

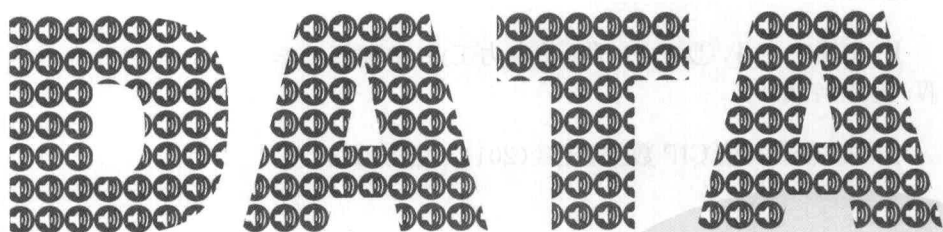
DATA

用数据说话： 迎接电力营销服务的大数据时代

克宜 薛冰 彭翎 王慧琴 合著



中国电力出版社
CHINA ELECTRIC POWER PRESS



用数据说话：

迎接电力营销服务的大数据时代

江克宜 薛冰 彭翎 王慧琴 合著



中国电力出版社
CHINA ELECTRIC POWER PRESS

内 容 提 要

本书是我国电力系统第一本从大数据角度剖析电力营销服务的著作。它结合精益管理方法和数据分析,从实际工作出发,对现阶段电力营销服务工作提出了新方法和新思路。结合大量的实际案例,充分体现了数据分析的成果和理论与实践结合的应用,图文并茂、浅显易懂。

本书适合电力行业管理人员阅读,尤其适合基层电力营销服务管理人员借鉴学习。

图书在版编目 (CIP) 数据

用数据说话:迎接电力营销服务的大数据时代/江克宜等著. —北京:中国电力出版社,2014.3

ISBN 978-7-5123-5494-4

I. ①用… II. ①江… III. ①电力工业-市场营销学
IV. ①F407.615

中国版本图书馆 CIP 数据核字 (2014) 第 013283 号

中国电力出版社出版、发行

(北京市东城区北京站西街 19 号 100005 <http://www.cepp.sgcc.com.cn>)

航远印刷有限公司印刷

各地新华书店经售

*

2014 年 3 月第一版 2014 年 3 月北京第一次印刷

710 毫米×980 毫米 16 开本 13 印张 209 千字

印数 0001—3000 册 定价 39.00 元

敬 告 读 者

本书封底贴有防伪标签,刮开涂层可查询真伪
本书如有印装质量问题,我社发行部负责退换

版 权 专 有 翻 印 必 究

前言

电子邮件、手机短信、微信、飞信等数字时代产物已经渗透到我们生活的方方面面，每个人的信息、语言、行为等产生的数据都不停地被存储、筛选、分析。企业亦是如此，不可避免地受到大数据浪潮的冲击和影响。2013年，被不少业内人士称作“大数据元年”。世界上存储的数据预计能达到约1.2ZB（约12亿TB），著名的未来学家阿尔文·托夫勒曾热情地将大数据赞颂为“第三次浪潮的华彩乐章”，一点也不为过。有种说法很贴切，其实“大数据比我们自己更了解我们自身”。

2012年3月22日，奥巴马政府宣布投资2亿美元拉动大数据相关产业发展，将“大数据战略”上升为国家战略，将大数据定义为“未来的新石油”，并希望引发教育和学习的变革。

2012年9月，日本西科姆集团宣布收购东京电力旗下数据管理中心——AT TOKYO公司50%以上股份，总投资额将超过500亿日元。西科姆的数据管理业务规模也将扩大到原来的10倍以上，其旗下各企业强化了对数据管理中心的升级改造，带动需求急剧上升。

越来越多的政府、企业正逐步意识到这隐藏在数据山脉中的金矿，表现出对大数据挖掘的超前意识和对大数据应用的极大关注。电力行业同样也有着丰富的数据资源，如何使用好大数据是关键。大数据时代提供给我们的将是更快的运算、更丰富的数据分析结果。

