

● 中国商业联合会数据分析专业委员会 主办 ●



# 数据分析

China Data Analysis

洞察损益 · 量衡天下

会·员·特·刊

中国数据分析行业核心刊物

2014年  
第3期  
总第十九期

04 徐子沛——第一个写大数据的中国人

05 数据分析助力文化产业实现跨越发展

12 数据分析师的利器：R语言

16 如何分析销售数据

26 浅谈事务所发展之出击

收获

[www.chinacpda.org](http://www.chinacpda.org)

欢迎登陆中国数据分析行业网

## 卷首语

- 01 机遇留给有准备的人

## 协会动态

- 02 北京、上海两地顺利办公益沙龙活动  
03 贵州、广东地区创业指导活动圆满落幕

## 行业精英

- 04 徐子沛——第一个写大数据的中国人

## 事务所风采

- 05 数据分析助力文化产业实现跨越发展

## 浅析数据分析

- 09 中国数据量有多大？  
10 一起说说到底什么是数据分析  
12 数据分析师的利器：R语言  
14 用故事告诉你什么是数据挖掘

## 技术解析

- 16 如何分析销售数据

## 数据分析应用

- 18 大数据与商业地理分析：中国视角  
20 如何通过数据挖掘帮助母婴店扩大生意  
23 用数据分析解读你的生活  
—— 情侣之间的那些事儿  
—— 养孩子过日子的成本  
—— “一瓶水”的价格构成

## 数据分析园地

- 26 浅谈事务所发展之出击



### 2014年第3期 总第十九期

主办单位：中国商业联合会数据分析专业委员会

编辑：臧萍、贺岩、马辰、侯建

美工：崔峻珩

地址：北京市朝阳区建国门外大街建华南路17号现代柏联大厦5层

电话：010-65670421 转 115

传真：010-65671802

邮件：xiehui@chinacpda.org



中商联数据分析委员会



微信号：wxchinacpda

欢迎每日关注**微博****微信**  
精彩行业信息等着你！

## 机遇留给有准备的人

近几日，国务院出台了《关于取消和调整一批行政审批项目等事项的决定》，大家熟悉的注册税务师、注册资产评估师等已取消了执业资格，行业准入、法定业务的限制也逐渐放开。这些信息引起数据分析行业从业者们的热议，大家都在讨论，这种变化给我们行业带来的影响如何？

就行业发展的市场化问题，我在2008年写了一篇文章，就我国的数据分析业务应向法定业务发展、还是坚持市场发展的道路，阐述了自己的观点。下面是文章内容的摘录：

协会都成立了，我们是不是终于能有自己的“法定业务”了？这是个老生常谈的话题，但“回头率”高，所以还是个热门话题。我要与大家沟通三个观点：

一、中国的经济将越来越市场化，“反垄断法”出台，很多原来政府垄断的行业现在也将面临全面市场化。所谓的法定业务，总让人不自觉地与“垄断”相联系，真正的法定业务只可能是应特种行业或特种业务的需要而产生的，它是因为某些行业不能完全放开市场而形成的特殊性质的“垄断”。政府有个很明确的观点：企业的市场行为应该让市场去鉴定和评价。所以，对于有着明确的市场特征的数据分析业而言，这种特殊的垄断是不可能形成的。

二、数据分析业务中很重要的一个特点是“预测”，它要求分析师具有很强的经验性和判断性，这与注册会计师行业的已发生的、历史的数据有着本质上的区别。科学性和艺术性的高度一致是数据分析师们追求的目标，但这也意味着它无法具备“法定业务”中所要求的特征。

三、法定业务是国家相关部门赋予行业的部分特殊权力，同时一定也意味着相应的监督和限制，比如收费的限制、业务范围的限制等等，清晰的规范意味着利润空间的固定和狭窄，而数据分析业恰恰由于其工作的复杂性和分析的多样性，造成了其高利润的行业特点。所以，从某个角度上看，“法定业务”并不像想象的那么美好。

时间已过去六年，我们欣喜的看到：政府推动行业市场化的力度越来越大，越来越多的行业、企业开始了自己市场化的发展。我一直认为：无论是数据分析行业发展、还是数据分析师事务所的发展，以高水准的服务和信誉赢得客户的信任，远远比借助行政干预或所谓的“法定业务”更为可靠。

随着国家经济改革深化，特权经济、行业垄断模式会进一步打破，我相信，会有一批行业和企业由于不适应这种改变而退出市场舞台；也会有一批行业和企业，借这个难得的发展机遇，强化人才培养、积极提升业务水平、更好的服务于市场客户，形成自己的品牌和竞争优势，进而取得更大的发展空间。

在这个机遇面前，你准备好了吗？机遇永远留给有准备的人！

中国商业联合会数据分析专业委员会



# 北京、上海两地 顺利举办公益沙龙活动

◎ 文 / 市场处 侯建 图 / 崔峻珩

2014年7月，北京、上海二地再次顺利组织了数据分析行业的公益沙龙活动。由2012年开始，沙龙活动的定期开展已经成为了我会的一大特色，全国各地根据当地的数据分析人才需求定期开展数据分析的沙龙活动，旨在为数据分析师提供更新、更深入的专业知识，并解决大家在工作中遇到的有关数据分析的问题。每次沙龙活动协会都会精选一类话题为沙龙的主要交流内容。沙龙活动一般分为主题宣讲和实战交流二个步骤，既从理论上奠定基础，又从实践中交流经验。同时，与会的数据分析师及对数据分析行业感兴趣的人员可以畅所欲言，进行思想碰撞，不仅可以更多的学习到多方面的知识和经验、同时可以分享自己的心得和经验，更可以增进参加活动人员之间的感情，促进行业内人员向专业化发展。

## 北京公益沙龙活动：

7月12日，中国商业联合会数据分析专业委员会在北京中国科学院国家科学图书馆内成功举办了数据分析公益沙龙。此次沙龙活动主题为电子商务运营数据分析，特别邀请了首都经济贸易大学数量经济学专业博士、注册会计师非执业会员、中级审计师、中国证券业协会会员李玉曼老师为大家分享数据分析在电子商务运营中的应用。

商家在收集和分析数据后，可以了解客户的购物模式及消费体验，从而改进产品设计，调整电子商务策略，可以说数据分析对于电商的运营来说是至关重要的。本次的沙龙活动就是针对从事电子商务人群所制定的。

李玉曼老师长期从事证券投资、国际金融、计量经济学等相关课程的教学及研究工作，主持并参与多个上市公司微观数据分析及行业预测项目，在实践中积累了丰富的数据统计及分析经验。而且李老师精通Excel、Spss、Amos、Eviews、matlab等专业化统计和计量分析软件，在本次沙龙的实际案例讲解过程中，同时融入了软件的操作技巧的演示和讲解。

李老师在宣讲时由预测数据分析、营销决策等知识带入，然后分享了其在实际工作中所应用的理论知识和总结的经验，最后，李老师现场操作了应用Excel做数据架构对函数嵌套的应用，这种实际问题解决方法的演示获得了大家的共



鸣，深入的交流和探讨在此次沙龙过程中引起了一次又一次的互动高潮。

## 上海公益沙龙活动：

7月26日下午，项目数据分析师（CPDA）上海授权中心在沪成功举办了以“空间大数据与微位置营销”为主题的数据分析沙龙活动。参会人员来自于空间大数据应用的从业人士及数据分析爱好者们。本次沙龙活动由徐连明博士为大家分享，徐连明博士是北京智慧图科技有限责任公司联合创始人、首席技术官，主管智慧图全国机场和商场行业LBS业务，从事无线定位研究工作多年，拥有多项知识产权，对无线定位、LBS服务、智慧商场营销等领域的市场发展状况和趋势等热点领域有广泛的研究。



徐博士通过自己多年的创业和定位研究经验，给大家分享了大数据时代下，他对智能定位引擎、时空大数据挖掘和时空商业分析的应用中多年的实践经验与案例，并分享对时空大数据商业应用的美好蓝图与商业计划。让与会人员更加深入的体会了在移动互联时代时空数据分析所能带来的新的购物生活改变。

FIN

# 贵州、广东地区 创业指导活动圆满落幕

◎ 文 / 市场处 侯建 图 / 崔峻珩

## 贵州省项目数据分析师创业指导活动

随着贵州省政府出台了《关于加快大数据产业发展应用若干政策的意见》等行业利好性政策，在数据分析产业发展规划中，引进大数据领军人才、引进和培养高端人才成为重要议题。大数据时代已经被引爆的今天，众多发达国家已经在多年前就普遍运用数据，而中国的数据分析行业还拥有非常广阔的发展空间。每天中国都会产生众多的数据需要分析，因此数据分析行业的人才培养问题、项目数据分析师事务所未来发展方向问题都成为了目前最为重要的议题。政策的利好，优质的就业前景，涌现出一大批想要进入数据分析行业的人才。



2014年7月2日，协会在贵阳成功举办了“2014年项目数据分析师创业指导活动”，就各地区行业的发展及事务所组建、业务开展等方面内容与大家进行了交流。贵阳日报、贵州电视台均针对此次活动做了专题报道，并在活动结束后对协会邹会长就此次活动及数据分析行业现状和发展趋势进行了专题采访。

采访中邹会长表示：“目前大数据已经在很多领域开始发挥重要作用，人们通过数据分析可以获得很多有用的信息，例如消费者在网上购物时弹出来的推荐物品都是商家通过数据分析得出的一些预见性的判断。在一些敏感性行业如互联网、电子商务、零售、医疗、金融等行业的大数据特征日趋明显，数据分析行业被越来越多的企业和政府加以关注和重视。”为了让有意成立事务所的个人少走弯路，协会会员处负责人特地讲解了事务所会员入会流程、筹备事务所需要做的前期准备等问题。

## 广东省项目数据分析师创业指导活动

2014年8月30日，广州举行了数据分析师从业交流活动。参与此次活动的数据分析师分别来自金融、投行、it、电商、互联网、通信、食品、家具等行业以及行政事业单位。广州市数据分析师俱乐部也在会上宣布成立，其中广州市铭数教育咨询有限公司董事长林浩钊担任俱乐部名誉主席。



“数据分析人才体量小、人才市场需求巨大，成为企业招聘遇到的显著问题。”中国商业联合会数据分析专业委员会副秘书长蒋文伟表示，近两年来，受“大数据”概念的影响，企事业单位和个人都越来越重视数据分析的重要性，但人才难觅成了发展的瓶颈。数据显示：目前广东数据分析持证人数约为500人，而该省数据分析人才缺口还很大。所以进入数据分析行业的人才在未来都会是向成功的方向发展的。蒋副秘书长特意为想要创业的参会人员详细介绍和指导了筹备事务所的流程等问题。

协会本着对每一个进入行业的人负责的态度而开展的创业指导活动，旨在帮助广大希望从事数据分析行业人员更清晰的了解行业，了解市场，明晰发展方向，避免从业人员盲目的“跟风”，协会衷心希望每一个进入行业的人在行业内都可以有所发展，一展所长，开创一片属于自己的数据分析天地。

FIN

## 关于协会会员处工作调整的通知

由于中国商业联合会数据分析专业委员会内部工作调整，现会员处工作由贺岩同志负责。特此通知！欢迎各位会员随时与会员处保持联系。

网址：[www.chinacpda.org](http://www.chinacpda.org)

电话：010-65670421转115 传真：010-65671802

QQ：2530404542

# 涂子沛 —— 第一个写大数据的中国人

◎ 文 / 市场处 侯建 图 / 崔峻珩

涂子沛，江西吉安人，生于1973年。本科毕业于华中科技大学计算机系，研究生分别就读于中山大学和卡内基梅隆大学，现居美国硅谷。2012年其著作《大数据：正在到来的数据革命》在中国社会开启大数据之先河。

