白：您刚才讲通过分析，隐藏在数据内部的关系和规律就会逐渐浮现出来，那么通过什么方式展现出这些关系和规律，才能让人一目了然呢？

Mr.林：不错，孺子可教也！这就是我接下来要说的数据展现了。众所周知，每个人看待事物都有自己的理解方式，所以数据分析师在展现结果的时候一定要保证绝大部分人的理解是一致的。



第1章 数据分析那些事儿



数据展现

一般情况下，数据是通过表格和图形的方式来呈现的，我们常说用图表说话就是这个意思。常用的数据图表包括饼图、柱形图、条形图、折线图、散点图、雷达图等，当然可以对这些图表进一步整理加工，使之变为我们所需要的图形，例如金字塔图、矩阵图、漏斗图、帕雷托图等。

大多数情况下，人们更愿意接受图形这种数据展现方式，因为它能更加有效、直观地传递出分析师所要表达的观点。记住，在一般情况下，能用图说明问题的就不用表格，能用表格说明问题的就不用文字。

小白：原来图表的运用还这么有讲究呀，我一定好好学习如何用图表说话！

1.2.6 报告撰写

Mr.林：数据分析报告其实是对整个数据分析过程的一个总结与呈现。通过报告，把数据分析的起因、过程、结果及建议完整地呈现出来，供决策者参考。所以，数据分析报告是通过对数据全方位的科学分析来评估企业运营质量，为决策者提供科学、严谨的决策依据，以降低企业运营风险，提高企业核心竞争力。

一份好的数据分析报告，首先需要有一个好的分析框架，并且图文并茂，层次明晰，能够让阅读者一目了然。结构清晰、主次分明可以使阅读者正确理解报告内容；图文并茂，可以令数据更加生动活泼，提高视觉冲击力，有助于阅读者更形象、直观地看清楚问题和结论，从而产生思考。

另外，数据分析报告需要有明确的结论，没有明确结论的分析称不上分析，同时也失去了报告的意义，因为我们最初就是为寻找或者求证一个结论才进行分析的，所以千万不要舍本求末。

最后，好的分析报告一定要有建议或解决方案。作为决策者，需要的不仅仅是找出问题，更重要的是建议或解决方案，以便他们在决策时作参考。所以，数据分析师不仅需要掌握数据分析方法，而且还要了解和熟悉业务，这样才能根据发现的业务问题，提出具有可行性的建议或解决方案。下面的图1-8给出了一个分析报告的基本框架示例。

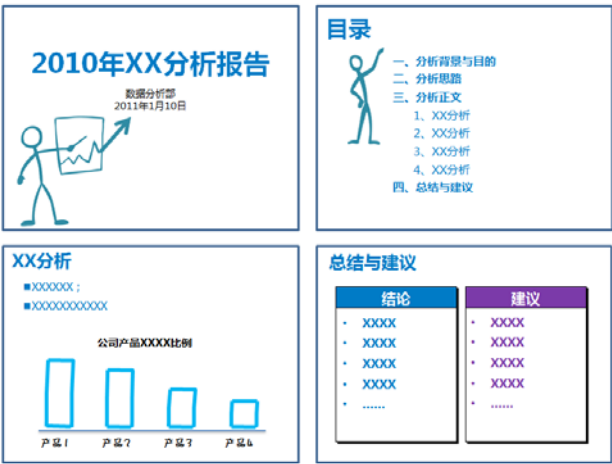


图1-8 数据分析报告示例

Mr.林：以上是数据分析的6个关键步骤，今天只是简要介绍，以后将逐一为你讲解。说了这么多，给你讲个段子轻松一下，顺便回顾下刚讲的内容。

小白一听有段子听，立马精神起来：好啊！好啊！

Mr.林：如何判断数据分析师正处于分析流程的哪个阶段？

- ★ 手托腮帮，沉思痛苦状——在思考分析思路。
- ★ 手放键盘上不动，表情呆滞——在数据处理。
- ★ 鼠标在飞快移动——在用数据透视表做分析。
- ★ 不断交替点击鼠标左右键——在画图表。
- ★ 断断续续敲击键盘，时而移动鼠标——在写PPT分析报告。

