软件系统开发设计说明文档-报告

软件1802 孟祥帅 120181080520

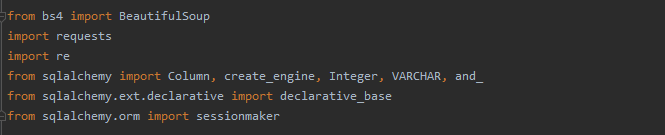
7 在51job网站上，爬取2020年发布的Python开发工程师的职位的薪酬，计算北京地区该职位的平均薪酬；

**一、课程设计(综合实验)的目的与要求**

要求在51job网站上，爬取2020年发布的Python开发工程师的职位的薪酬，计算北京地区该职位的平均薪酬。

**二、设计（实验）正文**

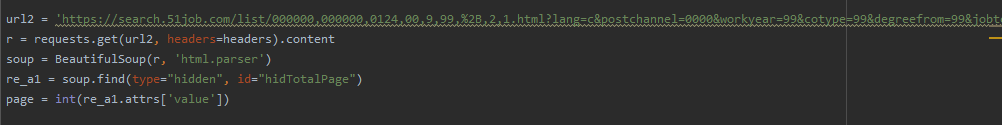
导入需要的模块



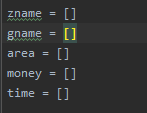
获取headers模拟浏览器访问

C:\Users\mac\AppData\Local\Temp\WeChat Files\ff88ec8cec46dc4f48eda0dc6c245cd.png