2012年，涂子沛出版《大数据：正在到来的数据革命》时，国内才刚刚兴起对大数据的讨论。而这个概念，他已经听了将近20年。

1992年，涂子沛考入华中科技大学计算机专业，从此与数据结缘。毕业后，他进入武警边防部队，设计了“中国第一套反偷渡系统”，据说这套系统前几年还在用。之后的10年，涂子沛做过缉私艇指挥官、外经贸局的公务员。2006年，他辞去公职赴美留学，就读于卡内基梅隆大学，获信息技术硕士学位。

本科时，涂子沛就知道大数据浪潮。但那时候，因为网络还不够普及，数据信息对人们的生活影响甚微，这个概念也不受人关注。

2010年之后，大数据逐渐为人熟知，凭借多年的研究，涂子沛成了第一个写大数据专著的中国。2012年，他出版了《大数据：正在到来的数据革命》。有评论认为，这本书开启了中国人对大数据的认识。而他的新书《数据之巅》则从历史的角度说明，对数据的理解和掌握是人类文明进步的标志。

《数据之巅》分为两个部分。第一部分以美国历史为主线，展现了大数据文化是如何形成并深入人心的。第二部分，涂子沛详细考证了大数据浪潮的来龙去脉，并结合3次工业革命分析了大数据对商业运营、社会治理的深远影响。他认为，更大的数据爆炸时代正在到来，大数据浪潮最终将引领人类社会迈进一个新的形态——智能型社会。

涂子沛认为：大数据的特点绝对

不是量大，而是多个源头。不仅仅包括数字，还包括文字、音频、视频等。人类知识的来源就是对过去经验的记录和整理，如果说传统的数据是人类部分知识的来源，那随着人类记录范围的不断扩大，知识的边界也在扩大，现代意义上的大数据将逐渐成为人类全部知识的来源。



另外，大量的、多元的数据汇集到一起可以更好地还原事实。人类在做决策的时候首先要从事实出发，而当所有决策都以数据为依据时，我们就必须达到一种“数据最优”状态，靠精准的记录、计算和预测来推动社会发展。如果你不能达到这个状态，而别人达到了，那你就出局。

针对大数据是否没有缺点的问题，徐子沛的观点是：大数据有很多不靠谱的地方，但关键是，一个东西在你手里看你怎么去用。针对中国大数据的现状，徐子沛提出中国在历史上缺乏这种数据文化。

中国数学历史学家张奠宙先生曾讲过，古希腊的民主政治推动了数学的发展。在古希腊的政治环境下，要想证明自身观点的正确性，需要在平等的基础上，用充分的理由说服对方，反映在学术上，就是“证明”。欧几里得的名作《几何原本》就是在这样的背景下产生的。相比之下，中国历史上数学家的工作则主要是丈量田亩、兴修水利、计算税收等，是“管理数学”和“木匠数学”，缺乏抽象的数学思维。

首先要建立数据意识，用数据去管理去决策，追求精确，第二要意识到数据是种权利，要用数据说话。数据不仅仅是种工具，它代表隐私，代表知情权，整个社会文明的生

活都是围绕数据而开展的。第三要知道数据是一种资源，对企业对个人都是一种资产，一种创新的资源。谷歌的无人驾驶汽车基于大数据，阿里巴巴为什么能在短时间可以贷款是基于数据，很多都是基于数据。

“数据就是静态的历史，  
历史就是动态的数据。”  
——徐子沛

对大数据的认识和使用，我们已经起步晚了。未来在这方面，政府应该发挥主导作用，首先在公共领域推行数据治国的理念，力争在全社会形成“用数据说话、用数据管理、用数据决策、用数据创新”的文化氛围和时代特点。

FIN

## 数据分析助力文化产业实现跨越发展

◎ 文 / 中颢润（北京）项目数据分析师事务所 万宾博士 图 / 崔峻珩

随着文化产业的蓬勃发展以及国际先进模式的借鉴和引入，作为现代企业管理核心要素之一的数据分析，逐渐受到国内文化行业的重视。文化产业与数据分析好比鱼和水，鱼离不开水，水为鱼提供良好的生存条件。文化产业引入数据分析具有得天独厚的优良条件，同时，数据分析的引入也将大大促进文化产业的发展。

在此大背景下，中颢润北京数据分析师事务所顺势所趋，秉承严谨科学的优良传统，积极投身于数据分析在文化产业的具体应用中，在文化企业如何借助数据分析提高量化经营管理水平方面摸索出了一条行之有效的道路，取得了良好的效果。

### 大数据时代下的文化产业

随着大数据时代的到来，数据分析已渗透进了生产生活的各个领域。数据分析这个强有力的工具不但大幅度提升了硬实力，同时也在慢慢地被应用于软实力的提升。文化与科技的融合已成为大数据时代文化产业发展的主题。

2011年国家“十二五”规划纲要明确指出，要进一步推进文化产业转型升级，推进文化科技创新，提高技术装备水平，改造提升传统产业，培育发展新兴文化产业。另外，规划纲要还提出以重大技术突破和重大发展需求为基础，促进新兴科技与新兴产业深度融合。在近几年国务院发布的相关文件中也多次提到，希望“新一代信息技术成为国民经济的



支柱产业”，其中包括“加快推进三网融合”、“提升网络增值服务等信息服务能力”、“大力发展数字虚拟等技术，促进文化创意产业发展”、“加快建设宽带、融合、安全的信息网络基础设施”。这标志着以数据分析带动文化产业发展的运动正如火如荼的进行开来。

#### 文化产业何以能够引入数据分析？

数据分析的本质在于通过数据更精准地挖掘用户的需求，科学合理的分析企业开展项目的未来收益及风险情况，为企业解决决策难题。而谁能掌握用户的需求，谁就可以引领其所在行业的发展。而文化产业也不例外，更进一步说，文化产业引入大数据技术有着先天优势。

首先，文化产业所拥有的数据具备较高的消费价值。文化产业属于第三产业，生产和消费一般同时进行，故文化产业所涉及的行业和产品大多和消费者直接相关，能够直接为用户所消费，其数据有着明确的直接消费价值。

其次，文化产业本身就是生产数据的行业，从中能够不

断地获得新的数据资源。根据美国的统计资料，文化传媒行业数据是仅次于政府信息数据的第二大数据来源。

第三，文化产业本身就拥有极其雄厚的客户资源。由于文化产业直接面向消费者，拥有广阔的用户基数和规模，而基于庞大的客户资源进行数据分析，则将成为文化产业未来基于大数据业务转型的关键性条件。

最后，文化产业大部分为一次性消费产品，各种文化产品零碎、分散，但整个行业又集中、综合，其消费频率之高，速度之快，足以使之敞开胸怀接纳数据分析这一强有力的武器。

#### 数据分析为文化产业开启了一扇大门

既然文化产业自身拥有引入数据分析的众多优势，那么数据分析将如何带动文化产业发展呢？下面我们将详细谈谈这个问题。

从宏观角度来看，首先，数据分析能有效解决文化产品供需脱节的矛盾，有益于解决原创文化内容水平不高的问

题。通过对各类人群的喜好数据进行分析，能够明确目标受众的品味和需求，创造出适销对路的文化产品。通过用户行为数据的挖掘和分析，确定用户的消费心理和消费习惯，为其定制复合型产品套餐，满足用户深层次的情感和人文需求，自然而然中激发用户的文化消费欲望，将使文化产业经济更好地转型为体验经济。

其次，数据分析能有效解决文化产品定价难题。通过数据分析，能够有效分析出用户的消费承受点，找到产品成本和运营收入之间的平衡点，从而在降低产品运营风险的前提下，覆盖最广的消费人群。同时，根据不同用户的消费类型和消费能力，制定更加灵活的文化产品定价策略。根据福布斯商业评论的研究结论，获得用户消费数据越多，覆盖的用户数量越广，采用差异化定价策略就能够为企业提供越高的利润。而文化企业在精准地大数据分析之后，采取降低定价、扩大覆盖、差异定价的经营策略，将有效降低经营风险、锁定产品利润、挤压竞争对手、提升市场竞争力。

第三，数据分析能有效延长文化产品的产业链，增加文化企业的赢利点。文化企业通过搜集整理消费者的情感体验数据，能够有效分析和提炼出市场的时尚和审美发展趋势，这种趋势数据将有助于工业制造企业实现“产业文化化、产品文化化”。

最后，数据分析合理优化文化产业结构，调节文化资源配置。对于大众化文化产品，首先要考虑广泛、全面，点、线、面全面结合，至于其产品，因为消费人群大众化，故采用粗放型的资源配置即可，由于这部分是消费的主体，故在数量上要首先达标。对于特定消费群体，则要考虑精细，这部分消费群体品位独特，对文化产品的质量要求较高，故数量上不要求过多，但质量一定要精，所以这部分适于采用集约型的资源配置。通过数据分析，可以清楚的知道粗放产品和集约产品的比例，以达到文化基金的合理使用。

从微观角度来看，首先，数据分析有助于企业合理市场细分，精准市场定位。例如电影《纸牌屋》就是数据分析结果的第一次成功的战略运用。用户只要登录Netflix网站，



对某一个视频的每一次点击、播放、暂停、快进、回放，看了几分钟就彻底关掉视频，或者停了一段时间又重启等一系列动作都会成为一个“事件”被记录下来并汇入后台进行分析。单独某一个人的这些行为还没有太大价值，但如果足够多的人在整段视频中的同一个地方做了相同的举动，那么对这一举动的解读将非常重要，此时数据分析就开始显露出意义了。

正是运用数据分析这一强大武器，所以《纸牌屋》也一改美剧的传统推送方式，还没有开播就一次性确定要拍两季共计26集，这在美剧界无疑是一次豪赌。但Netflix有其自己的理由：《纸牌屋》根本不需要攒口碑，因为在制作之初便掌握了观众的喜好，对观众的需求进行了“精确打击”，所以开播之后只需向标签为“喜爱凯文·史派西”、“喜爱大卫·芬奇”或“喜爱政治剧”的观众推荐一下就行了。了解了这些，也就明白为什么《纸牌屋》如此成功了。这也再次让我们见证了数据分析的力量。

其次，数据分析有助于企业定位目标观众，进行精准营销。西方国家很多剧场都会向观众发放会员证。会员除了购票优惠外还可知道有关电影戏剧方面的信息，并可参加一些专题讨论等。持会员证的条件是每月或每季至少要到该影院看一定次数的演出。由于观众的口味不同，需要针对不同观众的特点精准营销，而要精准营销，就要对上座率的影响因素进行数据分析。在国内，许多大型剧院也开始实行会员制。通过观众在办会员卡时留下的相关信息，可以对更多人口统计信息进行了分解。这些信息包括性别、学历、收入、职业、子女年龄以及生活方式等指标。对这些数据的分析可以使剧院能够了解其各类细分顾客群体所具有的不同特质，找出最有可能到庭院剧院看戏的顾客，并对其展开针对性的营销。这类数据分析工作可以使剧院避免将资金耗费在争取不可能看戏的观众上，而是集中更多资源吸引新观众观看演出，更有效地提高庭院剧院的上座率。在分析观众个人信息时再结合对地域信息的数据分析，可以了解不同地域的观众到剧院看戏的可能性。此项工作将帮助剧院确定其地区或全国巡演的最佳路线。

再次，数据分析有助于企业分析消费偏好，改善服务质量。例如，在美国，游戏产业正在充分利用大数据技术，并通过推动玩家的广泛参与，赚取更多的广告收入，进而优化游戏体验。与任何行业一样，360度全方位的用户视觉体验对于游戏产业来说是至关重要重要的。当玩家们玩一款游戏时，他们留下大量的数据，方便游戏公司进行跟踪。无论该游戏是通过Facebook连接的一款在线社交游戏，还是用PlayStation/Xbox玩的离线游戏，当玩家开始玩游戏时，都会

创建大量不同格式的数据信息。而他们在游戏过程中所进行的一切操作都会创造出的海量数据流。包括玩家之间是如何相互配合的、他们玩多久、与谁在什么时间玩了游戏、玩家在虚拟产品上的花费如何，游戏过程在与谁聊天了等等。如果游戏配置文件链接到社交网络或一个玩家被邀请参与数据统计，那么，关于该玩家在现实生活中喜欢什么样的游戏的信息可以更加丰富，而游戏公司则可以适当的根据玩家的数据档案信息优化游戏。

