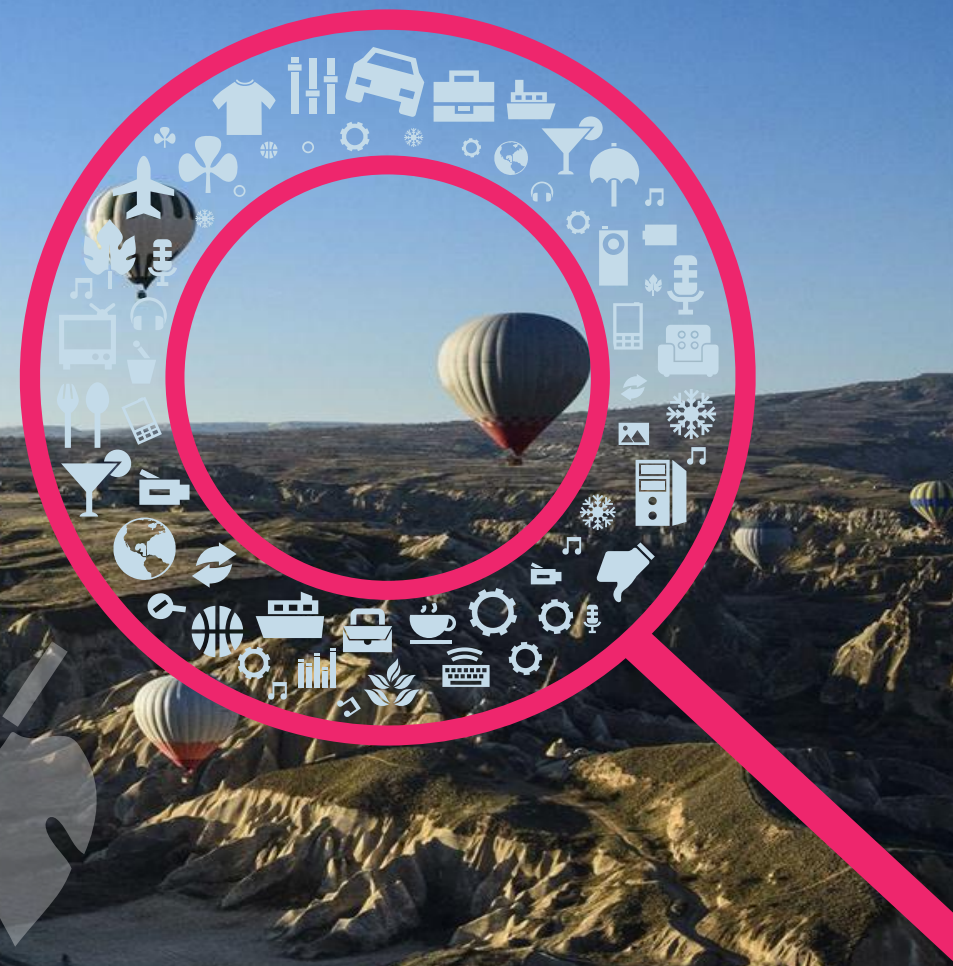


Python/期中作业

学生：凌俊杰

导师：邱儒

时间：2019年4月



目录\Contents

01

爬取豆瓣Top250

02

爬取华北地区天气

03

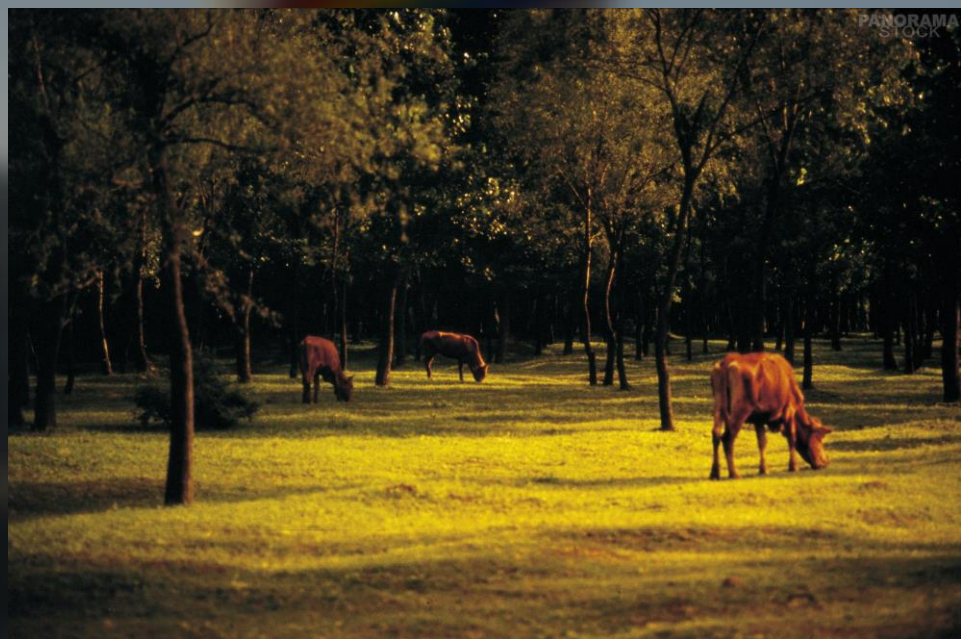
爬取京东Python图书

04

简单的环境配置



豆瓣电影Top250



用到的库



爬取方法



显示效果



豆瓣电影Top250

用到的库

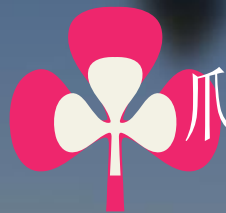
request re pymysql

爬取方法

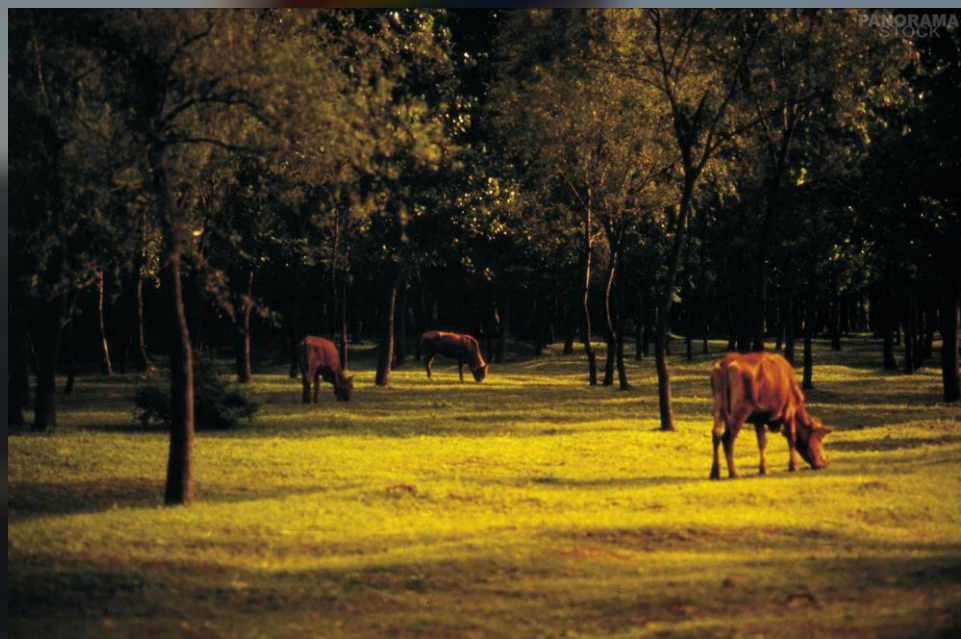
利用request库命令方法发送请求获取网页相应内容，利用re库进行字符串匹配，解析内容后保存数据到数据库表中

显示效果

