

● 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

2013年
第2期
总第十五期

- 05 2013年度行业会员执业教育
- 07 2013项目数据分析师新课程发布
- 09 用Excel对会员客户交易数据进行RFM分析
- 19 Netflix赌赢《纸牌屋》背后的秘密武器
- 26 数据分析是决定大数据成败的王道
- 35 禽流感考验数据分析能力

www.chinacpda.org

欢迎登陆中国数据分析行业网



卷首语

- 01 从奥巴马到奥斯卡，打造数据分析师精英团队

协会动态

- 03 2013年协会年检总结
- 04 关于推动各地举办数据分析行业推广活动的通知
- 05 关于举办2013年度行业会员执业教育的通知
附：关于远程执业教育课程学习的说明
附：关于参加面授执业教育课程学习的说明
- 07 2013项目数据分析师新课程发布
附：关于项目数据分析师培训新课程体系的说明

会客厅

- 09 用Excel对会员客户交易数据进行RFM分析

行业聚焦

- 17 数据科学家炙手可热
- 19 Netflix赌赢《纸牌屋》背后的秘密武器
- 21 大数据分析的三大演进方向

事务所风采

- 23 云南鼎臻项目数据分析师事务所
- 24 青岛易和盛世项目数据分析师事务所
华融成(北京)项目数据分析师事务所
- 25 湖南翰林项目数据分析师事务所
福州信达永道项目数据分析师事务所

数据应用

- 26 数据分析是决定大数据成败的王道
- 28 主动防御：大数据引发的信息安全技术革命
- 30 玩具外贸电商掘金欧美：数据分析爆款致胜
- 31 浅析数据分析在航空运输业中的运用

分析师园地

- 32 通道式数据与智慧决策时代
- 33 低成本，大收益的数据分析之思

生活驿站

- 35 禽流感考验数据分析能力
- 36 数据分析告诉你:Win8用户对开始按钮有多饥渴



2013年第2期 总第十五期

主办单位：中国商业联合会数据分析专业委员会
编辑：芦雁 关静
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
邮编：100022
电话：010-65670421 转 111
传真：010-65671802
邮件：xiehui@chinacpda.org
广告热线：010-65679535



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！

从奥巴马到奥斯卡 打造数据分析师精英团队

微软纽约研究院的经济学家David Rothschild采用大数据分析技术对第85届奥斯卡奖项进行了预测，答案令人称奇。除最佳导演外，其它各奖项的预测David全部命中，微软大数据分析堪称预言帝。事实上，早在2012年美国总统大选中，David就曾正确预测了51个选区中50个地区的选举结果，准确性高于98%。大数据分析支持下的决策为奥巴马的连任立下了汗马功劳，并且成为研究2012年总统大选的关键元素，更是开启了政界的大数据时代。

随着各个领域开始重视数据价值，企业间的竞争将日趋激烈，很多企业将面临以数据武装起来的更加咄咄逼人的竞争对手。谁能更好地利用数据带来的决策信息，做出更好更快的市场决策，谁才有可能最后的胜者。但是，在经营数据分析方面起步较晚的企业都会面临如技术和工具不足、人才短缺、外部信息库不健全、数据安全和隐私保护风险等挑战。面对这些挑战，企业必须借助专业的数据分析技术和数据分析师事务所的协助，逐渐提升数据分析的效用。

据Accenture对欧美发达国家的600多家企业的调查显示，有三分之二的公司在之前18个月任命了负责数据管理和分析工作的高管，而即使是尚未设立此类高管职位的公司，也有71%准备在不久的将来作出任命。而到2018年，仅在美国和英国，资深数据分析职位需求增长速度预计将是传统行业职位需求的五倍，是金融服务等信息密集型行业职位的四倍。

而数据分析行业，也需要不断跟进大数据时代的发展趋势，不断提高数据分析技术和能力，并提高为企业服务的专业素养，帮助企业在大数据时代更好地进行经营决策。为此，协会在2013年4月18日正式向全国发布了项目数据分析师全新课程体系，新课程将帮助学员更贴切地认识和理解数据分析在企业决策各个环节中起到的关键作用，具备解决企业各个决策环节的数据分析理论与方法，并掌握基本的数据分析实践操作技巧。

除此之外，协会在2013年还将通过“专业论坛”、“执业教育”、“互动沙龙”、“会员创业指导”、“专业数据库支持”等各式各样的活动，为广大项目数据分析师和行业会员搭建深度了解行业、强化业务知识水平和增加实操经验的平台，确保中国数据分析行业在快速发展的同时也保持更高端、更前沿的技术水平，打造国内顶尖的数据分析师精英团队。

中国商业联合会数据分析专业委员会





给自己一次
释放

别把心灵囚禁在狭小的数字迷宫中

迎接精彩未来

2013 培训课程全面升级

www.chinacpda.com
项目数据分析师唯一官方网站



给自己一次释放，迎接精彩未来！

——2013 培训课程全面升级

- 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！
- 项目数据分析师，继注会、注税后，又一专业权威认证！

全国咨询热线

400-050-6600

2013年协会年检总结

◎ 文 / 会员处 芦雁 图 / 崔峻珩

近年来随着数据分析行业进入快速发展阶段，全国各地的团体会员和个人会员迅速增加。为了适应行业飞速发展，自2012年开始，协会加大了会员年度审查力度。从2012年全年来看，加强监管和检查的效果明显，团体会员及个人会员的健康从业得到了有效的保证。

2013年年检工作协会在2012年年检工作的基础上，优化、完善了年检内容，并将年检工作标准化、制度化，真正发挥协会的职能作用、履行监管职责，并利用年检这一重要监督检查手段，了解会员真实的从业情况及存在的问题，帮助会员找出解决问题的方法，纠正和查处在备案管理、经营发展中存在的违规行为等。

2013年年检工作于3月1日正式启动，3月31日结束。纵观整个年检工作，本次年检与往年有较大的不同，主要表现在以下几个方面：

1、协会在2012年年检项目的基础上，结合了行业发展情况和特点，调整了2013年年检项目。其中，加大了会员对遵守行业守则、管理规定方面的检查力度，新增了对不正规资质、不正规从业行为的审查项目，并对积极参加

行业活动、为行业献言献策、维护行业利益的事务所和个人会员予以年检加分的表扬。

2、本次年检除了常规审查项目外，还特别设立了关于客户投诉率的审查条款，即有客户投诉的事务所，协会针对每条投诉都进行了详细地调查与核实，并进行了相应地登记和处理，凡年投诉率超过3起的事务所或个人，在年检审查时都给予了减分处罚。

3、对于长期从事不正当竞争、承接不正当业务、在经营过程中存在明显或严重违规行为、投诉较多等，以及未按时接受年检或违反协会相关规定的会员，协会做出了取消其从业资质并不再对其进行年检的决定。

本次年检被取消资质的团体会员如下：

- 南京摩尔项目数据分析师事务所有限公司
- 德特爱乃睐斯（上海）项目数据分析技术有限公司
- 陕西中精诚项目数据分析师事务所有限公司
- 江苏鹏举数据分析师事务所有限公司

4、从业务情况来看，在2012年已经有包括北京、山东、安徽、上海、广州、湖南、吉林、重庆、福建等地区的事务所会员单位从原来单一的投资类业务开始向从事企业经营决策业务转型，并有完整的数据分析报告备案。对于这些地区的会员单位，协会及年检评审小组给予了较高的评价。

年检对于行业来说是重中之重的大事，对于全国所有的团体会员和个人会员，协会都本着对行业发展负责、对会员负责的态度，将服务和监管相结合，通过优化年检流程和年检手续，但严格审查年检项目的方式，对每一个参加年检的会员给予客观、公正且有针对性的评价。协会希望通过年检，达到规范事务所经营行为、优化市场主体、

维护市场秩序的目的，切实加强事务所监督管理，进一步提高协会服务水平，促进我国数据分析行业健康发展。

写在最后：

随着数据分析行业在全国各地区、各行各业获得越来越广泛的认可，数据分析业务日益增加，这给从事数据分析行业的会员们带来机遇的同时，更是要求事务所会员和执业的个人会员们用真实、准确、客观、公正的专业态度和专业技能，获得需求方和社会的认可与信赖，这样才能保证数据分析行业和事务所都能健康且长久的发展。

fin

关于推动各地举办数据分析行业推广活动的通知

中商联数据分析委 [2013] 007号

各地项目数据分析师培训机构、项目数据分析师事务所：

近年来，随着全球进入大数据时代，数据分析的重要性越来越突显。全球500强企业中，90%以上的重要投资和经营决策都取决于充分的数据分析支持。在欧盟、美国、日本等发达地区，数据分析普遍被作为运营决策的前提要素，被视为影响政府和企业进行企业管理、投资分析、信贷融资、企业战略规划、决策管理等重大决策的依托。

随着中国市场经济发展过程深化，企业未来的科学性决策将越来越多的依赖数据分析，有远见的企业家逐渐认识企业数据化程度的重要，认识到未来成功经营的关键是专业的数据分析。特别是在金融、能源、制造、零售、IT、文化等热点行业，很多企业已经将数据分析作为企业发展最为重要的依托。

各地项目数据分析师培训机构、项目数据分析师事务所是行业发展的主干力量，近年来在所在区域开展了多种多样的业务宣传工作，形成了良好的市场宣传作用。但受推广力度、专业性或行业权威性等方面的限制，在各地并没有形成较大的反响。所以，不少机构提出如果由协会主导，在各地组织专业性活动，不仅在权威性、专业性上有保证，而且可以充分利于地方机构的资源优势，使活动产生更好的效应。

通过这种方式，会有力的推进地方政府、企事业单位对数据分析重视程度，激活当地的人才培养工作，拉动企业对经营数据分析的真实需求。

为此，我会在2013年年初对此建议进行了深入探讨，认为这种活动可以起到提升行业品牌的作用，有利于地方机构业务的顺利开展。故协会根据2013年工作的总体方针，响应行业呼声，采取协会补贴、地方策应的方法，正式启动各地举办数据分析行业推广活动，活动可以采取CPDA数据分析专业沙龙或高端行业论坛的形式。

联系人：杨浩然

联系方式：010-65670421转105

中国商业联合会数据分析专业委员会

二零一三年五月十四日

关于举办2013年度数据分析行业会员执业教育的通知

中商联数据分析委 [2013] 008号

所有协会个人会员：

近年来随着我国经济的快速增长，数据分析行业得到了各行各业的广泛认可，“数据决策”意识的日渐普及，数据分析应用领域不断扩大，企业依托数据分析支持经营决策的需求逐渐增加，市场对数据分析专业人士的要求越来越高。为了满足社会对数据分析人才专业化的要求，经过几年的专业研究及探索，项目数据分析师课程在2013年正式进行调整，新课程与老课程体系存在着非常大的差别，特别是新课程引入了量化经营这个全新科目。由于新课程更加贴近现今社会的需求和数据分析行业的发展，很多协会会员希望能够通过学习新课程提升执业技能水平。

鉴于会员的从业需求和行业发展的实际情况，根据《会员管理办法》和《会员执业教育管理办法》的相关规定，协会将面向全国所有个人会员举办“2013年数据分析行业个人会员执业教育”。本次执业教育将以量化经营和企业案例实操两部分内容为主，旨在通过执业教育提高个人会员职业素养和专业技术水平，帮助会员掌握企业经营决策数据分析技术。

2013年数据分析行业个人会员执业教育安排如下：

- 一、执业教育时间：2012年6月20日——9月30日
- 二、参加对象：全国范围内数据分析行业个人会员
- 三、举办形式：以远程课程为主，在重点地区提供面授课程
- 四、执业教育内容：

本期执业教育内容包括：企业经营决策知识体系与技术、业务营销实操两部分会员广泛关注的课程。

五、执业教育要求：

- 1、每名个人会员均需参加本次执业教育课程学习；
- 2、会员可以选择参加远程课程或面授课程，一般默认为学员选择参加远程课程，如希望参加面授课程须提前向会员处申请，由会员处统一安排；
- 3、会员应在2013年9月30日前完成年度执业教育课程学习；
- 4、2013年行业会员执业教育实施方案详见协会官网相关通知。

中国商业联合会数据分析专业委员会

二零一三年五月三十日

附件

关于远程执业教育课程学习的说明

- | | |
|--|---|
| <ol style="list-style-type: none">一、执业教育时间：2013年7月1日——9月30日二、参加对象：数据分析行业所有个人会员三、执业教育网址：www.chinacpda.org（从协会官网点击“2013年个人会员执业教育”窗口进入学习页面）四、执业教育内容：<ol style="list-style-type: none">1、数据分析针对企业经营决策环节中客户数据分析、企 | <ol style="list-style-type: none">业品牌研究、新产品开发决策及主要营销组合决策中的应用；2、企业生产管理环节中的生产数据分析需求、生产计划和产品混合决策、分厂选择模型、经济生产量模型EPQ、设备最优化使用寿命决策、经验曲线模型；3、从事企业经营决策业务的事务所的经验分享；4、新加入行业的事务所还可以补学基本知识：市场调研 |
|--|---|

方案设计、制作复杂结构的计划完成表、Excel图表实战技巧。

五、登录方式：

会员在内登录执业教育的远程课程页面，才可以参与学习，登录方式为：



用户名：个人会员证档案编号（协会会员证照片下方）

密码：初始密码为“000000”，进入后请及时修改个人登录密码。

六、远程课程学习流程：

- 1、会员按上述“登陆方式”登陆远程学习页面进行学习；
- 2、远程课程的视频可反复观看，但仅限会员学习，未经协会同意，会员不得录制或外传包括课程在内的所有执业教育资料；

