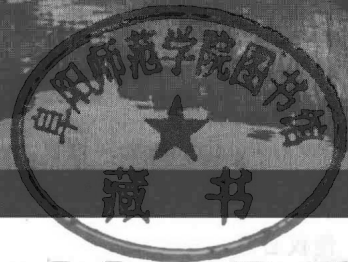


用Python3完成GIS开发及遥感分析的实用指南

Python 地理空间 分析指南 (第2版)

**Learning Geospatial Analysis
with Python**
Second Edition

[美] Joel Lawhead 著
邓世超 译



Python地理空间 分析指南 (第2版)

[美] Joel Lawhead 著
邓世超 译

人民邮电出版社
北京

图书在版编目 (C I P) 数据

Python地理空间分析指南：第2版 / (美) 莱哈德
(Joel Lawhead) 著；邓世超译. -- 北京：人民邮电出版社，2017.4

ISBN 978-7-115-44821-7

I. ①P… II. ①莱… ②邓… III. ①地理信息系统—
应用软件 IV. ①P208

中国版本图书馆CIP数据核字(2017)第039802号

版权声明

Copyright ©2015 Packt Publishing. First published in the English language under the title
Learning Geospatial Analysis with Python, Second Edition.

All rights reserved.

本书由英国 Packt Publishing 公司授权人民邮电出版社出版。未经出版者书面许可，对本书的任何部分不得以
任何方式或任何手段复制和传播。

版权所有，侵权必究。

-
- ◆ 著 [美] Joel Lawhead
 - 译 邓世超
 - 责任编辑 胡俊英
 - 责任印制 焦志炜
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
三河市海波印务有限公司印刷
 - ◆ 开本：800×1000 1/16
印张：21
字数：415 千字 2017 年 4 月第 1 版
印数：1-2 500 册 2017 年 4 月河北第 1 次印刷
著作权合同登记号 图字：01-2016-2083 号

定价：79.00 元

读者服务热线：(010) 81055410 印装质量热线：(010) 81055316

反盗版热线：(010) 81055315

广告经营许可证：京东工商广字第 8052 号

内容提要

Python 作为一种高级程序设计语言，凭借其简洁、易读及可扩展性日渐成为程序设计领域备受推崇的语言之一。

本书以地理空间分析为背景，介绍了 Python 在地理信息处理方面的应用技巧。全书共分为 10 章，分别介绍了 Python 与地理空间分析、地理空间数据、地理空间技术、Python 的地理空间分析工具、Python 与地理信息系统、Python 与遥感、Python 与高程数据、Python 与地理空间高级建模、实时数据、综合应用等内容。

本书内容结构清晰，示例完整，适合希望了解测绘数字化和分析的读者，也适合想使用 Python 进行空间地理分析、建模和 GIS 分析的开发人员及研究人员参考阅读。

作者简介

Joel Lawhead 目前是 NVision 公司的 CIO，该公司是业内屡获殊荣的地理信息集成和传感工程服务提供商。他还获得了管理专业人员资格认证 (Project Management Professional, PMP) 和地理信息专业资格认证 (GIS Professional, GISP)。

Joel 于 1997 年开始使用 Python，从 2000 年开始使用它进行地理空间软件的开发，也是 *Learning Geospatial Analysis with Python* (第 1 版) 和 *QGIS Python Programming Cookbook* 的作者，并且这两本书都是 Packt 出版的。他撰写的 Python Cookbook 系列图书由 O'Reilly Media 出版，其不同版本更是各具特色。除此之外，他还是广受赞誉的开源软件 PyShp (Python Shapefile Library) 的主要开发人员。他的技术博客是 <http://geospatialpython.com/>，Twitter 昵称是 @SpatialPython，通过上述方式，用户可以和他讨论地理空间分析与 Python 编程相关的话题。

2011 年，Joel 通过逆向工程处理了 Shapefile 文件的空间索引格式，帮助地理空间 Python 工程师 Marc Pfister 实现了上述索引算法，这一做法有利于世界各地的开发者开发出更易于集成、更健壮的地理空间应用。