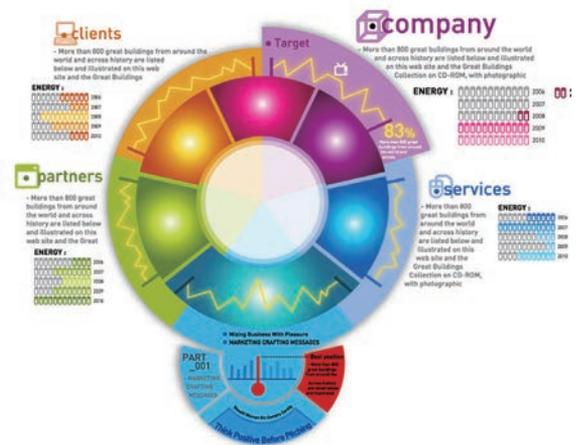
最后，数据分析有助于企业预警消费流失，提高消费粘性等，这些较浅显易懂，这里不详加赘述。例如2010年对中国动画电影市场的数据分析得出，动漫产品从电视转到电影，讲故事的方法不一样，电影要观众掏钱来看，故事卖点肯定是一大看点，《虹猫蓝兔》是当时上映的第一部国产动漫电影，可能在故事呈现还有些稚嫩，故事还不够精彩，关键是想象力还未足够丰富，这好像是今天中国所有动画片的通病，以致流失了许多消费个体。通过这次的经验，未来的《虹猫2》故事能在保留原有消费群体的同时，更吸纳一些边缘消化个体。之后上映的《喜羊羊与灰太狼之牛气冲天》也据此得到了良好的票房收入。



### 前景展望

大数据对于文化产业的发展，助益颇多。但万事都存在风险，还需要有清醒的认识。

文化企业对于大数据理解的核心在于对“数”的理解，是要获得满足业务需求的信息—“数”，从而得到企业决策的先发优势，理顺业务优化的逻辑—“据”，突破困扰企业发展的核心问题。因此，搜集高质量的数据，获得更多类型的数据，扩大数据的来源渠道和形态，往往比获得巨量的单



一数据更有意义。

面对大数据时代的到来，文化企业应该以更高的姿态迎接市场挑战，赢得市场先机。大数据的实施和应用，需要多个角色的参与。互利共赢是企业非常明智的选择，任何一个企业都无法保证能够适应所有的角色，必须和其他企业相互合作，共创商业价值。特别是对于文化企业，目前科技化、信息化、自动化水平还不高，更需要敞开心扉，积极投向于其他产业的大熔炉，引入行外企业，取长补短，帮助自身的业务成长。

在大数据时代，各文化企业要提高对数据的敏感度。吸纳更多专业人才，同时建立自己的数据处理中心。不但要存储大量的数据，还要对存储的信息加以分析，提高其利用率，而不能将其视为搪瓷花瓶。同时，要以宽容和开放的态度看待数据共享，本着访问就能产生价值的长远态度，做好数据的共享。

总之，数据分析是一把双刃剑，利用好会带来巨大收益，利用不好也会带来负面效果，我们要辩证的看待它。文化产业与数据分析完美的结合，还有很长的路要走，虽然道路是曲折的，但前途是光明的。随着大数据时代的到来，运用数据分析和数据挖掘实现文化产品创新和服务质量提升，将成为文化企业在信息时代的必然趋势。虽然文化产业的转型升级更需要的是体制机制的革新和转变，但是数据不会说谎，它将帮助文化产业走上市场化的快车道。中颀润事务所将继续在这一领域发挥自己的积极作用，为中国文化产业的腾飞贡献自己的一份力量。

FIN

# 中国数据量有多大？

© 转自 / 36大数据 编辑 / 臧萍 图 / 崔峻珩



从我们手头掌握的数据来看，2013年度，中国存储市场出货容量超过1个EB，2倍于2012年，相当于2009年全球的数据总量。存储总量IDC曾经发布的预测表明在未来的3-4年，中国存储总容量可能达到18个EB。从数据存储市场的需求来看，互联网、医疗健康、通信、公共安全以及军工等行业的需求是主要的，且上升态势明显。

鉴于存储和服务器的紧密相关，我们从已经获得的资料可以知道，目前全球运行的服务器总量超过5000万台，美国国内运行的服务器总体容量接近1000万台。从各种市场公开数据来看，2013年中国内地服务器销售总数接近为100万台。大体估算，截止到2013年底，中国内地整体在运行的服务器总数量在300万台以上。

从现有存储容量看，中国目前可存储数据容量大约在8EB-10EB左右，现有的可以保存下来的数据容量大约在5EB左右，且每两年左右会翻上一倍。这些被存储数据的大体分布为：媒体/互联网占据现有容量的1/3，政府部门/电信企业占据1/3，其他的金融、教育、制造、服务业各部分占据剩余1/3数据量。

公开数据显示，互联网搜索巨头百度2013年拥有数据量接近EB级别、阿里、腾讯声明自己存储的数据总量都达到了百PB以上。此外，电信、医疗、金融、公共安全、交通、气象等各个方面保存的数据量也都达到数十或者上百PB级别。

在目前被广泛引用的IDC和EMC联合发布的“2020年的数字宇宙”报告预测到2020年，全球数字宇宙将会膨胀

到40000EB，均摊每个人身上是5200GB以上，这个量将会如何被有效存储和应用，我们眼下还很难想象。然而我们看到该报告指出，从现在起到2020年，全球数字宇宙的膨胀率大约为每两年翻一番。事实上，根据上述调查结论和服务器容量调查，我们也能做出个相对合理的推断：目前，全球产生的数据量中仅有1%左右的数据能够被保存下来，也就是说今天全球能够被保存下来的数据也就是在50EB左右，而其中被标记并用于分析的数据更是不到10%。

作为全球人口和计算设备保有量的大国，我国每年所能产生的数据量也极为庞大，有数据说2014年甚至可能达到ZB级别，但是真正被有效存储下来的数据仅仅是其中极微小部分，中国保存下来数据占全球数据的比例大约在10%左右，也就是上面说的5EB。这些数据中，目前已被标记并用于分析的数据仅达到500PB左右，也是接近10%的一个比例。

随着各企业和部门对于数据资产保存和利用意识的增强，未来2-3年一定会有越来越多的行业、大企业步入到PB、百PB、甚至EB级别数据俱乐部，未来2-3年中国的数据总量也将呈翻倍上升态势，我们预测2015年中国就可能突破10EB数据保有量，被标签和分析利用数据量也将上升到EB级别。

FIN



# 一起说说到底什么是数据分析

◎ 摘自 / 数据分析精选 编辑 / 贺岩 图 / 崔峻琦



之前在微博上发起一个话题，让大家用一句话向外行说说什么是数据分析？有众多同学参与讨论和转发，其中一些有意思的、不错的说法，大家一起来看看。

## 最专业的数据分析：

@项目数据分析师：有针对性的收集、加工、整理数据，并采用统计和挖掘技术分析和解释数据的科学与艺术！

## 最简洁的数据分析：

@小苹果：简单的很，就是分析数据。

@咔嚓：描述数据特征，预测数据趋势，展示分析结果。

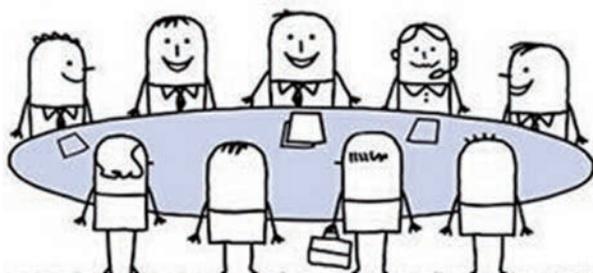
@默默：从一大堆数据中提取到你想要的信息，就是数据分析。

## 最浪漫的数据分析：

@ CPDA：数据分析就是：茫茫人海中，你通过观察、

跟踪、记录等手段找到你生命中最爱的那个人，进而根据对方的喜好，成功的展示了自己的优点，改进了自己的缺点，并且说服了对方的父母，承诺终身与ta为伴的一项工作。

@数据化管理：你喜欢上一个姑娘，你会搜集她的兴趣、爱好、星座、闺蜜、乃至三围... ..等等各种信息吧，然后



想自己怎么能搭讪上、约出来，碰壁了会继续找原因、想办法，这里面你都有在做数据分析呀。

@cnn: 从你每天发的微博研究你喜欢哪个明星、是哪个星座的、喜欢吃什么买什么、大概几岁会嫁出去。

#### 最IT的数据分析:

@穆浩: 所以我一般都说我是做IT的。

@Yico: 复制粘贴。

@仓鼠: 你应该对男的低调说自己是码农，对女的高调说自己是分析师！现在甭管是什么职业，加个师字就显得特牛B，比如策划师、揉奶师、面包师、搬砖师等等。

#### 最实用的数据分析:

@ETwise: 我告诉别人：你买了牙膏我还会推荐你买牙刷，我就是做这样的工作的。

@老仲琢磨: 可以知道刚当爹的男人买尿布时最可能顺便买点啥。

@茄子把: 每个月的月底，老婆问你，这个月的钱你都是怎么花的？你的回答就是数据分析。

#### 最神化的数据分析:

@张启仁: 数据分析是算命的。

#### 最霸气的数据分析:

@淘货郎: 数据分析就是企业的军统和中统。

@牧羊人: 这是一个很有前途特牛B的职业，具体细节你没必要知道。

#### 最恐怖的数据分析:

@游走的背包: 数据分析就是文革时的查人祖上三代，最后定义你这代什么成份！

@宝泥马冷死了: 从你爸的日常行为统计中推测你下次考试不及格他会用什么东西揍你。

#### 那到底什么是数据分析呢？

说说数据哥的理解：数据分析是基于商业目的，有目的的进行收集、整理、加工和分析数据，提炼有价值信息的一个过程。

实际上，数据分析过程的主要活动由识别信息需求、收集数据、分析数据、评价并改进数据分析的有效性组成。

#### 一、识别信息需求

识别信息需求是确保数据分析过程有效性的首要条件，可以为收集数据、分析数据提供清晰的目标。识别信息需求是管理者的职责管理者应根据决策和过程控制的需求，提出对信息的需求。就过程控制而言，管理者应识别需求要利用那些信息支持评审过程输入、过程输出、资源配置的合理性、过程活动的优化方案和过程异常变异的发现。

## 二、收集数据

有目的的收集数据，是确保数据分析过程有效的基础。组织需要对收集数据的内容、渠道、方法进行策划。策划时应考虑：

①将识别的需求转化为具体的要求，如评价供方时，需要收集的数据可能包括其过程能力、测量系统不确定度等相关数据；

②明确由谁在何时何处，通过何种渠道和方法收集数据；

③记录表应便于使用；

④采取有效措施，防止数据丢失和虚假数据对系统的干扰。

## 三、分析数据

分析数据是将收集的数据通过加工、整理和分析、使其转化为信息，通常方法有：

老七种工具，即排列图、因果图、分层法、调查表、散点图、直方图、控制图；

新七种工具，即关联图、系统图、矩阵图、KJ法、计划评审技术、PDPC法、矩阵数据图；

## 四、数据分析过程的改进

数据分析是质量管理体系的基础。组织的管理者应在适当时，通过对以下问题的分析，评估其有效性：

①提供决策的信息是否充分、可信，是否存在因信息不足、失准、滞后而导致决策失误的问题；

②信息对持续改进质量管理体系、过程、产品所发挥的作用是否与期望值一致，是否在产品实现过程中有效运用数据分析；

③收集数据的目的是否明确，收集的数据是否真实和充分，信息渠道是否畅通；

④数据分析方法是否合理，是否将风险控制在可接受的范围；

⑤数据分析所需资源是否得到保障。

你是怎么理解数据分析的呢，说说看吧。

FIN

# 数据分析师的利器：R语言

© 文 / 数据中心 何琳 编辑 / 臧萍 图 / 崔峻珩

## 工欲善其事，必先利其器!

近年来，随着大数据浪潮的到来，数据科学快速发展，数据分析师要处理的数据问题越来越复杂，传统的数据分析工具越来越显得力不从心。

R语言作为国外数据分析师常用工具在国外的数据分析业务中有广泛的应用领域，同时几乎也是国外高校统计专业的必修课，在学术届更是有着独一无二的权威性和应用领域。R语言的处理数据和数据可视化的强大能力，吸引了越来越多的数据分析师投入到R语言的阵营。

