

运用Python语言编写的Plotly可视化绘图专业图书  
同时提供Plotly在R语言、MATLAB和JavaScript中的应用案例



从基础使用到具体应用场景，帮助读者快速掌握Plotly的实战应用

# Python

## 数据分析

▶▶▶▶▶▶▶▶

### 基于Plotly的动态可视化绘图

孙洋洋 王硕 邢梦来 袁泉 吴娜 著

# Python

## 数据分析

基于Plotly的动态可视化绘图

孙洋洋 王硕 邢梦来 袁泉 吴娜 著



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

## 内 容 简 介

随着信息技术的发展和硬件设备成本的降低，当今的互联网存在海量的数据，要想快速从这些数据中获取更多有效的信息，数据可视化是重要的一环。对于 Python 语言来说，比较传统的数据可视化模块是 Matplotlib，但它存在不够美观、静态性、不易分享等缺点，限制了 Python 在数据可视化方面的发展。为了解决这个问题，新型的动态可视化开源模块 Plotly 应运而生。由于 Plotly 具有动态、美观、易用、种类丰富等特性，所以一经问世就受到开发人员的喜爱。本书主要介绍 Plotly 在可视化各领域的应用，包括基础绘图、数据处理、网页开发、程序 GUI、机器学习和量化投资等，方便读者对 Plotly 快速上手。

本书绝大部分代码用 Python 语言编写，同时也给出了 Plotly 在 R 语言、MATLAB 和 JavaScript 中的应用案例。

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

Python 数据分析：基于 Plotly 的动态可视化绘图 / 孙洋洋等著. —北京：电子工业出版社，2018.6  
ISBN 978-7-121-34113-7

I. ①P… II. ①孙… 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 (2018) 第 083114 号

策划编辑：黄爱萍

责任编辑：牛 勇

印 刷：三河市华成印务有限公司

装 订：三河市华成印务有限公司

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：25 字数：460 千字

版 次：2018 年 6 月第 1 版

印 次：2018 年 6 月第 1 次印刷

定 价：79.00 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 [zltz@phei.com.cn](mailto:zltz@phei.com.cn)，盗版侵权举报请发邮件至 [dbqq@phei.com.cn](mailto:dbqq@phei.com.cn)。

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



# 前 言

Python 是一门非常优秀的编程语言，其语法简洁、易学易用，越来越受到编程人员的喜爱；Python 也是一门非常“人性化”的编程语言，其各种语法规则的设计符合人们的思维方式，开发人员可以用最简单的方式实现自己的编程目的，降低时间成本；同时，Python 又是一门非常强大的编程语言，其在编程的各个领域都有非常不错的表现，比如在网页开发、程序 GUI 设计、网络爬虫、科学计算、数据可视化、机器学习与深度学习等领域，Python 都有非常好的解决方案来解决现实中的业务问题。

互联网的快速发展为我们积累了庞大的数据，计算机硬件的创新为存储与分析这些数据创造了硬件条件，编程语言的发展为分析这些数据创造了软件条件。在数据分析这个领域，Python 有着自己独有的优势，简单易用的特性与强大的开源模块的支持使其成为数据分析领域方便、好用的利器。

Python 在数据分析领域的广泛应用离不开其强大的开源模块的支持，大名鼎鼎的 NumPy、SciPy、StatsModels、Pandas 等模块的建立与发展奠定了 Python 在数据分析领域中的重要地位。这些模块简单又好用，它们提供的解决方案能够满足绝大部分业务需求。在人工智能领域，Python 也有非常棒的解决方案，如 Sklearn、TensorFlow、MXNet、Theano、PyTorch、Caffe 等都是非常好的开源模块。尤其在人工智能中最前沿的深度学习领域，Python 几乎占据了霸主的地位。Python 借助在数据分析领域中开源模块的优势，在量化投资领域逐渐占据了领头羊的地位。国内外主流量化投资网站大多支持 Python 语言，其在量化投资领域有一种逐渐淘汰其他语言，一统“江湖”之势。

对数据的分析离不开数据的可视化，相对于 Python 在数据分析、人工智能、量化投资等领域中的发展，在数据可视化方面的发展有些滞后。最经典的 Python 可视

化绘图库莫过于 Matplotlib 了，Matplotlib 就是 MATLAB+Plot+Library，即模仿 MATLAB 的绘图库，其绘图风格与 MATLAB 类似。由于 MATLAB 的绘图风格有些偏古典，为了绘出更漂亮的图像，Python 开源社区开发出了 Seaborn 绘图模块，它本质上是对 Matplotlib 的封装，绘图效果更符合现代人的审美观。尽管如此，由于 Matplotlib 是基于 GUI 的绘图模块，所以存在特有的缺陷。