小白：还真形象啊！Mr.林，您该不会说的就是您自己吧？呵呵！

Mr.林：哈哈，数据分析师差不多都是这样的状态。

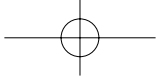
1.3 数据分析的三大误区

Mr.林：小白，刚才对数据分析进行了简要介绍，而在实际的学习、工作中，常常有数据分析人员陷入一些误区，现在我来讲一讲这些误区，你日后一定要注意。

小白：好的，请Mr.林指教。

① 分析目的不明确，为分析而分析

Mr.林：经常有人问：要用多少图？除了摆数据，还需要说些什么？在此我想说的是，数据分析不应为了分析而分析，而是应该围绕你的分析目的（了解现状、找出业务变动原因、预测发



第1章 数据分析那些事儿

展等)而进行分析。

只有对自己的目的有清晰的认识,你才知道要怎样去实现这个目的,需要通过哪些图表展现,才会知道这些图表是否能反映问题,自然而然地进行相应的问题分析,而不是连该说些什么都不知道。

① 缺乏业务知识,分析结果偏离实际

Mr.林: 目前现有的数据分析师大多是统计学、计算机、数学等专业出身,他们大多缺乏从事营销、管理方面的工作经验,对业务的理解相对较浅,对数据的分析偏重于数据分析方法的使用,如回归分析、相关分析等。

有的公司老板抱怨手下的数据分析师每天给他看几十个零散数据,虽然做出的报告很专业,图表也很漂亮,但所作的分析忽视了业务逻辑上的关联性,得不到全面、综合性的结论。

在企业中所作的数据分析不是纯数据分析,而是需要多从业务方面进行分析,不应停留在数据表面,要思考数据背后的事实与真相,使得分析结果更加切合实际,为老板的决策提供有力的支撑,否则就是纸上谈兵。

所以说,数据分析师的任务不是单纯做数学题,数据分析师还必须懂营销,懂管理,更要懂策略。

② 一味追求使用高级分析方法,热衷研究模型

Mr.林: 在进行数据分析时,相当一部分人都喜欢用回归分析、因子分析等高级分析方法,总认为有分析模型就是专业的,只有这样才能体现专业性,结果才是可信的。其实不然,高级的数据分析方法不一定是最好的,能够简单有效解决问题的方法才是最好的。

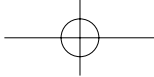
我们坚信,仅有分析模型远远不够,围绕业务发现问题并解决问题才是数据分析的最终目的!不论高级的分析方法还是简单的分析方法,只要能够解决业务问题,就是好方法,正如我们常说的“不论黑猫还是白猫,只要能抓老鼠就是好猫”。

小白: 好的,我在以后的工作中会注意这些的,做好一个数据分析师并不是一件容易的事,我要努力向Mr.林靠齐。

1.4 数据分析师的职业发展

1.4.1 数据分析的广阔前景

Mr.林: 数据分析作为一个新的行业领域正在全球迅速发展,它开辟了人类获取知识的新途径。



>> 谁说菜鸟不会数据分析（入门篇）

目前，数据库技术、软件工具、各种硬件设备飞速发展，在这些软硬件技术与设备的支持下，信息技术的应用已在各行各业全面展开，尤其是对通信、互联网、金融等行业的发展做出了巨大贡献，并且经过长期的应用积累了大量丰富的数据。但大部分企业对其存储信息的利用率极低。庞大的历史数据是否有价值？有何价值？是否可以综合利用分析？是否能够为领导决策提供参考依据？

回答是肯定的，数据分析这一项工作越来越受到决策层的重视，借助数据分析的各种工具，从海量的历史数据中提取、挖掘对业务发展有价值的、潜在的知识，找出趋势，为决策层的决策提供有力的依据，对产品或服务的发展方向起到积极作用，有力推动企业的科学化、信息化管理。

从20世纪90年代起，欧美国家开始大量培养数据分析师，直到现在，对数据分析师的需求仍然长盛不衰，而且还有扩张之势。根据美国劳工部预测，到2018年，数据分析师的需求量将增长20%。就算你不是数据分析师，但数据分析技能也是未来必不可少的工作技能之一。

小白：数据分析师如此抢手的原因何在呢？

Mr.林：一个简单的原因就是社会越发达，人们对数据的依赖就越多。无论政府决策还是公司运营，科学研究还是媒体宣传，都需要数据支持。那么，对数据有如此大的依赖，就必然导致对数据分析的大量需求。因此，将数据转化为知识、结论和规律，就是数据分析的作用和价值。