从发电到输电、配电，这一产业链处处都产生数据。经历近10年来的电力信息化发展，数字信息技术全面渗入到电力从生产到管理的整个流程，甚至在某种程度上已经触动并正在改变传统电力企业的生产模式，技术在进步，管理在转型。如何从海量数据中通过专业化的处理，挖掘出数据的潜在价值是电力企业面临的前所未有的新课题、新挑战。

根据近几年来结合营销管理的数据分析，从打击窃电、降低管理线损方面的工作实践中我们总结编写成《用数据说话——迎接电力营销服

务的大数据时代》一书。本书上篇、下篇和外篇三部分共十章，包括管理遇上大数据、供电企业应用大数据和利用大数据破获窃电案例集锦等内容。

我们对大数据的研究毕竟刚入门、刚起步，书中难免有疏漏之处，敬请广大读者不吝赐教。

编 者

2013 年 12 月

目录

前言

序幕

当大数据轻轻来敲门

大数据：改变正山雨欲来蓄势待发/3

变革服务：从“克强指数”到数据价值/10

变革经济：剖析历史可以洞察未来/13

变革管理：基于精益的数据管理体系/17

上篇

当“细”管理遇上大数据

第一章 数据的应用更胜于数字本身/22

第一节 小猪的故事：从目标管理到过程管理/22

第二节 海量数据：大量基础资源/25

第三节 “指标”：长期以来关注数据的方式/28

第二章 循“数”管理从PDCA开始/33

第一节 5W1H：获取并分析有效信息/34

第二节 PDCA：精细化管理的制胜法宝/38

第三节 5S：基于大数据的现场管理/44

第三章 大数据时代面临的机遇/49

第一节 莫里的导航图：大数据的最早实践之一/49

第二节 全数据：保存全部的“沙子”/52

第三节 混杂性：更接近“事物”的全貌/55

第四节 相关性：数据“说出”的事实/59

第五节 $1+1=10$ ：数据混搭与价值倍增/62



第四章 大数据时代面临的挑战/65

第一节 有效的组织：把想的人和干的人分开/65

第二节 做正确的事：后台支撑数据化/69

第三节 正确地做事：前端作业简单化/74



当供电企业嫁接大数据

第五章 网格化管理崭露头角/80

第一节 电网：网格技术的源泉/80

第二节 突破：网格管理的探索/86

第三节 创新：网格5S管理的实践/92

第六章 数据可视化图的华丽演绎/100

第一节 可视化：揭开数据的神秘面纱/100

第二节 美丽的证据：线损管理可视化/104

第三节 视觉解释：可视化图的广泛应用/108

第七章 借鉴精益的抄核收管理模式/115

第一节 发轫：抄核收管理的破窗理论/115

第二节 探索：从汽车生产到抄表管理/118

第三节 契机：减少抄表数据的“在制品”/122

第四节 实践：连续流与消除浪费/126

第八章 精益思想与营业厅排班/135

第一节 排队的博弈：服务需求与服务供给/135

第二节 服务的需求：从一味满足到合理引导/138

第三节 服务的供给：营业厅的“削减”艺术/141

第四节 排班的技巧：今天开放多少窗口/144

第九章 大数据时代的线损分析/149

第一节 线损立方：玩转数据建立模型/150

第二节 趋势分析：“福尔摩斯”查窃电/154

第三节 精确比对：锁定重点怀疑对象/159



外篇 利用大数据破获窃电案例集锦

第十章 基于大数据的防窃电技术/166

第一节 以静制动：后台监控破获遥控装置窃电案/169

第二节 抽丝剥茧：网格线损分析破获总分表窃电案/175

第三节 事半功倍：电量比对破获总分表窃电案/179

第四节 瓮中捉鳖：多次侦察破获大规模集体窃电案/182

第五节 知己知彼：数据监控警电联手成功破案/186

第六节 守株待兔：反侦察展开与窃电分子的博弈/191

第七节 欲擒故纵：数据比对突破窃电分子顽固防线/193

○ 序幕 当大数据轻轻来敲门 ○

我们信靠上帝。除了上帝，任何人都必须用数据来说话。

——W·爱德华兹·戴明 美国管理学家、统计学家



大数据：改变正山雨欲来蓄势待发

在您阅读这本书之前，让我们先做这样一个小测试！请先思考这个问题：到电力营业厅办理缴费业务的客户最多来自哪里？花一分钟时间来思考，如果已经锁定了自己的答案，请您开始阅读本书的内容。