横向对比各种数据分析工具，结合数据分析师需要面对的日益复杂的应用场景，数据分析师要提升专业能力和处理数据的能力，选择称手的工具，R语言是一个非常好的选择。

### 一、R语言是数据分析领域的通用语言

R语言是统计学家发明的工具，早期主要是学术界的统计学家在用，他们将其用在各种不同的领域，包括统计分析，生物信息学，应用数学，计量经济，金融分析，财经分析，人文科学，数据挖掘，人工智能，生物制药，全球地理科学，数据可视化等等。

近年来，随着大数据时代的到来，其他领域的数据分析人员，如互联网数据分析师、IT工程人员、广告数据分析人员等开始认识到R语言的强大能力，越来越多的其他领域的专业人员加入到R语言的使用者队伍。

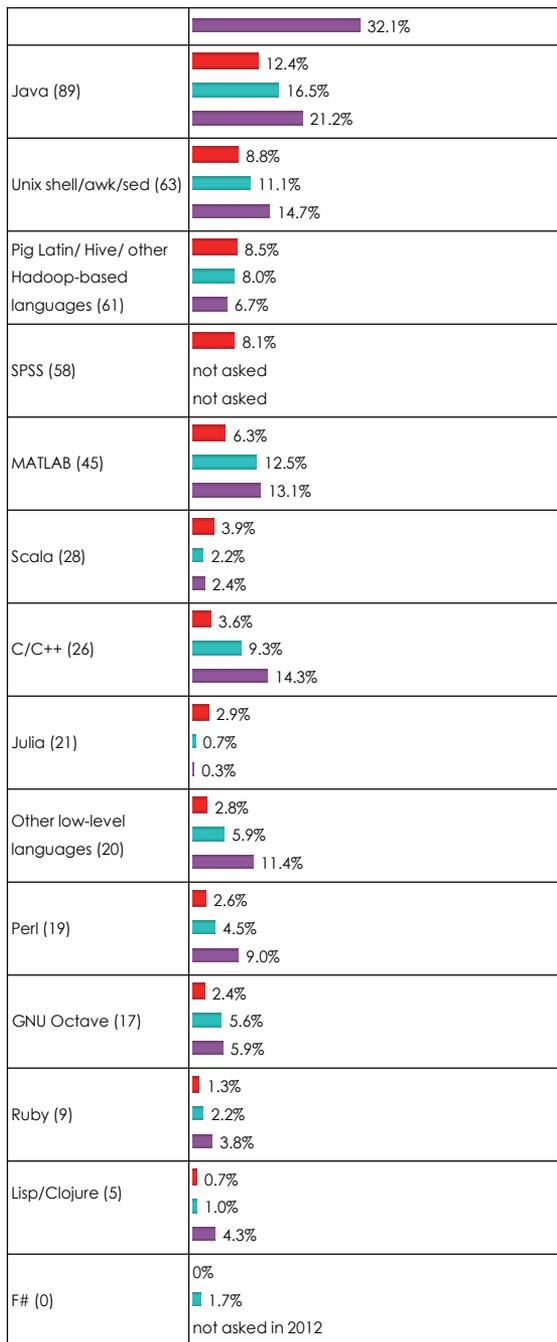
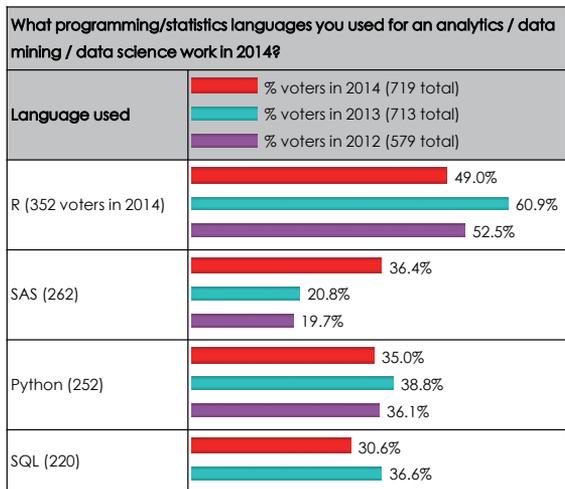


图 - Kdnuggets对数据分析/数据挖掘/数据科学领域编程语言的调查

根据.kdnuggets的调查显示,在2012、2013、2014连续三年,R语言都是数据分析/数据挖掘/数据科学领域排名第一的主流语言和工具。

Revolution Analytics的社区总监和新开源解决方案组的领头人指出:“R语言几乎成为了数据科学领域的通用语言。”

网络上更是流传数据分析师鄙视链的戏谑说法:R > SAS > stata > SPSS > EXCEL,从另外一个方面说明了掌握的工具与所处的行业地位的之间关系,掌握了R语言的数据分析师往往具有更强的数据分析能力和更好的行业地位。

## 二、R语言不仅仅是一门编程语言

### (1) R语言是一门编程语言。

维基百科中对R语言的定义:一种自由软件编程语言与操作环境,主要用于统计分析、绘图、数据挖掘。

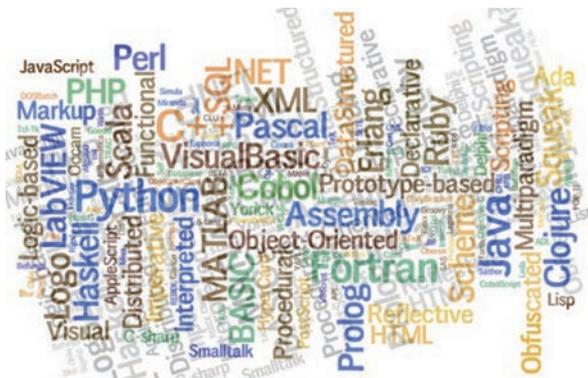
既然R语言是一种编程语言,我们对比其他编程语言,将R语言作为一种编程语言的特性进行对比分析如下:

种类	高级语言	面向对象	解释型	编译型	交互型	动态类型
R语言	✓	✓	✓		✓	✓
Java	✓	✓		✓		
C++	✓	✓		✓		
C	✓			✓		
Fortran	✓			✓		

表 - R/Java/C++/C/Fortran编程语言特性对比

(画✓表示这种编程语言具备的特性,空白表示不具备这种特性)

在这个表格中,我们能看出,R语言作为一种面向对象的高级编程语言,与Java、C++、C、Fortran这些编程语言相比,具有解释型、交互型和动态类型的特点。而后面这3个特点,特别适合数据分析师进行数据分析的业务场景,变量无需事先定义,直接拿来就用,每一步分析都能与系统交互,看到处理的结果,能快速排除错误和逐步深入数据分析业务本身,将重点放在分析解决问题上,将数据分析师从繁杂的



编程中抽离出来。面向问题而不是强调编程的特点,大大便利了数据分析师的分析业务。

此外,R语言以向量为基本运算对象,这不仅能有效降低代码的冗余度,也显著提升了代码的运算效率。简单的几行代码,不仅能实现其他语言一大段代码的功能,而且运算速度也很快。

(2) R语言不仅仅是一门编程语言,R语言更是一门专业的统计计算语言。



R语言是统计学家发明的语言,天生具有统计的基因。不需要很长的代码,不需要过多的思考程序设计,不需要过多的思考算法的效率,几个函数调用,几个参数传递,就能实现复杂的统计模型。这就意味着,数据分析人员将可以把主要的精力放在统计分析上,而不是放在编程上。

另外,R语言有强大的社区支持,累计到2014年9月,在R的综合档案网络(CRAN)中,有5859个通过严格审核的R包,其内容涵盖了从基本统计计算到机器学习,从金融分析到生物信息,从社会网络分析到自然语言处理,从各种数据库各种语言接口到高性能计算模型,可以说无所不包,无所不容,这些都是可以自由免费使用的。Google首席经济学家Hal Varian说:“R的最让人惊艳之处在于你可以通过修改它来做所有的事情,而你已经拥有大量可用的工具包,这无疑让你站在巨人的肩膀上工作。”

(3) R语言是开源的,免费的,具备跨平台的能力,对各种格式的数据文件有良好的支持,并且提供了丰富的数据接口。这些特点都让R语言有更宽广的使用领域和发展前景。

FIN



## 用故事告诉你什么是数据挖掘

◎ 编辑 / 臧萍 图 / 崔峻珩

现如今，到处都在提大数据、云计算、数据挖掘、数据分析等概念，那么这些听起来很深奥的概念到底是什么呢，今天就让我们用一个小故事来告诉你，什么是数据挖掘与机器学习吧？

某天，嘴馋的你想吃芒果了，于是你走到水果摊，挑了几个让老板过过秤，然后你再根据芒果的斤两付钱走人。

显然，买芒果你当然是挑着最甜、最熟的来买（因为你是根据重量而不是质量来掏钱的）。怎么个挑法才靠谱呢？

对了，你奶奶说过，金黄色的要比浅黄的更甜些。于是你就做了一个简单的规定：只挑金黄色的买，过磅、付钱、

回家。就这么简单？

不尽然。生活没那么简单

拎着芒果回到了家，尝了尝，你发现有些很对口味，有些则马马虎虎。显然，光凭你奶奶的智慧还是有所不足的。挑芒果不能只看颜色。

一番细品，再一番苦想之后，你得出了结论：那些大个儿的、金黄色的芒果一定是甜的，而那些小个儿的、金黄的就只有一半是甜的了。

比如，你买了 100 个黄灿灿的芒果，50 个个头较大，50 个个头一般，那么个头大的那 50 个就都是甜的，个头小

的 50 个当中就平均只有 25 个是甜的了。

你对自己的研究发现甚感满意，想着下次买芒果的时候能更靠谱些。终于你又来到了水果市场，发现你最常去的那个水果摊已经见不着踪影了。于是你就找了另一个摊档，而他们的芒果又产自不同的地方。你挠挠头，叹了口气：得了，“大个、金黄色的果子甜”的经验在这又不适用了，一切从头开始。尝了一遍后，你发现那些小个的、浅黄的芒果是当中最甜的。



你的表妹大老远的来看你，你决定给她整些美味的芒果尝尝。但她并不在乎果子甜不甜，而只关心是不是鲜嫩多汁。好吧，你又挠挠头，在果摊尝了一遍，发现越软的越多汁。

这次，你到了另一个国度。这儿的芒果味道与家乡的截然不同：绿色的竟然要比黄色的味道更棒。

你结了婚，发现她对芒果完全不感冒，而对苹果很是喜欢。这回，你又跑去买苹果去了。你又得学神农“遍尝百果”，摸索苹果的物理特征和味道之间都有怎样的联系——为了她，又有什么不可以的呢？满满的都是爱呀。

写段电脑程序来帮忙

为什么不用电脑程序来帮你挑芒果（或苹果）呢？你把规则写了出来：

if (color is bright yellow and size is big and sold by favorite vendor): mango is sweet.



if (soft): mango is juicy.

etc.