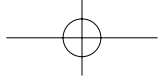
小白：那数据究竟会庞大到什么地步呢？

Mr.林：据国际知名咨询公司估计，到2020年，全球每年产生的数据量将达到3500万亿GB，打个比方，就是用普通的DVD一张一张地摞起来，可以从地球摞到月球两次。



全球每年产生的数据量

面对这样庞大的数据，对数据分析师的要求就不仅仅是单纯做分析了，更重要的是与相关业务部门进行合作，将数据真正应用到业务中，根据实际的业务发展情况识别哪些数据可用，哪些不适用，而不是孤立地在“真空环境”下进行分析。这就要求数据分析师不仅具备洞察数据的能力，还要对相关业务的背景有深入的了解，明白客户或业务部门的需求，从而将数据信息化、可



第1章 数据分析那些事儿

视化，最后转化为生产力，帮助企业获得利润，这就是整个数据“供应链”。当然，数据分析师也需要理解这个“供应链”。

1.4.2 数据分析师的职业要求

小白：那么怎样才能成为一名优秀的数据分析师呢？数据分析师的职业要求又有哪些？

Mr.林：你的问题可以用“五懂”来回答，即懂业务、懂管理、懂分析、懂工具，还要懂设计。

📍 懂业务

从事数据分析工作的前提就是需要懂业务，即熟悉行业知识、公司业务及流程，最好有自己独到见解，若脱离行业认知和公司业务背景，分析的结果只会是脱了线的风筝，没有太大的实用价值。

例如公司2011年的运营收入是1000万元，那么不熟悉业务的数据分析师看到的只是1000万这个数字，而熟悉业务的数据分析师，还能看到这个数字背后隐藏的信息，如1000万元是由哪几个业务收入构成，哪个业务收入占主要部分，哪个业务收入是最小占比，最高业务收入的地区又是哪个地区等。

这就是懂业务与不懂业务的数据分析师之间的区别。

从另外一个角度来说，懂业务也是数据敏感的体现。不懂业务的数据分析师，看到的只是一个数字；懂业务的数据分析师，则看到的不仅仅是数字，他明白数字代表什么意义，知道数字是大了还是小了，心中有数，这才是真正意义的数据敏感性。

📍 懂管理

懂管理，一方面是搭建数据分析框架的要求，比如刚才介绍的数据分析六步曲的第一步确定分析思路就需要用到营销、管理等理论知识来指导，如果不熟悉管理理论，那你如何指导数据分析框架的搭建，以及开展后续的数据分析呢？

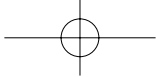
懂管理另一方面的作用是针对数据分析结论提出有指导意义的分析建议，如果没有管理理论的支撑，就难以确保分析建议的有效性。

所以数据分析师需要掌握一定的管理理论知识。

📍 懂分析

懂分析是指掌握数据分析的基本原理与一些有效的数据分析方法，并能灵活运用到实践工作中，以便有效地开展数据分析。

基本的分析方法有：对比分析法、分组分析法、交叉分析法、结构分析法、漏斗图分析法、综合评价分析法、因素分析法、矩阵关联分析法等。



>> 谁说菜鸟不会数据分析（入门篇）

高级的分析方法有：相关分析法、回归分析法、聚类分析法、判别分析法、主成分分析法、因子分析法、对应分析法、时间序列等。

不论简单还是复杂的分析方法，只要能解决问题的方法就是好方法。

📍 懂工具

懂工具是指掌握数据分析相关的常用工具。数据分析工具就是实现数据分析方法理论的工具，面对越来越庞大的数据，依靠计算器进行分析是不现实的，必须利用强大的数据分析工具完成数据分析工作。

常用的数据分析工具有Excel、Access、SPSS、SAS，建议先用好Excel分析工具。Excel就是一款非常实用的数据处理、分析工具，它能解决、满足你工作中80%，甚至100%的要求。有兴趣、时间、需要的话，再学习SPSS、SAS等统计分析工具。

同样，应该根据研究的问题选择合适的工具，只要能解决问题的工具就是好工具。

📍 懂设计

懂设计是指运用图表有效表达数据分析师的分析观点，使分析结果一目了然。

图表的设计是门大学问，如图形的选择、版式的设计、颜色的搭配等，都需要掌握一定的设计原则。

小白：颜色搭配都有要求？不能根据我自己的喜好选择么？

Mr.林：当然有要求，根据喜好选择也要在符合设计原则的基础上选择，例如公司的VI设计是以蓝色为主色调，属冷色系，那么你选的图表颜色就要与公司的VI设计相吻合，尽量避免使用红色、黄色等暖色系的颜色，这方面知识我会在讲到数据展现部分再为你详细介绍。

