

法律声明

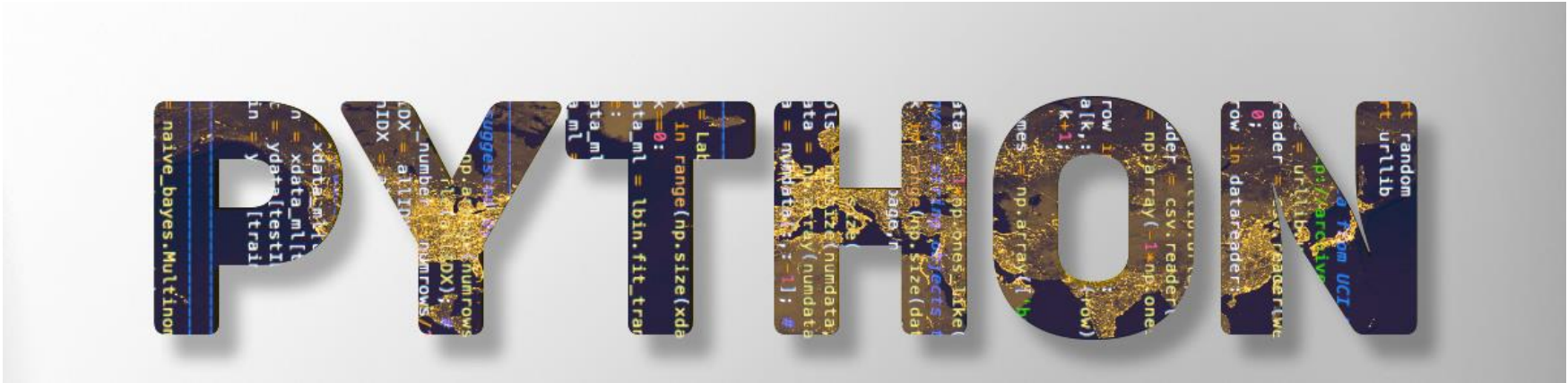
□ 本课件包括：演示文稿，示例，代码，题库，视频和声音等，小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。

□ 课程详情请咨询

■ 微信公众号：大数据分析挖掘

■ 新浪微博：ChinaHadoop





零基础Python入门

--梁斌

第九讲



空气质量指数计算 10.0

案例描述

- 为了能有效地提取并利用网络信息并工作提高效率，出现了网络爬虫
- 利用网络爬虫实时获取城市的空气质量
- 利用beautifulsoup4获取所有城市的空气质量
- 将获取的所有城市空气质量保存成CSV数据文件
- 利用Pandas进行数据处理分析
- 数据清洗；利用Pandas进行数据可视化

Pandas数据清洗

- 处理缺失数据
 - `dropna()` 丢弃缺失数据
 - `fillna()` 填充缺失数据
- 数据过滤
 - `df[filter_condition]` 依据`filter_condition`对数据进行过滤



Pandas数据可视化

- Pandas提供了内建的绘图功能（基于matplotlib）
- `plot(kind, x, y, title, figsize)`
 - `x, y` 横纵坐标对应的数据列
 - `title` 图像名称
 - `figsize` 图像尺寸
- 保存图片
 - `plt.savefig()`
- 更多例子请参考：<https://pandas.pydata.org/pandas-docs/stable/visualization.html>

第九讲 小结

空气质量指数计算器 (共10次课)



疑问

□ 问题答疑：<http://www.xxwenda.com/>

■ 可邀请老师或者其他人回复问题

小象问答邀请 @Robin_TY 回答问题



联系我们

小象学院：互联网新技术在线教育领航者

- 微信公众号：小象
- 新浪微博：ChinaHadoop