这就是你挑芒果用到的规则了。把它发给你的弟弟，想必他也能帮你买到对你口味的芒果。

但问题在于，每回你在试验之后得出了观察结果，你都把规则做一番修正。你还得明察秋毫地了解清楚都有哪些因素在影响芒果的品质。如果问题复杂起来的话，你在上面耗费的心血甚至都有可能为你拿个“芒果科学”的 PhD 下来（如果有的话）。但这样的大闲人还真的不多。

写个机器学习算法吧

机器学习算法是普通算法的进化版，它们让你的程序变得“更聪明”，能从你提供的数据里自动学到东西。

你在市场上随机选择了某个品种的芒果（training data），把每个芒果的物理特征都写进了一个表格——颜色、大小、形状、产地、所属果摊等（features），甜度、多汁程度、成熟度（output variables）也做了记录。你把这些数据都放进了一个机器学习算法（classification/regression），然后这个算法就会自动从芒果的物理特征和品质之间得出一个相关性模型。



等到下一回你到市场的时候，你把在售的芒果的特征信息都收集起来，再扔进你的机器学习算法，它就会利用之前计算出来的模型来预测哪些芒果是甜的、熟的、以及/或多汁的了。该算法可能会使用和你曾经手写的差不多的规则，也有可能使用的规则会更有相关性，不管怎样，在很大程度上你都不用操太多心了。

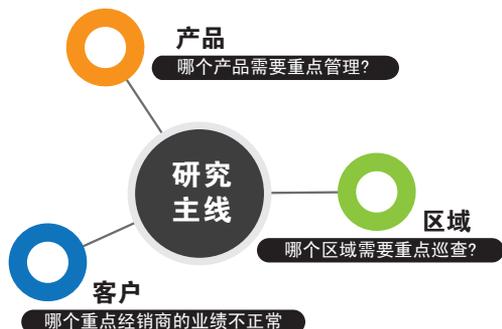
现在，算法在手，你就可以挺直身板去买芒果去了。更重要的是，你的算法还能继续演进（reinforcement learning），读取更多的训练数据，准确率也更高，每预测错误后再进行自我修正。更妙的是，你还能用同一个算法来训练不同的模型，预测预测苹果、橙子、香蕉、葡萄、樱桃、西瓜什么的。

FIN

# 如何分析销售数据

© 文 / 中颢润项目数据分析师事务所 许梁 图 / 崔峻珩

一、分析销售数据通常是从产品、区域、客户三条主线来研究



二、需要展示的数据指标有如下几个方面：

销售数据	
1. 当日整体业绩	2. 累计业绩达成
关键数据： • 当日业绩达成率 • 销售数据构成分析  实际操作意义： 1. 跟进弱势区域 2. 跟进弱势产品 3. 实时掌握销量	关键数据： • 月累计业绩达成进度要和进度进行对比 • 档期任务达成  注意： 1. 特殊时段任务档期管理
3. 发货情况	4. 销售质量
关键数据： • 发货客户数量 • 发货次数  重点关注 1. 当前前15天上面两项数据情况 2. 下半年及时跟进	关键数据： • 产品结构 • 区域结构 • 重点客户和渠道结构  衍生指标： 1. 平均价 2. 重点产品比率

三、整体销售分析的思路：

• 销售额/销量

分析近几年的总体销售额、销售量，与行业标准相比较从而分析企业的业绩状况并判断企业的业绩变化类型

• 季节因素

依据行业销售淡旺季规律，与销售数据中的销售行程进行对比，分析淡旺季发展规律；可以为客户提供渠道压货规划及生产运做规划

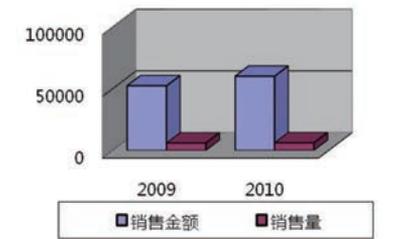
• 产品线

通过总体产品结构分析，了解整体产品结构分布和重点产品表现

• 价格体系

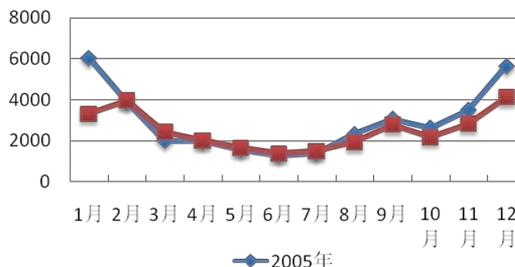
通过总体价格结构分析，了解企业的优势价位区间，提供价格结构调整的合理性建议

1、销售额 / 销售量



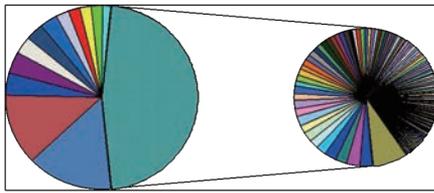
通过对销售额和销售量的增长趋势的把握，可以找出客户增长或下滑的本质；如销售额增长大于销售量增长，说明增长主要来源于产品平均价格的提高，它反映了市场平均价格的提高或者是客户产品结构升级，即结构性增长；反之，为容量性增长。

2、季节性分析



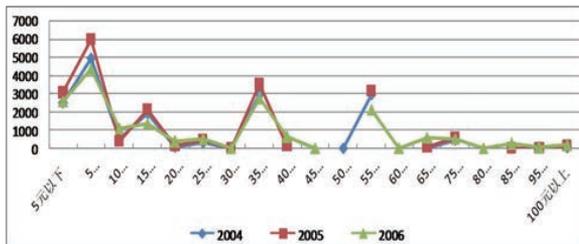
很多的消费品行业存在明显的季节性趋势；根据行业规律，为企业提出生产运作及渠道供货的合理性规划

### 3、产品结构



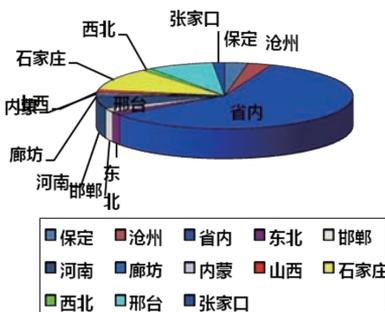
从产品结构看主导产品和产品成长合理性，企业的利润源和销售量是否对应，初步判断企业未来产品规划的调整方向。

### 4、价格体系



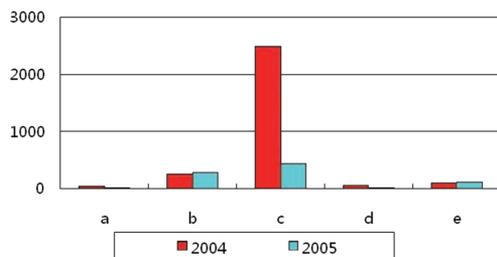
从价格结构看产品分布合理性，也可以判断目前企业现状与企业的战略发展方向是否一致。

### 5、区域分布



企业的销售区域分布看企业市场分布的合理性；企业区域分布与整体战略目标的一致性；明确下阶段企业区域布局的规划方向。

### 6、区域销售异动



通过对数据分析的分析，发现存在异动的产品或区域；并分析异动发生的原因。

## 四、具体分析方法：

### 1、“三维”数据分析法

数据分析的目的是为了找到问题，要逐层深挖“罪魁祸首”。

首先看整个公司的达成率、成长率、增长率、产品结构，发货客户数是否健康，发现任何一个指标异常，立刻要从“区域、客户、产品”三条线往下深挖：

业绩滞涨是那个产品造成的——这个产品业绩差是哪个区域造成的那些客户造成的，最终锁定问题产品的产生原因——问题产品的问题区域和问题客户。

业绩滞涨是那些区域造成的——这个区域的问题产品和问题客户是谁，最终锁定问题区域产生的原因——问题区域的问题客户和问题产品

哪些经销商业绩问题最严重——这个问题经销商业绩是哪个问题产品造成的，层层抽丝剥茧找到最终原因，管理才会有放矢。

### 2、6个月纵向分析法：

对锁定的问题产品，问题区域，问题客户可以纵向深度分析6个月数据。

该区域6个月内的总销量及主要品项的达成/增长/成长率是否稳定？

发货客户数和发货品项数有无异常趋势？

从6个月的纵向趋势分析更容易剔除当月偶然因素，做出完整结论。

比如：问题区域纵向分析示例：“该区域6个月以来连续4个月达成率小于100%，达成率近5个月低于公司水平，已经有三名经销商停止进货。高价产品A每月销量在减少（说明这个产品回转很差快死了），需要立刻去走访，尤其注意观察停止进货的经销商是什么原因？还有高价产品A的市场表现有什么异常？”

### 把全国市场分为三类：

管理类市场（如：市场波动大，成长率小于5% 甚至负成长）；

成长类市场（如：市场稳定成长率小于50%）；

攻击类市场（如：市场空白较多，成长率大于50%甚至更多）

对于不同类型的市场进行归类分析，除了所有市场都要求任务达成率之外，不同类型市场要关注不同指标；

对管理类市场，由于市场不稳定，有负成长趋势，所以要重点关注“零发货客户”指标，小心市场萎缩。

对攻击类市场，由于可开发空白较多（这类市场不应该有淡季，应该逐月增长，因为有很多空白网待开发），所以要持续关注逐月增长率，同时成长率低于20%说明进攻速度放缓，应予以关注。

对成长类市场，重点关注成长率，保证市场稳定。

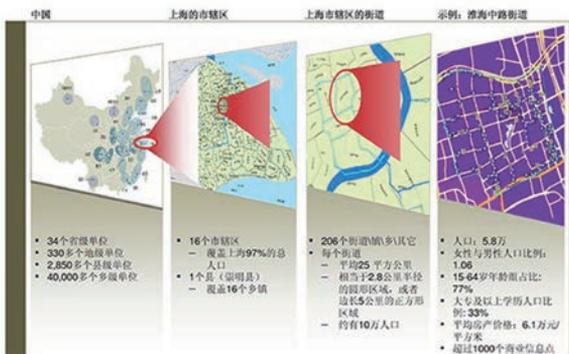
FIN

# 大数据与商业地理分析：中国视角

◎ 出处 / 麦肯锡的中国视角 文 / Justin Liu 编辑 / 会员处 马辰

从全球范围来看，采用商业地理数据进行商业选址及消费者地理细分在发达经济体已经非常普及。为更精准地服务不断升级的中国消费者，宜家家居、麦当劳、星巴克等专门成立了商业地理分析团队，来指导其在中国的店铺选址。麦肯锡的“解读中国”商业地理分析团队亦感受到来自客户方越来越强烈的需求。我们以下图来说明架构在大数据之上的商业地理分析。

商业地理分析使得逐步深入到微观层面的分析成为可能



“80%的商业数据都是带有地理信息的” “商业地理分析的目的就是把对的产品放在对的位置上”

“选址分析专家就是帮助客户找到最有利位置的‘风水’先生”

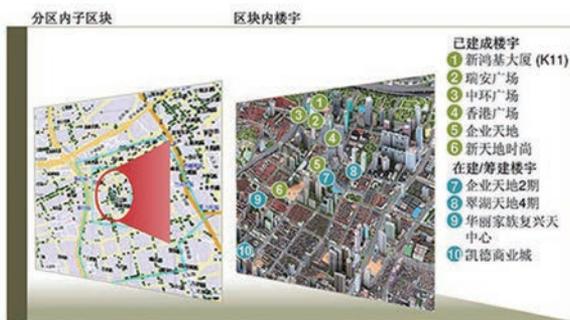
## 科学选出最优位置

我们服务过一家全国股份制商业银行，该银行希望规划未来在中国某一线城市的网点开设计划。这些网点须开设在（潜在）顾客集中的区域，方便个人及企业客户的业务办理，同时要避免选择过度竞争的区域，确保业务的健康增长。该如何科学地选出最优位置？

评估某一个特定地点是否具有商业价值，深入该地进行调查是传统的“笨”办法。若想从一百多个城市中选出每个城市的重点商圈，仅凭个人或者小团队的有限知识和商业直觉是远远不够的。我们认为，一个可行的方案是，利用这些城市的矢量地图并加载更细层面的经济、人口和地理数据，借助地理信息系统（GIS）来实现批量处理和定量分析。打个比方，风水先生一旦配备了现代化装备，就升级成为商业地理分析专家，他们凭借商业地理数据帮助客户寻找并确定城

