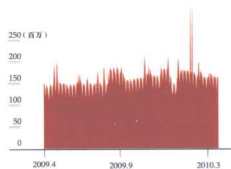
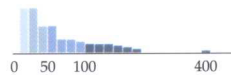
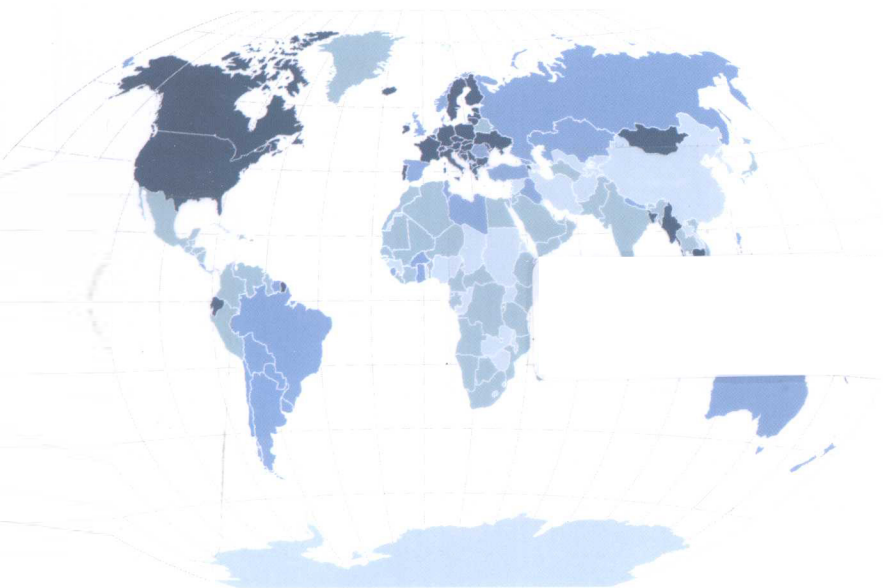


# 鲜活的数据

## 数据可视化指南

[美] Nathan Yau 著  
向怡宁 译



Sources: Mozilla, Internet World Stats, Design by: Nathan Yau, FlowingData

TURING

图灵程序设计丛书



# 鲜活的数据

## 数据可视化指南

[美] Nathan Yau 著  
向怡宁 译

人民邮电出版社  
北京

## 图书在版编目 (CIP) 数据

鲜活的数据：数据可视化指南 / (美) 邱 (Yau, N.)  
著；向怡宁译. — 北京：人民邮电出版社，2012.10  
(图灵程序设计丛书)  
书名原文: Visualize This: The FlowingData  
Guide to Design, Visualization, and Statistics  
ISBN 978-7-115-29381-7

I. ①鲜… II. ①邱… ②向… III. ①数据收集一指  
南 IV. ①TP311.13-62

中国版本图书馆CIP数据核字(2012)第220656号

## 内 容 提 要

本书是一本系统介绍数据可视化的图书。书中主要阐述了如何将冰冷枯燥的数据转换成易于理解、生动有趣、主题清晰的图表。作者根据数据可视化的一般顺序，先后介绍了如何获取数据，将数据格式化，然后用可视化工具（如 R）生成图表，最后在图形处理软件（如 Illustrator）中修改完善，使图表达达到最佳的可视化效果。本书详细介绍了柱形图、饼图、折线图和散点图等图表的绘制方法及各自的优缺点，还用专门的一章介绍与地图相关的数据可视化技巧。

本书实例丰富、图文并茂，适合数据分析师、视觉设计师和对数据感兴趣的开发人员学习提高。

图灵程序设计丛书

## 鲜活的数据：数据可视化指南

- 
- ◆ 著 [美] Nathan Yau
  - 译 向怡宁
  - 责任编辑 朱 巍
  - 执行编辑 罗词亮
  
  - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号  
邮编 100061 电子邮件 315@ptpress.com.cn  
网址 <http://www.ptpress.com.cn>  
北京艺辉印刷有限公司印刷
  
  - ◆ 开本：800×1000 1/16  
印张：18.5 彩插 16  
字数：437千字 2012年10月第1版  
印数：1-4 000册 2012年10月北京第1次印刷
- 著作权合同登记号 图字：01-2012-3283号

