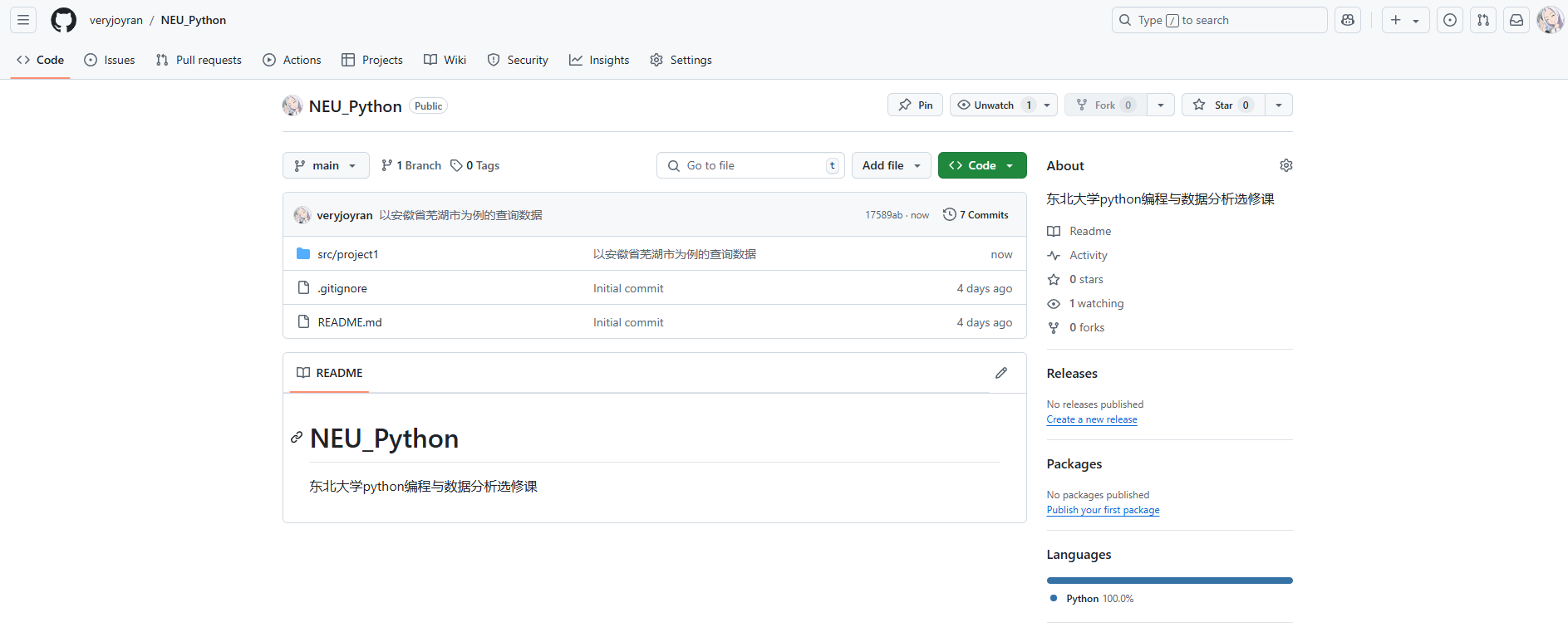
Project2：二手房查询与房价预测系统-GUI版本

人工智能2202班-刘庆然-20226494

所有代码以及README均已上传至github:

https://github.com/veryjoyran/NEU\_Python/tree/main/src/project2





目录

[一、 问题描述 3](#_Toc185634653)

[二、说明文档 4](#_Toc185634654)

[三、实验结果 5](#_Toc185634655)

[1.爬取数据 5](#_Toc185634656)

[2.爬取数据存储到数据库 6](#_Toc185634657)

[3.读取数据库数据 9](#_Toc185634658)

[复用project1中的代码进行排序 10](#_Toc185634659)

[价格升序 10](#_Toc185634660)

[价格降序 10](#_Toc185634661)

[四、源码解读 11](#_Toc185634662)