市中的最优位置。我们采取了抽丝剥茧层层深入的方法，从街道到商业楼宇，对可能的位置进行深入分析。综合该城市超过200个街道的人口统计信息、分区富裕程度、分区内各银行网点的分布及开业年限、各类商业信息点的分布等信息，将这些街道分区归纳为核心分区、次核心分区和避免分区三个大类。新设网点时优先考虑核心分区。接下来，深入到每一个街道分区内部，根据分区特征、商业信息点的分布与区域聚集度进行打分，结合该城市各分区内已建/在建/筹建楼宇列表选出网点的最优位置。

定分区后，结合商业信息点的分布以及楼宇列表，通过综合评分落实网点选址



## 沙盘上的商业地理

商业地理分析正如将军俯视沙盘，挖掘商业数据的地理纬度，将城市的战略高地和价值洼地一览无遗，运筹帷幄，决胜千里。

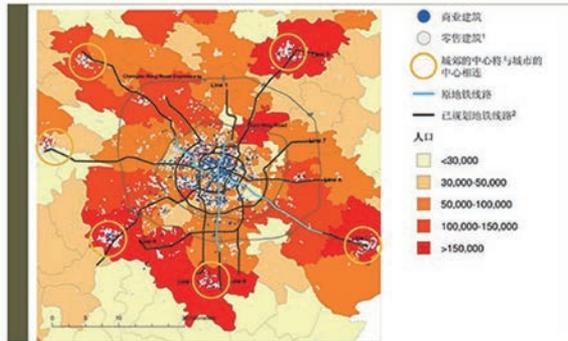
打开麦肯锡“解读中国”的22个城市集群，查访每一个城市，从市辖区到街道，从街道到居委会，乃至2km×2km的栅格，商业地理的分析工具使得“战略图景”的解析度和可视化程度大大提高。全新的高清影像不仅冲击着跨国公司、本地龙头企业，还有政策制定者。即使是城市轨道交通建设这样长期而浩大的工程，商业地理分析亦能提供独特的视角。

近期我们获邀为西南某省会城市的轨道交通发展把脉。从地理空间的角度来考察地铁规划再合适不过。该市地铁尚处于公共轨道交通建设初期，而未来10年间将从现在的2条线增加到10条线。

将该市的地铁规划、人口分布、商业网点分布及楼宇价格都放在GIS平台上，地铁的未来蓝图跃然纸上（见图2）。全面竣工后39%的城市人口将会在地铁站点周边800米内。

但与伦敦和莫斯科等国际都市相比，地铁站点密度仍然偏低。同时，半数以上的医院和学校超出了地铁站点800米覆盖范围，站点附近尚缺乏足够的配套公共服务设施。进一步分析发现，还有部分地铁站点周围人口稀疏且商业活动不频繁，可能是城市中的价值洼地。

铁路规划将主要城郊与市中心很好地连接起来



### 处于地理信息产业链的顶端

地理信息产业在中国方兴未艾。近年来，提供地理信息系统、数据和服务的公司呈现出跨越式增长。从地图测绘和遥感影像获取，到地图加工并构建地理信息数据库，再到数据的批发与零售，一个完整的地理信息产业链已经形成。借助成熟的GIS软件，通过图形化的二次开发界面，商业地理分析员可以根据客户需求快速地进行定制分析。

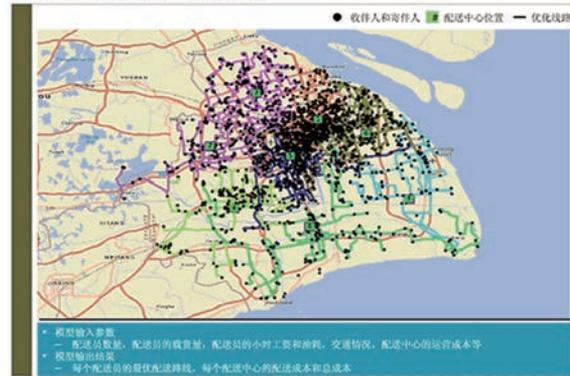
最大的赢家将是站在产业链顶端的定制商业地理分析产品的提供者。商业地理分析产品将涵盖：选址分析、销售区域分配、配送路径优化、潜在消费者空间分布、城市规划等。其中，如前所示的选址分析应用最广最深，涉及经济环境分析、交通条件分析以及竞争/互补店分析，称其为商业地理的灵魂亦不为过。

几乎所有的行业都需要商业地理分析：银行、快消、电信、医药、航运、家具等等，即便是电子商务这样虚拟的行业，也需要商业地理的帮助，来判断消费者的地理分布以及不同地区消费者的特点，从而有的放矢地发布网络或者平面广告，抑或根据不同地区制定相应战略。物流公司更是离不开商业地理分析的统筹规划，通过与全面系统的商业地理信息数据库相结合，传统的运筹学焕发出新的活力。

以我们曾经服务过的某物流公司为例，该公司希望知道在上海布局多少个配送中心才能使配送成本最小。利用GIS软件中的Vehicle Route Problem模型可以很好地解决这个问题，我们测试从8个配送中心开始，把配送中心地理位置、收寄件人的位置、配送员数量及载货量、配送员小时工资和油耗、交通情况等作为参数，求解得到总成本，再与其他数量的配送中心求解得到的总成本进行比较，得到一条配送中

心数量与总成本的曲线，从而求得配送中心的最优数量和地理位置。

8个配送中心情况下的最优配送路线



商业地理分析面对的客户往往是有着成千上万个网点的实体零售商或电商、数以千万计接货送货人的物流公司、网点遍布全国的大型银行。为这类大客户服务，决定了商业地理分析将处于地理信息产业链的顶端，并必然伴随着丰厚利润回报率。而中国目前专门从事商圈分析及商业选址分析的公司还只是少数，拥有大客户资源、掌握丰富数据且具备商业地理分析技术的公司将会成为新兴产业的领袖。

### “数据难”制约中国商业地理分析

从商业地理的全球实践来看，发达经济体可以将地理信息成熟运用于商业的各个环节，在印度、印尼这样的新兴经济体中，地理信息技术也日益得到重视。

相比国外完善的商业地理数据服务，目前在中国，地理信息数据的可获取性、准确性和全面性仍然制约着中国地理信息产业的发展。能够提供商业价值较高的街道及以下层次（如街道、邮编区域、居委会乃至小区）边界的地图供应商极为稀少，与之相配套的数据，如人口、收入、消费、住房房价和商业楼盘的租金，也不易获取。在二线以下城市，边界地图数据可能要从各地测绘机构零散地加以收集，获取覆盖全国的数据非常困难。

过去几年，在麦肯锡全球商业地理分析团队的帮助和多方努力下，我们已经构建了深入到街道级别的地理信息数据库，涵盖近千万的商业信息点，并已经应用于数十个客户项目的分析中，在中国处于领先地位。目前，全国人口普查数据、房价数据都已经整合进入这一空间数据库，更高层次（如居委会等）的数据整合工作正在进行中。

正如大数据改变着世界一样，地理坐标将推动新一轮大数据的进化。前瞻未来，也许重要的变革机遇就蕴藏在商业地理分析中。

FIN

# 如何通过数据挖掘帮助母婴零售店扩大生意

◎ 出处 / Intely 文 / ETwise 编辑 / 会员处 马辰 图 / 崔峻珩

随着中国新一轮的生育高峰期到来，母婴零售店遍地开花，市场竞争趋向白热化。目前越来越多的母婴零售店都采用了会员制度，并以“会员专享优惠”、“多倍积分”、“积分兑换奖品”等优惠方式吸引大批新的消费者留下个人手机号码、宝宝性别、宝宝月龄、家庭地址等相关信息。而消费者也会使用相应的会员卡来母婴店进行购物，以便获得更多的优惠。

2013年是电子商务逐步成熟的一年，网上购物越来越方便；而线下母婴零售店促销打折不断，竞争异常激烈，留给母婴零售店高层管理者两个不得不认真去思考的企业经营问题：

1. 如何去唤醒招募进来的沉默会员？
2. 如何让目前活跃的老客户进店购买更多的产品，最大限度地释放其价值？

因为大多母婴零售店采用了会员制度，积累了越来越多的会员数据，在数据化管理的时代，会员的消费数据意味着公司的财富。如何对会员数据进行深入挖掘，并辅以有效的营销活动执行，更好地帮助母婴零售商、母婴品牌商共同把生意做大做强呢？

本文结合所操作过的某高档母婴奶粉品牌联合某母婴零售店进行联合销售的数据挖掘项目经验，如图1所示，将从以下4个步骤深入浅出地介绍如何通过数据挖掘帮助母婴零售商、品牌商把生意做得更好、更强。



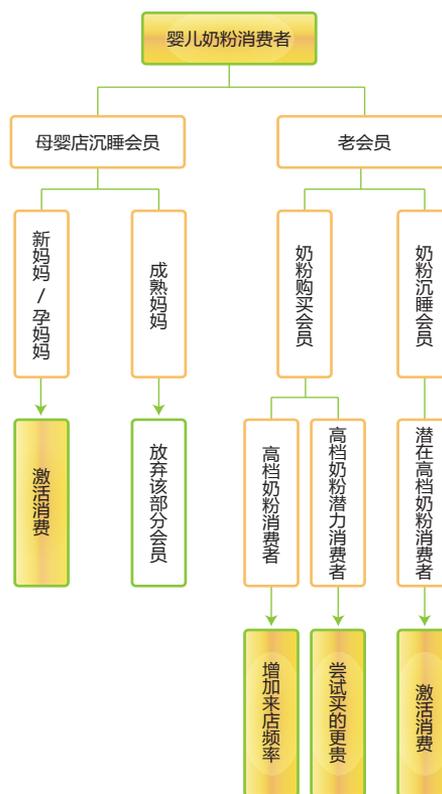
图1. 母婴行业数据挖掘总流程

## 一、进行客户细分，挖掘不同群体的潜力消费者

古语有云：“物以类聚，人以群分。”说明了不同的群体具备不同的行为特征，所采取的营销策略也不一样。母婴

零售行业也不例外，只有对会员有效地进行客户细分，然后针对不同的会员采取相符合的营销政策，才能更好地达到相应的营销效果。

结合某大型母婴零售店的会员信息及某高档母婴奶粉品牌的特征，对该母婴零售店的会员进行客户细分，主要分为4大类客户：



1. 新妈妈/孕妈妈：新妈妈主要指最近3个月新招募过来的且小孩年龄在0~3岁左右的妈妈，但在该母婴店却没有任何消费记录的会员；而孕妈妈是最近3个月新招募回来的且还在怀孕期的妈妈，但在该母婴零售店却没有任何消费记录的会员。

2. 成熟妈妈：指最近3个月招募回来的但小孩年龄已经超过了3岁的妈妈，在该母婴店却没有任何消费记录的会员。

3. 奶粉购买会员：是指小孩年龄在0~3岁的且在该母婴

店购买过奶粉的会员。

4. 奶粉沉睡会员：是指小孩年龄在0~3岁的且在该母婴零售店购买过非奶粉的会员。

结合商业目标是挖掘高档奶粉潜在目标客户，帮助母婴零售店及奶粉品牌商联合共赢，奶粉购买会员里面可以细分两种目标用户：高档奶粉消费者及高档奶粉潜力消费者。而奶粉沉睡会员众多，需要找出那些具有消费能力的客户进行激活消费，此类客户需要寻找潜在高档奶粉消费者。

通过对会员数据的有效细分，寻找潜在目标用户，并针对不同的客户群体进行营销，如对新妈妈/孕妈妈、潜在高档奶粉消费者进行激活消费；促进高档奶粉消费者买得更多及增加其来店频率，最大释放其消费价值；而对高档奶粉潜力消费者，则让其尝试买得更贵。