ISBN 978-7-115-29381-7

定价：69.00元

读者服务热线：(010)51095186转604 印装质量热线：(010)67129223

反盗版热线：(010)67171154

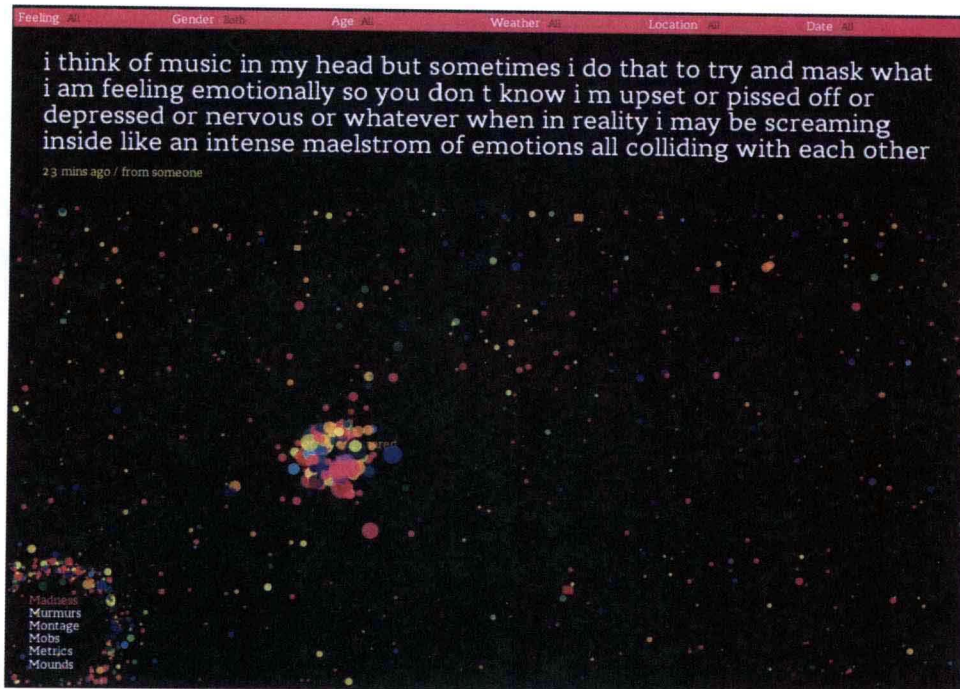


图 1-2 Jonathan Harris 和 Sep Kamvar 的作品 *We Feel Fine*

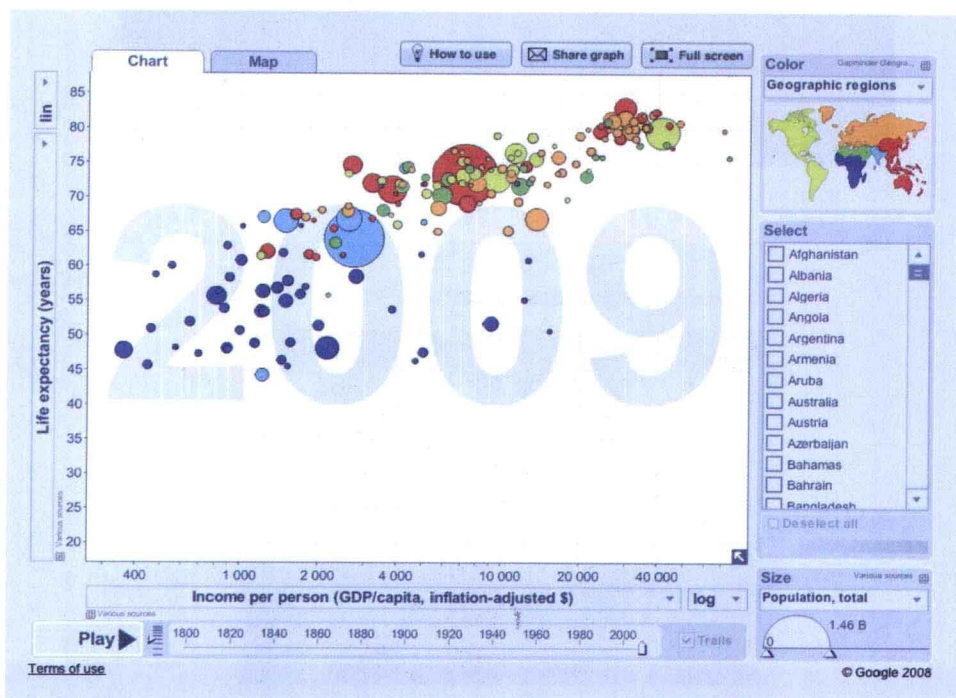


图 1-4 Gapminder 基金会开发的 Trendalyzer 软件



图 1-9 世界各国 Firefox 的使用情况



图 1-10 Firefox 使用区域数据图的说明

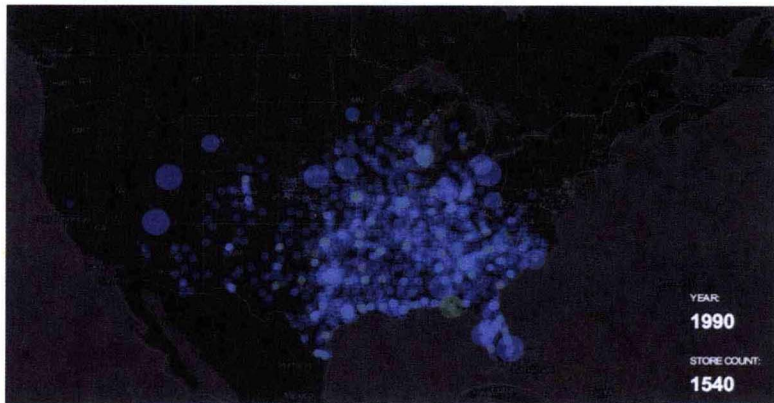


图 3-13 动画演示沃尔玛企业成长的地图，以 ActionScript 写成

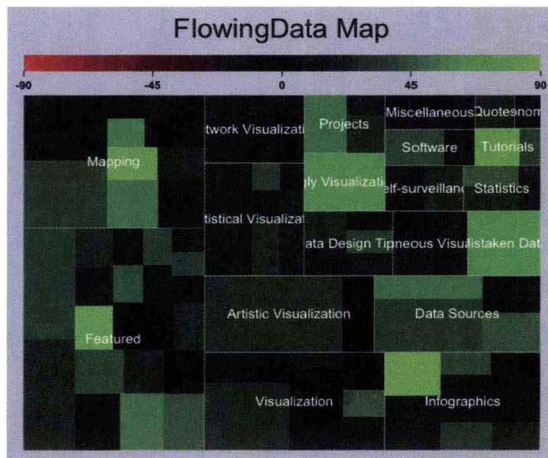


图 3-19 在 R 中利用 Portfolio 工具包生成的板块层级图

哪些与数据相关的领域你最感兴趣？

以下是FlowingData网站在2009年5月进行的一次投票的结果。参与者来自于多个领域，但统计学、设计和商务领域成为了主导。

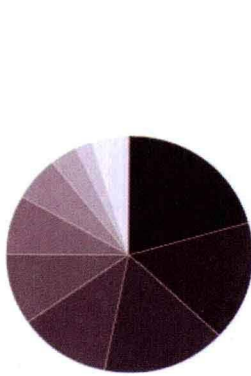


图 5-6 着色后的饼图

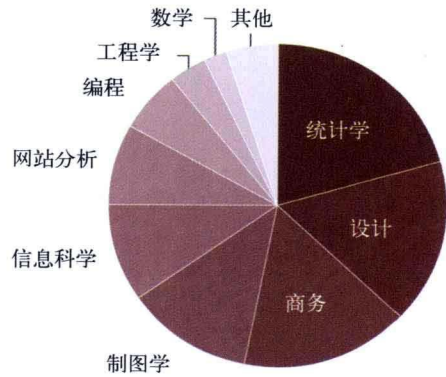


图 5-7 带有标签和说明文字的最终饼图

### 巴拉克·奥巴马的支持率

近期民意调查显示，民众对巴拉克·奥巴马在种族关系问题上的举措持有52%的支持率。这也是以下13个问题中他唯一获得多数支持的举措。在其中8个问题上他都获得了多数反对。

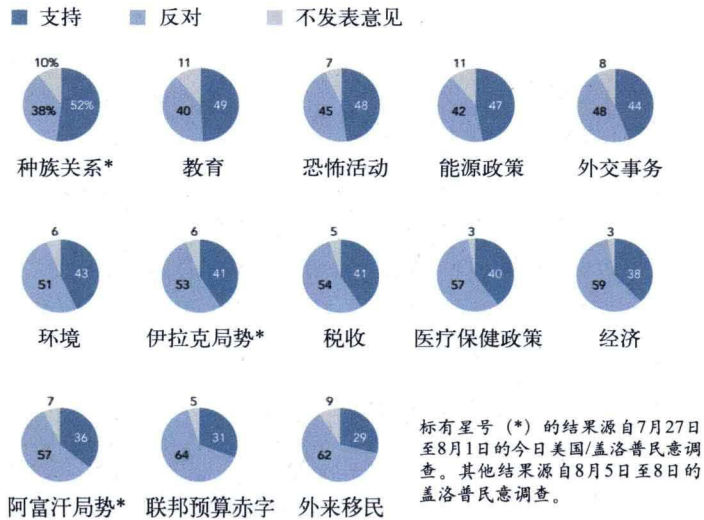


图 5-12 系列饼图

### 巴拉克·奥巴马的支持率

近期民意调查显示，民众对巴拉克·奥巴马在种族关系问题上的举措持有52%的支持率。这也是以下13个问题中他唯一获得多数支持的举措。在其中8个问题上他都获得了多数反对。