电力营业厅与电信、银行、邮政等营业厅性质十分相似，都是为了满足客户的业务需求。但是，经过仔细对比分析却又发现，电力营业厅与其他营业厅又不尽相同。假如您需要存缴电话费，是会选择去电信营业厅窗口办理缴费业务，还是选择购买电话卡、网上充值等方式呢？需要小额存取款时，您是会选择去银行窗口排长队办理，还是会选择 ATM 机自助存取款呢？仔细罗列下来，电力企业也提供了多种电费缴纳方式，例如银行托收、电费卡充值、易办事自助终端、网上缴费（支付宝、网上银行等）、电话缴纳、银行柜台等。那为何常常有客户仍选择排长队在营业厅窗口缴纳电费呢？

本书对某供电企业辖区内营业厅办理缴费业务的 381 736 条客流数据进行了分析，绘制成气泡式数据图^①如图 0-1 所示。气泡式数据图是一种常见的数据地图，它是在地图上相应位置显示相应的气泡，以气泡的大小代表各区域指标值的大小。该图中通过气泡的大小来代表该区域在营业厅办理缴费业务的客流量。

数据分析结果出乎我们的意料。在传统观念里，到营业厅办理缴费业务的客户大多来自人口流动快的城中村，事实上，根据图 0-1 中看出，到营业厅办理缴费业务的客户主要集中在片区 E 和片区 F 两个区域，而该供电企业辖区内共设立了两个营业厅，正位于片区 E 和片区 F。通过对这两个区域的客流做进一步分析，发现这两个区域前往营业厅办理缴费业务的客户都集中分布在营业

① 刘万祥. 用地图说话 [M]. 北京：电子工业出版社，2012.

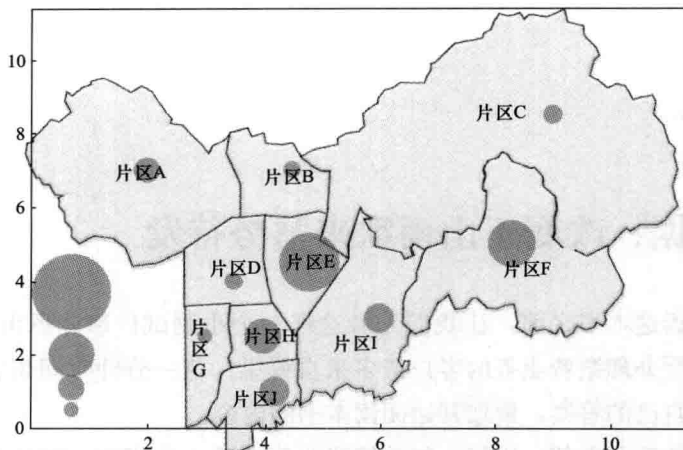


图 0-1 某营业厅缴费业务量分布气泡图

网点的周围。用分档填色型热力地图^❶分析片区 F 营业厅、片区 E 营业厅缴费业务量分布，将各区域按数值分为几个层级填充不同的颜色，如图 0-2、图 0-3 所示，图中五角星标注处为营业厅位置。图 0-2 和图 0-3 中由浅到深的颜色表示该区域到营业厅办理缴费业务的客流量分布由少至多。