3、为掌握会员参与执业教育的情况，请会员在学习过后，完成在线测试，测试可做三次，以最高分为最后成绩；成绩在60分以上视为通过本次执业教育，60分以下则需要参加补考；

4、本次执业教育远程课程对个人会员免费开放；如会员需要教材、书籍、讲义等，费用自付。

详见协会官网相关通知。

5、本次执业教育截止时间为2013年9月30日。

fin

附件

关于参加面授执业教育课程学习的说明

为了更好的帮助会员学习企业经营决策的数据分析技术、提高业务实操，除远程教育课程外，协会还协调了多方资源，在行业发展重点地区开办执业教育面授课程。协会希望开办面授班地区的会员能够尽量参加面授课程学习，以远程课程为辅助学习方式，以达到更好的学习效果。对于在事务所里执业的个人会员，以事务所为单位，由1-2名骨干人员参与面授课程学习，在面授结束后，骨干人员带领事务所其他分析师进行远程网络学习，以达到真正掌握新知识、提高业务水平、拓展业务领域的目的。

一、面授课程学习时间：2013年7月份

二、举办面授课程学习的地区：

北京、陕西、四川、上海、广东、云南、山东、湖北。具体请查看各地区面授课程开办时间表，详情请登陆协会官网查询。

三、申请面授课程学习的流程

- 1、查询当地或周边地区面授课程开班时间；
- 2、填写《面授课程申请表》，提交至协会会员处；

- 3、会员处根据申请班次的具体情况，安排会员参与学习；
- 4、参加面授课的会员须凭听课证参加学习；

5、会员参加面授课程后，可通过远程课程对知识点继续复习和再学习；

6、会员参加面授课程后，应在2013年9月30日前登陆远程学习页面，完成在线测试，测试可做三次，以最高分为最后成绩；成绩在60分以上视为通过本次执业教育，60分以下则需要参加补考；

7、所有参加面授课程的会员均须在2013年6月20日前将申请表提交到协会会员处；

8、本次执业教育面授远程课程对个人会员免费开放，如会员需要教材、书籍、讲义等，费用自付；

9、未经协会同意，会员不得录制或外传包括课程在内的所有执业教育资料。

fin



新课程发布会现场

2013项目数据分析师新课程发布

◎ 文 / 中国经营网 夏欣 图 / 崔峻珩

4月18日协会在京召开了“2013项目数据分析师新课程发布会暨渠道招募大会”。新课程打破了原有的量化投资分析技术体系，深化企业决策数据分析知识，细化量化经营的内容，强调企业决策全面数据化分析和量化经营的重要性，同时也更加注重实践操作，关注学员的实战体验。

新课程的知识体系接轨与国家数据分析技术，能够满足国内企业对数据分析的需求，为企业提供专业有效的数据分析服务。

项目数据分析师认证课程自2003年在全国各地区推广以来，培养了万余名专业的数据分析师人才。但随着大数据概念的推广，越来越多的企业开始关注数据的价值，并在企业经营决策的各个环节中运用数据分析技术，企业对数据分析人才的需求渐趋于井喷之势。为了满足企业运营中的实际需求，协会成立了专家组对项目数据分析师认证课程进行了

全面的更新与完善。

项目数据分析师培训认证项目的新课程作为国内唯一在将企业决策体系全面引入量化研究与分析的课程，能帮助企业应用先进的数据分析技术进行经营决策与管理，有效提升企业核心竞争力，也为协会推出的“2013企业经营决策数据分析服务年”活动提供了强大的技术支撑。

邹东生会长表示：企业在大数据时代，会面临更加咄咄逼人的竞争对手。谁能更好地利用好大数据带来的决策信息，做出更好更快的市场决策，谁才有可能最后的胜者。但是，在经营数据分析方面起步较晚的企业都会面临：技术和工具不足，人才短缺，外部信息库不健全，数据安全和隐私保护风险等挑战。面对这些挑战，企业必须借助专业数据分析和数据分析师事务所的协助，逐渐提升数据分析的效用。而我国目前注册的数据分析师事务所百余家，取得认证

的项目数据分析师一万多人，执业注册数据分析师近千人，与行业发展相比，人才培养工作出现滞后。

发布会上还公布了全国培训机构考培合作方案，协会希望在2013年以开放式的合作方式带动全国项目数据分析

师人才培训的推广，并将辅助各地区授权管理中心积极开展学员沙龙、论坛、技术研讨、案例研究等多类型活动，吸引更多人才成为项目数据分析师，促进地方数据分析行业的发展。 

附件

关于项目数据分析师认证培训新课程体系的说明



数据分析课程系列教材



远程学习中心



课程讲义



项目决策数据分析软件



远程学习卡

新的项目数据分析师认证培训课程体系包括三个模块：项目投资分析，企业经营决策分析，以及数据分析基础。

项目投资分析是对项目的可行性和盈利性的评估过程，可以帮助投资者和融资者更好地判断项目的优劣。任何项目开始投资启动前都必须经历可行性评估的过程，而这种可行性的评估需要遵循科学的方法。课程体系中的项目投资分析模块旨在系统地介绍投资项目分析的科学方法，从预测技术到经济效益指标分析，从现金流量表、资金平衡表的编制到项目数据分析报告的撰写。掌握了这套系统的方法，学员就可以为各种投资项目“诊脉”了。

企业经营决策分析是指运用量化的分析工具，去发现、分析、并解决企业在经营的各个环节中遇到的问题。企业经营包括战略、采购、生产、库存管理、物流管理、市场营销、人力资源管理、预算等多个环节。本模块除了

包括战略分析、生产决策、营销决策、采购决策、全面预算决策以外，还会介绍经营决策的评估方法。掌握了这个模块的内容，学员不但能全面了解企业经营的整个过程，还会使用量化分析工具帮助企业解决经营方面的问题，为企业的战略制定提供强有力的数据支持。

第三个模块涵盖数据分析的基本概念、主要模型、和分析方法，对数据分析的步骤、方法、模型进行系统介绍，包括如何对数据进行采集、处理、分析、报告，如何撰写项目建议书，以及如何使用Excel和SPSS等常用数据分析软件等等。





用Excel对会员客户交易数据进行RFM分析

◎ 文 / 赵兴峰 图 / 崔峻珩

随着越来越多的企业开始意识到数据分析对企业的价值，在企业需求与日俱增的同时，企业对项目数据分析师的从业素养的要求也有了较大提高。为了帮助从业分析师更好地理解和掌握企业经营决策数据分析技术，协会将开始就企业经营决策数据分析技术做系列讲解。

本期内容主题为《用Excel对会员客户交易数据进行RFM分析》。

背景：一个会员服务的企业，有近1年约1200个会员客户的收银数据。由于公司想针对不同类别不活跃客户进行激活促销；同时，为回馈重点客户，也计划推出一系列针对重点客户的优惠活动，希望保留这些客户，维持其活跃度。因此希望利用该数据进行客户分类研究。

根据客户的需求，RFM模型相对简单并且直接，按照R(Recency-近度)、F(Frequency-频度)和M(Monetary-额度)三个维度进行细分客户群体。由于该客户的数量较少(约1200个)，所以，采用 $3 \times 3 \times 3 = 27$ 个魔方(1200/27=44左右)较为合适，虽然平均每类客户数量较少，考虑到集中度分布情况，数量多的分类也能够有200-300左右，适合针对会员客户进行短期的电话、短信营销或者信函营销的数量。

RFM模型原理：

RFM模型是一个简单的根据客户的活跃程度和交易金额贡献所做的分类。因为操作简单，所以，较为常用。

近度R：R代表客户最近的活跃时间距离数据采集点的时间距离，R越大，表示客户越久未发生交易，R越小，表示客户

越近有交易发生。R越大则客户越可能会“沉睡”，流失的可能性越大。在这部分客户中，可能有些优质客户，值得公司通过一定的营销手段进行激活。

频度F：F代表客户过去某段时间内的活跃频率。F越大，则表示客户同本公司的交易越频繁，不仅仅给公司带来人气，也带来稳定的现金流，是非常忠诚的客户；F越小，则表示客户不够活跃，且可能是竞争对手的常客。针对F较小、且消费额较大的客户，需要推出一定的竞争策略，将这批客户从竞争对手中争取过来。

额度M：表示客户每次消费金额的多少，可以用最近一次消费金额，也可以用过去的平均消费金额，根据分析的目的不同，可以有不同的标识方法。一般来讲，单次交易金额较大的客户，支付能力强，价格敏感度低，是较为优质的客户，而每次交易金额很小的客户，可能在支付能力和支付意愿上较低。当然，也不是绝对的。

RFM的分析工具有很多，可以使用SPSS或者SAS进行建模分析，然后深度挖掘。IBM SPSS还有个Modeler，有专门的RFM挖掘算法供使用。本文为了普及，介绍使用Excel（2007版）做初步的RFM分析。

操作步骤：

第一步：数据的清洗

原始数据集：数据请参考数据集Excel,大家可以下载练习。

该数据集共有26600多条数据，包含记录ID（数据库的primarykey）、客户编号、收银时间、销售金额、销售类型共5个字段

	A	B	C	D	E
1	记录ID	客户编号	收银时间	销售金额	销售类
26645	00042394	040051	2010-07-03 20:29	299.00	正常
26646	00042491	010011	2010-07-04 16:54	449.00	正常
26647	00042492	030266	2010-07-04 16:54	0.00	赠送
26648	00042493	010240	2010-07-04 16:55	249.00	正常
26649	00042494	040144	2010-07-04 16:55	0.00	赠送
26650	00042495	030291	2010-07-04 17:00	20.00	正常
26651	00042496	030045	2010-07-04 17:00	20.00	正常
26652	00042501	030093	2010-07-04 17:56	449.00	正常
26653	00042514	030031	2010-07-04 20:08	449.00	正常
26654	00042808	020088	2010-07-05 16:02	-249.00	退货

通过简单的筛选，可以看到，在交易金额中有0消费额，有负数消费额，继续查看交易类型，发现为“赠送”和“退货”所造成。这些数据在本次分析中用不到，所以在数据处理时需要通过筛选除去。

Excel操作：

鼠标点击第一行的行标“1”以选中第一行

菜单栏点击“数据”，快捷按钮栏点选“筛选”

鼠标点击“销售类型”筛选下拉按钮，可以看到所有数据集中有的销售类型

点选“销售金额”字段的小角标也可以看到有负数出现。

第二步：数据处理

根据分析需要，R用客户最后成交时间跟数据采集点时间的时间差（天数）作为计量标准；F根据数据集中每个会员客户的交易次数作为计量标准（1年的交易次数）；M以客户平均的交易额为计量标准。通过Excel的透视表即可计算以上RFM数据。

Excel操作:

菜单栏点击“插入”

快捷按钮栏点击“透视图”

如右图所示

选择数据区域，确认所有的数据都被选择

选择在“新工作表”中插入数据，然后点击“确定”

将“客户编号”拖入“行标签”栏

将“收银时间”、“记录ID”、“交易金额”拖入数值计算栏

点击“收银时间”数值计算栏按钮，选择“值字段设置”

在“计算类型”中选择“最大值”

在对话框左下角，点击“数字格式”，设定时间格式为：yyyy-mm-dd，然后“确定”

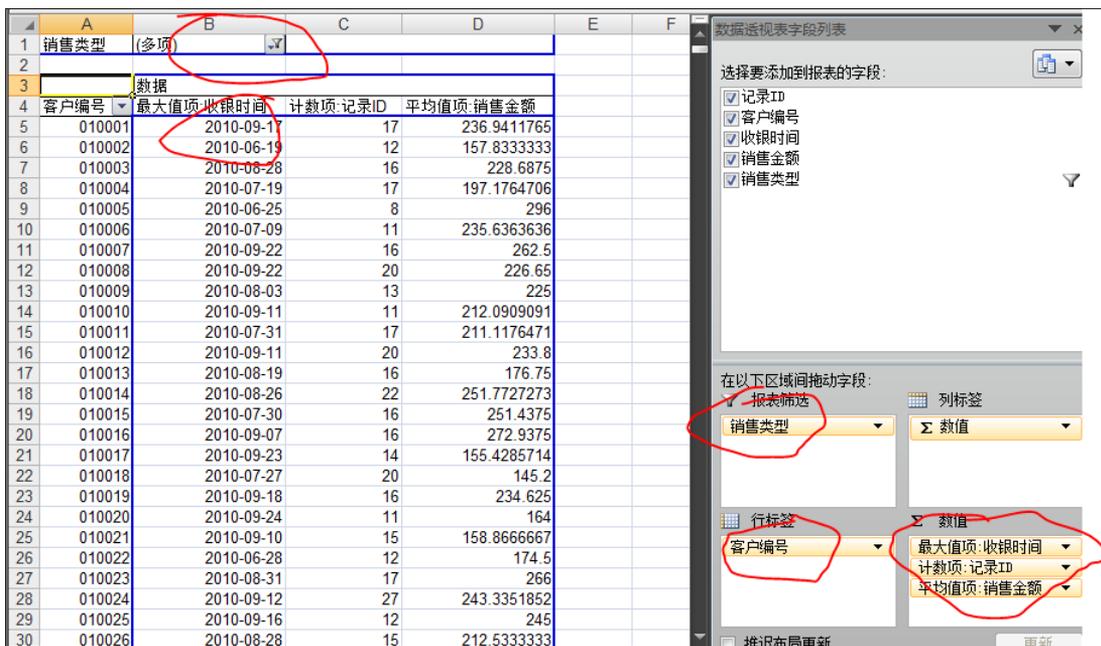
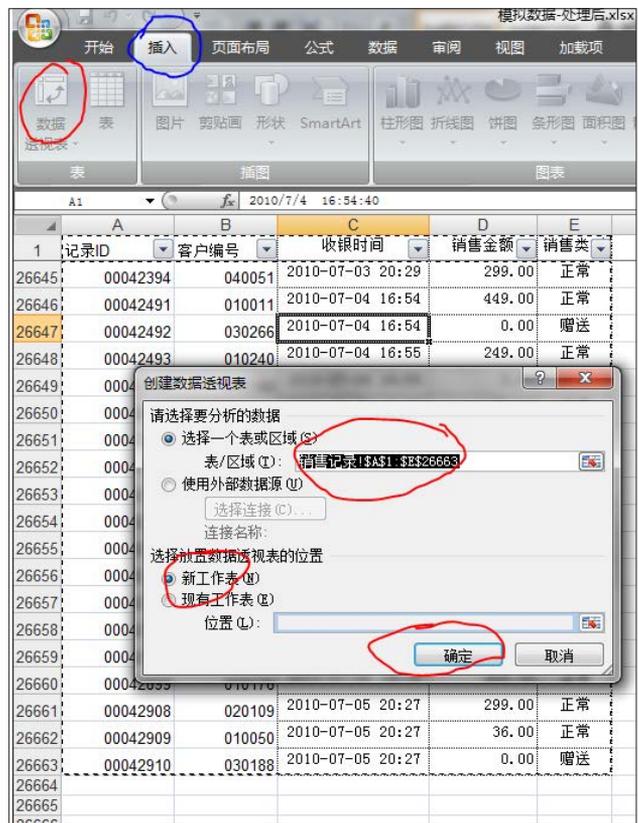
点击“销售金额”数值计算栏按钮，选择“值字段设置”