就笔者使用的经验而言，Matplotlib 主要存在两大缺陷：首先，Matplotlib 是一个静态的绘图模块，即我们绘出的图像是静态的，就像用看图软件打开图片一样，没有网页绘图的交互式效果；其次，Matplotlib 绘图结果的分享很不方便，在绘图结果分享给别人时只能以图片的方式分享，别人看到的绘图结果完全是静态的，分享体验很不好。Matplotlib 一直以来都是 Python 可视化的主力军，但是确实存在无法克服的缺陷，并且其他的 Python 绘图模块如 Ggplot、Bokeh、Pygal 等都比较小众，绘图功能比较单一，完成不了对 Matplotlib 的替代。

为了解决 Python 在可视化中存在的问题，Plotly 应运而生，它是一个基于 JavaScript 的动态绘图模块。Plotly 的绘图效果与我们在网页上看到的动态交互式绘图效果是一样的，其默认的绘图结果是一个 HTML 网页文件，通过浏览器就可以查看。我们可以把这个 HTML 文件分享给其他人，对方看到的效果与我们在本机上看到的效果完全一样。

Plotly 有着自己强大又丰富的绘图库，支持各种类型的绘图方案。Plotly 是基于 JavaScript 的绘图库，所以其绘图结果可以与 Web 应用无缝集成。总之，Plotly 在绘图模块上是 Matplotlib 强有力的竞争对手，Plotly 绘图的种类丰富、效果美观、易于保存与分享，因而越来越受数据分析人士的喜爱，至少笔者对 Plotly 的喜爱胜于对 Matplotlib 的喜爱。

Plotly 最初是一款商业化的绘图软件，在 2015 年 11 月 12 日，Plotly 开发团队决定把该模块的核心框架 plotly.js 开源，由此 Plotly 得到了快速发展。虽然在 2016 年 6 月，Plotly 开发团队才正式发布其 Python-api 文档，在 2017 年 1 月，Plotly 1.0 才正式发布，但是这些都阻止不了程序员对 Plotly 的喜爱。自 plotly.js 开源之后，我们可以使用本地的离线模式进行绘图，不依赖于官方的服务器，使得绘图速度更快，效果与在线绘图一样，这也是目前使用 Plotly 绘图的主流模式。

市面上有很多关于 Matplotlib 的可视化绘图教程，但是还没有 Plotly 的相关图书，作为一款非常优秀的可视化绘图模块，市面上急需一本科普性的图书。在本书创作之前，市面上就已经出现了电子版的对 Plotly 的简单翻译版本《Plotly 中文说明

1 期》，这是极宽量化开源团队在 2017 年 1 月的作品。极宽量化开源团队是一群研究“Python 量化投资”的爱好者自愿组成的一个团队，该团队成立的初衷是为国内量化投资领域做出自己的一份贡献，目前已经成功初步翻译 PyAlgoTrade、Seaborn、StatsModels、Plotly 等开源模块，并公开上传到网络上，《Plotly 中文说明 1 期》正是该团队的一个作品。

我作为极宽量化开源团队《Plotly 中文说明 1 期》项目的第一负责人，最初的想法只是单纯地把 Plotly 的基础内容简单翻译一下，以最简单、最快速的方式呈现给大家，方便大家使用。但是后来电子工业出版社的黄爱萍编辑找我沟通，请我编写一本 Plotly 数据可视化的相关图书，她认为 Plotly 发展很快，市场上需要一本 Plotly 的相关教材。经过一段时间的权衡，考虑到个人对 Plotly 的掌握程度、开源团队对 Plotly 的热情、个人在写《PyQt 5 快速开发与实战》时与黄编辑建立的良好关系，以及《Plotly 中文说明 1 期》存在的太多缺陷等，也为了能让更多的人接触 Plotly 这个优秀的绘图模块，于是决定再次抽出大量的时间来完成本书的创作，这就是本书的写作背景。