Joel 作为架构师、项目经理和地理空间应用合作开发者，多次为美国政府机构提供服务，其中包括 NASA、FEMA、NOAA、美国海军，以及其他商业和非营利性机构。2002 年，由于其领导开发的实时应急响应协作工具 (the Real-Time Emergency Action Coordination Tool, REACT) 表现杰出，他获得了 Esri 公司颁发的国际 GIS 特别成就奖。该软件的主要用途是使用地理空间分析技术进行应急管理。

审阅者简介

Mark Cederholm, GISP, 拥有超过 20 年开发 GIS 应用的经验, 他使用的 Esri 技术包括 ARC/INFO、AML ArcObjects、ArcGIS Runtime 和 Web SDKs。他目前住在亚利桑那州的弗拉格斯塔夫。

他还是 *Developing Mobile Web ArcGIS Applications* 一书的技术编辑, 该书由 Packt 发行。

Truc Viet Le 目前在新加坡管理大学攻读信息系统博士学位。主要研究方向是人类移动行为模式建模和轨迹预测, 以及根据海量 GPS 和传感器网络数据进行城市交通和车辆行驶策略智能学习。每天的工作都会用到 R 和 Python, 而且他发现 R 善于操作数据和分析可视化数据, 而 Python 则善于完成机器学习任务。他还对社交网络和全球协作开发方面的数据分析很感兴趣。工作之余, 他非常喜欢旅游和探险, 是一个很有抱负的旅行作家和摄影师。其作品可以通过个人网站 <http://vietletruc.com/> 找到。

他在宾夕法尼亚州匹兹堡市的卡内基梅隆大学度过了美好的一年, 然后获得了博士学位。此前他在南洋理工大学获得了计算机工程学士学位和数学科学的硕士学位。

John Maurer 目前供职于夏威夷檀香山的太平洋岛屿海洋观测系统 (the Pacific Islands Ocean Observing System, PacIOOS), 岗位职责是程序开发和数据管理。他开发和配置了多数据源的访问、可视化和海洋数据映射的 Web 接口和数据服务, 其中包括卫星遥感影像、预测模型、GIS 图层, 以及现场观测 (浮标、传感器、鲨鱼跟踪等), 覆盖范围包括整个太平洋岛。他在 Colorado 大学 Boulder 分校先后获得了遥感专业的研究生学位和地理学博士学位, 在此期间, 主要的研究方向是分析地面穿透雷达 (ground-penetrating radar, GPR) 对格陵兰岛冰盖的测量分析。在 Boulder 分校期间, 为美国国家冰雪数据中心 (the National Snow and Ice Data Center, NSIDC) 工作了 8 年, 这激起了他对地球科学和地理空间最初的

兴趣，这对于本科专业是音乐、科学和技术，毕业于斯坦福大学的他来说虽然有些意外，但也是非常不错的经历。

Julia Wood 目前是一名地理空间分析师，是华盛顿特区的一名项目承包商。她以优异的成绩毕业于佛吉尼亚州弗雷德里克斯堡的玛丽华盛顿大学，2014年春季获得了历史和地理学学士学位。虽然职业生涯刚刚起步，但是她仍然不断地学习，提升自己的专业技能。她认为参与编审本书有助于提高她的专业素养，并通过半工半读的方式获得了硕士学位。工作之余，她喜欢读书、手工制作、烹饪、旅游和品尝各地的特色美食。

审阅 Packt 发行的书是她第一次参与专业审阅图书工作，希望今后可以进行更多的尝试。

首先要感谢我的父母 John 和 Diana，一直以来在学习和工作中对我的鼓励；感谢我的姐姐 Sarrina 和哥哥 Jonathan 在我遇到挫折时真诚的建议和鼓励；感谢我的男友 Max 和爱猫 Coco 在审阅本书时的温馨陪伴。最后感谢 Packt 出版社让我能够经历这样一次难得的历练！

前言

本书以地理空间分析为背景，介绍了一批在特定领域流行且实用的技术，如地理信息系统（Geographic Information Systems, GIS）、遥感、高程数据、高级建模以及实时数据等。本书的重点是充分利用 Python 语言和框架的强大功能，高效地进行地理空间分析。为此，我们将精力主要放在了充分利用 Python 实现的工具、API 和算法上。读者将学会分析多种地理空间格式数据，以及实时数据的监测和实际应用。