在“计算类型”中选择“平均值”，然后“确定”

在“记录ID”数值计算按钮栏，选择“值字段设置”

在“计算类型”中选择“计数”，然后“确定”

在透视表顶部筛选项“销售类型”处，点击下拉按钮小角标，在“选择多项”前的小方框中打勾，然后点掉“退货”和“赠送”前的勾，然后“确定”会得到如下结果。



以上我们得到了：

1) F值：客户这1年共消费了多少次

2) M值：客户每次交易的平均消费金额

但是，R值还需要做些处理。目前R值只得到的是客户最近一次消费日期，需要计算距离数据采集日期的天数。

Excel操作：

鼠标拉选列标签ABCD,选中透视表所在的四列

按ctrl^C（复制），点击“开始”菜单栏下，快捷按钮栏“粘贴”下的小下拉三角标，选择“粘贴值”【或者点“选择性粘贴”，然后选择粘贴值】，用单纯的数据形式覆盖原有透视表。

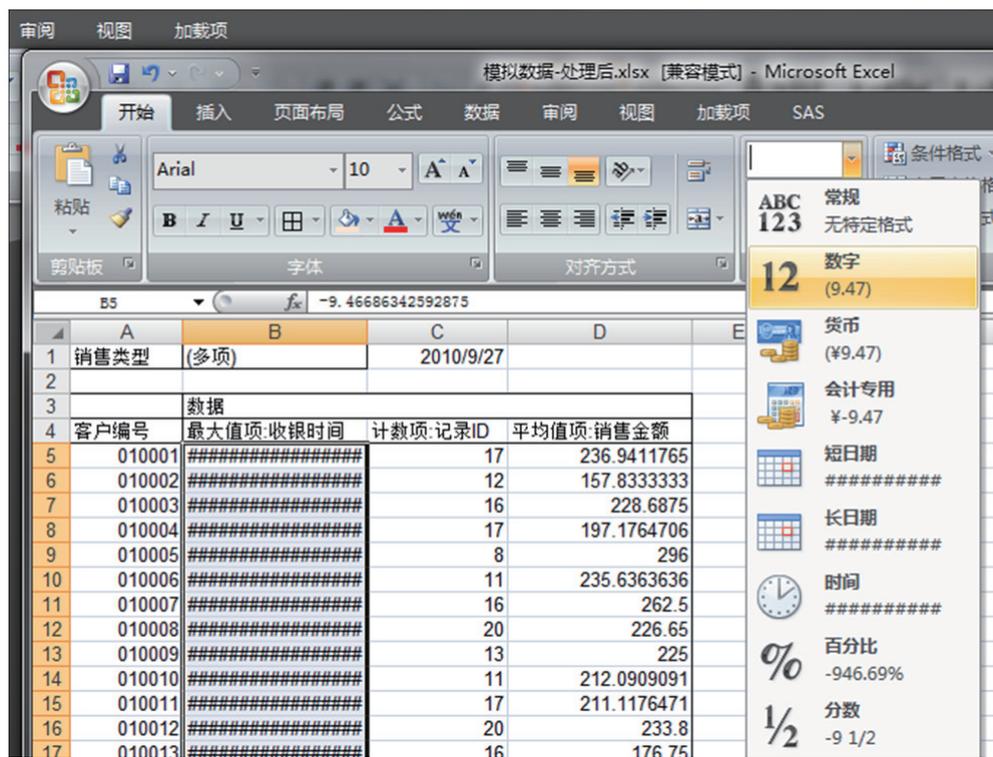
在C1单元格中输入数据采集日期2010-09-27，格式为yyyy-mm-dd

然后选中C1单元格，复制其中内容

选中B5：B1204【快捷操作：点中B5,同时按住Shift^Ctrl后点击向下箭头，松开ctrl键，继续按住shift键，按一次向上箭头，取消数据最后一行的汇总数据】

点击“开始”菜单栏下快捷按钮栏上的“粘贴”按钮下方的下拉箭头，选择“选择性粘贴”，在对话框中勾选“减”，然后“确定”

在不取消目前选择的情况下，选择“开始”菜单栏下快捷按钮栏上的格式化工具栏，选择“数字”



因为得到的数据为最后交易日期 减去 数据采集日期的天数，是负值，所以，还需要处理。

在D1单元格中输入-1

然后ctrl^C复制D1单元格中的值（-1）

然后选中B5：B1204【快捷操作同上】

“开始” - “粘贴” 下拉按钮 - “选择性粘贴” - 在计算部分选择“乘”，然后点击“确定”

最后得到：

	A	B	C	D	E	F	G
1	销售类型	(多项)	2010/9/27	-1			
2							
3		数据					
4	客户编号	最大值项: 收银时间	计数项: 记录ID	平均值项: 销售金额			
5	010001	9.47	17	236.9411765			
6	010002	99.12	12	157.8333333			
7	010003	29.33	16	228.6875			
8	010004	69.25	17	197.1764706			
9	010005	93.39	8	296			
10	010006	79.45	11	235.6363636			
11	010007	4.38	16	262.5			
12	010008	4.26	20	226.65			
13	010009	54.06	13	225			
14	010010	15.28	11	212.0909091			
15	010011	57.16	17	211.1176471			
16	010012	15.18	20	233.8			
17	010013	38.51	16	176.75			
18	010014	31.34	22	251.7727273			
19	010015	58.44	16	251.4375			
20	010016	19.14	16	272.9375			
21	010017	3.40	14	155.4285714			
22	010018	61.43	20	145.2			
23	010019	8.49	16	234.625			

到此，我们得到R,F,M针对每个客户编号的值

第三步：数据分析

R-score, F-score, M-score的值，为了对客户根据R,F,M进行三分，我们需要计算数据的极差（最大值和最小值的差），通过对比R（或者F,M）值和极差三分分距，来确定R（或者F,M）的R-score, F-score, M-score。

所以先计算R、F、M的最大值、最小值、极差三分分距

Excel操作：

	A	B	C	D	E	F	G	H
1	销售类型	(多项)	2010/9/27	-1	最大值	266.47	27	354.3
2					最小值	1.11	1.11	92
3		数据			极差三分分距	88.45164	8.629051	87.43333
4	客户编号	R	F	M	R-score	F-score	M-score	RFM-score
5	010001	9.47	17	236.9411765				
6	010002	99.12	12	157.8333333				
7	010003	29.33	16	228.6875				
8	010004	69.25	17	197.1764706				
9	010005	93.39	8	296				
10	010006	79.45	11	235.6363636				

F1到H1代表RFM的最大值，利用公式“=max(B5:B1204)”计算，（计算F时B换成C，M时B换成D即可）

F2到H2代表RFM的最小值，利用公式“=min(B5:B1204)”计算（计算F时B换成C，M时B换成D即可）

F3到H3代表RFM的极差三分分距，利用公式“=(F1-F2)/3”计算（计算F时F换成G，M时F换成H即可）

【以上快捷操作可用，先输入F1,F2,F3单元格里的公式，选择F1：F3三个单元格，然后拉动右下角的黑色小十字架，向右拖动复制F列公式到G和H列即可】

R-score的计算公式为：

E5单元格内输入：“=IF(ROUNDUP((B5-\$F\$2)/\$F\$3,0)=0,1,ROUNDUP((B5-\$F\$2)/\$F\$3,0))”

之所以使用IF判断函数，主要是考虑到当R值为最小值时，roundup(B5,0)为0，用if函数判断如果是0，则强制为1。

之所以用\$F\$2锁定引用的单元格，是为了后续的公式复制，最小值和极差三等分距不会发生相对引用而变化位置【锁定引用单元格除了手工添加\$符号外，快捷方式是选中引用的单元格按F4快捷键，此处都比较麻烦，手工输入\$符号还快些】

【另外一种简单的处理方式就是直接用公式“=ROUNDUP((B5-\$F\$2)/\$F\$3,0)”，然后用ctrl^H快捷操作，将0值替换成1即可，这个替换需要将公式复制-快捷粘帖为数值后进行】

F-score和M-score如法炮制。

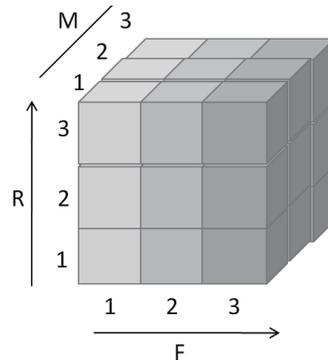
F5单元格公式为：=IF(ROUNDUP((C5-\$G\$2)/\$G\$3,0)=0,1,ROUNDUP((C5-\$G\$2)/\$G\$3,0))

G5单元格公式为：=IF(ROUNDUP((D5-\$H\$2)/\$H\$3,0)=0,1,ROUNDUP((D5-\$H\$2)/\$H\$3,0))

RFM-score的计算，利用分别乘以100-10-1然后相加的方式，让R、F、M分别为一个三位数字的三个百分位、十分位和个位表达，该三位数的三个位代表了3x3x3=27魔方三个维度上的坐标。

H5单元格的公式为：=E5*100+F5*10+G5

选中E4到H4区域，双击右下角小黑色十字架，复制E4到H4公式到所有客户数据中得到结果如下：

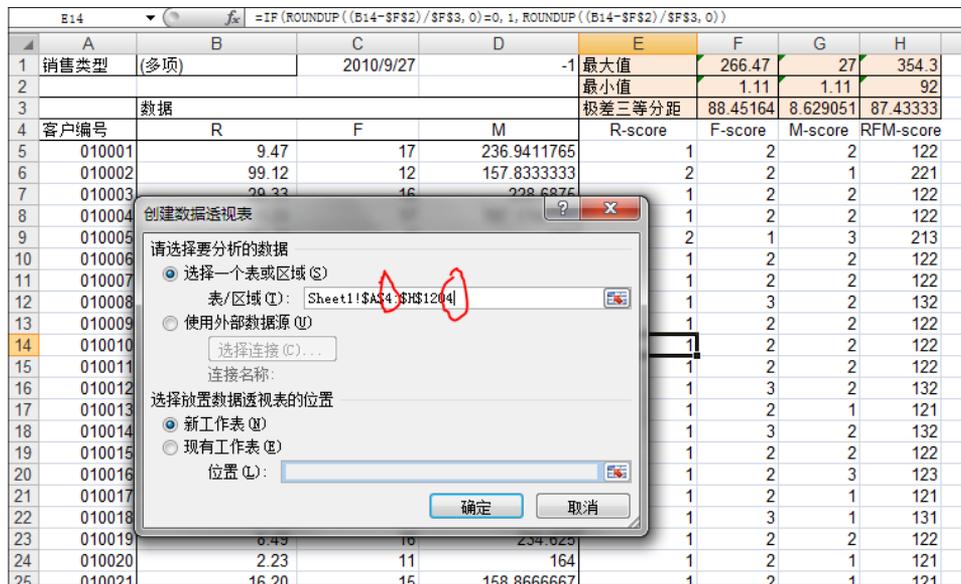


B5 =IF(ROUNDUP((B5-\$F\$2)/\$F\$3,0)=0,1,ROUNDUP((B5-\$F\$2)/\$F\$3,0))									
	A	B	C	D	E	F	G	H	I
1	销售类型	(多项)	2010/9/27	-1	最大值	266.47	27	354.3	
2					最小值	1.11	1.11	92	
3		数据			极差三等分距	88.45164	8.629051	87.43333	
4	客户编号	R	F	M	R-score	F-score	M-score	RFM-score	
5	010001	9.47	17	236.9411765	1	2	2	122	
6	010002	99.12	12	157.8333333	2	2	1	221	
7	010003	29.33	16	228.6875	1	2	2	122	
8	010004	69.25	17	197.1764706	1	2	2	122	
9	010005	93.39	8	296	2	1	3	213	
10	010006	79.45	11	235.6363636	1	2	2	122	
11	010007	4.38	16	262.5	1	2	2	122	
12	010008	4.26	20	226.65	1	3	2	132	
13	010009	54.06	13	225	1	2	2	122	
14	010010	15.28	11	212.0909091	1	2	2	122	
15	010011	57.16	17	211.1176471	1	2	2	122	
16	010012	15.18	20	233.8	1	3	2	132	
17	010013	38.51	16	176.75	1	2	1	121	
18	010014	31.34	22	251.7727273	1	3	2	132	
19	010015	58.44	16	251.4375	1	2	2	122	
20	010016	19.14	16	272.9375	1	2	3	123	
21	010017	3.40	14	155.4285714	1	2	1	121	
22	010018	61.43	20	145.2	1	3	1	131	
23	010019	8.49	16	234.625	1	2	2	122	
24	010020	2.23	11	164	1	2	1	121	
25	010021	16.20	15	158.8666667	1	2	1	121	
26	010022	90.29	12	174.5	2	2	1	221	
27	010023	26.25	17	266	1	2	2	122	
28	010024	14.23	27	243.3351852	1	3	2	132	
29	010025	10.25	12	245	1	2	2	122	
30	010026	29.27	15	212.5333333	1	2	2	122	