小白：好的，看来要成为一名合格的数据分析师还有很长的路要走啊。

1.4.3 数据分析师的基本素质

Mr.林接着说道：要成为一名优秀的数据分析师并非一件容易的事。虽然大学的专业与数据分析不相关，但你可以通过工作中的实践来学习数据分析，这需要付出大量的时间和精力，不经一番寒彻骨，怎得梅花扑鼻香？

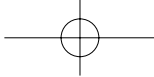
小白点点头：嗯，我会努力的。

Mr.林：刚才介绍的是数据分析师的职业要求，现在再介绍数据分析师的五大基本能力和素质。

📍 态度严谨负责

严谨负责是数据分析师的必备素质之一，只有本着严谨负责的态度，才能保证数据的客观、准确。

数据分析师可以说是企业的医生，他们通过对企业运营数据的分析，为企业寻找症结及潜在



第1章 数据分析那些事儿

问题。一名合格的数据分析师，应具有严谨负责的态度，保持中立立场，客观评价企业在发展过程中存在的问题，为决策层提供有效的参考依据。数据分析师不应受其他因素影响而更改数据，隐瞒企业存在的问题，这样做对企业发展是非常不利的，甚至会造成严重的后果。而且，对数据分析师自身来说，其职业生涯也会受影响，从此以后所做的数据分析结果都将受到质疑，因为你不再是可信赖的人，在同事、领导、客户面前已经失去了信任。

所以，作为一名数据分析师就必须持有严谨负责的态度，这也是最基本的职业道德。



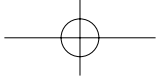
严谨负责

◎ 好奇心强烈

好奇心人皆有之，但是作为数据分析师，这份好奇心就应该更强烈，应该把自己当成数据福尔摩斯，要积极主动地发现和挖掘隐藏在数据内部的真相。



好奇心强烈



>> 谁说菜鸟不会数据分析（入门篇）

在数据分析师的脑子里，应该充满着无数个“为什么”：为什么是这样的结果，为什么不是那样的结果，导致这个结果的原因是什么，为什么结果不是预期的那样等等，只有这样才有突破点。

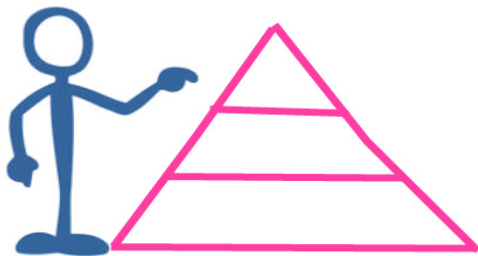
这一系列问题都要在进行数据分析时提出来，并且通过数据分析，给自己一个满意的答案。越是优秀的数据分析师，好奇心越不容易满足，回答了一个问题，又会抛出一个新的问题，继续研究下去。

只有拥有了这样一种刨根问底的精神，才会对数据和结论保持敏感，继而顺藤摸瓜，找出数据背后的真相。当然，你也会从中获取成就感。

① 逻辑思维清晰

除了一颗探索真相的好奇心，数据分析师还需要具备缜密的思维和清晰的逻辑推理能力。我记得有位大师说过：结构为王。何谓结构，结构就是我们常说的逻辑，不论说话还是写文章的时候都要有条理、有目的地做，不可眉毛胡子一把抓，不分主次。

通常，从事数据分析时所面对的商业问题都是较为复杂的，我们要思考错综复杂的成因，分析所面对的各种复杂的环境因素，并在问题的若干发展可能性中选择一个最优的方向。这就需要我们对事实有足够的了解，同时也需要我们能真正理清问题的整体以及局部的结构，在深度思考后，理清结构中相互的逻辑关系，只有这样才能真正客观地、科学地找到商业问题的答案。



逻辑思维清晰

② 擅长模仿学习

在做数据分析时，有自己的想法固然重要，但是“前车之鉴”也是非常有必要学习的，它能帮助数据分析师迅速地成长。因此，模仿学习是快速提高学习成果的有效方法。这里说的模仿主要是指参考、借鉴他人优秀的分析思路和方法，而非直接“照搬”。