在示例中大量使用第三方地理空间库的同时，我们尽量使用原生的 Python 代码，以减少依赖。本书不同于业内其他作品之处主要体现在随书源码是完全使用 Python 3 编写的。本书是目前市面上少有的只使用 Python 3 的地理空间类图书，同时我们还在上一版的基础上添加了若干流行软件包的介绍。

本书概要

第 1 章“Python 与地理空间分析”介绍地理空间分析在实际生活中的应用，以及 GIS 和遥感的区别。另外，还使用插图、公式、伪代码和 Python 演示了地理空间分析的基本流程。

第 2 章“地理空间数据”在介绍几种主流数据格式的同时，还会讲解时下非常流行的数据格式。地理空间数据格式多样，地理空间分析的难点之一就是进行分析之前的数据准备。熟悉这些数据格式将有助于进一步掌握地理空间分析技术。

第 3 章“地理空间技术概览”简要介绍地理空间技术的生态系统，它是由数以千计的软件包构成的，它们常会让地理空间分析的初学者不知所措。快速掌握地理空间分析技术

的秘诀就在于其核心的软件包屈指可数。大部分软件都继承自这些核心软件包。理解了地理空间软件的层次结构和基本用法之后，学习和评估其他地理空间工具就易如反掌了。

第4章“Python的地理空间分析工具”简要介绍了本书已经用过和将要使用的软件包。Python在地理空间领域扮演的角色是多元的，包括GIS脚本语言、语言间的胶水语言和全能型的编程语言。本章的示例代码演示了数据编辑以及在第1章中提到的若干地理空间的基本概念。

第5章“Python与地理信息系统”用Python向你展示如何构建一个简单实用的GIS系统，而且它可以处理多种问题。

第6章“Python与遥感”介绍了如何使用遥感地理空间数据。遥感中包含一些非常复杂并且说明文档极少的地理空间操作。本章将为你学习后续章节打下坚实的基础，并且会揭开Python进行遥感分析的神秘面纱。

第7章“Python与高程数据”介绍高程数据的基本使用和特性。高程数据单独成章足以显示其重要性。高程数据几乎存在于任何地理空间格式中，但是用途和其他地理空间数据差别很大。

第8章“Python与地理空间高级建模”用Python向你展示地理空间技术的真正威力。对地理空间数据进行编辑和处理帮助我们更好地认识世界。但是地理空间分析的真正威力在于抽象建模。地理空间模型帮助我们预测未来，获取最优解决方案，将自然界无法直接观察到的概念可视化。

第9章“实时数据”考察了地理空间分析在实际生活中的应用。一位睿智的地理空间分析专家曾说过：“一张地图被创建之后就过时了”。即使当前你收集了有关地球的数据，将其加工之后制造一款地理空间产品，它所反映的信息也和实际情况不同了。现代地理空间技术打破了这一局面，通过互联网更新数据的周期能以分钟甚至秒来计算。这就从根本上改变了我们进行地理空间分析的方式。

第10章“综合应用”结合前面几章介绍的技巧，按部就班地构建一个通用的企业系统来管理客户请求和现场支持人员，并且它几乎适用于任何组织。

读者须知

你需要安装Python(3.4或更高版本)，硬件最低配置为300MHz处理器、128MB内存、1.5GB硬盘；操作系统为Windows、Linux和OS X之一即可。

目标读者

本书的目标读者是希望了解测绘数字化和分析的读者，使用 Python 和其他脚本语言进行自动化/非自动化计算的读者，以及使用 Python 进行地理空间分析、建模和 GIS 分析的开发人员和研究人员。

排版约定

阅读本书时你会发现不同类别的信息使用了不同的文本样式，下面举例说明其中一些样式，及其表示的含义。

代码文本、数据库表名、文件名、文件扩展名、路径名、URL 地址、用户输入以及 Twitter 用户名是这样表示的：“Python 库 pycsw 实现了 CSW 标准。”