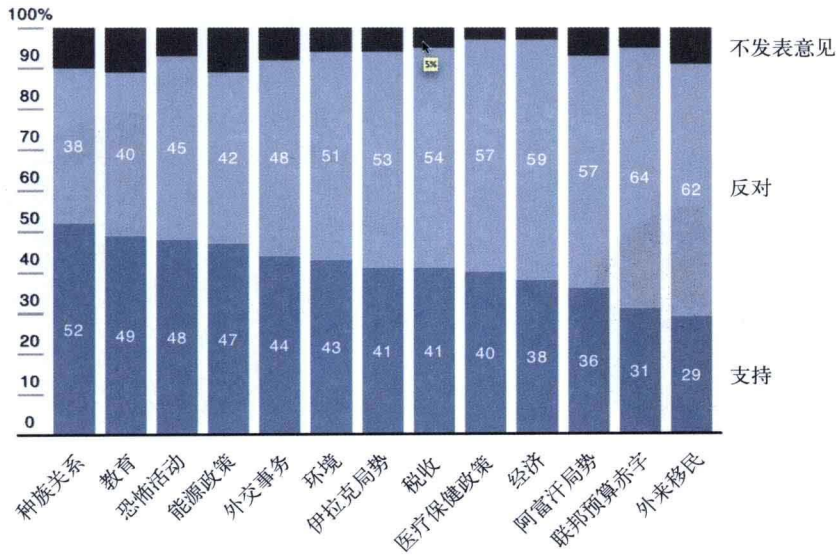


图 5-13 利用 Protovis 创建的可交互堆叠柱形图

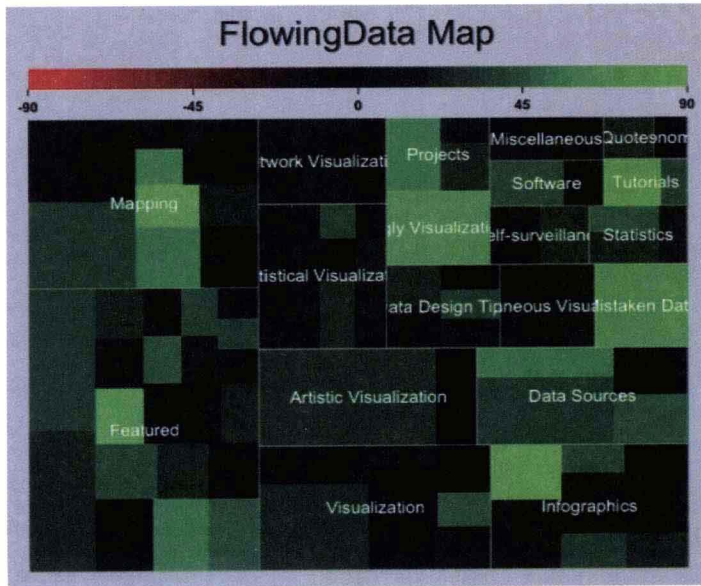


图 5-18 R 默认生成的板块层级图

Reported Occupations - U.S. Labor Force, 1850 - 2000 (source: <http://ipums.org>)