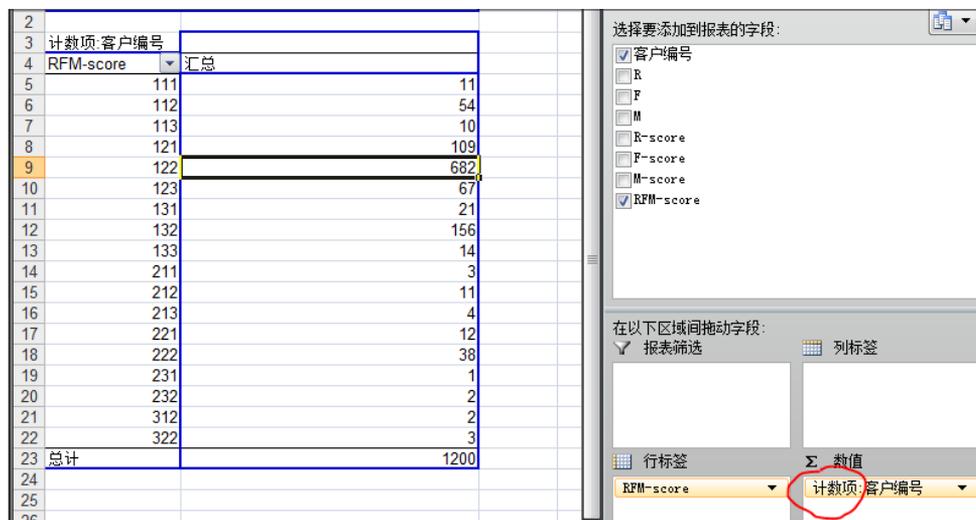
接下来的步骤就是统计各个魔方上的客户数量
再次利用透视表形成统计结果

Excel操作：

“插入”菜单栏下快捷按钮栏按“透视表”，在数据表区域中选择A4: H1204【确认这个选择，自动跳出来的区域要改一下的哦】，然后点击“确定”



将RFM-Score拖入“行标签”中，将“客户编号”拖入“数值计算”栏中，点击“数值计算”栏中的“客户编号”项，选择“字段数值设置”，选择计算方法为“计数”，得到处理结果如下：



第四步：数据分析结果解读和可视化

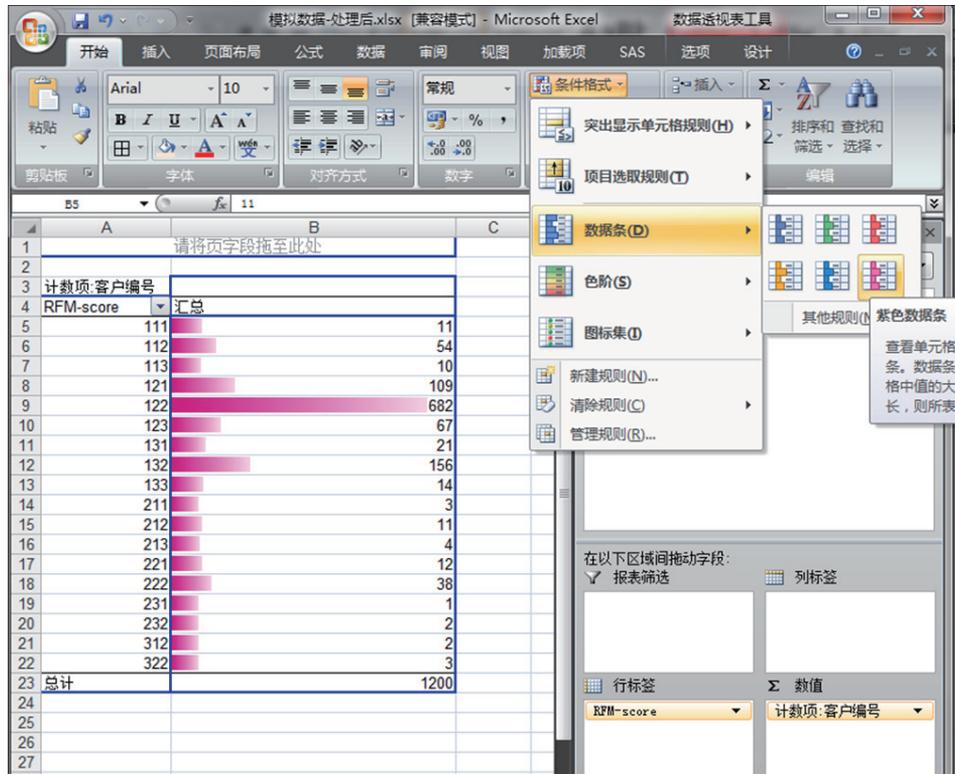
得到这个分析结果，利用Excel的条件格式功能可以对得到的数据分析结果做简单的可视化。

Excel操作:

将透视表中B列拉宽

选中B5:B22列

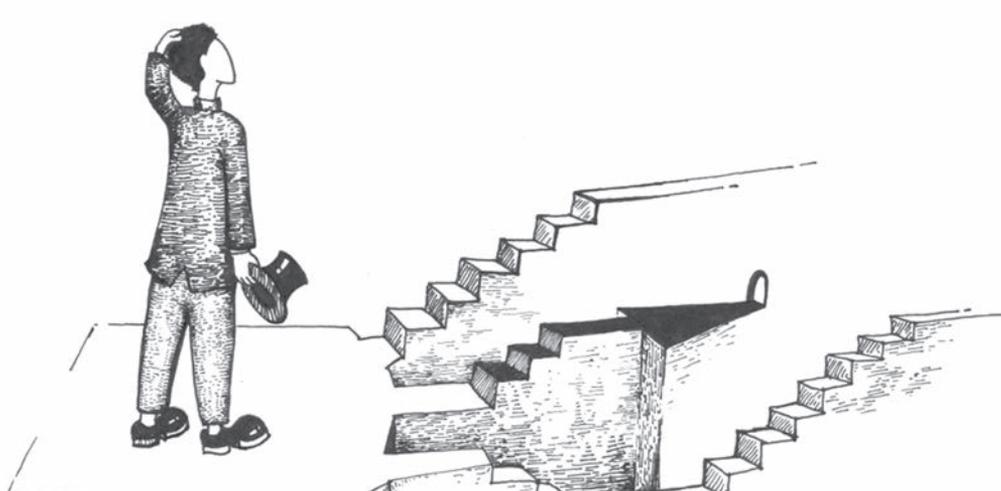
“开始”菜单栏下快捷按钮栏点击“条件格式”下拉菜单中选择“数据条”，然后选择一个颜色即可通过条形图的视觉化，可以直观地对比哪类客户数量较多。



第五步：数据分析结果的商业解读（略）

因为此数据为模拟数据，且数据分析过程中有许多需要商业活动参与执行者参与的过程。此案例仅仅是一种操作的演示。具体的分析结果解读，读者可以自行参考。

祝大家生活愉快！





数据科学家炙手可热

© 文 / 51Callcenter.com 图 / 崔峻珩

埃森哲称：

“如果对数据分析的需求加速升高，那么中国也将可能出现短缺。”

《哈佛商业评论》(Harvard Business Review)近期声称，21世纪最性感的工作是数据科学家。这一美国商学院期刊表示，数据科学家集“数据黑客、分析师、沟通大师和受信任的顾问”于一身，并指出，这种技能的结合极为罕见。

这正是全球各地诸多企业的问题所在。尽管公司经理深知大数据所能带来的效益，但他们难以找到拥有合适技能的人才。

利用大数据的潮流毫无放缓迹象。管理咨询公司埃森

哲(Accenture)去年调查了600家美国和英国公司，结果发现有三分之二的公司在之前18个月任命了负责数据管理和分析工作的高管。即使是尚未设立此类高管职位的公司，也有71%准备在不久的将来作出任命。

招聘顾问也表示，对数据分析专家的需求正在飙升。专注信息技术领域、在伦敦和阿姆斯特丹设有办事处的Cititec表示：“今年前六周，我们收到的大数据招聘请求与之前六个月一样多。我们估计，该数字今年将比去年高出100%，甚至更多。”

的确，大数据正快速成为IT招聘机构的重要专长。“我们现在拥有一位大数据专家，而一年前还没有，”Cititec补充道，“这是一块竞争激烈的市场，有很多公司竞相争夺资深人才。”

埃森哲北美金融服务数据分析部门执行董事布莱恩·麦卡锡(Brian McCarthy)认为，全球及各行各业的需求意味着，数据分析技能严重供不应求，尤其是在美国和英国。

他说：“背景适合从事数据科学家工作——计算机科学、统计学、机器学习——的毕业生正在走出校园，但他们的数量还不够。”

麦卡锡补充道，尽管许多公司转而聘请合同工——埃森哲调查表明，有近60%的公司求助于外部分析师和咨询顾问——但它们仍无法找到需要的人才。

Cititec称，这种紧缺十分严重，以至于英国承接IT外包工作的资深数据架构师或业务分析师可日赚500英镑至650英镑。在该领域资质最为优秀的一端，博士级别的数据科学家能够拿到9.5万英镑的高额年薪。

“甚至连海外外包也无法解决这个问题，”麦卡锡指出，“因为印度、中国和巴西等新兴经济体没有足够的资深人才。”

人才紧缺似乎还将持续数年。

埃森哲在今年发布的《数据分析在行动：通向高投资回报率之路的突破与壁垒》(Analytics in Action: Breakthroughs and Barriers on the Journey to ROI)报告中预测，到2018年，光是在美国和英国，需要具备高深科学、技术、工程和数学(STEM)知识的职位的增长速度将是其他职业的五倍，是金融服务等信息密集型行业职位的四倍。

报告称，新兴经济体培养出STEM人才的数量高于发达经济体，但仍无法满足全球的潜在需求。

确实，埃森哲卓越绩效研究院(Institute for High Performance)在去年的一次调查中研究了美国、中国、印度、英国、日本、巴西和新加坡对数据分析经验的需求。调查发现，到2015年，所有这些国家(除中国外)都将面临胜任分析科学家工作的博士毕业生数量净短缺的问题。

“美国、英国、日本、新加坡和巴西几乎将肯定遭遇高端人才的严重短缺，”埃森哲表示，“尽管印度的数据分析服务行业蓬勃成长，但它也将难以培养出足量的博士生来填补所有新的数据分析科学家岗位。”

例如，美国新增数据分析专家职位的数量将占全球新增总量的44%，但该国只能供应23%的人才，导致近3.2万



人的缺口。

只有中国似乎出现了少量的过剩，但埃森哲警告称：“如果对数据分析的需求加速升高，那么中国也将可能出现短缺。”

不过，有志学习数据分析技能的IT职业人士面临着充足的培训机会。IBM和甲骨文(Oracle)等大型IT厂商，以及独立的IT专业培训机构，均提供在线和课堂式课程。

大数据创业公司也提供培训。例如，基于Hadoop框架设计的软件和服务提供商Cloudera有自己的“Cloudera大学”，而竞争对手MapR也拥有“MapR学院”。

Cititec称：“优秀的数据分析师应当拥有从事数据分析的必要技能，可以通过学习或在职培训获得这些技能。IT专业人士开始意识到大数据领域与日俱增的重要性，他们在问：我如何获得这些技能？”

但埃森哲的麦卡锡认为，面向新的数据分析领域对IT专家进行再培训，并不像给现有员工上课那样简单。

“在职培训正在一定程度上缓解资深人才短缺的问题，”他说。

“但问题是，这不只是学习新工具——人们还需要改变思维模式……这对IT企业造成了更大的挑战，尤其是在它们的员工伴随90年代和新世纪头10年的技术成长起来的情况下。在很多情况下，引进新人才要比重新培训现有员工更加容易。”

fin

Netflix赌赢《纸牌屋》 背后的秘密武器：大数据分析

◎ 文 / 数据中心 郭伟峰 图 / 崔峻珩



继美剧《新闻编辑室》(Newsroom)走红之后，又一部描述美国政坛宫斗的《纸牌屋》(House of Cards)再次火了一把。投资方Netflix甚至在《纸牌屋》开播前就知道该剧一定会火，秘密就来自大数据分析。

作为全球最大的流媒体运营商，《纸牌屋》是Netflix首

次自行制作并且完全依靠网络发行的电视剧，借助于这部剧的热播，帮助Netflix在订阅用户数量上成功超越了传统的HBO电视网。Netflix新增用户数305万人，第一季度用户观看的视频总时长超过40亿个小时。