代码块的表示方法如下：

```
<?xml version="1.0" encoding="utf-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
  <Placemark>
    <name>Mockingbird Cafe</name>
    <description>Coffee Shop</description>
    <Point>
      <coordinates>-89.329160,30.310964</coordinates>
    </Point>
  </Placemark>
</kml>
```

命令行输入和输出的表示方法如下：

```
t.pen (shown=False)
t.done ()
```



警告或需要特别注意的内容。



提示或者诀窍。

读者反馈

我们非常欢迎读者的反馈。告诉我们你觉得这本书怎么样，以及你喜欢哪部分或不喜欢哪部分。有了读者的反馈，我们才能继续写出真正能让大家充分受益的作品。

如果你想反馈信息，很简单，发送电子邮件至 feedback@packtpub.com，并在邮件的主题中注明书名。

如果你也是某个领域的专家，并且有兴趣编写或者合作出版一本书，请参见我们的作者指南：www.packtpub.com/authors。

售后服务

很荣幸你是本书的读者，我们将提供让你物超所值的增值服务。

随书源码

凡是通过 <http://www.packtpub.com> 网站账户购买的 Packt 图书，随书源码都可以从网站上下载。如果你是从其他渠道购买了本书，也可以访问 <http://www.packtpub.com/support>，注册之后，我们会将随书源码通过邮件发送给你。

书内彩插

另外，我们也提供本书中快照和图表的彩色 PDF 格式文件，彩色图片有助于你理解输出结果的变化。你可以通过如下地址获取这些文件：https://www.packtpub.com/sites/default/files/downloads/24290S_ColorImages.pdf。

勘误

虽然我们会全力确保书中内容的准确性，但错误仍在所难免。如果你在某本书中发现了错误（文字错误或代码错误），而且愿意向我们提交这些错误，我们感激不尽。这样不仅可以消除其他读者的疑虑，也有助于改进后续版本。若想提交你发现的错误，请访问

<http://www.packtpub.com/submit-errata>，在“Errata Submission Form”（提交勘误表单）中选择相应图书，输入勘误详情。勘误通过验证之后将上传到 Packt 网站，或添加到现有的勘误列表中。若想查看某本书的现有勘误信息，请访问 <http://www.packtpub.com/support>，选择相应的书名。

关于盗版

对所有媒体来说，互联网盗版都是一个棘手的问题。Packt 很重视版权保护。如果你在互联网上发现我们公司出版物的任何非法复制品，请及时告知我们网址或网站名称，以便我们采取补救措施。

请通过 copyright@packtpub.com 联系我们，并提供疑似盗版材料的链接信息。

感谢你帮助我们保护作者的权益，使我们能够提供更有价值的内容。

疑难解答

如果你对本书有疑问，可以将问题发送到 questions@packtpub.com，我们会竭力为你解答。

目录

第 1 章 Python 与地理空间分析1	
1.1 地理空间分析与我们的生活	1
1.2 地理空间分析的历史	3
1.3 地理信息系统	6
1.4 遥感	7
1.5 高程数据	12
1.6 计算机辅助制图	13
1.7 地理空间分析和计算机编程	14
1.8 地理空间分析的重要性	15
1.9 地理信息系统基本概念	16
1.9.1 专题地图	16
1.9.2 空间数据库	17
1.9.3 空间索引	17
1.9.4 元数据	17
1.9.5 地图投影	18
1.9.6 渲染	19
1.9.7 影像数据	20
1.9.8 遥感和颜色	20
1.10 GIS 中矢量数据的基本概念	21
1.10.1 数据结构	21
1.10.2 缓冲区	22
1.10.3 融合	22
1.10.4 简化	23
1.10.5 叠置	24
1.10.6 合并	24
1.10.7 点包容性	25
1.10.8 联合	25
1.10.9 连接	26
1.10.10 地理空间中的多边形 规则	26
1.11 栅格数据的基本概念	27
1.11.1 波段运算	27
1.11.2 变化监测	27
1.11.3 柱状图	28
1.11.4 特征提取	28
1.11.5 监督分类	29
1.11.6 非监督分类	29
1.12 牛刀小试	29
1.12.1 Python 简介	30
1.12.2 构建 SimpleGIS	31
1.13 小结	38
第 2 章 地理空间数据39	
2.1 数据格式概览	39
2.2 数据结构	41