## 二、如何有效判断会员价值

每个妈妈在对不同品类母婴产品的选择上，有很大差异。比如有的妈妈倾向于购买最贵的纸尿裤，但认为奶粉只要中档品牌就可以；也有妈妈购买低廉价格的纸尿裤，却买最贵的玩具和服饰。因此进行客户细分之后，还要对他们是否对我们目标商品有价值进行预判。

结合高档奶粉目标用户的商业目标出发，消费者价值的预判主要针对其购买历史记录，从以下三个角度进行判断会员价值：

### 1. 是否经常购买高档奶粉关联产品

通过数据挖掘的购物篮关联分析去研究买过高档奶粉的购买者行为特征，找出高档奶粉的亲缘品类。如图3所示，通过购物篮分析可以发现，很多客户在购买一段雅培高档奶粉的同时也会购买某品牌的婴儿专用洗衣液，则该品牌的婴儿专用洗衣液为该高档奶粉的亲缘品类。结合亲缘品类的购买情况，可以判定“未购买目标品类的会员”是否有购买目标品类的潜力。

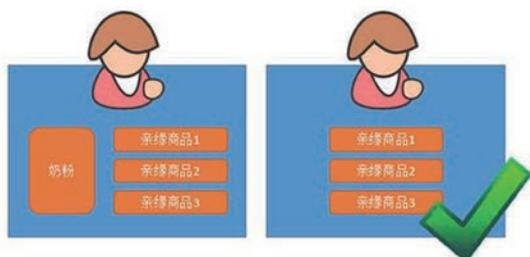


图3 - 通过亲缘品类购买行为判定消费者是否有购买目标商品的潜力

### 2. 是否曾经购买过高端标杆产品

并不是有钱的家庭一定会购买高档商品，也不是经济条件差的家庭就一定会选择最低价商品。但购买过高端标杆产

品的会员就肯定有一定的能力消费得起高档产品。

在母婴零售行业中，有几类高端标杆产品能够凸显会员的消费能力。如表1所示，经常购买高档纸尿裤、购买过高档推车（如价值上千的GoodBaby推车）或者购买过高档奶瓶（如购买过日本原装进口 betta贝塔智能型玻璃奶瓶），这些会员有经济能力消费得起高端母婴产品，就说明其经济能力能够消费得起高档婴儿奶粉，这些会员则是高档婴儿奶粉的潜在会员。

高档纸尿裤	高档推车	高档奶瓶	高档奶粉
花王	Inglesina	betta	合生元
大王	Goodbaby	贝亲	惠氏启赋
帮宝适特级棉柔	康贝	新安怡	雅培菁智
好奇金装	Quinny	NUK	多美滋金盾

表1 高端标杆产品品牌

### 3. 结合登记地址辅助纠正会员的价值

在新会员招募的时候，母婴零售商家会要求会员留下相应的家庭联系地址，以方便后期能够更好地进行精准营销。通过登记地址信息梳理，辅助纠正会员的消费价值。

综上所述，在进行会员价值判断的时候，必须结合会员的历史消费记录、登记地址梳理，才能更有效地识别会员的价值，从而找到潜在的高档婴儿奶粉的潜在目标。如图4所示，某会员在该母婴零售店购买了中档的婴儿奶粉及相关产品，从购买该奶粉的历史记录来看，该会员消费能力仅为中等水平。但再结合其过去购买产品记录及其登记地址进行梳理，可以发现该会员有能力买得起高档奶粉，为高消费能力的潜在会员。



图4 会员价值判断举例说明

结合会员的历史购买记录数据及注册信息，通过数据挖掘可以将所有潜在购买者划分为的4个象限（如图5所示）：

第1象限：不但购买过高档奶粉，而且又有较高消费能力的群体，是我们的核心顾客，应该重点维系。

第2象限：虽然购买高档奶粉但消费能力比较低的群体，对促销优惠会更加敏感，该象限会员为优惠敏感顾客。

第3象限：既没有消费能力又不买高档奶粉的顾客，该象限会员为低消费群体。在营销经费有限的情况下，可以选择性放弃。

第4象限：消费能力高却只买低档奶粉的顾客，通常比较容易被忽略，但却是非常具有潜力的群体。

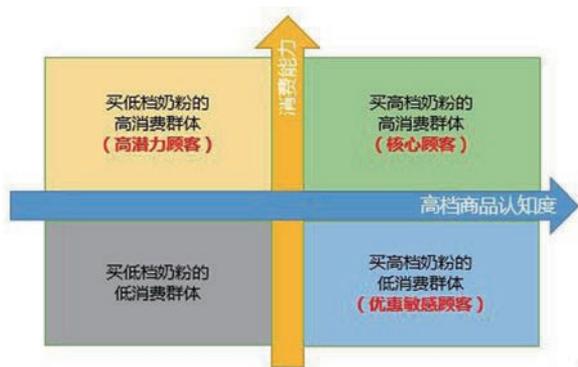


图5 母婴商品消费者价值判断模型

### 三、如何设计合理的营销活动

在有效进行客户细分及有效判断会员价值的前提下，如何设计合理的营销活动成为了关键？

对会员进行营销推广有很多不同的形式，其推广的目的也不一样。

推广目的	购买点促销	赠送优惠券	赠送样品	价格折扣	举办竞赛	奖品	附加赠送包装
引起试用产品	*	*	*	*	*		
转变使用者为重复购买者		*		*			*
鼓励增加购买数量		*		*		*	*
刺激非计划购买	*	*		*		*	*
提高合作协议的可接收性				*	*		
加强广告促销的效果					*		
巩固品牌形象					*	*	

表2 消费者营销推广的目的

表2列出了母婴零售企业对消费者进行营销推广的各种目的，以及每种目的可以使用哪些营销推广方式。

结合高档奶粉数据挖掘的商业目的“唤醒沉睡会员及让老会员购买得更多”，针对不同群组的会员设计了不同的营销活动方案，如表3所示。

结合不同的婴儿奶粉细分客户群，针对不同的营销目的，匹配相应的营销活动设计，能够更有效地提高营销活动成功率。

细分客户群	具体目的	营销推广激励机制
激活新会员	拉动新/孕妈妈	短信优惠券、试用装、试用装&短信优惠券
激活新会员	刺激成熟妈妈	放弃该部分客户
买过高档奶粉	提高来店频率	短信优惠券、多倍积分
买过奶粉的潜在高档用户	买得更贵	短信优惠券、多倍积分、试用装&短信优惠券
激活沉睡会员	激活沉睡	短信优惠券、多倍积分

表3 不同细分客户群采用不同的营销活动设计

### 四、如何进行营销效果评估

在进行客户细分、价值判断、营销活动设计等一系列工作之后，需要进行科学的营销效果评估，以帮助母婴零售店能够及时了解营销活动的效果，同时发现营销执行中存在的问题，不断优化营销活动方案设计。

如图6所示，目前我们主要采用了“五度”评估体系，从营销目标完成度、营销推广响应度、营销推广成功率、参加客户成长度、参加客户满意度等五个方面去评价营销活动的效果。



图6 营销效果评估体系

围绕着五度评估体系，针对母婴零售行业的特性，我们主要从以下几个方面对营销效果进行科学评估：

1. 营销活动前后的整体销售额、客户数大小变化情况；
2. 营销活动前后的老客户响应人数、购物频次、购物篮大小等关键指标的变化情况；
3. 沉睡会员转化率、参与活动的客户数、贡献的销售额分析；
4. 各个营销活动的投入产出比（ROI）分析。

FIN

# 用数据分析解读你的生活

## ——情侣之间的那些事儿

© 文/CPDA上海授权中心 张曼 编辑/会员处 马辰

曾经有个朋友和我说，不想谈恋爱的原因是“太贵”。情侣之间七七八八的花费真的不少，就让我们来看看这些钱到底是花到哪里去了？

说到情侣之间那些事儿，我们可以先来说说宾馆酒店的价格成本。按照汉庭等最新一季财报分析，以一晚199元经济房为例，空置房间成本54元(27.1%)、物业租赁40.6元(20.4%)、人员开支23.3元(11.7%)、利润18.4元(9.3%)、水电易耗品15.85元(7.9%)、折旧14.5元(7.3%)、管理11.6元

(12%)、管理成本7.8元(12%)、纯利润5.85元(9%)、研发建设4.55元(7%)、销售费用3.25元(5%)、营业税1.95元(3%)。

谈恋爱固然不便宜，但是只要不求婚，成本还是可以控制的。一旦到了追求“钻石恒久远一颗永流传”那时候，大擦的钞票花出去真的只是分分钟的事儿。不过钻戒虽贵，好在你也能够买到实实在在的东西，制造成本占到了41%，由裸钻费、戒托材料费、加工费三部分组成。如果你不是非要追求像Tiffany这样的名牌的话，整体的价格还可以更低。



(5.8%)、税收5.8元(2.9%)、市场营销2.9元(1.5%)、其它7.3元(6.1%)。也就是说四分之一的房钱你是帮昨天订了房却没人付掉了。酒店空置的房间就像是服装店积压的库存，空耗成本却没有收益。更糟的是，服装店今天卖不掉的衣服明天可以接着卖，酒店的空房如果当晚卖不掉就只能呵呵了。

2014年2月上旬，广州海珠区法院受理一名“80后”白领要求前女友返还交往期间其网购避孕套的费用的案件。可见，安全套虽小，到底也是一笔花销。按照杜蕾斯最新一季财报分析，以一盒零售价65元的18片装杜蕾斯为例。其中原材料成本20.8元(32%)、经销商分成13元(20%)、广告投放7.8元

在曾经的某个月里，没有鲜花和钻戒，情侣之间互赠一本精神食粮就已经是足够浪漫的事情了。据了解，一本售价20元的书价格构成是怎样的，作者稿酬，按8%版税率算，约1.6元；印制成本，即印刷费和纸张成本，占书价的25%，约5元。出版社大约用6折的价格卖给经销商，中间出版社挣了大概5元；经销商利润约2.4元；零售商约赚6元。如果你要是买同样的一本盗版书，没有版税、更劣质的纸张和印刷、更简便的经销渠道，加起来就只要10元钱了。不过对方要是因此觉得你人品有问题而就此分手，那你的损失就太大了。所以，盗版书还是少买的好。 FIN

# 用数据分析解读你的生活

## ——养孩子过日子的成本分析

© 文/CPDA上海授权中心 张曼 编辑/会员处 马辰

如果你生活在魔都上海，而你又是一个追求生活品质的  
人，而更要命的是你又有孩子要养，那么你的生活成本将  
会是以下样子的：

其中可能某部分是可以压缩的，但都会影响生活品质。  
总之上海的品质生活绝对不轻松！当然啦，图上的生活品质已  
经不错啦，一家三口在上海也是有“经济版”生活账单的。

- 租房：1500 无房贷 无保姆
- 孩子：幼儿园500 课外班100
- 上班成本（两人交通、中饭、早饭）：600
- 伙食水果：500 水电煤：300
- 老婆的衣服、化妆品：800
- 总计：4300

爸妈最舍得在小朋友身上花钱，一套乐高玩具动辄几百  
块，说买也就买了。除了经销商费用占比略大之外，玩具的

材料和人工费用也是实实在在的。

卫浴产品属于耐用消费品，在大众消费中所占比例不  
高，其产品的平均消费周期在8年左右。以TOTO为例，其标  
价3000元左右的马桶，出厂价格通常在售价的60%，也就  
是说经销商的毛利率占到40%。不过，由于房产政策的压力以  
及原材料价格的上升，加上高涨的门店租金，经销商们的净  
利润实际在10%以下。

在花钱大手大脚，挣钱小打小闹的日子里，你是不是曾  
经想过一张彩票发家致富？以一张售价为2元的体育彩票为  
例，45%都作为奖金又发出去了，30%成为了福利收益，国  
家的管理发行费用仅占1%。这样看来福利彩票还是有些正经  
用处的。