# 问题描述

1. Project2 需求：

1）基于tkinter开发GUI；

2）面向对象设计并复用Project1中功能函数；

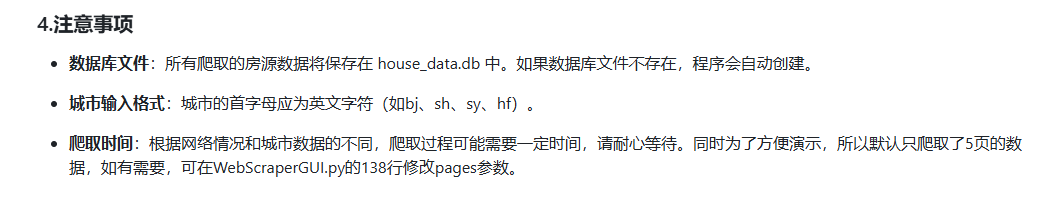
3）获取省会城市的二手房信息（位置/户型/面积/朝向/装修/楼层/楼龄/结构/价格等），注意访问的频率，可以不全部加载；

4）房屋信息存储到数据库；

# 二、说明文档





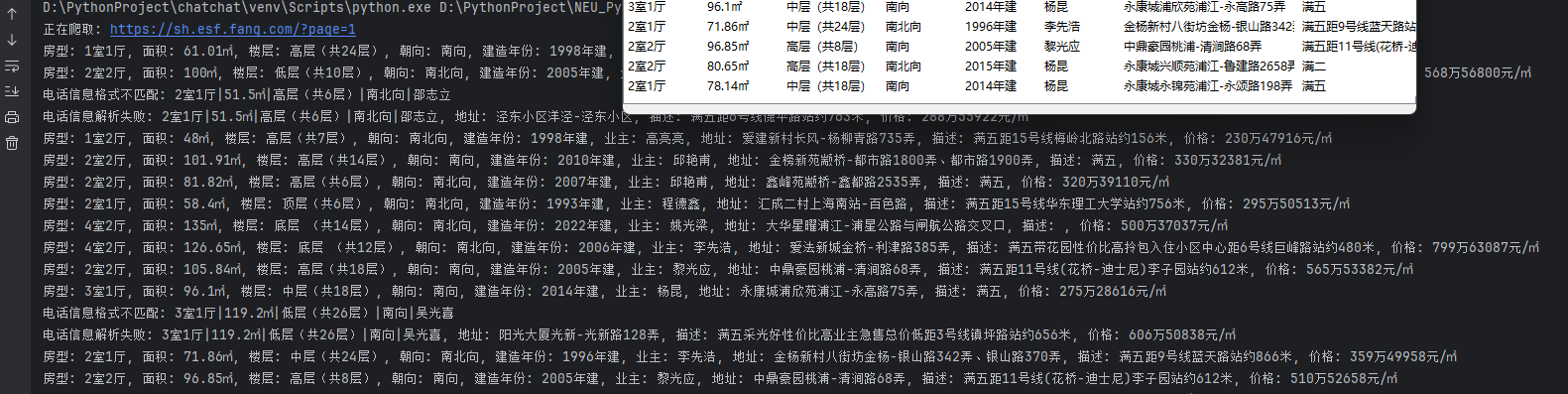


# 三、实验结果

## 1.爬取数据



以上海房价为例:



在爬取时会有相应信息在终端输出，当5页的内容全部爬取完之后，会在界面上加载，中间最麻烦的过程是寻找一个合适的网址，在此处使用的房天下，其二手房信息除北京外格式为：