Plotly 是一个非常优秀的顶级绘图模块，如此优秀的开源模块在国内的知名度却不是很高，这对国内开发人员来说是一个很大的遗憾。顶级模块在特定的领域达到家喻户晓的程度是一个必然的趋势，Plotly 正是这种模块，它在可视化绘图领域的表现终会大放异彩。虽然目前 Plotly 在国内知名度不是很高，但其在可视化绘图领域做到家喻户晓是一个必然的趋势，只是需要有人加速这种趋势的演化过程，这就是本书存在的意义。

## 本书结构

本书的框架结构如下。

第 1 章是本书的快速入门部分，介绍 Plotly 的安装环境，对在线绘图与离线绘图做了简要的介绍。

第 2 章是基础绘图部分，对 Plotly 的一些常见的基础图形如线形图、柱状图、饼图、气泡图和直方图等做了一些介绍。

第 3 章是高级图形部分，对 Plotly 的时间序列绘图、表格绘图、多个坐标轴绘图、多子图绘图、SVG 绘图等做了一些介绍，是 Plotly 绘图的高级应用。

第 4 章是 Pandas 部分，介绍 Pandas 这个顶级数据分析模块使用 Plotly 进行绘图的方法。

第 5 章是金融绘图部分，主要为金融领域的特殊图形尤其是 K 线图的绘制提供了解决方案。

第 6 章是 Matplotlib 部分，主要介绍如何把 Matplotlib 绘图迁移到 Plotly 中。

第 7 章是网页开发部分，主要介绍 Plotly 在 Python 网页开发框架 Django 和 Flask 中的应用。

第 8 章是 GUI 开发部分，主要介绍 Plotly 在 GUI 开发框架 PyQt 5 中的应用。

第 9 章是机器学习部分，主要介绍 Plotly 在机器学习框架 Sklearn 与 PyTorch 中的应用。

第 10 章是量化投资部分，主要介绍 Plotly 在量化投资领域中的可视化应用。

第 11 章是其他语言应用部分，主要对其他语言如 R、MATLAB、JavaScript 的 Plotly 绘图做了简要的介绍。

如果你仅仅对 Plotly 的基础绘图有兴趣，那么前两章的内容就能满足你的需求；如果你对 Plotly 更高级的绘图有兴趣，那么可以参考第 3 章的内容；对于本书其他章节的内容，可以根据自己的实际情况有选择地阅读，毕竟 Plotly 绘图所涉及的范围特别广泛，并不是每个人对这些领域都同时感兴趣。

## 本书代码与交流

本书的所有代码都将保存在 GitHub 上，后续代码更新也会以 GitHub 地址为准，网址是 <https://github.com/sunshe35/PythonPlotlyCodes>，读者可自行下载。另外，为了方便读者交流、学习 Plotly，笔者建立了 QQ 群（群号：72203080）。

## 致谢

Plotly 虽然只是一个绘图模块，但是其应用场景非常广泛，除有 Matplotlib 的基本绘图功能外，其在 Web 开发、GUI 开发、机器学习、量化投资等领域也有很好的应用场景。由于其应用场景特别广泛，结合笔者自身知识的局限性，所以写好一本书需要多个人的共同努力，非常感谢下面这些作者对本书的创作所付出的努力：王硕负责本书的 Flask、PyTorch、JavaScript 部分；邢梦来负责本书的 Matplotlib 部分；



袁泉负责本书的基础绘图与 Sklearn 部分；吴娜负责本书的 Pandas 部分，其他都由本人孙洋洋负责。在 Plotly 的写作过程中，还有一些人为本书的顺利出版做出了贡献：首先，感谢《Plotly 中文说明 1 期》的开源组成员们，你们的贡献是本书基础部分的雏形，相关人员的网名和 QQ 号分别是 youngle sunny 1535327967、余勤 441499022、华子 32509167、啦啦啦 505512828、禛 948280670、L. 1248515039、Rikimaru 11766429、iris 704699640、信平 759949947、吴娜 2184934、周涛 510548099、zw 木子 719735825、非洲兔 85011284、十二月 378258849、大朱 775941748、我爱作文你信吗 571171954。其次，感谢山西证券金融科技部的陈亦苏、成都数联铭品科技有限公司的刘赣，以及极宽量化开源组的梁勇对本书网页开发部分提供的帮助，感谢西南财经大学统计学院王彦锋博士对本书 R 语言部分提供的帮助，感谢北京大学汇丰商学院硕士研究生扶禄城对本书 MATLAB 部分提供的帮助。再次，感谢黄爱萍与戴新两位编辑对本书的出版所付出的努力。最后，感谢我的父母与兄弟姐妹对我的关心与照顾，我现在取得的成果离不开你们对我的付出。