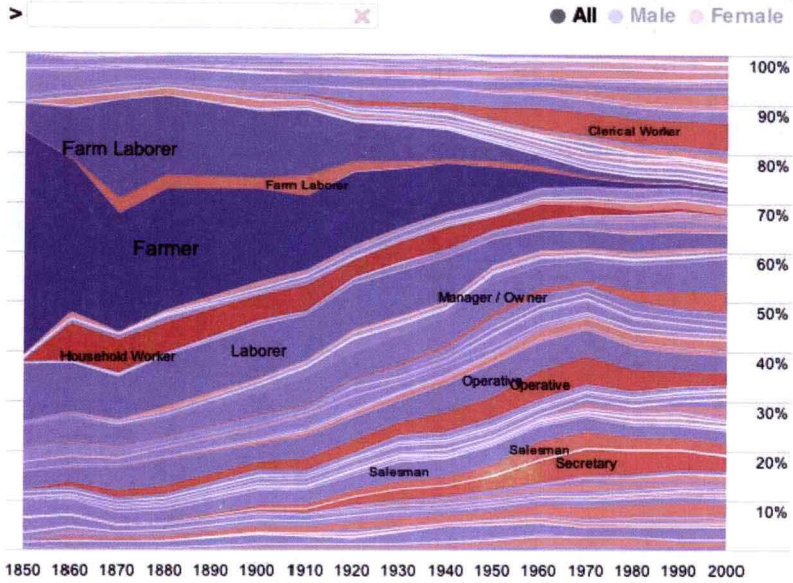


图 5-31 JobVoyager 应用

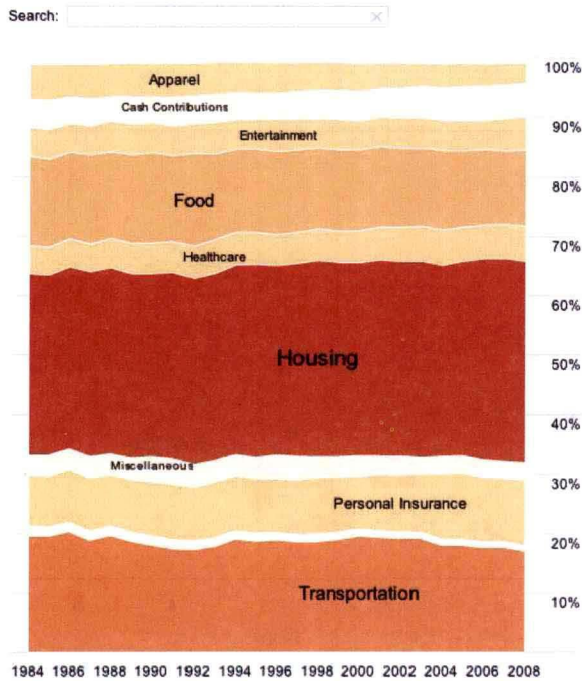


图 5-32 有关消费开支的可交互式堆叠面积图



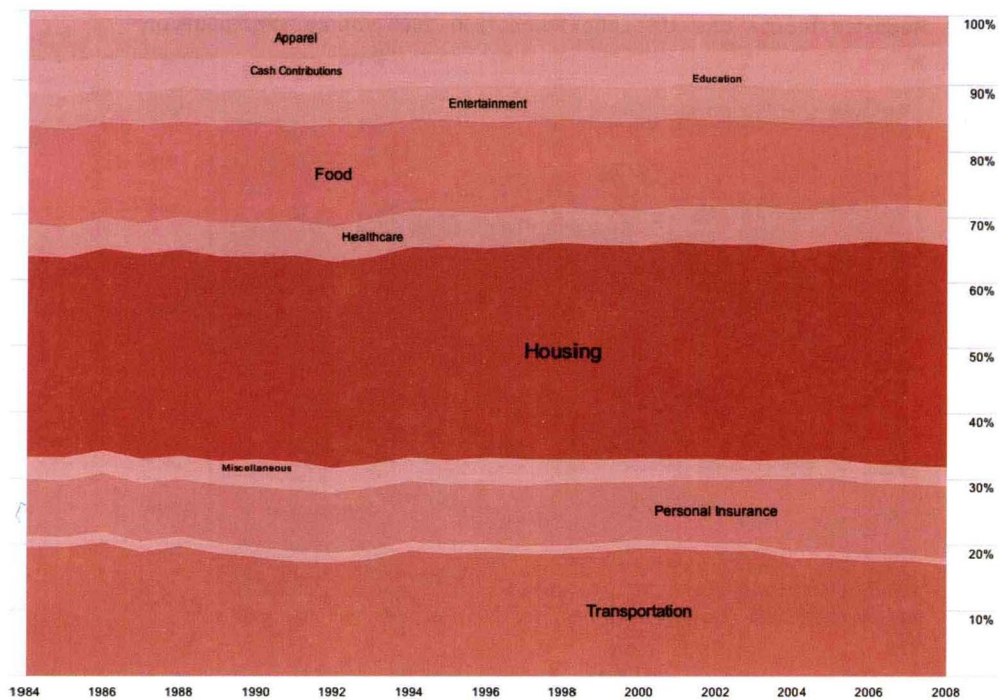


图 5-33 基础色调的堆叠面积图