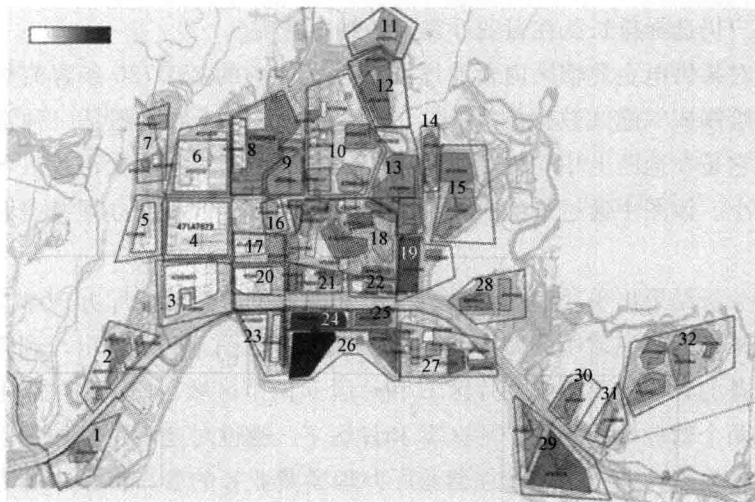


图 0-2 某区片区 F 营业厅缴费业务量分布图

❶ 刘万祥. 用地图说话 [M]. 北京: 电子工业出版社, 2012.

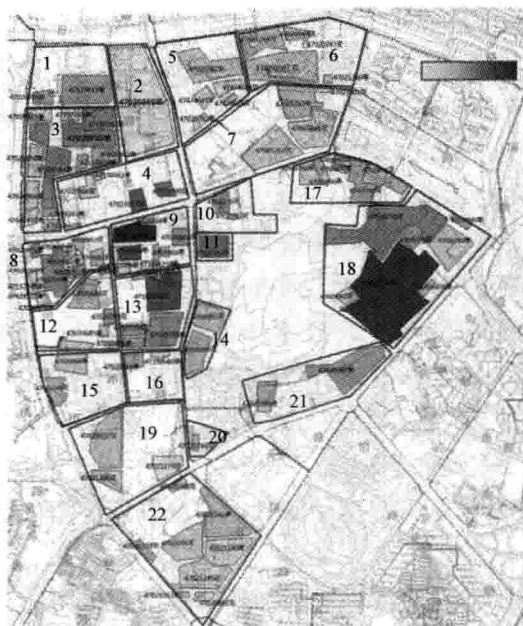


图 0-3 某区片区 E 营业厅缴费业务量分布图

与缴费业务相比，营业厅过户业务办理客流量呈现出了截然不同的信息，所有的过户业务客流量均匀分布在各个区域，如图 0-4 所示。

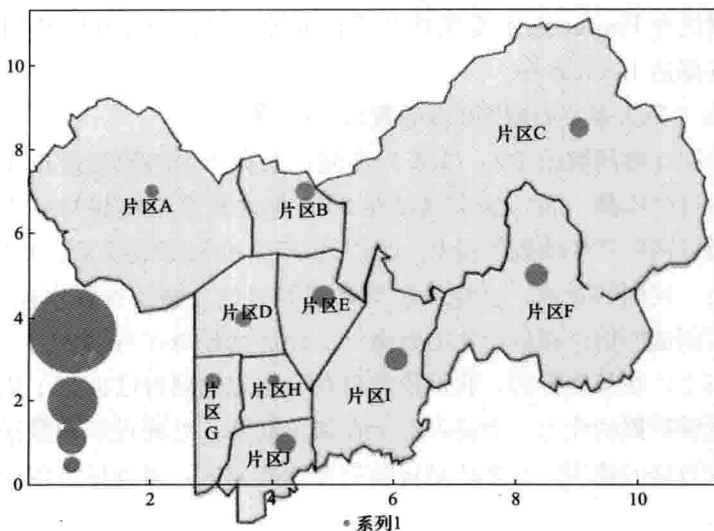


图 0-4 某区电力营业厅过户业务量分布气泡图



由此可以看出，排除停电催费模式（停电催费的模式下，前往营业厅办理缴费业务的客流量分布与各区域欠费情况呈正相关关系）的影响，前往营业厅办理缴费业务对客户而言并非刚性需求。我们大胆地设想，能否在正确的引导下，以自助终端、销售电费卡等方式取代电力营业厅收费业务窗口？