成功的模仿需要领会他人方法的精髓，理解其分析原理，透过表面达到实质。万变不离其宗，要善于将这些精华转化为自己的知识，否则，只能是“一直在模仿，从未超越过”。



第1章 数据分析那些事儿

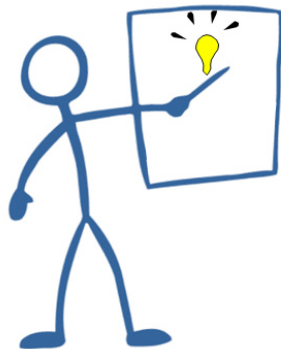


擅长模仿学习

◎ 勇于创新

通过模仿可以借鉴他人的成功经验，但模仿的时间不宜太长，并且建议每次模仿后都要进行总结，提出可以改进的地方，甚至要有所创新，不断总结分析方法、分析思路、分析流程，在总结中前行。

创新是一个优秀数据分析师应具备的精神，只有不断的创新，才能提高自己的分析水平，使自己站在更高的角度来分析问题，为整个研究领域乃至社会带来更多的价值。现在的分析方法和研究课题千变万化，墨守成规是无法很好地解决所面临的新问题的。



创新

听到这里，小白就掰着手指头算自己有几条符合优秀数据分析师的素质和能力。

Mr.林继续说道：学习数据分析需要时间和经验的积累，这些素质与能力不是说有就有的，需要慢慢培养形成，不能一蹴而就。在工作中运用不同的分析方法对数据进行分析，并与业务部门的同事积极沟通，加深自己对整个行业或研究内容的理解，相信在两到三年内，你就能成为一名合格的数据分析师。

>> 谁说菜鸟不会数据分析（入门篇）

小白：好的，我会努力学习的。

1.5 几个常用指标和术语

Mr.林：一个优秀的分析师要有扎实的数据解读功底，因为在进行分析时，经常会遇到一些分析指标或术语，对这些指标或术语的理解不够充分就没法开展工作。还有点时间，我先给你简单介绍一些基础的分析指标和术语。

小白：太好了！拿支笔先~~

平均数

Mr.林：我们在日常生活中提到的平均数，一般是指算术平均数，就是一组数据的算术平均值，即全部数据累加后除以数据个数。算术平均数是非常重要的基础性指标，它的特点是将总体内各单位数量差异抽象化，代表总体的一般水平，掩盖了总体内各单位的差异。

例如现有某学期学生的数学考试成绩，通过计算成绩平均数，可得如图1-9所示的结果。将每位同学的数学成绩与平均数相比较，就能发现哪些同学的数学成绩高于平均数，需要保持；哪些同学的成绩低于平均数，需要继续努力。

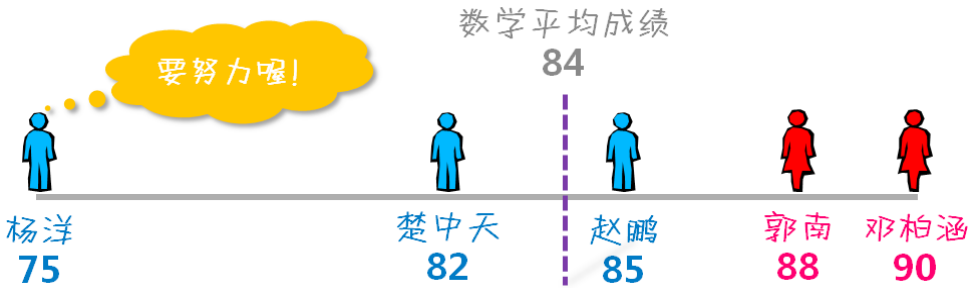


图1-9 数学平均成绩

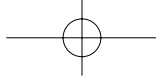
当然，在平均数这个指标中，除算术平均数以外，还有其他平均数，如调和平均数和几何平均数。

小白：一个小小的平均数都有这么多学问呀。

绝对数与相对数

Mr.林：绝对数是反映客观现象总体在一定时间、地点条件下的总规模、总水平的综合性指标，也是数据分析中常用的指标，如GDP、总人口等。此外，绝对数也可以表现为在一定时间、地点条件下数量增减变化的绝对数，比如A国人口比B国人口多1000万人。