1	肖申克的救赎	月黑高飞(港) / 刺激1995(台)	弗兰克·德拉邦特	1994	美国	犯罪 剧情	9.600000	1398913	希望让人自由。
2	霸王别姬	再见，我的妾 / Farewell My Concubine	陈凯歌	1993	中国大陆 香港	剧情 爱情 同性	9.600000	1035891	风华绝代。



爬取华北地区天气



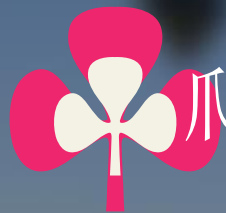
用到的库



爬取方法



显示效果



爬取华北地区天气

用到的库

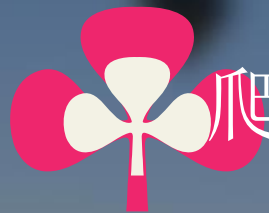
beautifulsoup requests pymysql

爬取方法

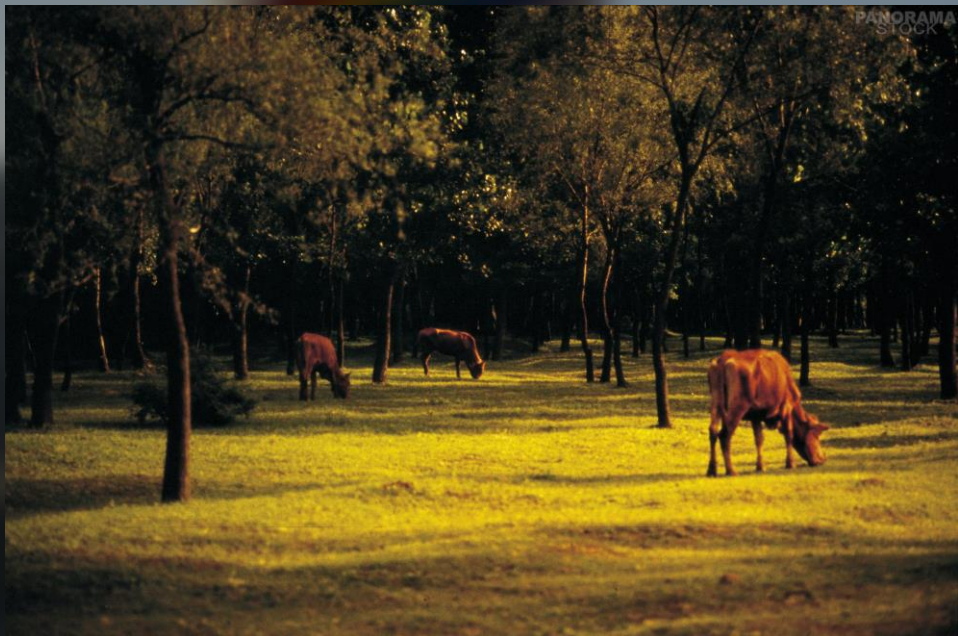
提交request请求，利用beautifulsoup库进行解析，enumerate返回元素位置及内容，获取天气text并存入数据库表