与读者相识于 Plotly 是一种缘分，能够看到本书说明读者对 Plotly 是感兴趣的，感谢读者愿意花费时间阅读本书，希望每一位读者都能够通过阅读本书有所收获，真心祝愿你们都能够学习顺利、事业有成。

孙洋洋

2018 年 4 月

轻松注册成为博文视点社区用户 ([www.broadview.com.cn](http://www.broadview.com.cn))，扫码直达本书页面。

- **下载资源：**本书如提供示例代码及资源文件，均可在 [下载资源](#) 处下载。
- **提交勘误：**您对书中内容的修改意见可在 [提交勘误](#) 处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- **交流互动：**在页面下方 [读者评论](#) 处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：<http://www.broadview.com.cn/34113>





# 目 录

第 1 章 快速开始	1
1.1 Plotly 简介	1
1.2 安装与安装环境	6
1.3 在线初始化	6
1.4 在线绘图隐私说明	8
1.5 开始在线绘图	8
1.6 使用离线绘图库	11
1.7 参数解读	13
1.8 查看帮助	14
第 2 章 Plotly 基础图形	16
2.1 基础案例解读	17
2.2 基本绘图流程	19
2.3 散点图	20
2.3.1 基本案例	20
2.3.2 样式设置	22
2.3.3 应用案例	24
2.3.4 参数解读	26
2.4 气泡图	28
2.4.1 基本案例	28

2.4.2	样式设置	29
2.4.3	缩放设置	30
2.4.4	参数解读	33
2.5	线形图	33
2.5.1	基本案例	33
2.5.2	数据缺口与连接	35
2.5.3	数据插值	38
2.5.4	填充线形图	41
2.5.5	应用案例	45
2.5.6	参数解读	49
2.6	柱状图	49
2.6.1	基本柱状图	49
2.6.2	柱状簇	50
2.6.3	层叠柱状图	52
2.6.4	瀑布式柱状图	54
2.6.5	图形样式设置	56
2.6.6	应用案例	58
2.6.7	参数解读	60
2.7	水平条形图	61
2.7.1	基本案例	61
2.7.2	应用案例	64
2.7.3	参数解读	70
2.8	甘特图	70
2.8.1	基本甘特图	70
2.8.2	甘特图（按数字索引）	71
2.8.3	甘特图（按类别索引）	72
2.8.4	应用案例	74
2.9	面积图	76
2.9.1	基本面积图	76
2.9.2	内部填充面积图	78
2.9.3	堆积面积图	80

# 原 书 缺 页



原  
书  
缺  
页

4.2.6 更多案例 .....	193
<b>第 5 章 金融绘图 .....</b>	<b>194</b>
5.1 快速绘制 K 线图 .....	194
5.1.1 检查 Plotly 版本 .....	194
5.1.2 快速绘制 OHLC（美国线）图 .....	194
5.1.3 快速绘制蜡烛图 .....	197
5.2 K 线图的优化 .....	199
5.2.1 过滤非交易时间 .....	199
5.2.2 设置形状、颜色和注释 .....	203
5.2.3 添加技术指标 .....	207
5.3 使用自定义数据的金融绘图 .....	211
5.4 高级金融绘图 .....	214
5.4.1 入门案例 .....	214
5.4.2 综合案例 .....	215
<b>第 6 章 Matplotlib .....</b>	<b>217</b>
6.1 Matplotlib 简介 .....	217
6.2 安装 Matplotlib .....	218
6.3 调整 Matplotlib 参数 .....	220
6.4 常用的 API 功能 .....	222
6.5 线性函数 .....	223
6.6 增加子图 .....	225
6.7 确定坐标范围 .....	228
6.8 概率图 .....	229
6.9 散点图 .....	232
6.10 柱状图 .....	235
6.11 更多扩展 .....	239
<b>第 7 章 Plotly 与网页开发 .....</b>	<b>240</b>
7.1 Plotly 在 Django 中的应用 .....	240
7.1.1 安装环境搭建 .....	240