然而，《纸牌屋》的走红并非偶然，也不是意料之外的

事，它是大数据分析在具体行业成功应用的经典案例。投资方Netflix甚至在《纸牌屋》开播前就知道该剧一定会火，秘密就来自大数据分析。美国新闻网站Salon.com曾这样描述，用户只要登录Netflix，其每一次点击、播放、暂停甚至看了几分钟就关闭视频，都会被作为数据进入后台分析。这样一来，Netflix就能精确定位观众的偏好，比如“最爱Kevin Spacey”，或者“最爱政治剧”。Netflix在拍摄前事先分析了订阅用户们的观影数据和操作习惯，保证里其首部原创剧集可以精确命中大量的潜在观众。



《纸牌屋》电影改编自Michael Dobbs所著的同名政治惊悚小说，从1990年开始被BBC分三部曲搬上银幕，每一部分分为4集，每集时长57分30秒。首次被搬上银幕后即获得巨大成功，该剧2000年British Film Institute (BFI)评选的“100部英国最优秀的电视节目”中排名第84位。2010年英国《卫报》评选的“史上五十佳英美剧”中排名第18位。

Netflix在下决定投资翻拍《纸牌屋》电视剧版前至少做了两件与大数据分析紧密相关的事：

- (1) 挑选制作人和演员
- (2) 确定播放形式

对Netflix订阅用户数据的追踪和分析并不是听起来那么简单，对基于基础数据派生的扩展数据量比通常人能想到的大得多。这一过程绝不仅仅只是分析观众喜欢看哪些主题的电影和偏好，Netflix还会统计观众如何观看电影和观影过程，观影过程中暂停的次数，会在看到几分钟的时候关闭视频等等，这些操作都会被作为数据进入后台分析。过去，

Netflix只是用这些数据来做影片推荐。如今，Netflix会投其所好，根据这些内容拍摄用户感兴趣的电影。

通过数据分析，Netflix发现喜欢观看1990版《纸牌屋》的影迷们同时喜欢看导演David Fincher的作品，另外，他们会经常观看奥斯卡影帝Kevin Spacey的作品。因此新版《纸牌屋》邀请了David Fincher(制作人)和Kevin Spacey(男主演)加盟这部作品的翻拍并不是凭空想象，而是基于影迷数据分析得出的结论。

为此，Netflix击败HBO、AMC等竞争对手，为26集的原创新剧提前支付了1亿美元，以作为剧组预算。协议规定，Netflix拥有《纸牌屋》的独家播放权两年。这意味着，这部美剧用户只能在Netflix上付费观看。

另一个选择在播放形式方面，按照传统连载美剧的习惯，基本都是每周播放一集，而Netflix根据相关数据的分析，更多人不喜欢在固定时刻收看电视剧，而是“攒起来”，直到全集播放完毕再一次性看完。因此，Netflix这次选择了一次性的播放13集《纸牌屋》。

目前，在笔者看来，这部剧集还引发了亚马逊、微软等公司对投拍互联网剧集、消除正规电视剧与在线电视剧差异的兴趣。据悉，亚马逊正在制作至少11部试播剧集，这些剧集仅可通过互联网观看，原因是这家公司正在与Netflix展开“战争”，竞相利用人们对于在智能手机、平板电脑和互联网电视上观看电视节目的兴趣，以扩大自身在流媒体播放服务这一领域中的占有率。与此同时，微软也正在制作自己的电视连续剧，而且有报道称，谷歌、苹果、英特尔和Twitter也正在考虑开发自己的类似产品。

相关背景：

Netflix, Inc. 是一家在线影片租赁提供商，该公司成立于1997年，总部位于加利福尼亚州洛斯盖图。在美国、加拿大提供互联网随选流媒体播放，定制DVD、蓝光光碟在线出租业务。

《纸牌屋》(House of Cards)是一部以政治为题材的美国电视连续剧，根据Michael Dobbs的同名小说改编，由Beau Willimon制作。第一季于2013年2月1日在流媒体服务商Netflix首播。剧背景设置在在现今的华盛顿哥伦比亚特区，主要讲述了众议院多数党党鞭Frank Underwood，在晋升为国务卿的希望破灭后，决定对那些背叛他的人报复的故事。

fin



大数据分析的三大演进方向

◎ 文 / 互联网分析 图 / 崔峻珩

最近Gartner发布了2013年度BI和分析的魔力四象限图，同时Wikibon也发布了2013年大数据市场预测，两份报告都明确指出，随着分析正在成为企业IT的核心，昔日的BI-ETL-EDW分析范型已经完全落伍，不再适用。

2013开年不久，一连串的重大事件的发生标志着大数据和分析领域正在加速演进，对于数据分析专业人士和企业管理者来说，2013年是大数据进入企业应用的关键一年。

近日Alteryx公司总裁乔治马修(George Mathew, Twitter帐号@gkm1)与大数据领域的著名专家Mayank Bawa、Mike Olson和Scott Yara就数据分析的传统范型(BI-

ETL-EDW)即将被新的分析范型取代达成共识，几位专家认为全新的数据分析平台将消除当前分析软件在设计和实施方面的延迟和低效率，从根本上重新思考和定义三大阻碍企业数据分析应用的关键问题：数据管理、分析透明度以及用户应用。

以下是马修在博客中对新数据分析范型三大演进方向的

解读，IT经理网编译整理如下：

一、数据管理

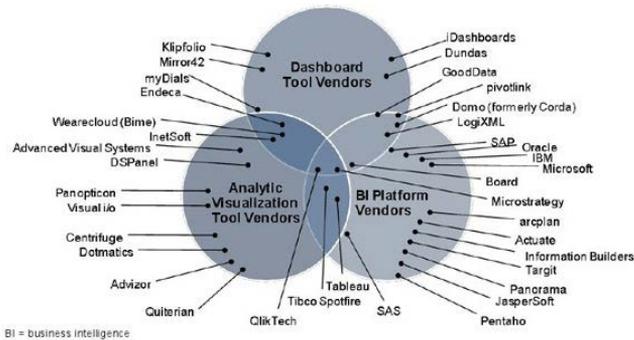
Hadoop已经成为企业管理大数据的基础支撑技术。最近随着Greenplum Pivotal HD、Hortonworks Stinger和Cloudera的Impala的发布，Hadoop的技术创新速度正在加快，上述Hadoop项目传递出一个非常明确的信号：主要的Hadoop发行商想要在Hadoop HDFS之上提供实时、互动的查询服务。这个趋势将两个领域的杰作整合到了一起：众所周知的SQL查询处理与具备指数级扩展能力的HDFS存储架构。

参考阅读：Hadoop发行版升级，NoSQL的未来是SQL?

二、去黑箱化

预测分析是管理者进行数据化决策的关键。目前预测和统计分析领域已经有很多技术可以帮助企业洞察不远的未来。但预测分析眼下面临的最大的问题是“黑箱”化。

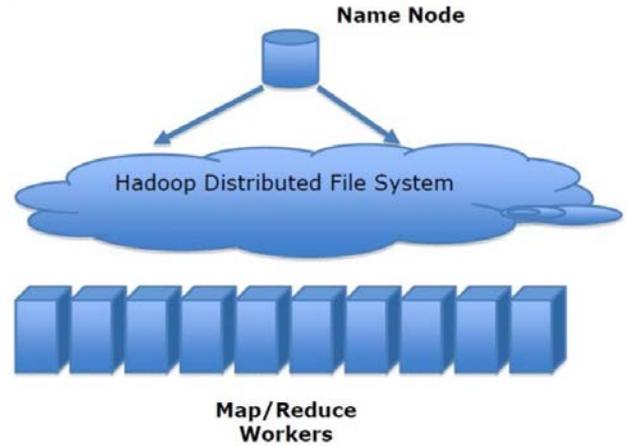
随着企业领导越来越多地以来预测分析技术做出重大商业决策，预测分析技术需要去黑箱化：包括应用自描述数据沿袭，增加对底层数学和算法解释等。“去黑箱化”有利于企业管理者学会彻底驾驭数据分析工具，不但看到数据分析结果，还知道分析是如何得来的，分析工具的设计原理等，这有助于管理者增加对预测分析的信心，而不是过去那样完全依靠“信仰”。



三、应用普及

即使实现了分析的去黑箱化，企业数据分析应用在企业中的部署依然面临以下几个方面的挑战：发布可复用应用，创建最佳实践、组织范围内的横向协作，无缝重组模型等。在最终用户(员工)中的应用普及是数据分析成功的关键。

例如建设一个专门提供分析应用的企业移动应用商店App



Store往往能大大加快数据分析的应用普及。

新数据分析范型的重要特征：

新的数据分析范型是目标导向的，不关心数据的来源和格式，能够无缝处理结构化、非结构化和半结构化数据。能够输出有效结果;能够提供去黑箱化的预测分析服务，能够面向更广泛的普通员工快速部署分析应用。

最近Gartner发布了2013年度BI和分析的魔力四象限图，同时Wikibon也发布了2013年大数据市场预测，两个报告都明确指出，随着分析正在成为企业IT的核心，昔日的BI-ETL-EDW分析范型已经完全落伍，不再适用。



新的分析范型正在崛起，以下是我们能看到的未来趋势：

Hadoop(和NoSQL)正在颠覆我们对PB级别大数据的管理方式。

R和Stata的崛起，正在冲击传统分析学术圈的的黑箱式分析方法，这也代表着商业世界的发展趋势。

分析应用将不再是数据科学家的专利，更多分析应用将以预先打包的内容和应用发送到分析人士和企业员工的手中。



云南鼎臻项目数据分析师事务所



云南鼎臻项目数据分析师事务所（以下简称“鼎臻事务所”）是经昆明市工商行政管理局登记注册，经中国商业联合会数据分析专业委员会备案核准，具有独立法人资格的专业性强、富有影响力的项目数据分析的服务机构。

鼎臻事务所是由注册项目数据分析师、高级会计师、投资银行业务、证券事务代表、注册会计师、注册资产评估师、房地产估价师、注册税务师、投资(咨询)工程师、律师、教授、博导等专业人士组成的专家团队，具有丰富的实战经验和专业知识经验。

鼎臻事务所主营是投资项目分析（评估、分析、规划、策划），企业经营分析，企业战略分析。致力于服务于生物制药、金融行业、公路桥梁、交通运输、房地产、煤炭行业、电子信息、有色金属。以数据分析为契机为客户提供增值的附产品，包括为企业提供融资（银行、信托、企业债、私募债、上市服务等）、内控构建规划、企业ERP工程等到经济咨询业务。

鼎臻的核心价值观是革故鼎新，臻于至善；鼎臻的使命是对客户负责，对员工负责，对社会负责，对股东负责；鼎臻的宗旨是以最低的成本为客户创造最大的价值，因为专业、因为执着、因为专注……

鼎臻事务所汇集了一批企业高管、金融高管、资深专家

团队，服务的行业企业定位准确，自己拥有多个行业的数据仓库，以实战为主，具有领先的数据分析处理工具与能力。出具的项目分析报告实用、客观、简练。为企业家智慧决策提供优质服务。

借助云南被纳入国家桥头堡战略规划建设蓝图的契机，云南也是国家与东南亚国家经济交流的窗口，云南产业升级及改扩建项目将会成倍增长，为公司业务开拓带来巨大商机。

鼎臻事务所具有强大的业务推广能力，借助自身团队、外部力量突出专业地位与优势，快速占领并赢得市场，战略定位是立足云南及周边省市，辐射全国，走向世界……以最低的成本为客户创造最大的价值。

办公地址：昆明经开区云大西路科技孵化区
A幢3楼322-324（650101）

联系人：何文雄

联系方式：13887761244 0871-66358667

青岛易和盛世项目数据分析师事务所



青岛易和盛世项目数据分析师事务所致力于以实现国内中小企业拥有可持续发展的核心竞争力为己任，秉承以“凝聚中外管理智慧，以数据解决企业实际问题”为宗旨，致力于为客户提供投资项目数据分析、投资项目评估、经济效益评价、项目数据分析研究、数据处理、投资项目策划、社会经济咨询、经营决策分析、人力资源分析报告等专业、系统的服务，并以良好的职业操守为项目投融资双方提供一份具有经济性、权威性、客观性、公正性、实用性的项目数据分析报告。

事务所资源网络依托易和盛世前身公司，将业务遍布全国10多个省市自治区，与多家银行、金融机构、上市公司、商会等建立业务合作意向及业务往来，业务涉及地产、制造、快消、物流、互联网、IT、农业、林业、金融、化工、建材、通信、电子、纺织、轻工、医药及医疗器械、机械、教育、环保、旅游等诸多行业领域，在为客户提供服务的过程中，积累了广阔的眼界和丰富的经验，促使公司已经形成一支执业经验丰富、人员结构合理、高素质的专业队伍，能承担各种类型项目数据分析及相关业务的复合型数据分析师事务所。

事务所以诚信、科学、实效、客观、公正的理念与态度，坚持高品质的数据分析，为客户提供最有价值的服务，努力成为客户发展历程中值得信赖的合作伙伴，力求成为项目数据分析行业的骨干精英，为推动中国项目数据分析事业的发展贡献力量。

办公地址：青岛市市南区徐州路98号丙 班芙春天大厦

联系人：杨海波

联系电话：0532-89093017

华融成(北京)项目数据分析师事务所



华融成（北京）项目数据分析师事务所注册资金300万，拥有高素质的数据分析人员；良好的工作作风，认真负责、严谨求实的工作态度。针对项目的多样性和复杂性，公司建立了自己的人才库，涵盖了数据分析师、注册税务师、注册会计师、纳税筹划师、资产评估师、矿产资源评估师、房地产评估师、精算师、律师、高级经济师、造价工程师、建造师、高级工程师等行业的专业人才。