显示效果

北京	多云	东南风<3级	28	15
海淀	多云	东南风<3级	28	15
朝阳	多云	东南风<3级	28	15



爬取京东Python图书



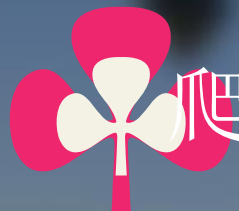
用到的库及方法



爬取方法



显示效果



爬取京东Python图书

用到的库及方法

webdriver WebDriverWait expected_conditions by keys random time MySQLdb

爬取方法

所用原理 selenium

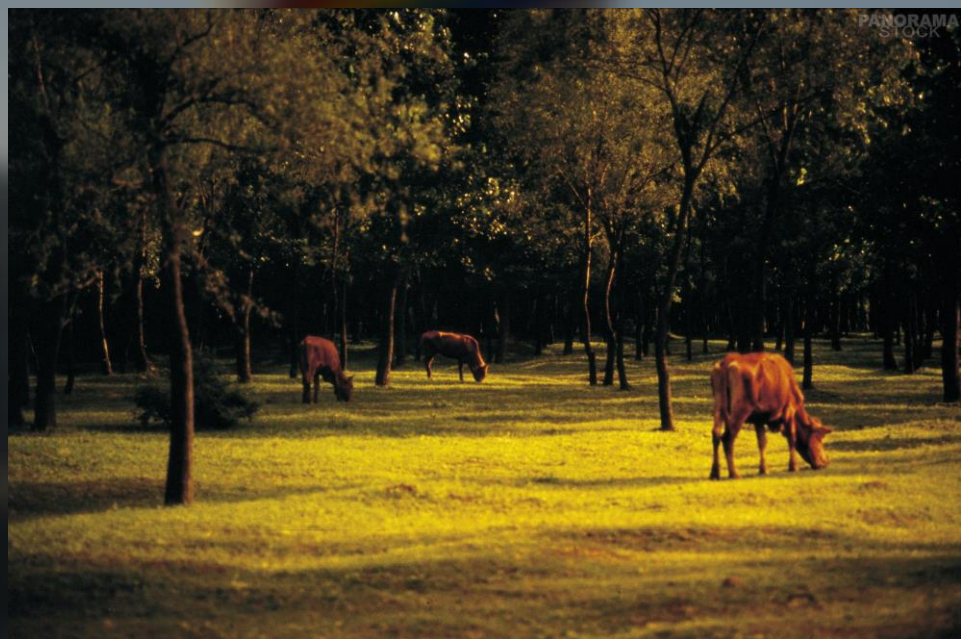
测试代码发送请求给浏览器的驱动 驱动解析代码并发送给浏览器 浏览器实行收获指令并完成效果
代码执行爬取图书命令 由谷歌浏览器驱动解析 打开京东页面 谷歌浏览器执行指令 爬取后存入数据库

显示效果

凤凰新华书店旗舰店	1100+	None	70.00	正版Python 3网络爬虫开发实战 Scrapy数据分析处理手册 数据抓取指南python程序设计
人民邮电出版社	1.7万+	自营	99.00	Python 3网络爬虫开发实战
人民邮电出版社	1.7万+	自营	41.70	用Python写网络爬虫 第2版



简单的环境配置



VSCode中文插件安装



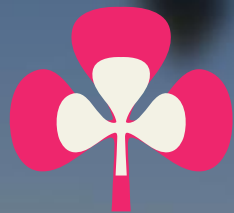
Python以及Django安装



数据库的配置问题



所要用到的命令符



简单的环境配置

VSCode中文插件安装

Ctrl+shift+P Configure Display Language 将locale属性内容改为zh-CN 保存 下载Chinese安装包

Python以及Django安装

Python使用安装包下载功能即可安装 Django使用pip命令pip install django==2.0安装

数据库的配置问题

mysql环境 在settings.py中databases中配置库名密码等数据

所要用到的命令符

python manage.py makemigrations	python manage.py migrate	python manage.py runserver
记录models.py改动并生成迁移文件	将改动作用到数据库	启动web服务器

A wide-angle photograph of a vast, rugged canyon landscape under a clear blue sky. Numerous hot air balloons of various colors and patterns are scattered across the sky, some appearing as small specks in the distance and others as larger, more detailed objects. The canyon floor is filled with dark, rocky terrain and some sparse vegetation. The overall atmosphere is serene and majestic.

感谢聆听

学号：16219111412