7.1.2	安装环境测试	241
7.1.3	入门案例一	243
7.1.4	入门案例二	247
7.1.5	更多案例扩展	254
7.1.6	应用案例一	256
7.1.7	应用案例二	258
7.2	Plotly 在 Flask 中的应用	267
7.2.1	安装 Flask	269
7.2.2	最小的 Web 应用	269
7.2.3	模板渲染	270
7.2.4	入门案例一	272
7.2.5	入门案例二	275
7.2.6	应用案例	283
<b>第 8 章</b>	<b>Plotly 与 GUI 开发</b>	<b>293</b>
8.1	PyQt 的安装	295
8.2	案例解读	295
8.3	设置提升的窗口部件	298
8.4	Plotly_PyQt 5 的使用	300
8.5	更多扩展 (Plotly)	304
8.6	Plotly 与 PyQt 5.6 的结合	305
8.7	更多扩展 (Matplotlib)	309
8.8	应用案例: 展示产品组合信息	309
<b>第 9 章</b>	<b>Plotly 与机器学习</b>	<b>316</b>
9.1	Plotly 在 Sklearn 中的应用	316
9.1.1	分类问题	316
9.1.2	回归问题	319
9.1.3	聚类问题	321
9.2	PyTorch 可视化工具	326
9.2.1	Visdom 简介	326



9.2.2	安装 Visdom	327
9.2.3	Visdom 与 Plotly	328
9.2.4	Visdom 基本概念	328
9.2.5	Visdom 经典案例	329
9.2.6	Visdom 与 PyTorch	345
<b>第 10 章</b>	<b>Plotly 在量化投资中的应用</b>	<b>346</b>
<b>第 11 章</b>	<b>Plotly 在其他语言中的应用</b>	<b>355</b>
11.1	Plotly 在 R 语言中的应用	355
11.1.1	安装 R 语言	355
11.1.2	安装 Plotly 模块	356
11.1.3	Plotly 应用分析	356
11.1.4	更多扩展	362
11.2	Plotly 在 MATLAB 中的应用	363
11.2.1	下载与安装	363
11.2.2	基础入门	365
11.2.3	经典案例	367
11.2.4	更多扩展	376
11.3	Plotly 在 JavaScript 语言中的应用	377
11.3.1	基础入门	377
11.3.2	散点图	380
11.3.3	条形图	383
11.3.4	扇形图	384
11.3.5	更多扩展	386

# 1

## 第 1 章

# 快速开始

### 1.1 Plotly简介

Pandas（潘达思）数据分析软件和 Plotly 交互式绘图模块可以称为 Python 数据分析和 Python 量化分析两大重量级模块库。

事实上，这两者的影响力早已超越 Python 语言领域。

- Pandas 目前已经被业界公认是大数据工程一线最好的解决方案。
- Plotly 被称为史上最好的绘图工具之一，是 GitHub 开源网站最“火”的两大新项目之一（另外一个 Netflix 的 Spinnaker 云计算项目）。

为了更好地展示金融数据的复杂性，下面首先介绍 Plotly 绘图模块。如图 1-1 所示是 Plotly 网站截图，网址是 <https://plot.ly>。

Python 量化的关键是金融数据可视化，无论是传统的 K 线图，还是现在的互动策略分析，都需要大量的可视化图表。具体到编程代码，就是各种 Python 绘图模块库，传统 Python 绘图模块库有 Matplotlib、Ggplot、Seaborn、Bokeh 等。

Plotly 绘图底层使用的是 plotly.js，plotly.js 基于 D3.js、stack.gl（WebGL 组件库，由 Plotly 团队的 Mikola Lysenko 带领开发）和 SVG，用 JavaScript 在网页上实现类似 MATLAB 和 Python Matplotlib 的图形展示功能，支持数十种图形，包括 2D 和 3D 图形，交互流畅，可以满足一般科学计算的需要。该项目成功后，开始向其他语言

