



Top Quant
CHRD前海智库
CHINA FRONTIER RESEARCH AND DEVELOPMENT CENTER
金融工程 · 量化策略



金融科技
丛书

零起点 Python大数据 与量化交易

何海群 著

Win Or Home

读完本书内容和配套的教学代码，
就能够编写简单的量化策略函数。

100%零基础，无须任何编程、交易经验，也不需要具备超强的数据分析能力，只要会使用Excel就可以轻松学会本书讲解的知识点。

 中国工信出版集团

 电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>



零起点 Python大数据 与量化交易

何海群 著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书是国内较早关于 Python 大数据与量化交易的原创图书，配合 zwPython 开发平台和 zwQuant 开源量化软件学习，是一套完整的大数据分析、量化交易的学习教材，可直接用于实盘交易。本书有三大特色：第一，以实盘个案分析为主，全程配有 Python 代码；第二，包含大量的图文案例和 Python 源码，无须专业编程基础，懂 Excel 即可开始学习；第三，配有专业的 zwPython 集成开发平台、zwQuant 量化软件和 zwDat 数据包。

本书内容源自作者的原版教学课件，虽然限于篇幅和载体，省略了视频和部分环节，但核心内容都有保留，配套的近百套 Python 教学程序没有进行任何删减。考虑到广大入门读者的需求，笔者在各个核心函数环节增添了函数流程图。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目（CIP）数据

零起点 Python 大数据与量化交易 / 何海群著. —北京：电子工业出版社，2017.2
（金融科技丛书）
ISBN 978-7-121-30659-4

I. ①零… II. ①何… III. ①软件工具—程序设计 IV. ①TP311.561

中国版本图书馆 CIP 数据核字（2016）第 308385 号

责任编辑：黄爱萍

印 刷：三河市双峰印刷装订有限公司

装 订：三河市双峰印刷装订有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编：100036

开 本：787×980 1/16 印张：27.5 字数：528 千字

版 次：2017 年 2 月第 1 版

印 次：2017 年 2 月第 1 次印刷

定 价：99.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zltts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：（010）51260888-819，faq@phei.com.cn。

丛书编委会

主编：何海群

编委：

欧耘华，CHRD 前海智库创始人，Python 产业联盟发起人。科技金融、消费金融、艺术金融与产业政策专家。北京亚欧科技、深圳“中国科谷”等多家机构特聘专家。

刘志明，免费开源金融数据接口 TuShare 创始人，专业从事量化投资支持与服务。

杨适安，中信建投证券金融产品与创新业务部副总裁，对期权、期货等衍生品种有较深入的研究，擅长用中低频量化的方法配置大类资产和金融产品。

吴尚谦，金融交易实盘培训专家，纯实盘首席金融交易教练，15 年股票、期货、外汇实盘交易经验，金融交易混元合一交易系统发明人，业内领先的现场实盘逻辑推演教学法资深讲师。北京正道阳光投资管理有限公司总经理。

曹嘉和，美国普林顿大学荣誉博士，北大当代企业文化研究所研究员。现任北大中国持续发展研究中心政府与社会资本合作研究所所长，未名湖智库秘书长。

王丁杰，QuantDigger 量化软件开发者，擅长机器学习、NLP 自然语言处理、人工智能。

王黎中，资深金融交易心理专家，北京正道阳光投资管理有限公司副总经理，幸福紫薇俱乐部董事长。

王家苍，经济学博士，李约瑟（《中国科学技术史》作者）隔代弟子，产业政策

专家。长期致力于科技和文化的融合研究，参与科技金融、消费金融和艺术金融等课题的研究。

蔡磊，原高通公司工程师、项目经理，精通 Python 数据挖掘、数据库技术、机器学习、量化交易理论，熟悉 4G 无线技术、手机芯片架构及 VOIP 技术，极宽量化开源团队核心成员。