事务所建立了完善的组织管理制度和数据分析流程，从制度和规范上保证所承接业务的质量。事务所设有专门的数据采集部门，专注于各行业的市场、客户数据调查、采集和挖掘，积累了丰富的数据资料。保证为项目提供详实丰富而完整的数据资源。保证我们真正能洞察客户的需求，了解市场动态，各行业的发展方向，辅助企业做出正确的决策，为客户提供有价值的数据分析，帮助客户科学定位，推动企业智能决策，规避风险，创造更大的价值。

华融成（北京）项目数据分析师事务所一直严把质量关，遵守“诚信、公平、公正”的职业道德守则，按照“服务及时、数据准确、分析完整、建议中肯，客观实际”的工作原则，实现：“诚信、高端、权威”的发展目标。

办公地址：北京市朝阳区朝阳路都会国际大厦15层

联系人：于经理

联系方式：010-58282186 18610868600

湖南翰林项目数据分析师事务所



湖南翰林项目数据分析师事务所有限公司由具有项目数据分析师、资产评估师、房地产估价师、土地估价师、注册税务师、会计师、司法鉴定人等多种执业资格的复合型专业人员申请发起，报中国数据分析行业的主管单位中国商业联合会数据分析专业委员会审核批准，经湖南省工商行政管理局名称核准，经邵阳市工商行政管理局核准登记的一家专业从事项目投资融资咨询、各类数据分析及管理咨询的专业机构。事务所在会计审计、各类价值评估等业务领域弥补项目数据分析师事务所在传统法定业务上的短板，同时也改善了

各类中介机构在新兴数据分析业务上的不足，构筑数据分析大平台，编织一张数据分析的大网，拓宽数据分析宽度，增强数据分析的深度，提升企业自主了解自身数据的意识。

事务所在运营规划上着眼于建立和发展“翰林数据分析”平台，重点筹建三个部门，建立三支队伍：“翰林数据储备中心”、“翰林数据分析中心”、“翰林数据营销中心”。建立“翰林数据”品牌，逐步向社会推销自己，树立自己品牌形象，使社会各界知晓“翰林数据”，垂询“翰林数据”，签约“翰林数据”。建立学习型事务所，培育“质量取胜、规范执业、合作共赢、开拓进取”的事务所文化，秉承“低调做人，高调做事”的处世态度。

办公地址：湖南省邵阳市大祥区西湖路南端市人民银行对面万基银座小区1单元402室。

联系人：卿启伟

联系方式：13187299268
0739-5189006

福州信达永道项目数据分析师事务所



福州信达永道项目数据分析师事务所，是福建省首家项目数据分析师事务所，具多年数据分析经验，尤其在通信行业与互联网行业有深厚的行业基础，提供企业经营决策分析服务与项目投资咨询服务。

部门设置：信达永道下设经营决策咨询部、投资咨询部

与市场调查中心三个部门。

团队构成：咨询团队包含本地团队与顾问团队。本地团队以经济学、统计学专业人士为主，是包括有数学、法学、会计学、计算机、信息系统等专业成员在内的复合型团队。顾问团队除立足于中数委专家团队，还包括福建省本地工农建商各行业、福建高校科研机构、金融财政行业等领域专家及团队组成。

行业经验：电信业-增值业务数据分析及战略研究；增值业务产品长尾分析；流量业务分析及战略研究；客户数据建模；电子商务-数据分析型团队组建；客户精准营销；网站分析与优化。

办公地址：福建福州东街7号信息广场西区12层

联系方式：0591-87559119

网 址：<http://www.chinadica.com>

数据分析是决定大数据成败的王道

◎ 文 / 罗布·恩德勒 图 / 崔峻珩



企业的IT高管们已然开始意识到，如果没有强大的分析系统，大数据几乎没什么价值。而借助分析系统，他们可以处理这些数据，以便给企业的重要决策者(他们的老板)提供易于消化的信息。但在市场上仅有如此少的解决方案的前提下，这真的是说起来容易，做起来难。

近几年来，我一直在与大数据公司的高管们探讨如何解决他们所遇到的问题。一些有趣的问题仍然持续存在。很明显，我们仍然还处在对于某些问题的初步了解阶段，要充分了解并使用有效的方案来解决这些问题，我们可能仍然还有很长的路要走。

Facebook和谷歌等公司获取了大量的信息。他们普遍遭到侵犯隐私的控诉，尽管我们并不知道这些企业利用他们收集的数据信息可以计算出关于我们的什么信息内容。我们总是假设他们使用这些数据信息是对我们不利的，即使他们很可能尝试的是使用这些信息为我们的谋利益。

没有更好的保护您的数据的方案

历来，围绕着大型数据存储库的许多问题均涉及到如

何管理他们。这主要是指确保那些需要访问这些数据的人员的范围权限：从管理报告到遵守一切管理的制度规范，以便让这些人员在需要访问这些数据时可以得到他们所需要的信息。这还意味着必须确保数据存储的安全。这在历史上已然成为了供应商们铁一般坚不可摧的服务准则。

上述这些历来管理数据的方法说明我们一直以来对待这些数据就像海盗的宝藏一样，只是努力寻找创造性的、廉价的方法来埋葬他们。但却没有拿出同样有创意的方案来及时分析、得到他们。

我们可以肯定的是，宝藏是确实存在的，但我们不知道其究竟在何处。甚至有些数据信息已经真的非常老了，其索引和存储往往是如此糟糕，以至于我们有时会认为如果我们当初没有将其存储着首要位置会不会更好些。

近几年来，我一直在与大数据公司的高管们探讨如何解决他们所遇到的问题。一些有趣的问题仍然持续存在。很明显，我们仍然还处在对于某些问题的初步了解阶段，要充分了解并使用有效的方案来解决这些问题，我们可能仍然还有很长的路要走。

Facebook和谷歌等公司获取了大量的信息。他们普遍遭到侵犯隐私的控诉，尽管我们并不知道这些企业利用他们收集的数据信息可以计算出关于我们的什么信息内容。我们总是假设他们使用这些数据信息是对我们不利的，即使他们很可能尝试的是使用这些信息为我们的谋利益。

没有更好的保护您的数据的方案

历来，围绕着大型数据存储库的许多问题均涉及到如何管理他们。这主要是指确保那些需要访问这些数据的人员的范围权限：从管理报告到遵守一切管理的制度规范，以便让这些人员在需要访问这些数据时可以得到他们所需要的信息。这还意味着必须确保数据存储的安全。这在历史上已然成为了供应商们铁一般坚不可摧的服务准则。

上述这些历来管理数据的方法说明我们一直以来对待这些数据就像海盗的宝藏一样，只是努力寻找创造性的、廉价的方法来埋葬他们。但却没有拿出同样有创意的方案来及时分析、得到他们。

我们可以肯定的是，宝藏是确实存在的，但我们不知道其究竟在何处。甚至有些数据信息已经真的非常老了，其索引和存储往往是如此糟糕，以至于我们有时会认为如果我们当初没有将其存储着首要位置会不会更好些。

新兴的公共云资源承诺低成本的存储与未来高可能性的访问。任意数据信息宝藏资源均被一排排整齐的存储。唯一需要权衡的，当然就是，安全、管理和遵守合规性的问题。

随着数据的不断增长和企业IT预算压缩，上述这些因素的权衡似乎已经不成其为问题了。也就是说，直到犯罪分子找到获得并发布了这些数据，才会造成风险。我们的企业甚至有了风险经理的职位，但这一职位也慢慢随着金融市场的崩溃而日渐变得对于保护企业资产没有什么意义。

数据分析才是决定大数据的成败的王道

现在，我们意识到，最关键的并非大数据本身。相反，是数据和移动设备接入的分析与结果报告。企业高管们越来越多地发现，如果他们能够从他们收集的数据中得到真实的信息，他们可以做出更好的决策，避免痛苦的重复错误，并提升他们在企业乃至整个行业中的地位。

事实证明，知识就是力量。因此，今天的成功人士已然开始把重点放在更多地了解他们的客户、合作伙伴、员工和企业环境，而不再是他们的竞争对手。

新时代的企业执行人员使用工具进行更强大的数据同步。这保证了被分析的数据的准确度和及时性。他们提供移动客户端，可以在智能手机和平板电脑上显示结果。他们利用云服务，可以同时解决企业的成本和安全性的要求。

Hadoop已经成为最大的数据分析平台，供应商之间正在使用Hadoop，进行提供最好的工具的竞争。然而，随着供应商开始花更多的时间开发附属品，而不是优化整个解决方案;或选择合作伙伴时无意造成了瓶颈，这些优化可能亏一篑。

明智地选择你企业的数据分析解决方案

最后，当我听到奥巴马总统连任的CIO分析得出结论，大数据本身的部分并不重要。为您的管理人员提供他们所需要的答案才是最重要的。

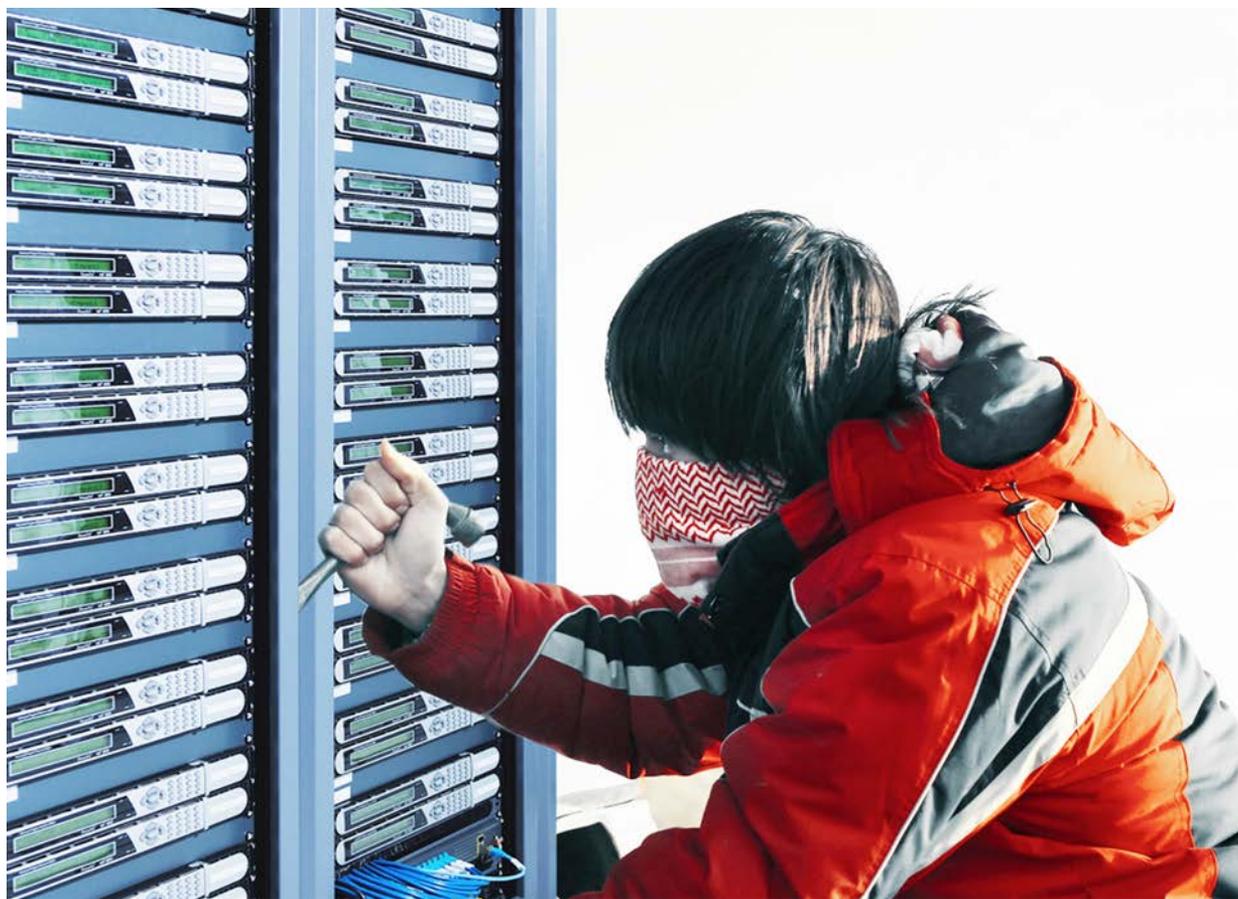
这可能听起来很简单，但它确实需要供应商必须满足下列条件：

- 具有您企业和行业方面的大量经验
- 愿意承担整个解决方案
- 在达到您企业的期望方面有良好的记录
- 具有公共和私有云资源的经验
- 有能力处理传统的数据存储和实时数据流

总之，这不是一个自己动手做的问题。您需要有经验、有信誉、有可靠性和值得信赖的供应商。只有少数供应商能满足上述这些标准。您需要进行明智的选择。

本文作者罗布·恩德勒是恩德勒集团的总裁兼首席分析师。之前他曾担任Forrester Research和Giga信息集团的高级研究员。在此之前，他曾供职于IBM担任过内部审计、市场竞争分析、市场营销，财务和安全管理等相关职务。目前，恩德勒为各种出版物撰写关于新兴技术、安全和Linux相关专栏文章，并受邀参与美国全国性的新闻电视节目录制：包括CNBC、FOX、彭博社和全国公共广播电台。





主动防御：大数据引发的信息安全技术革命

◎ 文 / 数据仓库 图 / 崔峻珩

社会化网络时代，黑客的攻击手法日趋多样化和高级化，企业面临的信息安全风险正成倍放大，IBM和思科等IT巨头也纷纷通过收购提供大数据信息安全解决方案。大数据预测分析技术能把被动的事后分析变成主动的事前防御，这也是大数据分析带给信息安全领域的最大创新。