2.3 空间索引	42	3.2.7 SpatiaLite	80
2.3.1 索引算法	43	3.2.8 路径分析	80
2.3.2 网格	44	3.3 桌面工具（包括可视化）	82
2.4 概要化	45	3.3.1 Quantum GIS	82
2.5 元数据	46	3.3.2 OpenEV	83
2.6 文件结构	46	3.3.3 GRASS GIS	84
2.7 矢量数据	48	3.3.4 uDig	85
2.7.1 Shapefile 文件	48	3.3.5 gvSIG	86
2.7.2 CAD 文件	51	3.3.6 OpenJUMP	87
2.7.3 标签和标记格式	51	3.3.7 谷歌地球	87
2.7.4 GeoJSON	53	3.3.8 NASA World Wind	89
2.8 影像数据	54	3.3.9 ArcGIS	91
2.8.1 TIFF 文件	55	3.4 元数据管理	92
2.8.2 JPEG、GIF、BMP 和 PNG 文件	56	3.4.1 GeoNetwork	92
2.8.3 压缩格式	56	3.4.2 CatMDEdit	92
2.8.4 ASCII 网格	56	3.5 小结	93
2.8.5 全球文件	57	第 4 章 Python 的地理空间分析工具	95
2.9 点云数据	59	4.1 安装 Python 的第三方模块	96
2.10 Web 服务	61	4.2 Python 的网络库	100
2.11 小结	62	4.2.1 Python 的 urllib 模块	101
第 3 章 地理空间技术概览	63	4.2.2 FTP	103
3.1 数据访问	65	4.2.3 ZIP 和 TAR 文件	104
3.1.1 GDAL	65	4.3 Python 的标记、标签解析器	107
3.1.2 OGR	67	4.3.1 minidom 模块	107
3.2 计算几何	68	4.3.2 元素树	109
3.2.1 PROJ.4 投影库	68	4.3.3 WKT 文本	115
3.2.2 CGAL	69	4.4 Python 的 JSON 库	117
3.2.3 JTS	71	4.4.1 json 模块	118
3.2.4 GEOS	72	4.4.2 geojson 模块	119
3.2.5 PostGIS	73	4.5 OGR	119
3.2.6 其他支持空间分析的 数据库	76	4.6 PyShp	120
		4.7 dbfpy3	121

4.8	Shapely	122	5.7.2	等值区域图	169
4.9	Fiona	123	5.8	使用电子表格	171
4.10	GDAL	125	5.9	使用 GPS 数据	173
4.11	NumPy	125	5.10	地理化编码	174
4.12	PIL	127	5.11	小结	176
4.13	PNGCanvas	129	第 6 章 Python 与遥感		177
4.14	GeoPandas	130	6.1	图像波段变换	178
4.15	PyMySQL	132	6.2	创建直方图	180
4.16	PyFPDF	133	6.3	图像剪裁	187
4.17	Spectral Python	134	6.4	图像分类	191
4.18	小结	134	6.5	图像特征提取	195
第 5 章 Python 与地理信息系统		135	6.6	变化监测	200
5.1	距离测量	135	6.7	小结	205
5.1.1	勾股定理	138	第 7 章 Python 与高程数据		206
5.1.2	半正矢公式	140	7.1	ASCII 网格文件	206
5.1.3	Vincenty 公式	141	7.1.1	读取 grids	207
5.2	方位计算	144	7.1.2	写入 grids	208
5.3	坐标转换	144	7.2	创建地形阴影	210
5.4	重投影	146	7.3	创建等高线	215
5.5	Shapefile 文件编辑	148	7.4	使用 LIDAR	219
5.5.1	Shapefile 文件访问	150	7.4.1	使用 LIDAR 创建网格	220
5.5.2	Shapefile 文件属性读取	151	7.4.2	使用 PIL 让 LIDAR 支持可视化	226
5.5.3	Shapefile 文件几何图形读取	154	7.4.3	创建不规则三角网络 (TIN)	230
5.5.4	Shapefile 文件修改	155	7.5	小结	234
5.5.5	Shapefile 文件合并	157	第 8 章 Python 与地理空间高级建模		235
5.5.6	Shapefile 文件分割	160	8.1	归一化植被指数 (NDVI) 应用	235
5.6	查询优化	161	8.1.1	建立框架	237
5.6.1	点包容性公式	161	8.1.2	载入数据	239
5.6.2	边框查询	163			
5.7	空间信息可视化	166			
5.7.1	点密度计算	166			