32 相对数是指由两个有联系的指标对比计算而得到的数值，用以反映客观现象之间数量联系程



第1章 数据分析那些事儿

度的综合指标。计算相对数的基本公式是：

$$\text{相对数} = \frac{\text{比较数值（比数）}}{\text{基础数值（基数）}}$$

分母是用做对比标准的指标数值，简称基数；分子是用做与基数对比的指标数值，简称比数。相对数一般以倍数、成数、百分数等表示，它反映了客观现象之间数量联系的程度。

使用相对数时需要注意指标的可比性，同时要与总量指标（绝对数）结合使用。关于绝对数与相对数，我举几个例子，你一看就明白了，如图1-10所示。

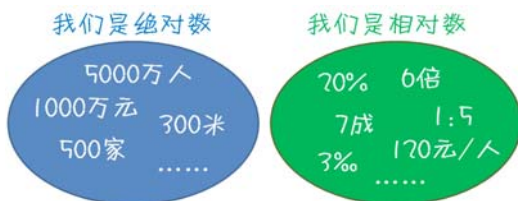


图1-10 绝对数与相对数

◎ 百分比与百分点

Mr.林：百分比是相对数中的一种，它表示一个数是另一个数的百分之几，也称百分率或百分数。百分比通常采用百分号（%）来表示，如8%，50%，168%等。由于百分比的分母都是100，也就是都以1%作为度量单位（如图1-11所示），因此便于比较，在数据分析中的应用非常广泛。

$$\frac{10}{100} = 10\%$$

图1-11 百分比的表示形式

百分点是指不同时期以百分数的形式表示的相对指标的变动幅度，1个百分点=1%。例如，图1-12中某公司发言人就混淆了百分比与百分点的概念。表示构成的变动幅度不宜用百分数，而应该用百分点。因此，这位公司发言人正确的说法应该是“公司今年的利润是45%，比去年的28%提高了17个百分点”。



图1-12 百分比与百分点



>> 谁说菜鸟不会数据分析（入门篇）

◎ 频数与频率

Mr.林：频数是指一组数据中个别数据重复出现的次数。如图1-13的左图所示，某校A班共50名学生，按性别进行分组，分为男与女两个组别，男同学的频数为30，女同学的频数为20。

频率是每组类别次数与总次数的比值，它代表某类别在总体中出现的频繁程度，一般采用百分数表示，所有组的频率加总等于100%。还是用某校A班的例子，30个男同学在50个同学中出现的频率为60%，即 $(30 \div 50) \times 100\%$ ；而20个女同学在50个同学中出现的频率为40%，即 $(20 \div 50) \times 100\%$ ，如图1-13的右图所示。

小白：所以频数是绝对数，频率是相对数。

Mr.林赞许地说：对，活学活用，你领悟得很快嘛。

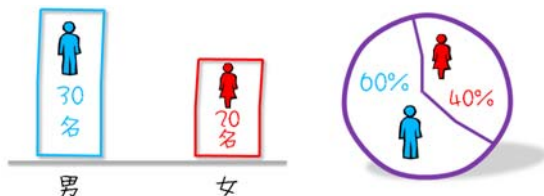


图1-13 频数与频率

◎ 比例与比率

Mr.林：比例与比率都属于相对数。

比例是指在总体中各部分的数值占全部数值的比重，通常反映总体的构成和结构。比如A班共有学生50人，男生30人，女生20人，则男生的比例是30：50，女生的比例是20：50。由此可以看出，比例的基数（也就是分母）都是全体学生人数，即为同一个基数。

比率是指不同类别数值的对比，它反映的不是部分与整体之间的关系，而是一个整体中各部分之间的关系。比如刚才的例子，男生30人，女生20人，则男生与女生的比率是30：20，如图1-14所示。这一指标经常会用在社会经济领域，比如我国的人口性别比就是用每100名女性数量相对的男性数量来表示的。

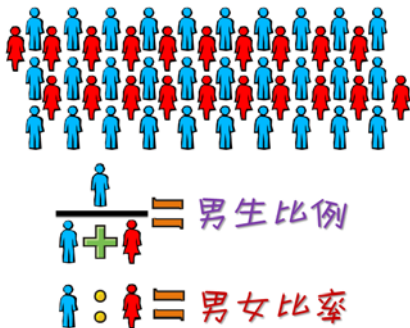
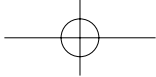


图1-14 比例与比率



第1章 数据分析那些事儿

◎ 倍数与番数

Mr.林：倍数与番数同样属于相对数，但使用时容易混淆。倍数是一个数除以另一个数所得的商。比如 $A \div B = C$ ，就是说A是B的C倍。需要注意的是，倍数一般表示数量的增长或上升幅度，而不适用于表示数量的减少或下降。