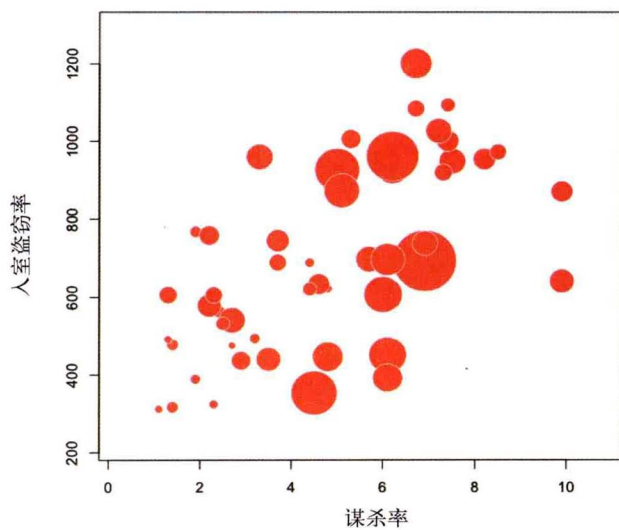


图 6-18 气泡缩小后的气泡图

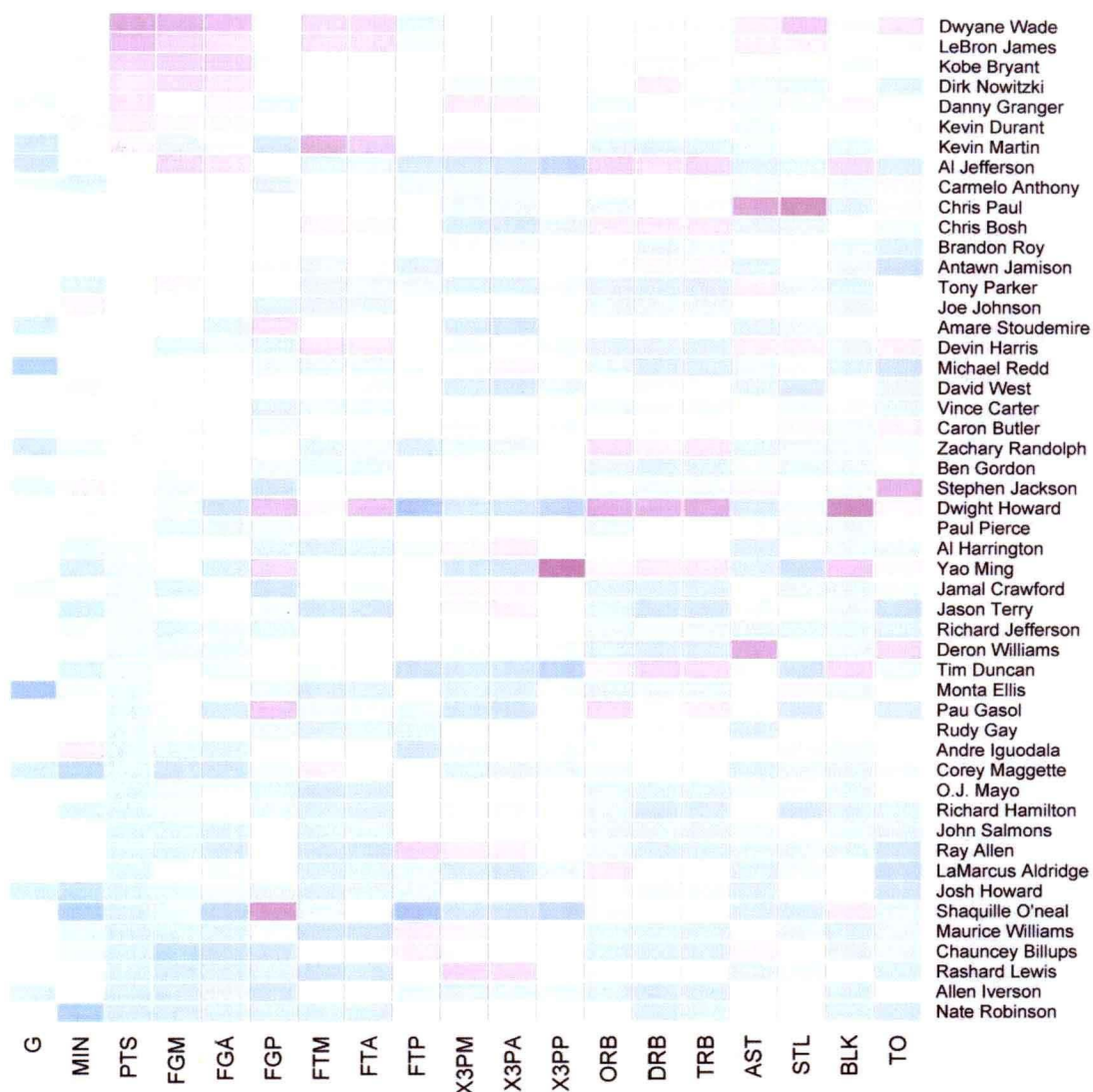


图 7-3 按 PTS 排序的默认热点图

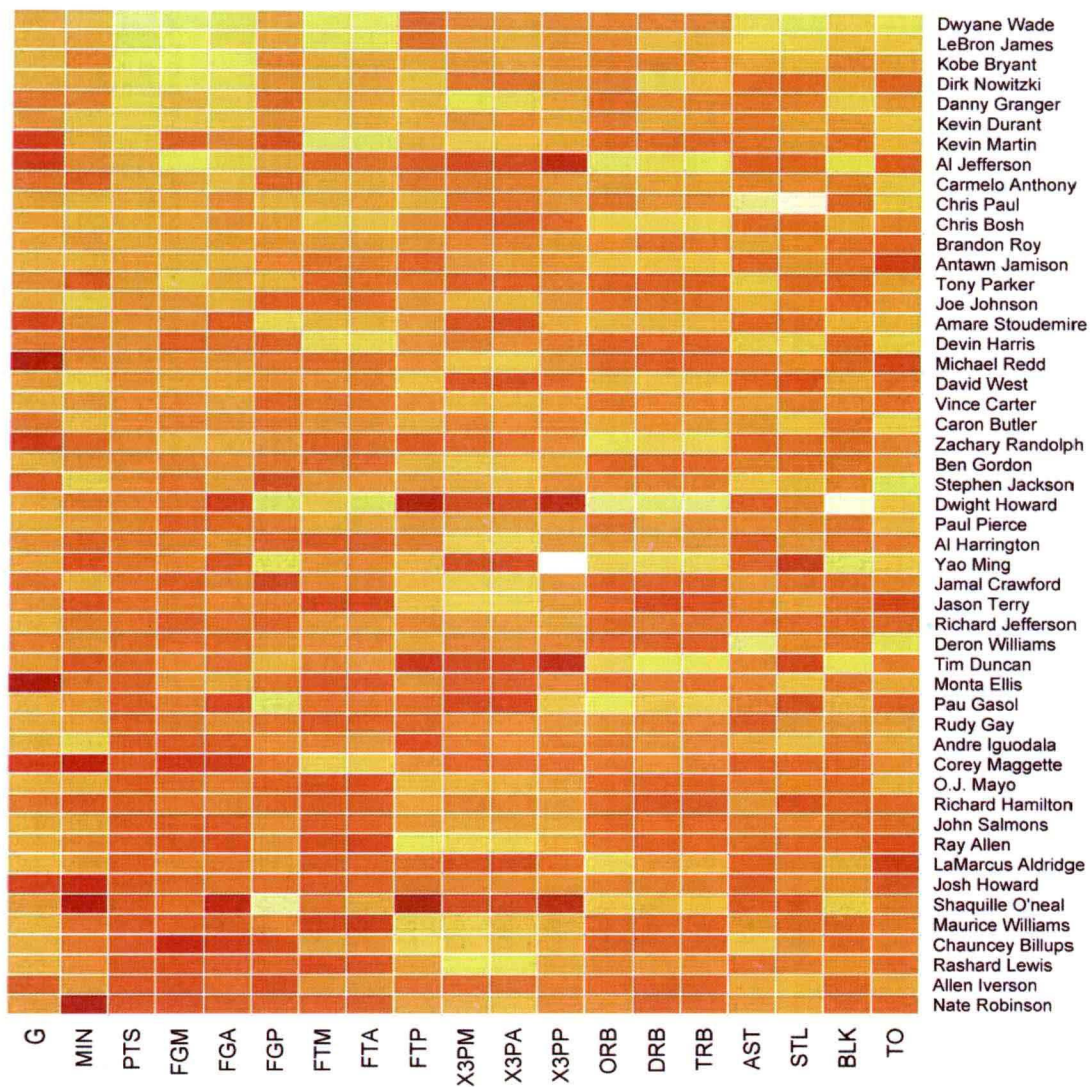


图 7-4 红-黄色系的热点图