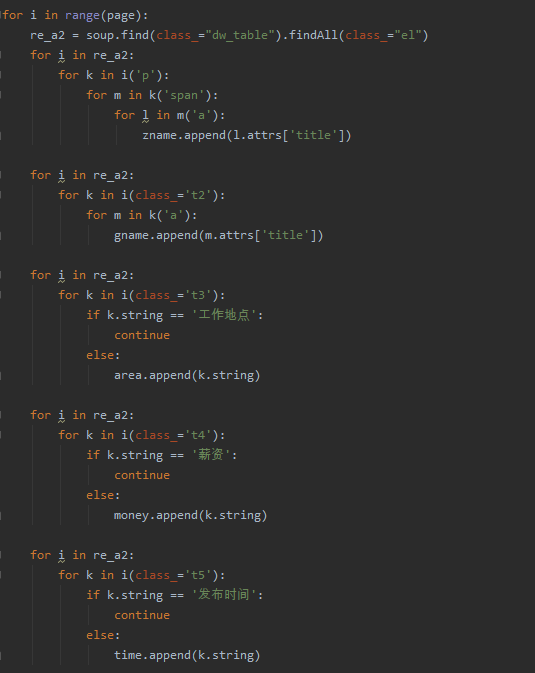
爬虫与网址链接

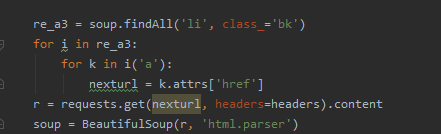


用于暂时存储爬取数据的列表

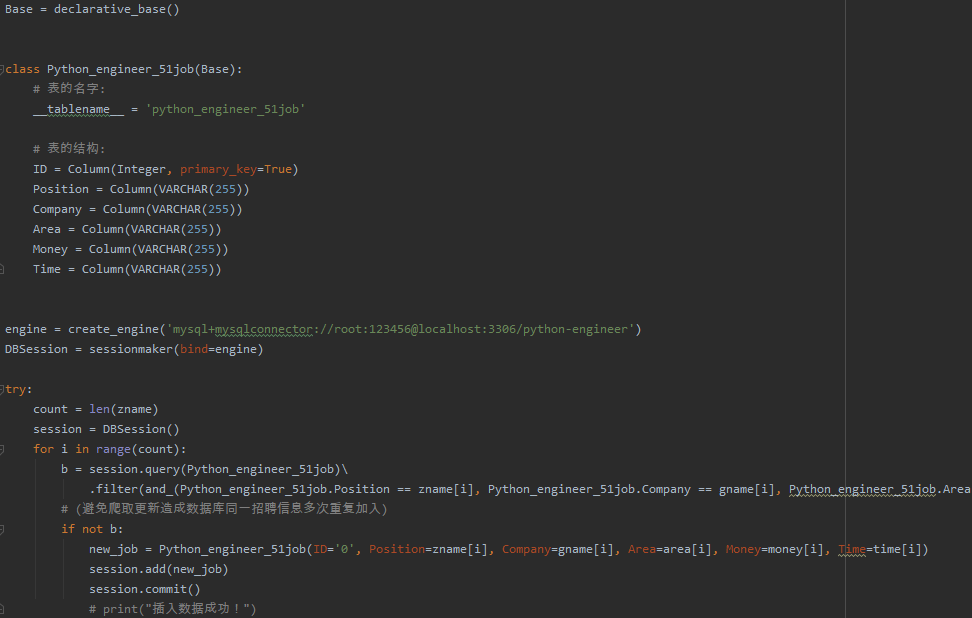


利用循环逐页爬取网站上所需要的数据，并保存与列表

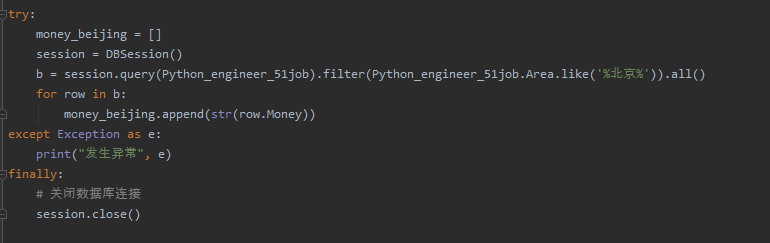




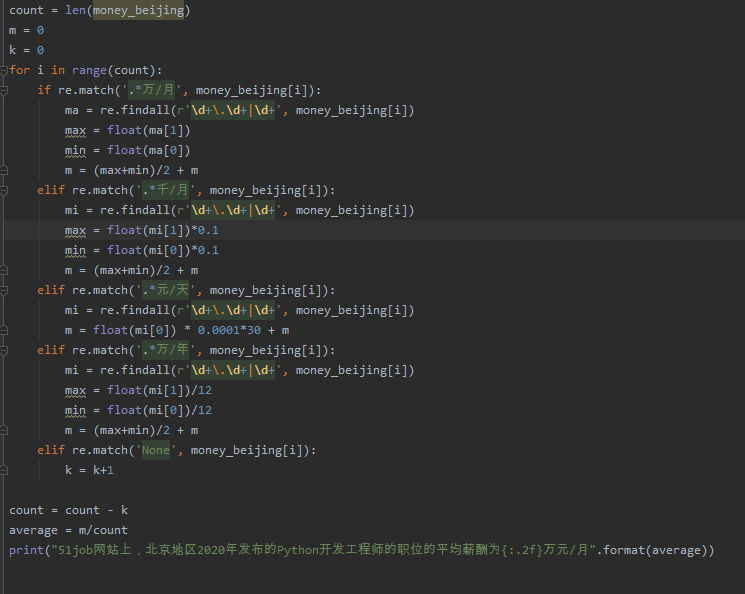
链接数据库，将保存在列表中是数据存入数据库



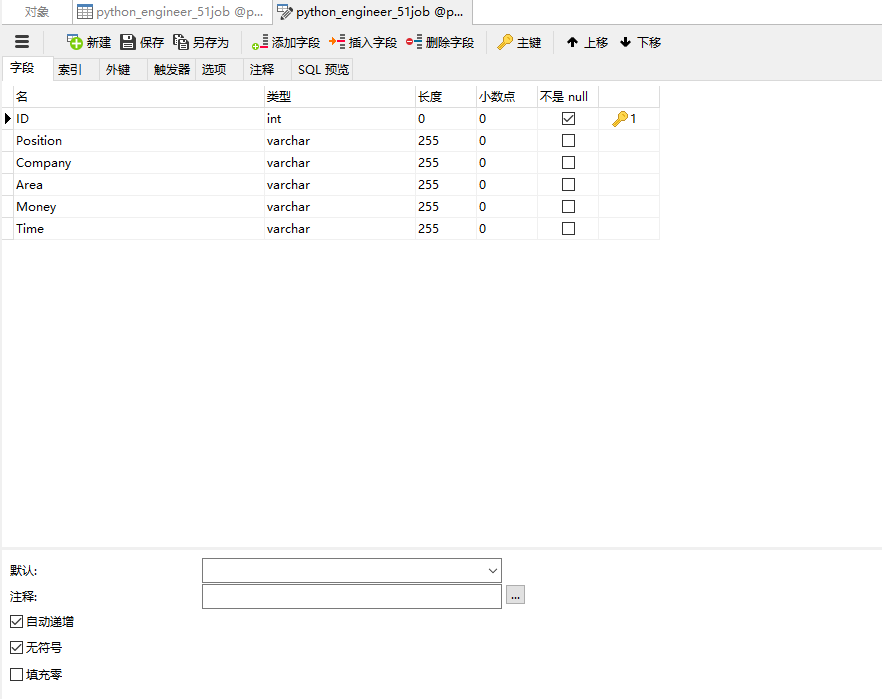
从数据库中筛选出北京地区的该职位的薪资



利用正则表达式，对数据进行处理（统一单位），并计算平均薪资



数据库的设计



ID:存入数据的序号

Position: 职位名

Company:公司名

Area:工作地点

Money:薪资

Time:发布时间