移植，目前已经有 Python、MATLAB、R 语言、Jupyter 等多种版本的 API 接口。

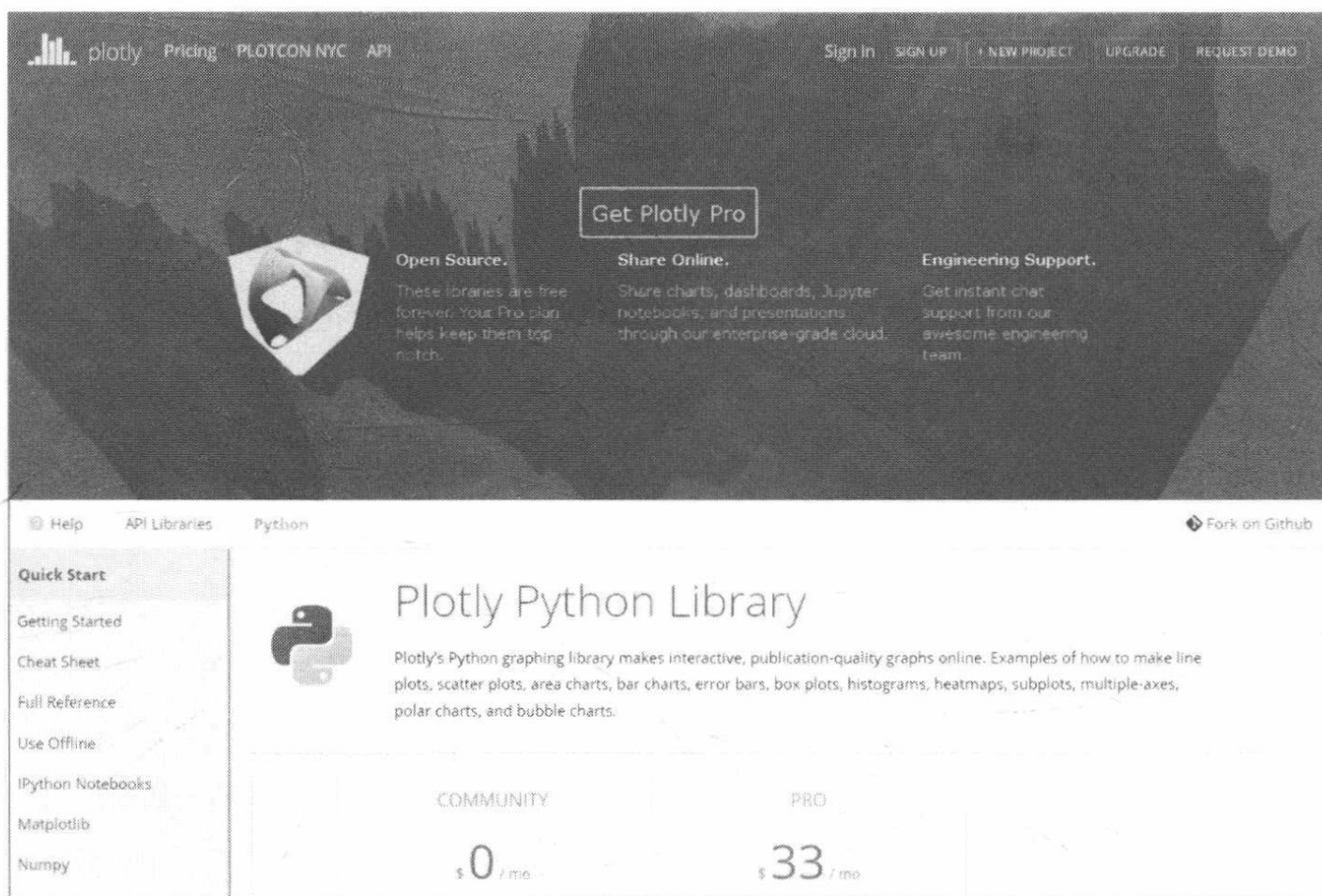


图 1-1 Plotly 网站截图

Plotly 项目的创始人是 Alex Johnson，是哈佛大学物理博士，曾经做过燃料电池。该项目的参与者中有一位华裔女生，叫 Baobao Zhang（张宝宝），是耶鲁大学政治学博士。Fernando Perez（IPython Notebook 创始人）是该项目的顾问。

Plotly 项目团队在 2013 年就已经拿到数百万美元的投资，并在加拿大蒙特利尔注册了公司，公司名称是科学数据云服务公司，号称是“首个科学数据平台”，提供线上和联机服务，方便用户使用分享、评论等功能，其投资商有 MHS 资本、西门子风险资本、Rho Ventures、Real Ventures 和硅谷银行等。

Plotly 公司有很多大牌客户，如西门子、Google、美国国家航天局（NASA）、美国空军、华盛顿邮报、美国国家卫生研究院等。

Plotly 公司的联合创始人 Jack Parmer 这样说过：对于开发者来说，绘制出漂亮的、交互式的图表并将之分享是一件很容易的事，但是对于一般的用户则很困难。Plotly 改变了这种局面，借助互联网的媒介、利用世界级的绘图分析工具，每个人都可以参与其中。

Hacker News（极客新闻）网站曾经专门讨论过各种图形模块库，如 Matplotlib、