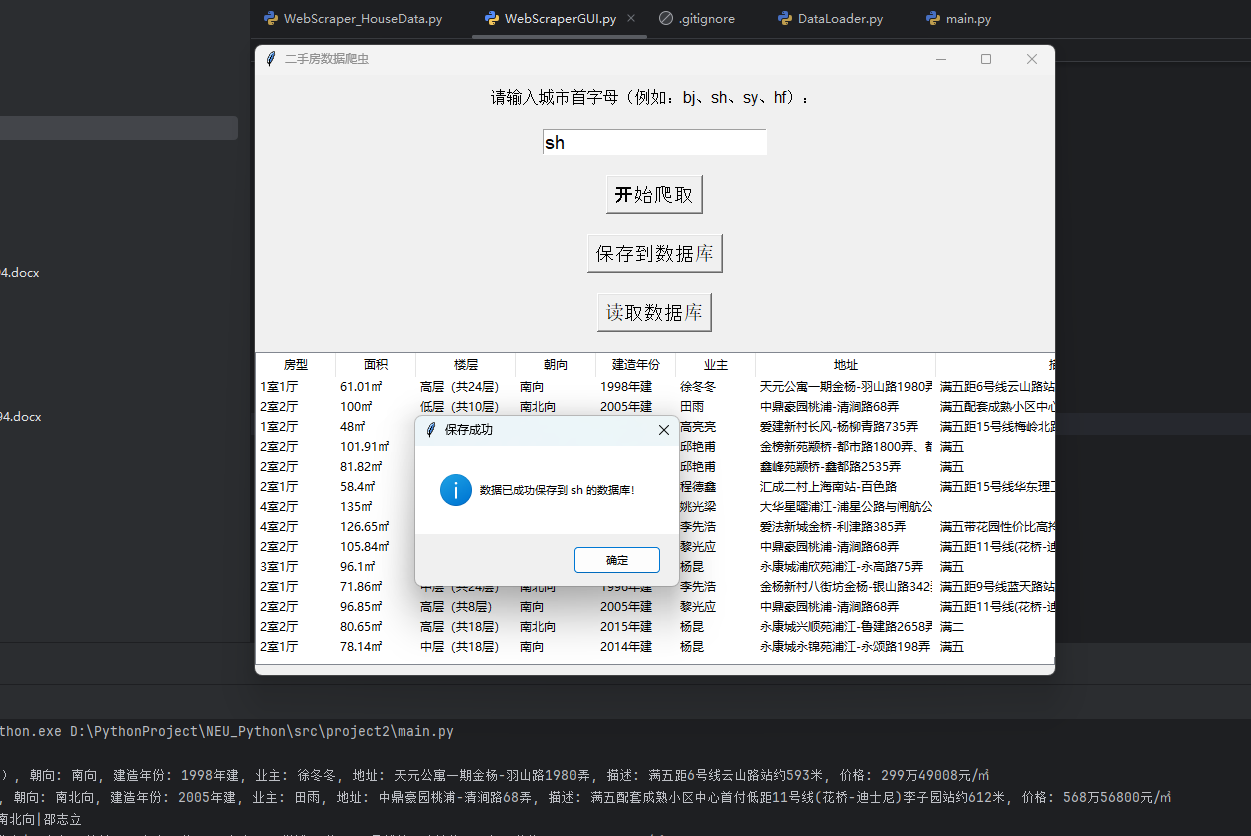
f"https://{city\_code}.esf.fang.com/"