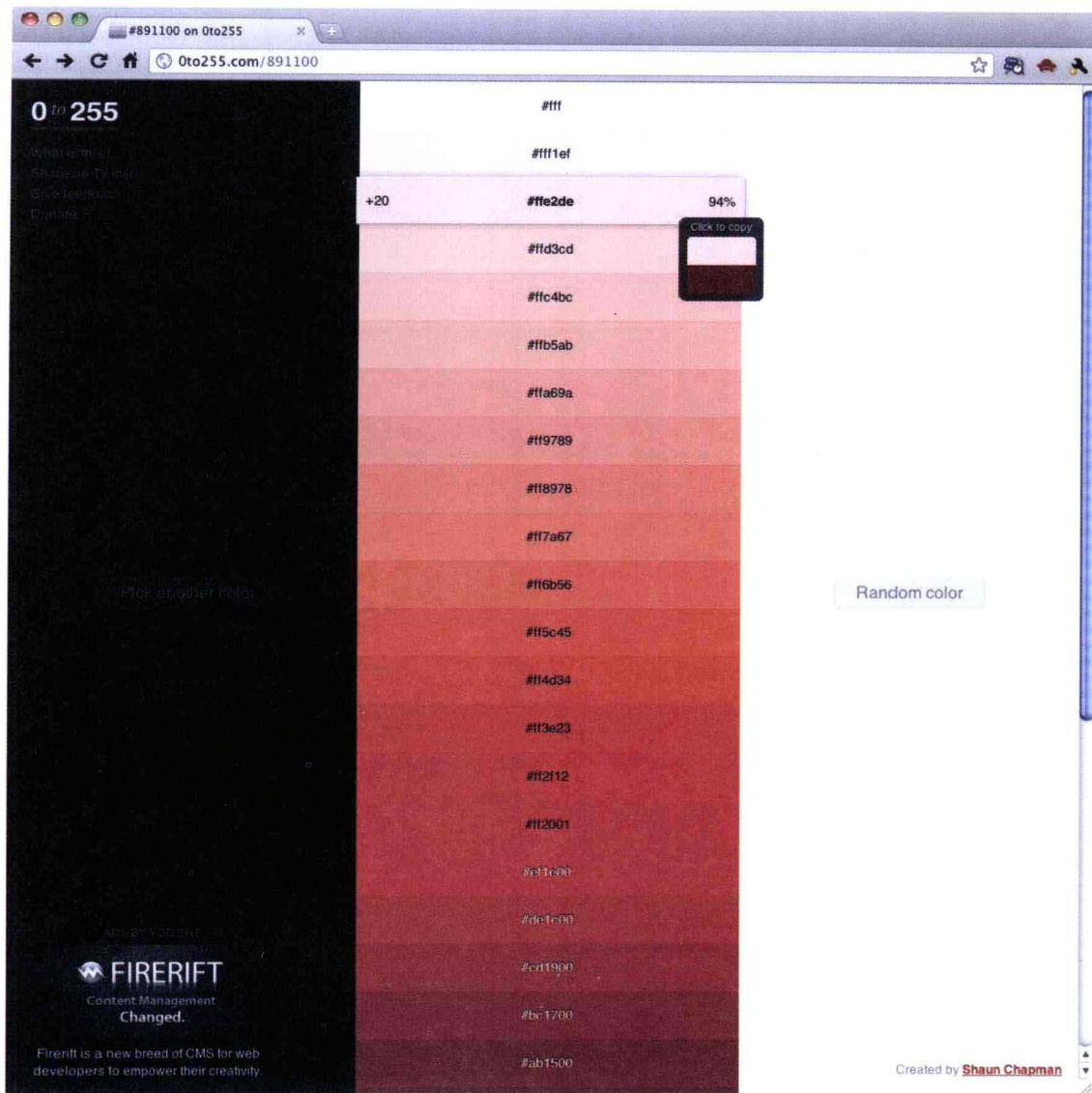


图 7-5 Oto255.com 里面的红色系色阶

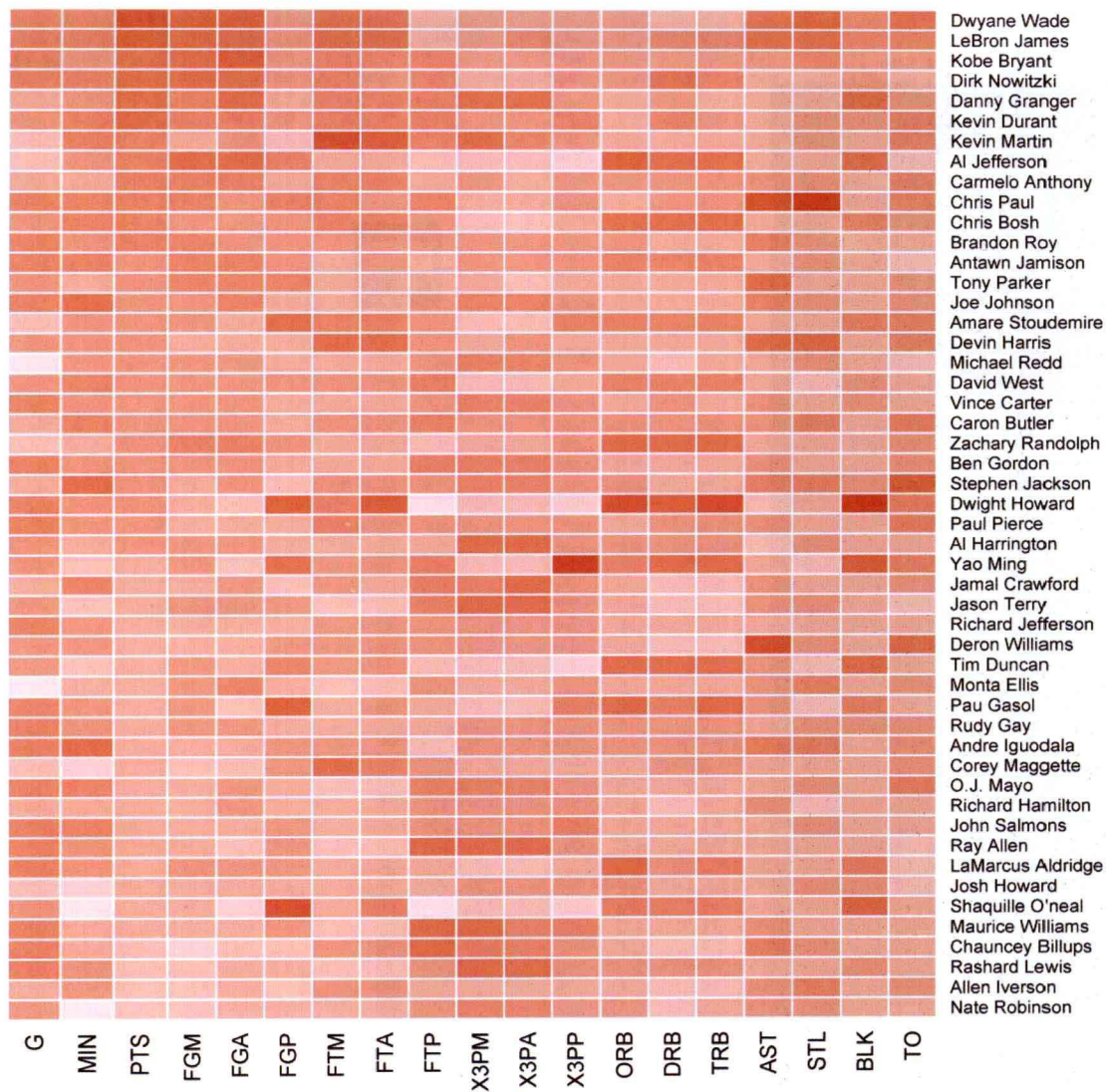


图 7-6 使用了自定义红色系的热点图

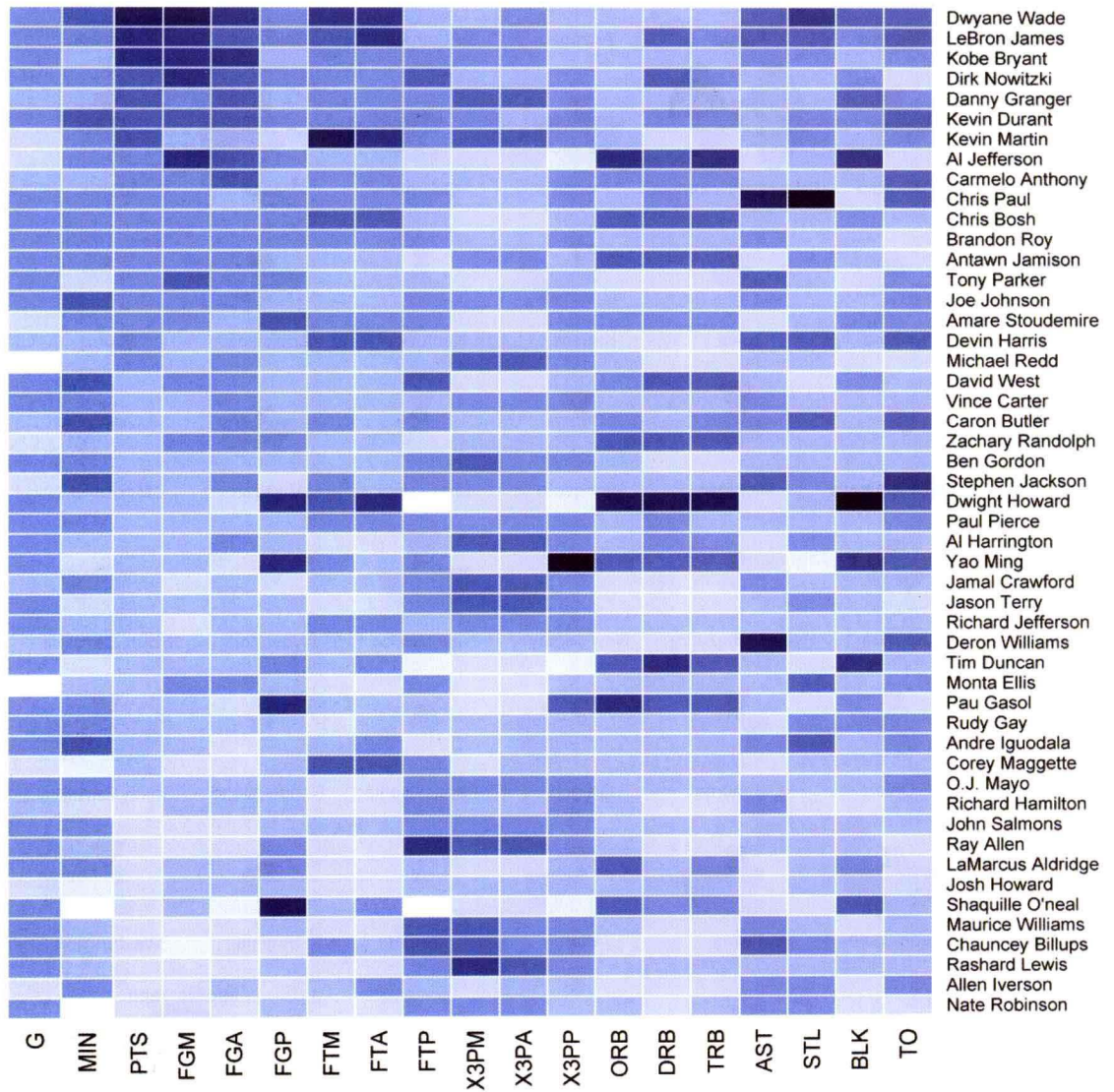