番数是指原来数量的2的N次方倍。比如翻一番为原来数量的2倍（ 2^1 ），翻两番为4倍（ 2^2 ）。如图1-15所示，这位公司发言人在讲话中就混淆了倍数与番数的概念。正确的说法应该是“公司产品销量翻一番（ $6.4 = 3.2 \times 2^1$ ），从去年的3.2万件提高到今年的6.4万件。此外，我们成本控制也很好，由20万元下降了一倍，今年成本为10万元”。

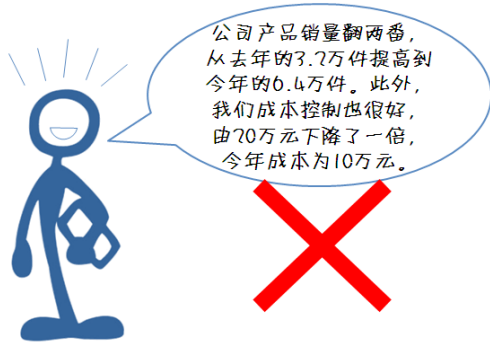


图1-15 倍数与番数

◎ 同比与环比

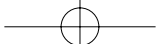
Mr.林：同比是指与历史同时期进行比较得到的数值，该指标主要反映的是事物发展的相对情况。例如2010年12月与2009年12月相比，如图1-16的左图所示。

环比是指与前一个统计期进行比较得到的数值，该指标主要反映的是事物逐期发展的情况。例如2010年12月与2010年11月相比，如图1-16的右图所示。



图1-16 同比与环比

Mr.林：讲了这么多，我估计你都听晕了。没关系，这些内容现在听起来虽然枯燥，但能够为你以后做数据分析打下坚实的基础。其中一些内容需要通过实际操作才能有更深刻的体会。



>> 谁说菜鸟不会数据分析（入门篇）

小白：是啊。我明白您讲的这些都是基本功，回去后我一定认真复习、加深理解，绝不辜负您的教导。

1.6 本章小结

Mr.林：今天讲了不少内容，估计你有点应接不暇了，我带着你回顾下这些内容。

- ★ 什么是数据分析以及数据分析的三大作用。
- ★ 数据分析的六步曲：首先是明确分析目的和思路，然后是数据收集，第三步是将收集回来的数据进行处理，第四步是根据分析目的和思路进行数据分析，第五步是将分析出的结果通过图表的方式展现出来，最后一步是撰写数据分析报告。
- ★ 认识常见的数据分析误区，明白数据分析师的要求与基本素质，了解如何成为一个合格的数据分析师。
- ★ 数据分析中经常用到的指标和术语，除掌握最基本的平均数、百分比等概念以外，还需要将容易混淆的指标进行明确区分，避免在以后阐述分析结果时出现错误。

最后送数据分析心法给你，也叫做数据分析三字经。

- ① 学习：先了解，后深入；先记录，后记忆；先理论，后实践；先模仿，后创新；
- ② 方法：先思路，后方法；先框架，后细化；先方法，后工具；先思考，后动手；
- ③ 分析：先业务，后数据；先假设，后验证；先总体，后局部；先总结，后建议。

当然你现在看这个还不会有太深刻的理解，没关系，正如三字经里所说，“先了解，后深入；先记录，后记忆”，等到积累一定经验，再回头看看你就会深有感触。

小白：谢谢Mr.林的指点，您说的我都记下了，我会加倍努力，早日出师。

数据分析技能提升十大建议

小蚊子

<http://blog.sina.com.cn/xiaowenzi22>



blog.sina.com.cn/xiaowenzi22

一、熟悉行业、公司业务

熟悉行业、公司业务及流程，若脱离行业认知和公司业务背景，分析的结果只会是脱了线的风筝，没有太大的实用价值；



<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22

二、明确分析目的

常常有朋友问我这些数据可以做什么分析？这是典型的为了分析而分析。数据分析的前提需要先明确分析目的，这样分析才有意义；



<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22

三、运用营销、管理等理论

营销、管理等理论是数据分析的指导思想，使分析思路系统化。例如4P理论等，从哪几个维度去分析？考虑哪几个方面？故而使数据分析变得有血有肉有脉络，真正做到理论指导实践；