该供电企业根据数据分析结果大胆尝试关闭部分营业厅后，客户满意度的指标并未受到负面影响。他们的决策是建立在大数据的基础之上的，这是当今社会所独有的一种新型能力：通过对海量数据进行分析，获得有巨大价值的产品和服务，或深刻的洞见。这样的技术理念和数据储备，对于供电企业乃至整个社会的发展而言，将成为一种更好的辅助决策工具。

随着互联网技术及信息技术的飞速发展，在 Web2.0 的时代，人们从信息的被动接受者变成了主动创造者。

“全球每秒钟发送 290 万封电子邮件，一分钟读一篇的话，足够一个人昼夜不息地读 5.5 年……

每天会有 2.88 万个小时的视频上传到 Youtube，足够一个人昼夜不息地观看 3.3 年……

Twitter 上每天发布 5000 万条消息，假设 10 秒钟浏览一条信息，这些消息足够一个人昼夜不息地浏览 16 年……

每天亚马逊上将产生 630 万笔订单……

每个网民在 Facebook 上要花费 7 千亿分钟，被移动互联网使用者发送和接收的数据高达 1.63EB……

Google 上每天需要处理 24PB 的数据……”^①

根据美国互联网数据中心（IDC）检测，人类产生的数据量正在呈指数级增长，大约每两年翻一番，这个速度在 2020 年之前会继续保持下去。这意味着人类在最近两年产生的数据量相当于之前产生的全部数据量。我们的国家、我们的社会、我们的企业，已经从包含数据的时代迈进了被数据包含的时代。这些由我们创造的信息背后产生的数据早已经远远超越了目前人力所能处理的范围。事实上，在某些方面，我们依然没有完全意识到自己拥有了能够收集和處理更大规模数据的能力，而是习惯于在制定制度、处理过程和激励机制中尽可能地减少数据的使用。大多还是遵循着传统的方式，通过尽可能少的样本数

① 李志刚. 大数据——大价值、大机遇、大变革 [M]. 北京：电子工业出版社，2012.

据获取预期（即提前设定好并带有主观意识）的信息，通过多次筛选多方验证以获取我们认为最精确的数据，通过人工选择一个关联物或者一小部分相似数据来逐一分析事件产生的原因。

如今，计算和制表不再像过去那样困难。感应器、手机导航、网站使用收集了大量数据，而计算机可以轻易地对数据进行处理，我们进入了大数据时代。

那么，为什么是大数据而不是数据？什么是大数据？它从何而来？大数据有何不同？大数据又有什么特征呢？

关于大数据的含义，目前主要有如下几种：

“增长如此之快，以至于难以使用现有的数据库管理工具来驾驭，困难存在于数据的获取、存储、搜索、共享、分析和可视化等方面。”

——维基百科

“任何超过了一台计算机处理能力的庞大数据量。”

——约翰·罗瑟 亚马逊大数据科学家

“大数据=海量数据+复杂类型的数据。”

——但彬 Informatica 中国区首席产品顾问

早在1980年，著名未来学家阿尔文·托夫勒便在《第三次浪潮》一书中，将大数据热情地赞颂为“第三次浪潮的华彩乐章”。不过，大约从2009年开始，“大数据”才成为互联网信息技术行业的流行词汇。美国互联网数据中心指出，目前世界上90%以上的数据是最近几年才产生的。此外，数据又并非单纯指人们在互联网上发布的信息，全世界的工业设备，汽车、电能表等生活设备上有着无数的数码传感器，随时测量和传递着有关位置、运动、震动、温度、湿度乃至空气中化学物质的变化，也产生了海量的数据信息。大数据技术的战略意义不在于掌握庞大的数据信息，而在于对这些含有意义的数据进行专业化处理。换言之，如果把大数据比作一种产业，那么这种产业实现盈利的关键在于提高对数据的“加工能力”，通过“加工”实现数据的“增值”。

大数据是用传统方法或工具不能处理或分析的数据信息。随着时代的变化，供电企业工作人员每天都记录着位置、电费、用户电话等各种各样的数据信息，庞大的信息构成了大数据。大数据是如此的复杂，需要有足够的想象力，以至于到现在也没有一个统一的定义。人们对大数据的理解还不够全面和深入，犹如盲人摸象，每个人都只认识到了大数据的一部分。