图 7-7 利用 RColorBrewer 作为调色盘的热点图

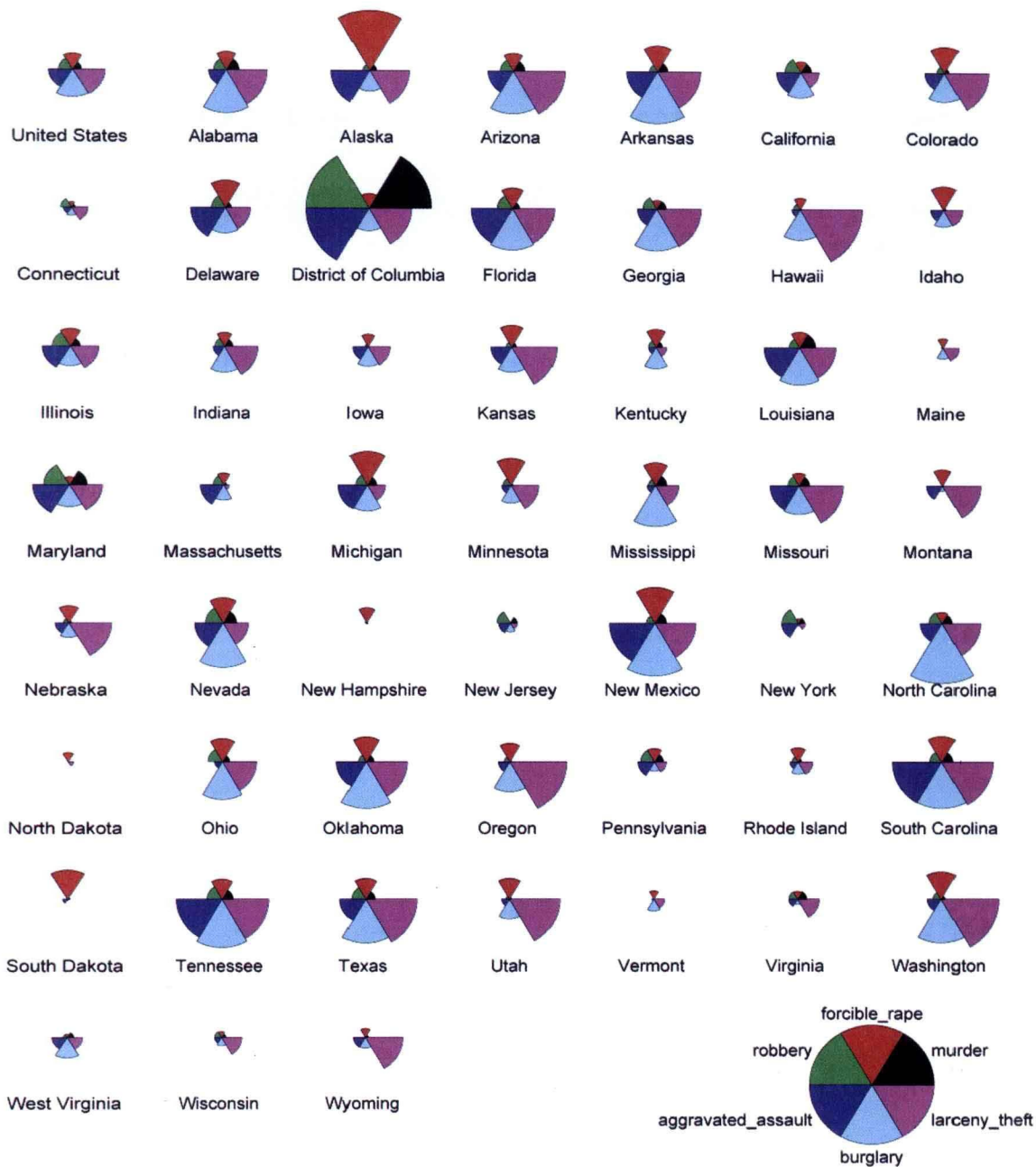


图 7-18 南丁格尔图显示的犯罪率

## 美国的教育

SAT考试中阅读科目平均得分较高的州，一般在数学及写作科目也会得到较高的分数。然而，这并不意味着这些州的学生获得了更好的教育。它似乎只能指出这些州有多大比例的毕业生参加了SAT考试。