Gartner认为，2013年将是企业大规模采用大数据技术的一年。42%的IT主管表示其所在的企业已经投资大数据技术或者将在一年内进行相关投资。从海量的低价值密度的结构化和非结构化数据中获取有价值的信息，已经成为企业IT收益的重要组成部分。

大数据对安全厂商而言，意味着海量日志、黑客攻击更加隐蔽，同时也是安全技术水平提升的有效手段。

企业IT管理人员一定不会对以下这个场景感到陌生：一名员工在集团上海分公司刷卡进入公司内部，五分钟后台系统显示该员工在北京分公司登录企业OA系统。孤立地看，这两件事都不属于安全事故，但如果将它们联系起来，IT人员就会立刻意识到问题的严重性，一个人怎么能在五分钟内从上海飞到北京企业信息正面临泄露风险。

以前，如果集团的IT系统复杂，各地分公司每天产生的日志数量繁多，并且不能集中管理，类似的安全威胁就可能淹没在几十万条安全日志里。现在，借用大数据分析，SIEM(安全信息和事件管理)正在让这些安全隐患无所遁形。

大数据在信息安全领域的应用包括宏观上的网络安全态势感知和微观上的发现安全威胁，尤其是APT攻击上。

某业内人士认为：“APT攻击往往潜藏数年，要想对其进行分析，就要调出企业一两年内的安全日志，并且进行对比分析，SOC(安全管理中心)的功能再强大也难以完成这个任务。”他表示，如果没有大数据分析，任何一个安全解决方案都很难实现每隔一段时间就对上百亿条信息做一次关联分析。

一些企业认为应该加强对大数据本身的隐私保护，业内人士却认为完全没有必要，“大数据是高价值低密度的数据，安全厂商没有必要保护大数据的安全，而是应该利用大数据分析来发现更多安全威胁，这是安全厂商难得的机会”。在他看来，大数据分析的技术难度并不大，安全厂商也可以通过购买或合作获得，“重要的是分析的逻辑，包括查询条件、查询时间的起止点等，这些考验的还是安全厂商的传统思维”。

“大数据给信息安全防护带来的最大改变就是我们通过自动化分析处理与深度挖掘相结合，可以将之前很多时候亡羊补牢式的事中、事后处理，转向事前自动评估预测、应急处理，让安全防护真正可以主动起来”。

数据仓库协会（简称为TDWI）近期的研究显示，随着技术难题被攻克，大数据已经成为企业重要的业务资产。在此项研究中，TDWI共访问了325名IT经理人、技术顾问和企业用户。这些受访者来自不同规模公司。在最终得到的报告中，TDWI指出，新技术的出现使得企业能够对不同类型的海量数据集进行更为复杂的数据分析。

超过三分之一的受访者表示，他们目前正在对大数据进行某种形式的高级分析工作，其中大部分应用在BI（商业智能）、预测分析、数据挖掘和统计分析任务。调查发现，近45%的受访者认为，大数据分析可以让商业洞察力变得更为精准。38%的受访者期待，能通过该技术更为敏锐地发现更多的销售和市场商机。超过60%的受访者希望，大数据分析能快速提升公司在社交媒体上的营销能力。

TDWI的调查还显示，大数据分析应用案例中增长最快的当属高级数据虚拟化。为了让数据的可视化程度更高，越来越多的企业开始使用高级分析工具对大数据集进行分析。

该报告的作者、TDWI分析师Philip Russom说：“大数据在过去曾经是一个技术难题，企业不得不想方设法应对海量数据的管理挑战。现在，如果使用分析的方法处理这个问题，你就能从大数据中获得诸多好处，这些好处是传统BI和数据仓库技术所无法给予的。”

如今，新技术与成本大幅下降的存储和硬件结合起来，让企业可以高效、快速地存储、管理和分析各种类型的海量



数据。越来越多的企业开始筛选这些复杂的海量数据，以进行建模。这些都是他们过去不知道或是无法理解的。目前，Aster Data、GreenPlum、Teradata、Netezza、ParAccel、Vertica和SAP等公司所提供的数据库和数据分析技术能够帮助企业完成这些工作。

Russom指出，这些厂商提供的产品采用了新的技术并进行了一系列创新，能够解决传统数据库技术中的一些不足。这些新技术包括内存数据库、海量并行处理分析技术，以及允许人们以过去无法想象的简单、高效的方法来存储、管理和查询大数据。Hadoop和MapReduce等开源工具则使企业能够以一种全新的方式来管理和跟踪大数据。

Russom在报告中称：“分析工具和数据库现在能够应对大数据带来的挑战。它们能够以更快的速度处理海量检索和解析表。新一代工具和平台将性能水平提升到了一个新的层次，这对于大数据应用非常重要。”

这份报告还提到，在充分发挥高级大数据分析能力的同时会面临挑战。报告指出，超过45%的受访者表示，大数据分析面临的最大的挑战就是严重缺乏技术熟练的专业人才。更糟的是，新一代分析工具与传统BI和数据仓库所需要的熟练技能不尽相同。报告还指出，缺乏商业支持和与大数据分析有关的整体成本是大数据分析所面临的另外两大障碍。

fin

玩具外贸电商掘金欧美：数据分析爆款致胜

◎ 文 / 敦煌网 图 / 崔峻珩



他被誉为外贸电商玩具行业的“爆款发起者”，在中国最早挖掘出“愤怒的小鸟”这款玩具的价值，同时还是敦煌网玩具行业2012年销售额第一名的卖家。记者日前联系到外贸电商玩具卖家章海剑，听他讲述了如何依靠数据分析做到上述一切。

4月10日，章海剑对记者表示，之所以能将自己的店铺在一年多的时间做到敦煌网玩具行业2012年销售额排名第一，最主要得益于团队超强的数据分析能力。数据分析的两大作用在于挖掘潜在爆款和锁定并留住平台大买家。

2012年1月，章海剑正式踏入外贸电商行业，目前通过敦煌网和阿里速卖通向欧美和澳洲出口毛绒玩具和塑料摆件。他此前有十几年的IT从业经验，主要从事数据采集和计费系统开发。

章海剑介绍，外贸电商非常重要的一环是选品。早期他的选品，主要通过分析目前市场上的流行因素，比如最近哪些电影流行、动漫流行、游戏流行，这种方式主要依据卖家的经验，需求花费大量的时间摸索且缺乏科学性。

后来，章海剑改变了选品策略，直接搜集亚马逊等国外电子商务平台玩具品类的销售数据并进行分析，模拟其销售轨迹，挖掘其中可能会成为爆款的产品。而分析对象，章海剑一般会选择国外卖家刚上架不久的商品，而不会选择已经成为爆款的产品。”

“已经是爆款的，证明它卖得非常成熟了，利润空间和盈利时间都会很有限，去分析它没有太大意义了。如果把分析对象放在没成为爆款之前可能会成为爆款的产品，这样的

话，相对于国内卖家我们永远是领先的，所以国内外贸平台上一些爆款通常是由我们发起。”章海剑说道。

记者了解到，目前外贸电商平台的同质化竞争严重，新产品的盈利周期很短，一般只有两到三个月。能够首先发现爆款，一方面可以帮助卖家在商品的高利润时期赚得足够的利润。另一方面则有助于增加买家的粘性。因为中国外贸电商的买家一大部分是国外的零售卖家，更早地购买到爆款，也有利于他们在当地的再销售。

“如果我永远只看敦煌平台、速卖通平台上哪些卖得好，等我发现这些卖得好再开始卖的时候，实际上毛利率已经非常低了，竞争对手也多，很难一下做大。”章海剑对记者说道。

据章海剑介绍，目前通过数据分析，他已可以做到新推出的10件商品中，至少有3件能够成为爆款。

数据分析的另一大作用在于可以锁定和帮助维护外贸电商的大买家。章海剑表示，他们团队每周都会花很大的精力去分析平台上买家的购买信息，以找出潜在的大客户。“速卖通比较好分析，因为它所有的购买轨迹都在，敦煌操作起来比较慢，得分析它的好评店铺，然后把相关的买家找出来。”

而一旦判断出该买家是大买家，章海剑团队则会继续细致研究其购买轨迹和购买习惯。“看他买了一些什么东西，以什么价格买的，我们就会做针对性的报价，所以说我们抢大买家，维护大买家的能力比较强。”

记者还了解到，目前在国内的外贸电商平台上，如果买卖双方没有发生过交易，卖家将无法获得买家任何的联系方式，这使得章海剑无法利用已有的数据主动和客户沟通。

“好在现在不管在速卖通还是敦煌，玩具行业相对来说我们做得还行，一般大买家只要是做玩具的，总有一天会到我们店铺里来的。已经我已经知道他是大卖家了。所以我有信心，只要他到了我们店铺我们就不会放他走。”章海剑说道。

注：爆款是指在商品销售中，供不应求，销售量很高的商品。即通常所说的卖的很多，人气很高的商品。

fin

浅析数据分析在航空运输业中的运用

◎ 文 / 昆明5期学员 王豪杰 图 / 崔峻珩



在交通运输业中，航空运输与铁路、公路和水路运输相比具有运输成本高、时间短、票价贵等特点。如何更好地助推航空公司在交通运输业中的发展，笔者结合实际工作，深刻地感受到了数据分析在航空运输业中的作用。

航空公司的竞争力主要体现在运输航线的优劣，确定出优势航线和劣势航线是提高航空公司核心竞争力的关键。运用数据分析理论，结合自身的工作实践，摸索出来了一套评估航线优劣的方法。

首先，基于所有航线的历年成本和收入数据，采用时间序列的方法评估预测出来每条航线未来一年内每个月的现金流量。其中预测每条航线未来一年内各月的成本和收入数据，主要采用了t=tttt的时间序列模型。

其中，对t的处理要特别注意到国庆节、端午节、元旦节等法定节假日采用公历日期，而春节要采用农历日期，这样才能和实际的客流变化情况对应。

其次，运用现金流量的贴现模型测算出来每条航线的净现值和内部收益率。所用到的公式主要有：

$$\sum (CI - CO)_t (1 + IRR)^t = 0$$

其中，为计算月份，为第t月的净现金流量，为折现率，IRR为内部收益率。

最后，对所有航线的净现值和内部收益率采用决策树模型，运用属性选择度量算法进行优化分类标准，进而对所有

的航线进行科学合理的分类。

其中属性选择度量主要采用了最大信息增益率作为分类的标准，如下：

$$GainRatio(A) = \frac{Gain(A)}{SplitInfo(A)}$$

上式中SplitInfo(A)表示属性A的分裂信息，其为

$$SplitInfo(A) = - \sum_{j=1}^v \frac{s_{1j} + \dots + s_{mj}}{s} \log_2 \left(\frac{s_{1j} + \dots + s_{mj}}{s} \right)$$

表示按照属性进行分类产生的信息增益量，其为

$$Gain(A) = Entropy(S) - Entropy(S, A)$$

运用上述模型和算法，准确合理地对航线进行了分类，尤其是对分类结果中的优势航线和劣势航线，可以针对性地采取不同的经营策略，以便更好地提升整个公司的竞争力和盈利率。

上述是数据分析在航线评估和经营策略制定的一个简单运用，其能带来巨大的商业价值。相信随着时代的发展，必将更加凸显数据分析在各行各业中的作用，也必将更加凸显数据分析师在各行各业中不可估量的价值。

fin

通道式数据与智慧决策时代

© 文 / 山东海洋项目数据分析师事务所 柴方明 齐文忠 图 / 崔峻珩

对于决策效能而言，
比大数据价值更高的是“通道式数据”！

现在大部分企业财务数据还停留在从固定点提取特定数据上，而不是对整个经营决策所需数据进行系统提取、分析。

“大数据系统——当我们还没有搞清楚信息时代是怎么回事的时候，已经进入了数据时代。信息时代和数据时代是有很大变化的，是完全不同的概念。信息的出发点是我认为我比别人聪明，数据的出发点是认为别人比我聪明；信息是你拿到数据编辑以后给别人，而数据是你搜集数据以后交给比你更聪明的人去处理”。

——马云

数据分析的价值核心所在

大数据时代，一些传统的商业思想正在被颠覆。这其中最为重要的，就是必须将数据作为企业的核心资产。

在大数据时代之前漫长商业社会进化中，企业无法脱离人才而独自经营，也正因为如此，人才变得异常重要，

并一度被视为企业的核心竞争力：一方面，企业智商被分布存储在这些人才的大脑中；另一方面，企业需要借助人才的商业智商，提升自身的商业智商。一定程度上，企业智商的高低，完全取决于人才的商业智商。但由于企业智商被分布存储于人才的大脑中，信息的分享与价值挖掘受到极大制约，很难完全发挥。

“在大数据时代，人才固然重要，却并非企业智商最重要的载体，数据才是企业智商真正的核心载体。这些能够被企业随时获取的数据，可以帮助和指导企业全业务流程的任何一个环节，进行有效运营和优化，并帮助企业做出最明智决策。在大数据时代的企业智商，才是真正被企业完全掌控的智商，而这一切的基础就是形形色色的数据。”

——IBM

数据除了位置数据，时点数据外，从截取的角度不同，目前通常被划分为“截面数据”（横向截取）、“时

间序列数据”（纵向截取），陈哲老师进一步提出了“面板数据”概念（横向截取+纵向截取）。

作为数据分析行业从业的分析师，我们认为对决策用数据的新定义是：“通道式数据”