王硕，高级软件工程师，精通 Python 数据分析，擅长 Java、JavaScript、HTML5 和数据库技术，极宽量化开源团队核心成员。

吴娜，电信数据挖掘工程师，精通缠论和江恩波动理论，率先研发江恩波动法则量化模型，著有《游戏数据分析的艺术》，极宽量化开源团队发起成员。

余勤，AMD 验证工程师，擅长数据分析，热爱 Python 量化分析，极宽量化开源团队发起成员。

孙洋洋，西南财经大学金融工程硕士，擅长机器学习、网络爬虫，有私募机构量化分析系统开发实盘经验，极宽量化开源团队发起成员。

李政隆，上海交大计算机硕士，分布式系统架构师，擅长金融数据抓取与分析，极宽量化开源团队发起成员。

前 言

2014 年，美国银行、美林证券的“石英”项目、摩根大通的“雅典娜”项目都不约而同地选择了 Python 作为金融行业的标准编程语言。

全世界的金融工程行业全部重新洗牌，这为中国的金融工程从业人员带来了前所未有的机遇。资本的力量是强大的，也是冷酷无情的。

2016 年 5 月，《华尔街日报》报道，目前华尔街的三大编程语言是：C、Java 与 Python。其中，C 与 Java 成为三大语言之一有两方面原因：一方面是由于历史积累，另一方面是系统架构设计的需要。而在应用领域 Python 更胜一筹，因为 Python 已经成为金融行业量化领域的标准编程语言。

本书是国内较早关于 Python 大数据与量化交易的原创图书。本书配合 zwPython、zwQuant 开源量化软件学习，是一套完整的大数据分析、量化交易的学习教材，可直接用于实盘交易。

本书内容包括：

- 近 50 万字的图文课件；
- 数十套结合课件的 Python 教学代码；
- 全套 zwPython 开源平台；
- 业内首套面向初学者的开源量化系统 zwQuant；
- 国内较大的开源金融数据包 zwDat，包括 tick 数据。

100%零基础，无须任何编程、交易经验，也不需要具备超强的数据分析能力，

只要会使用 Excel 就可以轻松学会本书讲解的知识点。读完本书内容和配套的教学代码，就能够编写简单的量化策略函数。

本书的内容源自笔者的原版教学课件，虽然限于篇幅和载体，省略了视频和部分环节，但核心内容都有保留，配套的近百套 Python 教学程序没有进行任何删减。

考虑到广大入门读者的需求，笔者在各个核心函数环节增添了函数流程图。

量化新人学习指南

1. 多参考笔者的字王量化网站 (<http://www.ziwan.com>) 和笔者博客 (以原创为主，网址是 <http://blog.sina.com.cn/zbrow>)。

2. 本书配套程序可在百度网盘下载，网址是 <http://pan.baidu.com/s/ljIg944u>。本书读者 QQ 群 (zwPython 量化总群): 124134140。网盘和群共享包含很多资源，读者可自行下载，也可以上传。建议先看“漫画学××系列”，该系列比较经典，而且简单；zwQuant 开源量化软件在群共享和网盘都可下载。

3. 强烈建议初学者先下载 zwPython 集成版，阅读 zwPython 中文手册，再开始学习 Python，这样可以少走很多弯路。

资源下载

为配合本书出版，方便广大读者学习 Python 量化系统，笔者特意将与本书相关的教学资源打包成一个独立的教学版本压缩包，便于读者使用。

- 教学版为一个独立压缩包，解压即可，并配有说明文档。
- 教学版内置了 zwPython2016M10 版开发平台 (Python 3.5 版)、Python 2.7 版开发平台、zwDat 金融数据包、zw_down25 金融数据下载更新程序包和 zw_k10 配套量化教学课件程序。

教学版软件和配套资源下载地址如下。

- 极宽公司网盘下载地址: <http://pan.baidu.com/s/1kVO6T19>，密码: v2ub。

- 出版社网站下载地址：<http://www.broadview.com.cn/30659>。

以上网址如果发生变动，请读者浏览字王网站（<http://www.ziwan.com>）或者极宽公司网站（<http://www.zquant.cn>）获取最新信息。