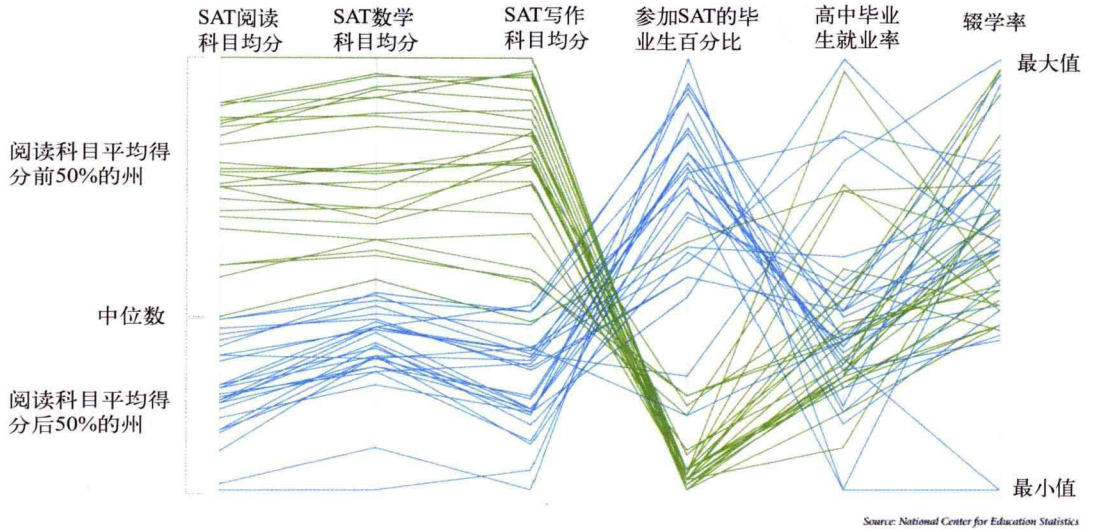


图 7-26 关于 SAT 得分的平行坐标图表

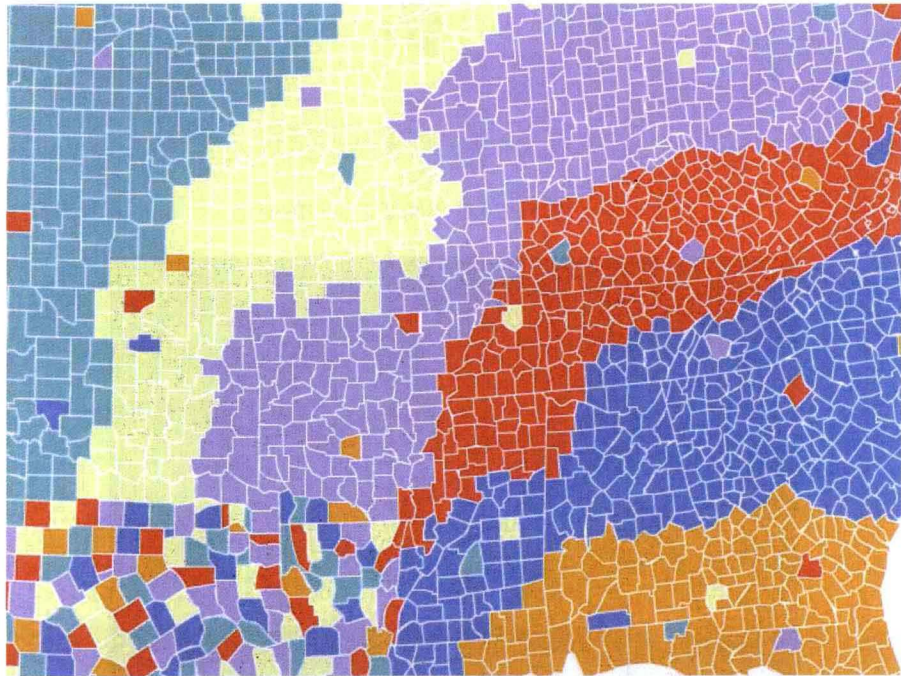


图 8-14 ColorBrewer 实现的定性配色



### 安全饮用水资源的获取

大多数人都能获得洁净、可饮用的水资源，但并不是每个人都这么幸运。这张地图显示了2009年各个国家能“合理”获得安全引用水资源的城市居民百分比。

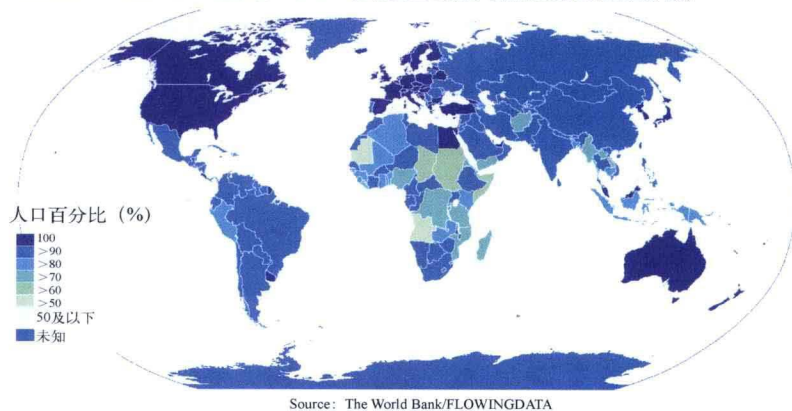


图 8-23 完成的世界地图

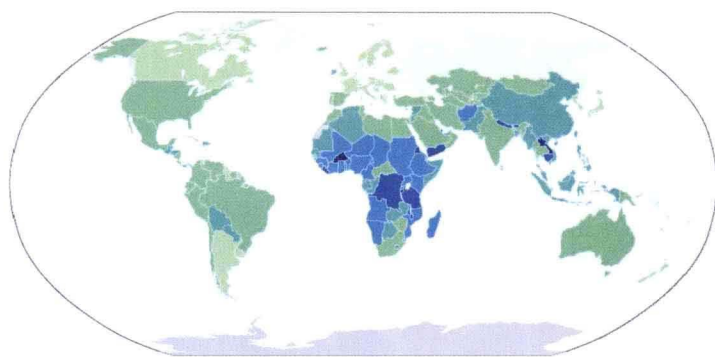


图 8-30 2005—2009 年间城市人口的变化

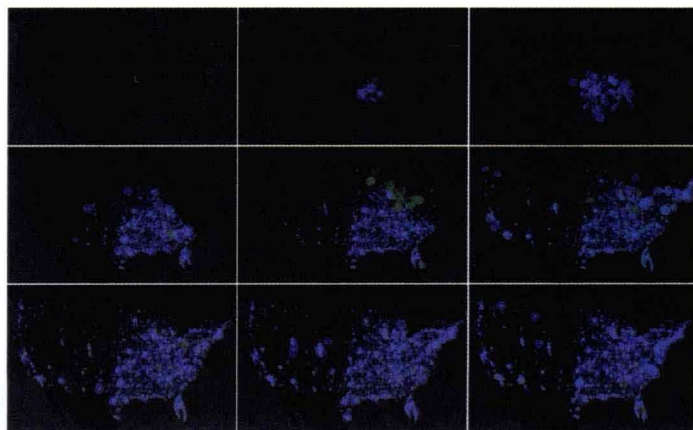


图 8-34 表现沃尔玛增长的动画地图