# 《数据驱动安全》

## 图书基本信息

书名：《数据驱动安全》

13位ISBN编号：9787111512677

出版时间：2015-9

作者：杰.雅克布,鲍布·鲁迪斯

页数：291

译者：薛杰,王占一,张卓

# 《数据驱动安全》

## 内容概要

本书由世界顶级安全专家亲笔撰写，深入剖析了安全领域中的数据分析及可视化方法，包含大量真实案例和数据。从安全数据收集、整理、分析、可视化过程，详细讲解如何设计有效的安全数据可视化，并走向数据驱动的安全研究。主要内容包括：第1章展示信息安全领域数据分析与可视化的基础知识，以及安全数据科学工作者需要掌握的技能概览。第2、3、4章分别介绍一些安全数据科学工作者需要掌握的软件工具、技术知识、使用技巧，涉及Python语言、R语言为主的实用分析方法。第5章介绍创建图表的技术以及一些核心的统计学概念。第6章讲解数据可视化的基础知识，以及有效展示的技巧。第7章介绍如何对安全漏洞进行分析和可视化，包含大量安全事件的真实数据。第8章涵盖现代数据库的概念，包含在传统数据库基础上新增的数据展示技巧以及NoSQL解决方案。第9章将带你进入机器学习领域，包括机器学习的核心概念，探索机器学习实现技术等。第10章及第11章介绍创建有效的可视化产品技巧，以及如何让这些信息展示得更加丰富有形。在第12章呈现如何将所学的知识应用到实际的安全环境中。

# 《数据驱动安全》

## 作者简介

Jay Jacobs在IT以及信息安全领域拥有超过15年的经验，主要致力于密码学、风险学以及数据分析方面的研究。作为Versizon RISK团队的一名高级数据分析师，他参与编纂年度《Data Breach Investigation Report》，并投入大量精力进行安全相关数据的分析与可视化。Jay也是Society of Information Risk Analysts的创立人之一，现在是该组织董事会的成员。他是一名活跃的博客纂稿人与演讲者，他还是Risk Science播客的主持人并且曾经是2014 Metricon安全指标/分析大会的联席主席。可以通过@jayjacobs在推特上找到他。他拥有美国康卡迪亚大学科技管理的学士学位以及美国宾夕法尼亚州立大学的应用统计学毕业证书。

Bob Rudis拥有超过20年的利用数据来帮助维护全球财富100强企业的经验。作为Liberty Mutual Insurance的企业信息安全及IT风险管理部门的主管，他负责协调与管理Advanced Cyber Security Center的多部门大范围安全分析计划。Bob是一名高级推特撰写人（@hrbrmster）、活跃的博主（rud.is）、作家、演讲者以及开源社区的投稿人（github.com/hrbrmstr）。他当前正任职于Society of Information Risk Analysts（SIRA）的董事会，是SANS Securing The Human方案的编委，同时，还是2014年Metricon安全指标/分析会议的联合主席。他拥有斯克兰顿大学的学士学位。

技术编辑Russell Thomas是一名Zions Bancorporation的安全数据科学家，还是一名乔治梅森大学社会计算科学的在读博士研究生。他拥有在计算机行业超过30年的技术、管理以及咨询方面的经验。Thomas先生是Securitymetrics.org的长期社区会员和Society of Information Risk Analysts（SIRA）的创始成员之一。

# 《数据驱动安全》

## 书籍目录

# 《数据驱动安全》

## 精彩短评

1、烂翻译
2、本书从数据分析和数据可视化的角度来定位、分析安全问题，主要内容包括数据驱动安全的基础原理，常用的数据分析编程语言Python和R，数据分析代表性的软件工具，数据可视化和机器学习的基础算法，以及如何有效的设计交互式安全仪表盘。本书从数据的角度来分析安全，这是网络空间安全的一个热门领域，如何采用人工智能、机器学习从海量的网络数据中发现潜在的威胁，已经发生的安全事件，攻击踪迹，甚至威胁情报信息，这是非常值得关注的问题。

3、360团队翻译的，并且把数据驱动安全作为公司宣传语，但这本书讲的真没意思，翻译的不好，对不起这个名字

# 《数据驱动安全》

## 章节试读

1、《数据驱动安全》的笔记-第12页

询问我们身边人们的看法会导致错误地肯定自身的观点，因为我们很自然地和志同道合的人聚集在一起，且想法趋同一致。

2、《数据驱动安全》的笔记-第17页

分析工作的背景以及目的是根据研究问题来制定的，而不是源于我们可获取的数据。

3、《数据驱动安全》的笔记-第9页

Excel作为临时的解决方案，其能很好地快速处理一次性任务。但是如果你有一个需要重复分析的任务或者反复使用的模型的话，最好用某种结构化编程语言来处理。
作为一种数据清理工具，使用电子表格初看起来是一个不错的解决方案（尤其是对一些熟悉这方面技能的人来讲），但是电子表格是事件驱动的，意味着它们需要通过点击、打字、拖拽来工作。如果想用来转行一行数据，你就不得不点击表格，选中该行数据，然后再转换数据。这适合一些小的数据集或者快速的任务，但是相信我，你将会（比预期的还频繁）不得不回溯原始数据然后重新清理它。某一天，也许你有一些新的日志文件需要处理，也许你会意识到应该再从原始数据中提取另外的数据关系，也许（累得喘息）你在数据清理过程中发现了一个错误。也许不止一次地，某个点、某个处理细节会导致你重新回溯原始数据，然后重复数据清理以及转换的过程，利用电子表格的话，意味着你需要更多的无数次点击。然而，写一个脚本来运行的话，就可以很轻易、灵活以及一致地执行数据清理过程。

# 《数据驱动安全》