FIN



# 用数据分析解读你的生活

## ——“一瓶水”的价格构成

© 文/CPDA上海授权中心 张曼 编辑/会员处 马辰



你以为买了一瓶迪奥香水就是奢侈，路边买一瓶矿泉水就是朴实吗？那你就可能大错特错了。香水的成本占到销售价的比例是很低，但是你想过矿泉水中水的成本有多低吗？都知道香水过度包装，但是你知道矿泉水的包装成本更让人“呵呵”吗？至于女人抱怨税收让国外化妆品价格飞涨，但是为了控制国外酒品，税收部门有多努力你知道吗？下面就让我们一起来看看“一瓶水”的价格构成。

以一瓶550ml在社区零售店售价为1.5元的矿泉水为例。水成本为1分钱；瓶子+盖+喷码+胶带约为0.17元；营运和广告费约为0.22元；经销商平均以0.6元每瓶的出厂价拿货，再以平均1元每瓶的价格批发给社区零售店，零售店再以1.5元每瓶的价格出售。

以一瓶售价为268元的进口波尔多红酒，有38.5元花在原

料和酿造成本上，22.5元花在了人工上，16元花在运输储存上，7.5元买了瓶子和软木塞，5.5元为它做了市场推广，24.3元给了代理商，19.7元给了酒庄，还有134元，用来缴纳了关税、增值税和消费税。所以说酒喝多了伤身，也费钱。

最后来看看香水，以Christin Dior Miss Dior香水为例市场价格780元，免税后销售价格520元。零售商利润很不错，占到了10%，也就是78元。几乎和制造商净利润差不多。至于真正原材料成本只占2%，也就是15.6元，不过和矿泉水比起来已经是多了的。

FIN

## 浅谈事务所发展之出击

◎ 文/会员处 贺岩 图/崔峻珩

随着大数据时代的到来，企业也正朝着智慧型的方向发展，我们的事务所只要可以迎合市场的需求，脚踏实地的为企业量身定制切实可行的数据分析报告，那么事务所的春天才能真正的到来。



我们都在说大数据时代来临，信息和数据大爆炸。从2013年初开始，对于大数据爆发的焦虑感，紧迫感，不由自主地驱动众多行业、企业和团体去关注、接触、了解大数据，自觉或不自觉的，主动或被动地去融入这波洪流。

但是，真的说到大数据，我们身边到底有多少数据量，它们都分布在哪些行业，哪些数据是目前可用的，哪些行业已经在使用数据依旧不是十分清晰。因此，鉴于目前中国数据领域的“数据”难等现状，项目数据分析师事务所作为专业从事数据分析领域研究的从业者，更应该怀着希望，以第一个吃螃蟹并期待来自行业的矫正的态度，首先尝试对于国内各个领域，行业以及机构的数据拥有情况，使用情况以及未来路径做一个粗犷地调研、梳理和判断。主动寻找业务，而非坐等政策业务，这是事务所挖掘各行各业数据的基本要求。

所谓术业有专攻，我相信各个事务所也是一样。大到事务所所在地区的政策扶植，当地的主要经济命脉，中到事务所本身拥有的优势资源，小到分析师本身的专业倾向，这些都是事务所锁定自己目标客户的重要依据。例如，事务所在一个旅游业比较发达的城市，那么我们就可以重点研究一下数据分析在旅游行业里的应用。比如，我们的事务所固有资源主要是来自医疗行业，那么我们就可以把目标锁定为数据分析在医疗领域的应用，诸如此类。当然，我在这里说的术业有专攻，并不是让事务所只承接一个方面的业务，而放弃整片的森林，只是要在研究方面有自己标志性的特长。如果能让某个领域的企业在有意做数据分析报告的时候，一下就能想起某家或某家事务所的名字来，我相信这将是事务所在业内得到的最高荣誉。

其次，锁定了目标之后就是如何与目标企业进行接洽。所谓万事开头难，“大数据”这个概念对于很多企业还是一个懵懵懂懂的未知数，究竟什么是“大数据”？“大数据”跟我的企业有什么关系？它能让企业赚钱吗？我相信绝大部分的企业老板脑子里都有这些疑问。那么解决这些疑问，无疑将是事务所敲开企业大门的敲门砖。

事实胜于雄辩，关于大数据营销的神奇之处，有一个故事在大数据这个名词还没有红透互联网营销领域时就已经被津津乐道。故事是这样的：2012年初，一个男人冲进一家位于明尼苏达州阿波利斯市郊的Target超市兴师问罪：为什么超市不停地向他的还是高中生的女儿邮寄婴儿尿布样品和配方奶粉的折扣券？“你们是在鼓励她怀孕吗？”愤怒的父亲质问Target超市经理。几天过后，超市经理打电话向这位父亲致歉，这位父亲的语气变得平和起来，他反过来说道歉说，他的女儿确实怀孕了，预产期在8月份。这是一个零售商如何应用大数据进行营销的故事，这个故事被《纽约时报》报道





后，大数据的威力轰动全美。

再如，海尔的大数据营销案例。海尔SCRM会员大数据平台是一个帮助企业转换视角，在网络化时代为用户提供精准营销与互动服务的平台。这其中的一个小事件是，SCRM会员大数据平台从亿万人群中精准的预测出，住在北京景泰西里小区的外企高级经理陈然可能需要购买能除PM2.5的海尔帝樽空调，随即海尔便向这位潜在的客户发送了精准的个性化服务推荐信息，当然结果可想而知，陈然在几天后欣然的购买了这款海尔的帝樽空调。

从以上例子不难看出，其实任何一个行业或是一个企业在日常的经营中都有着大量的数据源，而这些看似简单的数据，往往正是决定企业如何发展，是否能够突破瓶颈去争取更大利益的基础依据。一些大的企业可能早就意识到了数据分析的重要性，他们组建了自己的数据分析团队，对企业自身的经营数据及行业中的大量数据进行收集、探讨、清理、分析，取其精华，去其糟粕，为己所用。但大部分的企业，尤其是中小型企业，还是不具备自己组建数据研究团队的实力，他们不是不需要数据分析来帮助他们做出正确的决策，只是无从下手。所以抓住这部分潜在客户，将是事务所立足的根本。

至于如何调动企业对数据分析行业的兴趣呢，我以为不外乎以下两点：

请进来——事务所个体或者几家事务所可以联合起来经

常性的举办一些涉及各个领域的数据分析沙龙，邀请当地的相关企业参加，在会上事务所可以通过真实的案例为他们展示数据分析在各个领域的应用，也可以现场收集一些企业的需求，为他们排忧解难。

走出去——如果事务所能够摒弃坐等生意上门的想法，能够沉下心来走进企业，为企业义务的讲解一下数据分析对他们最切合实际的帮助，甚至可以无偿的为他们私人定制一些适合本企业的简单数据分析报告，帮助他们做出一些简单的经营决策，使他们从中获利。我们有理由相信，商人都是以盈利最大化为目的的，小投入，大回报的事，他们自然乐意去做。

总之，随着大数据时代的到来，企业也正朝着智慧型的方向发展，我们的事务所只要可以迎合市场的需求，脚踏实地的为企业量身定制切实可行的数据分析报告，那么事务所的春天才能真正的到来。

FIN

数据分析培训公开课

BIG DATA

# HR的春天

## 大数据推动人力资源管理变革

公开课  
11月1日相约北京

免费公益  
机会难得

- 让人事管理不再尴尬
- 人力资源管理大部分是艺术，小部分是科学！

# 以前

人力资源管理的价值难以被客观评估；  
人力资源的专业性难以得到充分认同；  
人力资源的话语权模糊而不确定。

在大数据技术的支持下，  
人力资源的选、育、用、留都可以纳入到量化范畴，  
使人力资源管理更加高效、更加精准并且更有话语权。

在选拔人才时确保效率与质量、  
既帮助企业培养储备人才也帮助员工进行职业生涯管理、  
使企业对人才的吸引力始终保持在业界平均水准之上。

# 大数据时代

人力资源管理的游戏规则变了！

## 未来选拔人才的流程

充分释放人才的价值！

### 1 进入人力资源管理数据系统

输入搜索关键词 | 业务名称、管理层次、工作经验

### 2 增加搜索关键词，控制人数降到个位数

增加搜索 | 关键词数量

### 3 点开猎头公司，输入搜索关键词进行遴选

点击更多数据服务项，分析排序 | 岗位胜任评估模型测试，排序

### 4 获得六份简历

查看应聘者社交网络信息 | 将应聘者关键维度数据进行对比

### 5 最前的三位应聘者，上报上级领导



利用大数据技术进行招聘的过程，相较于当前的企业招聘模式，无疑是革命性的进步。大数据时代的企业招聘，从技术方法角度看，是持续的数据挖掘过程；从信息角度看，是供需信息持续匹配，关联信息的不断交叉组合的过程；从人力资源专业来看，是对于人才价值评估、岗位胜任力理解的专业展现过程。

大数据时代，数据将贯穿人力资源管理全部过程。在这里，行业权威将从岗位选拔、薪资激励、领导力发展等几个方面为你诠释数据在人力资源管理中的作用。



## 对于企业

给你展现出一个在人力资源的选、育、用、留的各个环节中，以大数据技术为依托，实现全过程的量化、科学化，像其它管理职能模块一样，达到了科学管理所要求的可测量、可记录、可分析、可改善，全新的人力部门。

在大数据的支持下，使人力资源部门在包括人才选拔、任命、激励等，表现会越来越好，人力资源部门开始有更多机会成为业务部门不可或缺的合作伙伴，CEO也将越来越多求助人力资源部门作为战略达成的主要支撑部门。

## 对于业务部门



## 时不待我， 你还在犹豫吗？

给人力资源从业者创造极大机遇的同时，也带来了极大挑战。一方面，人力资源从业者通过运用大数据技术、移动互联网技术，对供职企业的介入深度大大增加，创造的价值也大大增加，从而也丰富了自己的职业生涯可选项。

另一方面，大数据体系的构思、构建、运用，以及如何把大数据技术与人力资源管理所衔接，对从业者来讲几乎是颠覆性的变革。



主办单位：中国商业联合会数据分析专业委员会  
联系地址：北京市朝阳区建国门外大街建华南路17号  
现代柏联大厦505

全国热线：400-050-6600  
咨询电话：010-65675375 、010-65678307  
咨询 QQ：1954496428 、1626045798  
报名邮箱：cdachina@chinacpda.org

培训官网：www.cdachina.cn  
协会官网：www.chinacpda.org

联系我们，  
给自己一个机会！

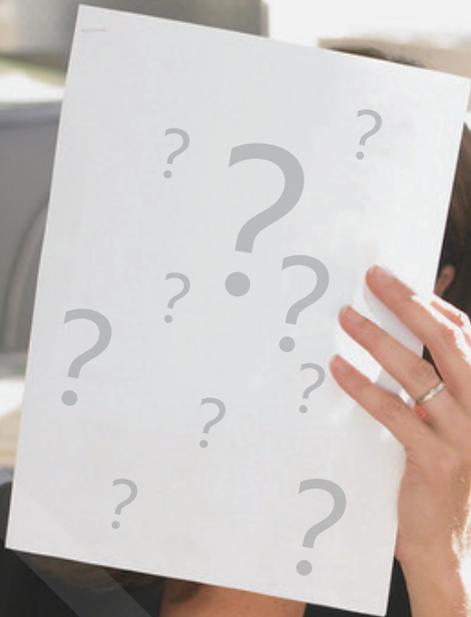


官方微博



官方微信

CPDA



# 漫无目的 不如 精准定位

项目数据分析师  
引领大数据时代的  
知识体系变革

早培训

早升职

早加薪



全新课程  
无限精彩

## 给自己一次机会， 迎接无限精彩的未来！



★ 关注微博微信参与互动 赢好礼!

提升量化经营技能，掌握数据分析必备的科学方法论！  
正确理解、灵活运用于企业决策全面数据化分析中！  
成功融入大数据时代，释放数据分析师全局视野！  
行业协会权威认证，承载坚实的启航平台！

 [www.chinacpda.com](http://www.chinacpda.com)

 400-050-6600