致谢

虽然很多网友在笔者博客留言，要求购买本书，但本书的创作和正式出版还是经历了许多波折。

如今本书终于出版，在此，要特别感谢电子工业出版社的黄爱萍和戴新编辑，感谢她们在选题策划和稿件整理方面做出的大量工作。

同时，在本书创作过程中，极宽开源量化团队和培训班的全体成员，提出很多宝贵的意见，并对部分课件程序做了中文注解。

特别是吴娜、余勤两位同学，为极宽开源量化文库和 zwQuant 开源量化软件编写文档，以及在团队成员管理方面做了大量工作，为他们的付出表示感谢。

何海群（字王）

北京极宽科技有限公司 CTO

2016年11月25日

目 录

第 1 章 从故事开始学量化	1
1.1 亿万富翁的“神奇公式”	2
1.1.1 案例 1-1: 亿万富翁的“神奇公式”	2
1.1.2 案例分析: Python 图表	5
1.1.3 matplotlib 绘图模块库	7
1.1.4 案例分析: style 绘图风格	10
1.1.5 案例分析: colormap 颜色表	12
1.1.6 案例分析: 颜色表关键词	14
1.1.7 深入浅出	17
1.2 股市“一月效应”	18
1.2.1 案例 1-2: 股市“一月效应”	18
1.2.2 案例分析: “一月效应”计算	19
1.2.3 案例分析: “一月效应”图表分析	24
1.2.4 案例分析: 颜色表效果图	26
1.2.5 “一月效应”全文注解版 Python 源码	27
1.2.6 大数据·宏分析	34
1.3 量化交易流程与概念	36
1.3.1 数据分析 I2O 流程	36
1.3.2 量化交易不是高频交易、自动交易	37
1.3.3 小资、小白、韭菜	38

1.3.4	专业与业余	38
1.4	用户运行环境配置	42
1.4.1	程序目录结构	43
1.4.2	金融股票数据包	44
1.5	Python 实战操作技巧	46
1.5.1	模块检测	46
1.5.2	Spyder 编辑器界面设置	47
1.5.3	代码配色技巧	48
1.5.4	图像显示配置	50
1.5.5	Python2、Python 3 双版本双开模式	51
1.5.6	单版本双开、多开模式	52
1.5.7	实战胜于一切	54
1.6	量化、中医与西医	54
第 2 章	常用量化技术指标与框架	56
2.1	案例 2-1: SMA 均线策略	56
2.1.1	案例要点与事件编程	58
2.1.2	量化程序结构	61
2.1.3	main 程序主入口	61
2.1.4	KISS 法则	63
2.2	Python 量化系统框架	64
2.2.1	量化行业关键词	64
2.2.2	国外主流 Python 量化网站	65
2.2.3	我国主流 Python 量化网站	67
2.2.4	主流 Python 量化框架	70
2.3	常用量化软件包	78
2.3.1	常用量化软件包简介	79
2.3.2	案例 2-2: 模块库列表	80
2.4	常用量化技术指标	82
2.4.1	TA-Lib 金融软件包	83
2.4.2	案例 2-3: MA 均线函数调用	84

2.4.3	TA-Lib 函数调用	86
2.4.4	量化分析常用指标	88
2.5	经典量化策略	90
2.5.1	阿尔法 (Alpha) 策略	90
2.5.2	Beta 策略	92
2.5.3	海龟交易法则	93
2.5.4	ETF 套利策略	95
2.6	常用量化策略	95
2.6.1	动量交易策略	96
2.6.2	均值回归策略	97
2.6.3	其他常用量化策略	98
2.7	起点与终点	100
第 3 章	金融数据采集整理	101
3.1	常用数据源 API 与模块库	102
3.1.1	大数据综合 API	102
3.1.2	专业财经数据 API	103
3.1.3	专业数据模块库	104
3.2	案例 3-1: zwDatX 数据类	104
3.3	美股数据源模块库	108
3.4	开源文档库 Read the Docs	109
3.5	案例 3-2: 下载美股数据	110
3.6	财经数据源模块库 TuShare	113
3.6.1	沪深股票列表	115
3.6.2	案例 3-3: 下载股票代码数据	116
3.6.3	CSV 文件处理	119
3.7	历史数据	121
3.7.1	历史行情	121
3.7.2	案例 3-4: 下载近期股票数据	124
3.7.3	历史复权数据	130
3.7.4	案例 3-5: 下载历史复权数据	131