大数据既包括数据量的激增，同时也包括数据复杂性的增加。大数据的数据类型丰富多样，既有像原有的数据库数据等结构化信息，又有文本、视频等非结构化信息，而且数据的采集和处理速度要求也越来越快。

大数据有 4 个典型的特征，定义为 4V^①，包括大量（volume）、多样性（variety）、速度（velocity）、价值（value）。大量（volume）是指数据量是非常庞大的，主要体现在数据存储量大和计算量大，全球在 2010 年正式进入 ZB 时代，IDC 预计到 2020 年，全球将总共拥有 35ZB 的数据量。速度（velocity）一是指数据在不断更新，增长的速度快；二是指数据存储、传输等处理速度快，这也是大数据区别于传统数据最显著的特征。在海量的数据面前，处理数据的效率就是企业的生命。多样性（variety）指大数据包含结构化的数据和半结构化、非结构化的数据，包括了订单、日志、音频、文本、视频图像等，数据之间的交互变得非常频繁和广泛。价值（value）既是特征也是对大数据应用的描述，大数据对于企业而言价值密度低，是沙里淘金。以视频为例，一部一小时的视频，在连续不间断监测过程中可能有用的数据仅时长一两秒，如何通过强大的机器算法更迅速地完成任务的价值“提纯”，是数据汹涌背景下亟待解决的难题。

大数据开启了一次重大的时代转型，就像望远镜让我们能够感受宇宙，显微镜让我们能够观测微生物一样，大数据正在改变我们的生活以及理解世界的方式，成为新发明和新服务的源泉，而更多的改变正蓄势待发。

随着大数据时代的到来，不断增长的数据是社会各企业和团体所面临的挑战。

从 IBM 的大数据战略到美国中央情报局（FBI）的超级数据库；从盖洛普的总统投票结果预测到零售帝国沃尔玛的数据挖掘技术；从淘宝网的年度数据发布到德勤数据分析研究所：数据不但影响社会经济，同时也在政治、文化等方面产生深远的影响。大数据可以帮助我们开启循“数”管理的模式，这也是当下“大社会”的集中体现。三分技术，七分数据，得数据者得天下。

事实上，企业并不缺数据。现在一般的大型企业都存储了 200TB 以上的数据，计算机存储信息大小最基本的单位是字节（Byte），一个汉字占两个字节空间，字母和数字占一个字节。存储容量的单位从小到大依次是：字节（B）、

① 李志刚. 大数据——大价值、大机遇、大变革 [M]. 北京：电子工业出版社，2012.

KB、MB、GB、TB、PB、EB、ZB、YB^❶。它们之间的关系见表 0-1，可以直观感知数据之大。企业的这些数据可以用来了解谁在何时、何处、以何种方式购买了他们的产品。当企业能够了解客户消费行为背后的动机，就可以更好地为客户服务。如来自于微博、视频、点击流及其他非结构化数据源，是关于客户意愿的大数据，现在已经拥有的技术和工具来分析和理解这些数据可获得相应价值。

表 0-1 单位换算关系

单 位	换 算	形象描述
KB (Kilobyte)	1KB=1024B (Byte)	信息量相当于一个笑话
MB (Megabyte)	1MB=1024KB	信息量相当于一部微型小说
GB (Gigabyte)	1GB=1024MB	信息量相当于一部广播级质量电影
TB (Terabyte)	1TB=1024GB	五万棵树制成的纸
PB (Petabyte)	1PB=1024TB	NASA EOS 对地观测系统三年数据
EB (Exabyte)	1EB=1024PB	5EB 全人类说过的所有的话
ZB (Zettabyte)	1ZB=1024EB	使用现在最快的宽带，下载 1ZB 的信息需要至少 110 亿年
YB (Yottabyte)	1YB=1024ZB	互联网的全部信息量加起来大概是 1YB

❶ Bill, Franks. 驾驭大数据 [M]. 黄海, 车皓阳, 王悦, 等. 译. 北京: 人民邮电出版社, 2013.