8.1.3 栅格化 Shapefile 文件	239	9.1.2 NextBus 路径列表	277
8.1.4 波段剪裁	241	9.1.3 NextBus 设备地址	277
8.1.5 使用 NDVI 公式	241	9.1.4 NextBus 地址映射	279
8.1.6 NDVI 分类	243	9.2 气象跟踪	283
8.2 创建洪水淹没模型	246	9.3 区域报告	291
8.2.1 洪水填充功能	248	9.4 小结	292
8.2.2 创建洪水	250	第 10 章 综合应用	294
8.3 创建彩色晕渲地形	253	10.1 GPS 报告示例	294
8.4 最佳路径分析	255	10.2 使用 GPX-Reporter.py	295
8.4.1 建立测试网格	256	10.3 程序概要	296
8.4.2 A*算法简介	257	10.4 初始化	297
8.4.3 生成测试路径	258	10.5 功能函数调用	298
8.4.4 查看测试结果	259	10.6 解析 GPX	302
8.4.5 实际应用	260	10.7 获取边框	303
8.5 街道路网规划	267	10.8 下载地图和高程影像	304
8.6 照片地理定位	271	10.9 创建地形	306
8.7 小结	273	10.10 创建地图	307
第 9 章 实时数据	274	10.11 高程测量	310
9.1 实时路况监测	275	10.12 距离测量	311
9.1.1 NextBus 代理列表	276	10.13 获取气象数据	313
		10.14 小结	317

第 1 章

Python 与地理空间分析

本章是对地理空间分析的概述。我们将以有史以来最严重的流行病之一为案例来阐述地理信息技术对当今社会的重要影响，在案例中地理空间分析为阻断致命病毒传播发挥了重要作用。接下来，我们将介绍地理空间分析的历史，它甚至比计算机和纸制地图的历史还要久远。然后我们将解释一下为什么你更想使用一门编程语言进行地理空间分析，而不是仅仅使用一款地理信息系统（Geographic Information System, GIS）应用程序。我们将认识到向大众普及地理空间分析技术的重要性，还将在本书中逐步介绍地理信息系统和遥感技术的基本概念和术语。在本章的最后，我们将使用 Python 从头构建一个尽可能简单的 GIS 应用。

本书假定读者了解 Python、信息技术的基本知识，并且至少对地理空间分析有一定的认识。本章的内容是有关地理空间分析的基础知识，需要重点关注的概念是遥感和 GIS，当然也包括本书其他章节的相关内容。

1.1 地理空间分析与我们的生活

2014 年 3 月 25 日，联合国世界卫生组织（World Health Organization, WHO）公布消息宣称，在西非发现一种致命的埃博拉病毒正处于爆发初期，它的传播非常迅速。在即将到来的夏季，三大洲 6 个国家将会受到影响，其中包括美国和欧洲。

政府和人道主义机构必须争分夺秒地和爆发的疫情作斗争。受感染患者未经治疗的话，短短 6 天之内就会病情加重而导致死亡。最关键的信息是现有病例和新感染病例的相对位置，他们遇到的挑战是以有限的基础设施报告农村地区新感染病例。确定病毒所在区域是应急响应机构制定疫情控制策略的基础。病例的位置界定了疫情爆发的范围，它可以让政府确定资源和医疗用品分配的优先级，同时也可以跟踪到受疫情影响的首个受害者。最终还可以让政府知道实行的措施是否减缓了疫情爆发。