3.8 其他交易数据	134
3.9 zwDat 超大股票数据源与数据更新	143
3.9.1 案例 3-6: A 股基本概况数据下载	144
3.9.2 案例 3-7: A 股交易数据下载	146
3.9.3 案例 3-8: A 股指数行情数据下载	150
3.9.4 案例 3-9: 美股交易数据下载	151
3.10 数据归一化处理	153
3.10.1 中美股票数据格式差异	153
3.10.2 案例 3-10: 数据格式转化	154
3.10.3 案例 3-11: A 股策略 PAT 实盘分析	156
3.10.4 案例 3-12: 数据归一化	158
3.11 为有源头活水来	160
第 4 章 PAT 案例汇编	162
4.1 投资组合与回报率	163
4.1.1 案例 4-1: 下载多组美股数据	163
4.1.2 案例 4-2: 投资组合收益计算	165
4.2 SMA 均线策略	168
4.2.1 SMA 简单移动平均线	168
4.2.2 案例 4-3: 原版 SMA 均线策略	169
4.2.3 案例 4-4: 增强版 SMA 均线策略	173
4.2.4 案例 4-5: A 股版 SMA 均线策略	174
4.3 均线交叉策略	175
4.3.1 案例 4-6: 均线交叉策略	176
4.3.2 案例 4-7: A 股版均线交叉策略	178
4.4 VWAP 动量策略	181
4.4.1 案例 4-8: VWAP 动量策略	182
4.4.2 案例 4-9: A 股版 VWAP 动量策略	183
4.5 布林带策略	183
4.5.1 案例 4-10: 布林带策略	185
4.5.2 案例 4-11: A 股版布林带策略	186

4.6	RSI2 策略	188
4.6.1	案例 4-12: RSI2 策略	190
4.6.2	案例 4-13: A 股版 RSI2 策略	190
4.7	案例与传承	194
第 5 章 zwQuant 整体架构		196
5.1	发布前言	196
5.2	功能简介	197
5.2.1	目录结构	197
5.2.2	安装与更新	198
5.2.3	模块说明	199
5.2.4	zwSys 模块: 系统变量与类定义	200
5.2.5	zwTools 模块: 常用(非量化)工具函数	201
5.2.6	zwQTBox: 常用“量化”工具函数集	201
5.2.7	zwQTDraw.py: 量化绘图工具函数	203
5.2.8	zwBacktest: 回溯测试工具函数	203
5.2.9	zwStrategy: 策略工具函数	203
5.2.10	zw_TA-Lib: 金融函数模块	204
5.3	示例程序	207
5.4	常用量化分析参数	208
5.5	回溯案例: 对标测试	209
5.5.1	对标测试 1: 投资回报参数	209
5.5.2	对标测试 2: VWAP 策略	211
5.6	回报参数计算	214
5.7	主体框架	220
5.7.1	stkLib 内存数据库	220
5.7.2	Bars 数据包	221
5.7.3	案例: 内存数据库&数据包	222
5.7.4	qxLib、xtrdLib	227
5.7.5	案例 5-1: qxLib 数据	228
5.7.6	量化系统的价格体系	230

5.7.7	数据预处理	231
5.7.8	绘图模板	234
5.8	新的起点	236
第 6 章	模块详解与实盘数据	237
6.1	回溯流程	238
6.1.1	案例 6-1: 投资回报率	238
6.1.2	代码构成	242
6.1.3	运行总流程	243
6.2	运行流程详解	244
6.2.1	设置股票数据源	244
6.2.2	设置策略参数	247
6.2.3	dataPre 数据预处理	249
6.2.4	绑定策略函数	253
6.2.5	回溯测试: zwBackTest	253
6.2.6	输出回溯结果数据、图表	258
6.3	零点策略	260
6.3.1	mul 多个时间点的交易&数据	263
6.3.2	案例 6-2: 多个时间点交易	264
6.4	不同数据源与格式修改	270
6.4.1	案例 6-3: 数据源修改	272
6.4.2	数据源格式修改	274
6.5	金融数据包与实盘数据更新	275
6.5.1	大盘指数文件升级	276
6.5.2	实盘数据更新	277
6.5.3	案例 6-4: A 股实盘数据更新	277
6.5.4	案例 6-5: 大盘指数更新	279
6.6	稳定第一	281
第 7 章	量化策略库	282
7.1	量化策略库简介	282