数据分四维，除了空间的三维外，加上时间为四维，有人讲颜色等为另一维，则有些牵强。

“通道式数据”，具有多重含义：

1. 通道可以是直的，可以是弯的，可以是方的，可以是圆的，可以是一条，可以是多条，可以是闭合的，也可以是开放的，可以是连续的，也可以是截段的，可以是稳定的，也可以是位移的，可以是并行的，也可以是穿梭的……

2. 通道可以指运营渠道，更进一层可以指运作模式，再上一层可以指整个决策体系；

3. 通道可以指数据本身特性，也可以指数据的“时空”价值；

4. 企业家从重视通道数据起始，在推开系统决策之门的同时，也往往可以通过这条通道，从运营的死地走向生门；

5. 通道是始终开放的，但只对有思想准备的个体及组合有效。

“大”数据时代，数据分析趋势，面对通道数据，分析工具趋向高度智能化，而分析思维趋向却直指事物本质和人性“根源”，对于决策效能而言，比大数据价值更高的是“通道式数据”。

我们认为，通道式数据背景下的智慧决策时代，已正式开启！

中国的CPDA们，您准备好了吗？



小成本,大收益的数据分析之思

◎文 / 云南4期学员 杨一庚 图 / 崔峻珩

依靠信息技术的不断进步，在大量数据靠着互联网实现串联，共享的过程中，无论是个人、企业还是国家都可以依靠“大数据”实现更为良好有序的发展。而数据分析业务，也将因为以云计算为代表的技术措施的帮助，更为广泛的开展，将促进人类社会各类生产生活过程中劳动效率的提高。

数据产生效用，无非就是获取数据、分析数据和指导实践的过程。

其实，无论互联网产不产生，人类的生产生活实际都离不开数据，从古至今，各行各业，因为有了数据定量分析的支撑，才有了人类社会更为秩序性的发展，只不过互联网的出现，将数据的作用提上了一个新的高度。过去我们只有依靠传统载体分享的信息在今天能够通过互联网广泛的传递，还算是封闭的消息也能够在一定程度上实现传播，乃至今天生活中，每个人因为与手机、身份证、互联网帐号等这些符号的关联，每个人每天的所作所为实际上已经能够被更为有效的知晓，当然，这里有个人隐私



的保护问题，只不过当这个途径被打开之时，诸如腐败问题、社会矛盾问题等也能够在更为公开的场面进行讨论，最终反倒是成就了真相的大白，这就是大数据的力量所

在，宏微观层面各类信息的全方位公布，使得把握当下相关领域发展特点，成为了可能，在便捷，低成本的获取数据的同时，能为生产生活产生较大收益奠定基础。与之对应的是，一些虚假不实的信息也会因此而伴生，各类传言也会因为更加难以辨析，在庞大数据面前，在信息爆炸的局面之下，或许带给人们的最大挑战也就将来源于数据真实可靠的鉴别方面。



即便是拥有了真实的信息基础资源，真正能够帮助人类社会进步的推动还来源于信息的分析，这需要一定的方法论，在以理性的构建信息分析模型的基础上才能够导出卓有成效的实践，这也就意味着，在传统的劳动分工不断细化的过程中，专业的数据分析岗位应运而生，这类职业，不仅需要准确的判断各类信息的真实性，更需要以专业的眼光，结合要分析的事物特点，做出客观理性的决策。

毫无疑问，这个劳动分工的促成，是积极的，就好比做一个投资项目，过去，要广泛的进行市场调查及定量分析，因为受制于空间时间的限制，其分析力度一定是有限的，在耗费大量时间及各类投入并做出分析结果的同时，实际上，也完全有可能因为实践背景发生了变化而导致分析本身意义不大。

但反观大数据时代，在数据获取较为容易的情况下，低成本的数据获得加上专业的数据分析，应该能够更为准确的为各类数据分析的本质目的产生应有的贡献，这个过程里，数据分析同样也暗含了较低成本博取更大价值创造的道理，是非常及时有效的。

更进一步的是，分析的结果还要指导实践，这就是专业的数据分析与被分析事物的结合问题，例如企业的

投资计划，战略选择，量化的结果是否意味着企业就一定要按数据指导意见按部就班进行实践呢？这应该打上一个问号。

在今天这个时代，过去还如日中天的企业也许会因为一定的战略选择失误而走向落日黄花，因为竞争太激烈，环境变化太快，那是否意味着搜集过往信息，科学分析得出的结论就毫无意义呢？

也并非完全如此，数据分析的过程更为重要的是一个理性推导的过程，有其科学借鉴意义，但如何预知未来，假如以数据分析过去的路径判断为依托，那显然这个路数是可靠的，只不过，未来发展变化超乎想象，如何决策呢？数据分析在科学分析过往数据的同时，最大层面的意义还在于建立理性分析之思路，因为获取信息便捷有效，对于呈现变化规律的对象又能够行之有效的判断其发展路径，在面对不确定因素时，也有相应的应对策略，科学建立的数据模型在演绎过往的数据路径之时，也更加在强调要建立面向未来导向的思维判断，在这个过程中重点的也并不仅仅是过往的单一可靠路径，而是在搜集、分析、判断过程中分析数据对于基本规律和把握机会能力思路的全方位构建，这也就成为了数据分析系统能力的关键所在。

在这里，系统性的构建处理应对事物变化的能力也能够最大程度的规避单一思维路径提出的解决方案带来的风险，从而，尽管在数据指导实践的过程中，需要更的精力全盘谋划，但其付出之成本，要远远低于因为考虑不周全而产生的损失之痛，也就暗合了与其他过程相一致的以小博大之理。

所以，无论是搜集数据过程中的信息甄别，数据分析中选择恰当数据模型，还是真正投入实践过程中的有选择性判断，都毫无保留的投射出数据分析整个流程中的理性决策之道，也正是因为如此，数据才能有效的发挥其小成本，大收益的积极作用，换句话讲，以较低的人财物投入换取高价值的创造，不就是社会劳动生产力提高，社会财富更有效创造的全面体现吗？

fin

禽流感考验数据分析能力

◎ 文/金融投资报 周姝妤 图/崔峻珩



舍恩伯格的著作《大数据时代》讲了这样一个故事：2009年出现了一种新的流感病毒。这种甲型H1N1流感在短短几周之内迅速传播，全球的公共卫生机构都担心一场致命的流行病即将来袭。有的评论家甚至警告说，可能会爆发类似1918年在西班牙爆发那场影响了5亿人口并夺走了数千万人性命的大规模流感。而最终2009年这场流感在美国被控制的原因，却是因为一家知名的互联网公司提前预测了这场流感。

这家公司曾表示，他们不仅可以预测到这场流感在美国甚至全球的爆发，甚至可以具体到特定的地区和州。这家公司是通过观察人们在网上搜索记录来完成这个预测的。他们每天都会收到来自全球超过30亿条的搜索指令，通过分析人们的搜索记录来判断这些人是否患上了流感。他们不是简单地以“哪些是治疗咳嗽和发热的药物”之类检索词条来判断，而是通过检索词条的组合建立了一个数学模型。所以，这一年甲型H1N1流感爆发的时候，与习惯性滞后的官方数据相比，这家公司建立了一个更有效、更及时的指示标。

确实如此，如果想要预防控制传播性疾病，必须先知道病源，但糟糕的是，人们一般感冒不会去医院，等到病重去医院时，基本上已经是一到两周之后了，医院搜集到

这个信息再传到疾控中心，延迟的时间将更久。对于一种可能飞速传播的疾病，信息滞后两到三周的后果将是致命的。就如这次禽流感上海的死者吴亮亮，在病亡后21天，才被确诊为H7N9流感。死者家属以“诊断失误”向法院索赔百万巨款，这种“医疗事故”是患者的悲哀，也不应该医院来买单，如果疾控中心早发现，早预测，早播报，也就不会出现这种悲剧。

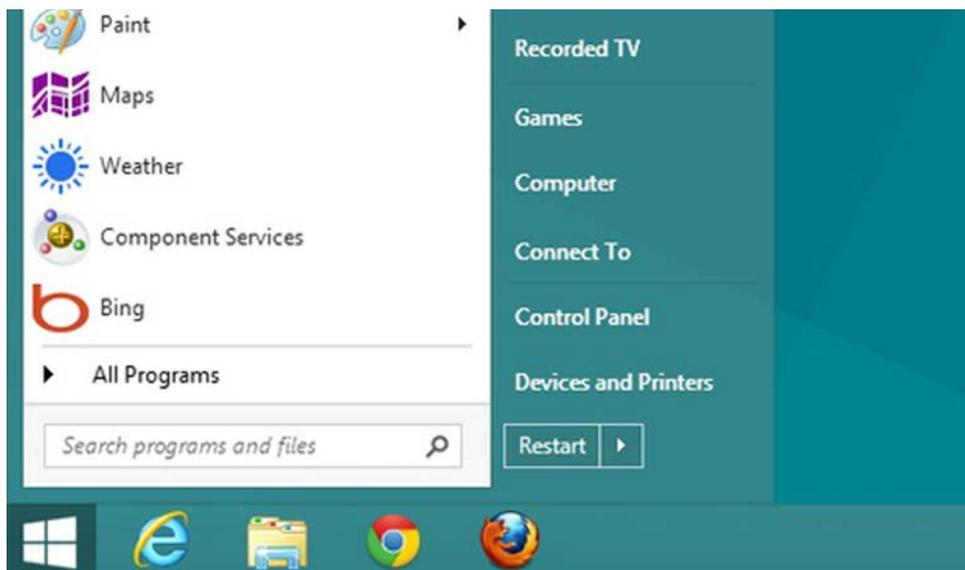
而在H7N9流感确认以后，政府的公开信息仅是将患者数量、死亡数量进行通报，并未把真实数据、就诊情况进行公示，而农业部一句“不排除更大范围检出H7N9禽流感的可能”，更是将市民推向了恐慌。板蓝根断货，口罩断货，医药股涨停，农副业风声鹤唳，死猪死鸭满河漂流，一波又一波的社会事件冲击着人们脆弱的神经。

舍恩伯格的故事是想告诉大家，大数据深刻地改变了公共卫生，改变了人们的思维方式。试想：如果我们有一个“公众大数据模型”，有能力对海量数据进行有效的分析，让疾控中心的疫情预报走在疾病传播的前面；不仅能预测，更能在流感传播前进行有效预防，最大限度减少公众恐慌。如今，是否建立一个“公众大数据模型”，显然应引起有关部门的重视。

in

数据分析告诉你： Win8用户对开始按钮有多饥渴

© 文 / 赛迪网IT分析 图 / 崔峻珩



昨日，微软Windows部门总裁Julie Larson-Green暗示，Windows 8在更新（Windows Blue）后可能会重新拥抱开始按钮，在过去一段时间，用户对开始按钮的呼声很高，微软也在用心聆听。

Windows 8用户对开始按钮的消失很是不习惯！

确实，不少Windows 8用户对开始按钮的消失很是不习惯，他们迫切希望开始按钮的回归，有多迫切？看看下面这组数据你就知道了。

伴随着Windows 8出现了很多第三方界面修改软件，这些软件帮助用户在Windows 8中重现开始按钮和开始菜单。外媒AllThingsD今天选取了其中两款较为知名的开始按钮找回软件进行试用，其中一款是来自Stardock的Start8，售价4.99美元，另一款是来自SweetLabs的免费Pokki。

这两款软件的功能相似，都是为Windows 8带来与Windows 7相近的开始按钮和开始菜单，这里就不再赘述。更令我们感兴趣的是其下载量。

Stardock首席执行官Brad Wardell不久前表示，

Start8的下载量突破了400万次，不过AllThingsD发现，当前Start8的下载量已经逼近500万次。1月份的时候，SweetLabs透露，Windows 8版Pokki下载量达到了150万次，而AllThingsD获悉的最新数据是300万次。

也就是说，Start8和Pokki的下载量加起来已经达到了800万次。与之相对比，微软本周宣布售出了1亿份Windows 8授权，其中还包括一些尚未激活的。这意味着，至少有8%的Windows 8用户曾下载了Start8或Pokki。不过，Windows 8开始按钮找回软件并非只这两款，受欢迎的还有ClassicShell、RetroUI、StartisBack等。

看来，微软确实应该认真考虑一下是否在Windows 8桌面中重新提供开始按钮了。

fin



还在为企业决策

左右为难



“用数据说话，做理性决策！”



项目数据分析师
CERTIFIED PROJECT DATA ANALYST

- 项目数据分析师以专业数据分析技能成就职场精英!
- 项目数据分析师事务所发展迅猛，遍布全国!
- 项目数据分析广泛运用于企业运营和投资的各个环节!
- 数据精准化，决策科学化，**数据分析行业在崛起!**



搜索

项目数据分析师





中颢润(北京)项目数据分析师事务所
Zhong Haorun (Beijing) Certified Projects Data Analyst Firm

- 中商联数据分析专业委员会常务委员单位
- 第二届中国数据分析行业峰会协办单位
- 2012年度优秀项目数据分析师事务所

洞察损益 实现企业价值

至诚致信 精益求精

立足新起点 开创新局面

态度决定一切 细节决定成败

- 以定量研究为主，定性、定量相结合。
- 以高水准的专业品质助您洞察项目，决策未来！
- 为您提供 **最具价值的数据分析服务！**



中颢润(北京)项目数据分析师事务所
Zhong Haorun (Beijing) Certified Projects Data Analyst Firm

北京市东城区建国门内大街7号光华长安大厦3座2层(100005)

Tel: 010-65186066、18611098373、13001995337

Fax: 010-65188088

Email: zhangem @ chinacpda.cn

wangmx @ chinacpda.cn

Web: www.chinacpda.cn