

招聘网站数据分析案例需求

Copyright © ruanko Technologies, Co., Ltd. . All Rights Reserved

1.案例介绍

运用python语言以及第三方技术完成招聘网站的数据信息爬取，再利用数据可视化技术讲将数据呈现，并进行数据分析，多维度展示招聘统计信息。

2.开发环境要求

案例研发中需要使用的开发环境配置如下表

名称	版本	说明
python 3	3.7及其以上版本	开发语言
PyCharm	无要求	开发IDE
requests	2.27	

3.功能需求

功能结构

模块	子模块	说明
职位数据爬取	爬取天气	爬取职位数据
	数据分析	分析数据并保存
用户交互	根据行业分析职位	根据行业的维度分析各个行业招聘岗位，薪资待遇，需求学历的报表
	根据薪资分析职位	根据薪资的维度分析热门岗位，薪资范围，学历关系的版本
	根据城市分析职位	根据城市统计该城市各项职位信息

4.项目效果展示

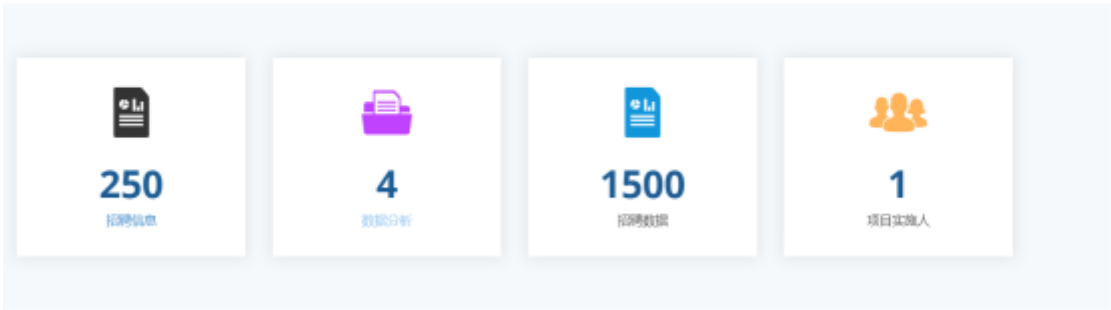
4.1 职位查询

招聘网站数据分析

[首页](#) [数据总览](#) [词云](#) [公司类型](#) [薪水](#) [工作地点](#)

51job招聘网站

应用python爬虫, flask框架, Wordcloud等技术实现数据分析



© Copyright 招聘网站数据分析 . All Rights Reserved

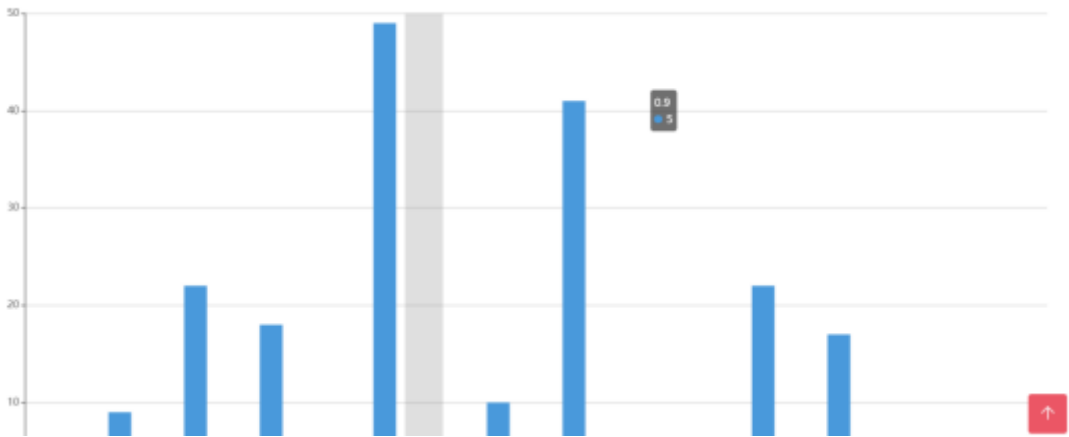
请输入想查询的信息

岗位名称:	<input type="text" value="请输入你想要搜索的内容, 如: 开发工程师"/>
公司地区:	<input type="text" value="请输入你想要搜索的内容, 如: 北京/重庆/上海"/>
学历要求:	<input type="text" value="所有"/>
经验要求:	<input type="text" value="所有"/>

一键查询

4.2 统计分析

工资分布图(单位: 万/月)



5.项目验收标准

任务	子任务	说明	功能分数
主程序设计	主程序代码的编写	程序代码必须无误	6
函数以及第三方模块的应用	模块使用目的明确且无语法错误	6	
目标数据分析	目标网站数据检测	能成功抓取到目标网站的数据	8
数据结构分析	从抓取的数据中获取有效的目标数据	8	
数据抓取	获取目标数据	BeautifulSoup解析目标网站，成功抓取到目标网站的数据	8
反爬虫技术突破	如果目标网站有反爬虫技术，解决问题并获取数据	8	
数据分析	目标数据清洗	从抓取的数据中获取有效的目标数据	10
数据分析	利用 Pandas 进行数据分析	10	
可视化操作	可视化库的使用	Python可视化库matplotlib的使用	8
招聘数据的可视化显示	使用图表完成招聘信息的可视化显示	8	
可视化多样性	完成其他可视化图形	多样化的显示招聘相关要求、职位等信息	20
总计			100

除了上述评分标准外：代码质量和注释率也是影响评分的标准之一

6.技术要点

- 1、使用爬虫功能爬取到正确且符合预期的信息，其中可能涉及到爬虫框架如scrapy，pyspider的使用、或者使用requests库进行页面信息请求等。
- 2、对获取到的数据进行数据分析，通过使用正则表达式，xpath，bs4等方法对数据进行分析，提取项目所需要的信息。
- 3、可视化界面，提取信息后需要进行数据展示，可以使用第三方库tkinter，pyqt，matplotlib等库实现数据展示，也可以使用flask框架将数据展示在浏览器上，利用前端技术可以实现更加美观的效果
- 4、如果爬取的网站有反爬机制，则需要使用反爬虫技术进行突破，比如代理池，在请求头中添加字段等操作。