7.1.1	量化系统的三代目	283
7.1.2	通用数据预处理函数	283
7.2	SMA 均线策略	286
7.2.1	案例 7-1: SMA 均线策略	286
7.2.2	实盘下单时机与推荐	289
7.2.3	案例 7-2: 实盘 SMA 均线策略	290
7.3	CMA 均线交叉策略	294
7.3.1	案例 7-3: 均线交叉策略	294
7.3.2	对标测试误差分析	296
7.3.3	案例 7-4: CMA 均线交叉策略修改版	299
7.3.4	人工优化参数	300
7.4	VWAP 策略	301
7.4.1	案例 7-5: VWAP 策略	301
7.4.2	案例 7-6: 实盘 VWAP 策略	303
7.5	BBands 布林带策略	304
7.5.1	案例 7-7: BBands 布林带策略	305
7.5.2	案例 7-8: 实盘 BBands 布林带策略	306
7.6	大道至简 1+1	307
第 8 章	海龟策略与自定义扩展	309
8.1	策略库	309
8.1.1	自定义策略	310
8.1.2	海龟投资策略	310
8.2	tur 海龟策略 v1: 从零开始	311
8.3	案例 8-1: 海龟策略框架	311
8.4	tur 海龟策略 v2: 策略初始化	312
8.5	案例 8-2: 策略初始化	312
8.6	tur 海龟策略 v3: 数据预处理	313
8.7	案例 8-3: 数据预处理	314
8.8	tur 海龟策略 v4: 策略分析	317
8.9	案例 8-4: 策略分析	317

8.10	tur 海龟策略 v5: 数据图表输出	320
8.10.1	案例 8-5: 图表输出	320
8.10.2	参数优化	324
8.10.3	案例 8-6: 参数优化	324
8.11	tur 海龟策略 v9: 加入策略库	325
8.12	案例 8-7: 入库	326
8.13	庖丁解牛	328
第 9 章	TA-Lib 函数库与策略开发	329
9.1	TA-Lib 技术指标	329
9.1.1	TA-Lib 官网	329
9.1.2	矩阵版 TA-Lib 金融函数模块	330
9.2	MACD 策略	331
9.2.1	MACD 策略 1	331
9.2.2	案例 9-1: MACD_v1	335
9.2.3	MACD 策略 2	336
9.2.4	案例 9-2: MACD_v2	338
9.3	KDJ 策略	340
9.3.1	KDJ 策略 1	340
9.3.2	案例 9-3: KDJ01	343
9.3.3	KDJ 策略 2	346
9.3.4	案例 9-4: KDJ02	347
9.4	RSI 策略	350
9.4.1	RSI 取值的大小	351
9.4.2	RSI 策略	351
9.4.3	预留参数优化接口	356
9.4.4	案例 9-5: A 股版 RSI 策略	357
9.5	基石、策略与灵感	358
第 10 章	扩展与未来	360
10.1	回顾案例 2-1: SMA 均线策略	360

案例 10-1: SMA 均线策略扩展	363
10.2 大盘指数资源	365
10.2.1 大盘指数文件升级	366
10.2.2 大盘指数内存数据库	367
10.2.3 扩展 zwQuantX 类变量	368
10.2.4 大盘指数读取函数	368
10.2.5 案例 10-2: 读取指数	369
10.2.6 大盘数据切割	370
10.2.7 案例 10-3: inxCut 数据切割	372
10.3 系统整合	373
10.3.1 案例 10-4: 整合设置	375
10.3.2 案例 10-5: 修改指数代码	376
10.3.3 修改 sta_dataPreOxtim 函数	377
10.3.4 案例 10-6: 整合数据切割	380
10.3.5 修改绘图函数	381
10.4 扩展完成	384
案例 10-7: SMA 均线扩展策略	384
10.5 其他扩展课题	386
10.5.1 复权数据冲突	386
10.5.2 波动率指标 DVIX	386
10.5.3 修改回溯主函数 zwBackTest	387
10.5.4 案例 10-8: 波动率	390
10.5.5 空头交易	392
10.5.6 虚拟空头交易	392
10.5.7 修改检查函数	393
10.5.8 案例 10-9: 空头数据	396
10.6 终点与起点	397
附录 A zwPython 开发平台用户手册	398
附录 B Python 量化学习路线图	423