<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22

四、掌握有效数据分析方法

了解数据分析流程，掌握数据分析基本原理与方法，并灵活运用到实践工作中，不论简单还是复杂的分析方法，只要能解决问题的方法就是好方法；



<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22

五、玩转数据分析工具

数据分析工具，建议先玩转EXCEL数据透视表，有兴趣、时间、需要的话，再学习SPSS、SAS等统计分析工具，同样，只要能解决问题的工具就是好工具；



<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22

六、学会用图表说话，玩转PPT等工具

学会如何用图表有效展现分析结果，PPT有助于数据分析结果展现达人必备；水晶易表亦对分析结果的展现有很大帮助，选择性使用思维导图可帮助理清分析思路，根据需要选用；



<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22

七、勤思考、多动手、多总结

- 需要经常发问为什么是这样的？为什么不是那样的？**只有这样才有突破点；**
- 光靠脑袋想是不够的，需要多动手实践，不要怕错，大不了错了重来，**数据分析就是一个不断假设、验证的过程；**
- 不断总结分析方法、分析思路、分析流程，在总结中前行；



<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22

八、关注行业动态—1

- 关注数据分析行业动态，如可关注业内人士博客，如沈浩老师的博客、小蚊子乐园、数据挖掘与数据分析、数据化管理等；
- 开新浪微博，加入数据分析微博圈 <http://q.weibo.com/233510>
- 关注数据分析业内人士微博（见下页），学习他人经验分享，也可分享你的经验心得；



<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22

八、关注行业动态—2



数据分析

@沈浩老师 @文彤老师
@刘德寰 @车品觉
@数据化管理
@张磊IDMer
@小蚊子乐园
@数据挖掘与数据分析
@郑来铁
@庖丁的刀
@sxlion
@文心武行
@夏之朝露
@朝阳小魔

数据分析与展现

@刘万祥ExcelPro
@MrExcel
@演绎生动Excel图表
@每日图表
@图表汇
@数据知天下
@EasyTrace网站分析
@全国数据分析业协会
@SAS中文论坛
@数据挖掘研究院
@中国统计网
@网站分析星期三

PPT制作

@蝎子
@栗子面
@秋叶语录
@西岭夜雪
@曹将PPTao
@Lonely_Fish
@刘俊0730
@小田_PPT
@我爱PPT
@PPT精选
@无敌的面包
@教育技术秋记

注：以上微博排名不分先后，欢迎大家继续补充数据分析方面有价值的微博。

<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22

九、收藏几本分析秘籍

通过不同合法途径（购买、赠书等），收藏一些实用的分析工具书以便随时查阅，如：《用图表说话》、《EXCEL图表之道》、以及近期（5月底）将上市的《**谁说菜鸟不会数据分析**》；



<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22

十、锻炼身体、享受生活

在工作之余，还要不忘锻炼身体，保持身心健康，普华永道的小表妹就是我们的前车之鉴！**做数据的伤不起呀！！谨记！！！！**



<http://blog.sina.com.cn/xiaowenzi22>



blog.sina.com.cn/xiaowenzi22

数据分析不仅是个工具，而且是门艺术，优秀的数据分析师不光要懂业务、懂管理，懂分析、还要懂创意、懂设计、懂生活，所以数据分析师也是个艺术家~~~~



<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22



<http://blog.sina.com.cn/xiaowenzi22>

blog.sina.com.cn/xiaowenzi22

为方便广大网友的分享阅读，特意整理出这份数据分析和专业图表方面的资料，资源来自 www.dozan.cn/bbs，由 Michael 整理。



网店地址:

当当 (<http://t.cn/zj7uDWX>)

卓越 (<http://t.cn/zjUESJy>)

京东 (<http://t.cn/zjzyQtF>)

China-pub (<http://t.cn/zjUEEA1>)

看完书的朋友还可以到豆瓣上写写书评:

<http://book.douban.com/subject/20373356/>

谁说菜鸟不会数据分析—配书数据和工具下载，具体包括如下 4 方面的内容，欢迎大家下载使用。

- 1、谁说菜鸟不会数据分析--数据下载
- 2、EXCEL 文本标签添加神器: JWalk Chart Tools
- 3、EXCEL 文件拆分、合并、汇总神器: 常用工具 6.5 版 (彭希仁)
- 4、屏幕动态录制神器: GGG

下载链接地址:

http://blog.sina.com.cn/s/blog_49f78a4b01017uny.html

还有菜鸟的 10 宗最，也挺有意思,写了读者与作者之间的趣事:

<http://www.dozan.cn/bbs/forum.php?mod=viewthread&tid=1&extra=page%3D1>