如 <https://sh.esf.fang.com/>

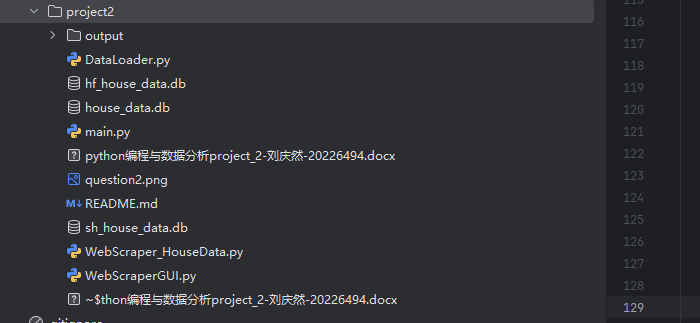
北京则为：https://esf.fang.com/



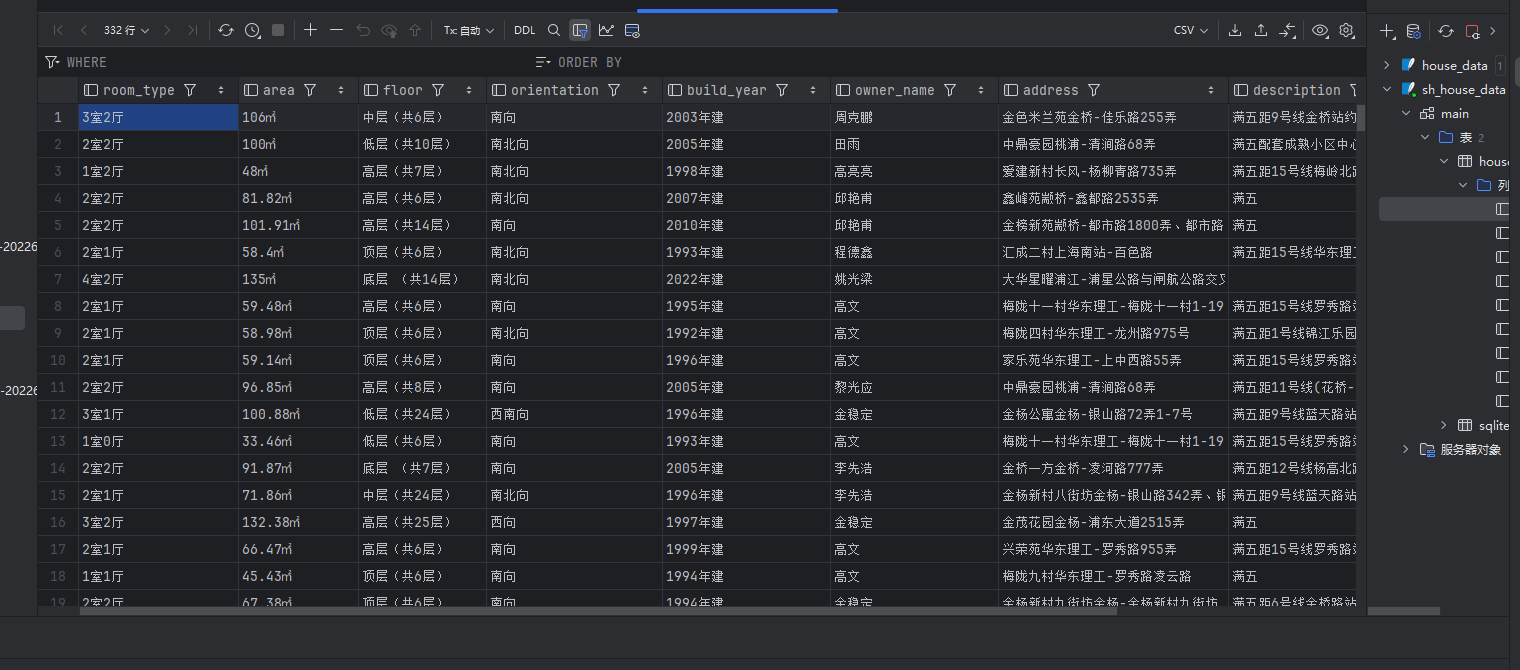
## 2.爬取数据存储到数据库

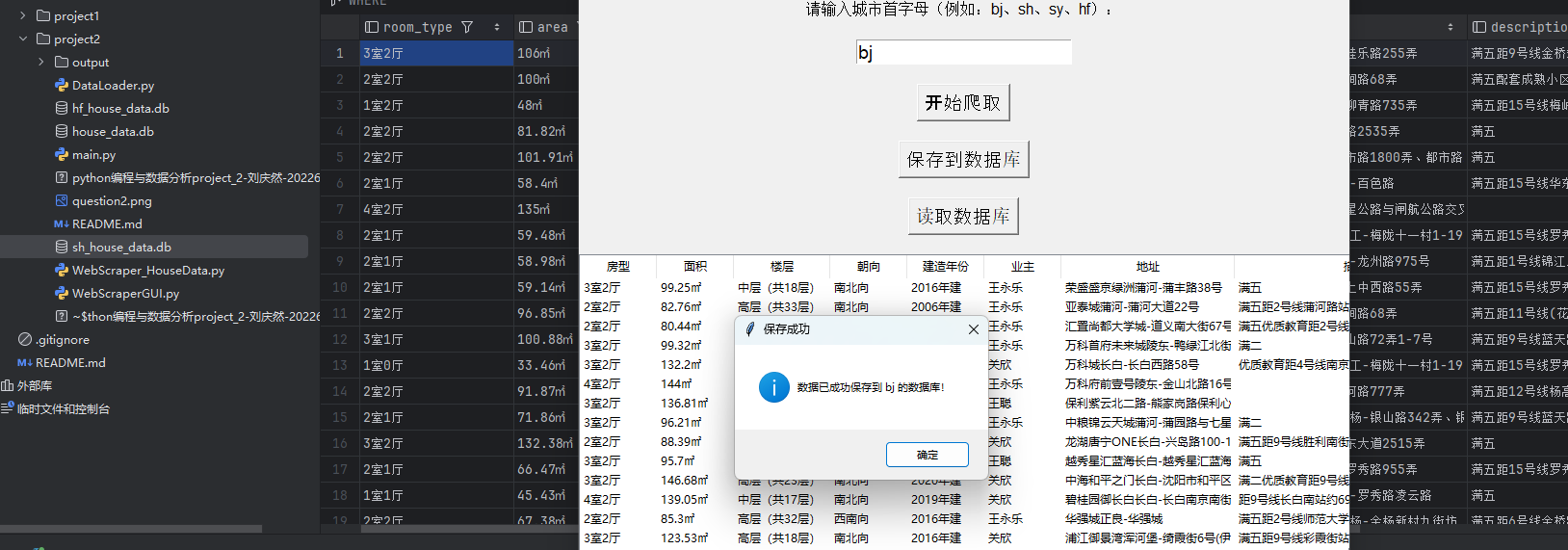


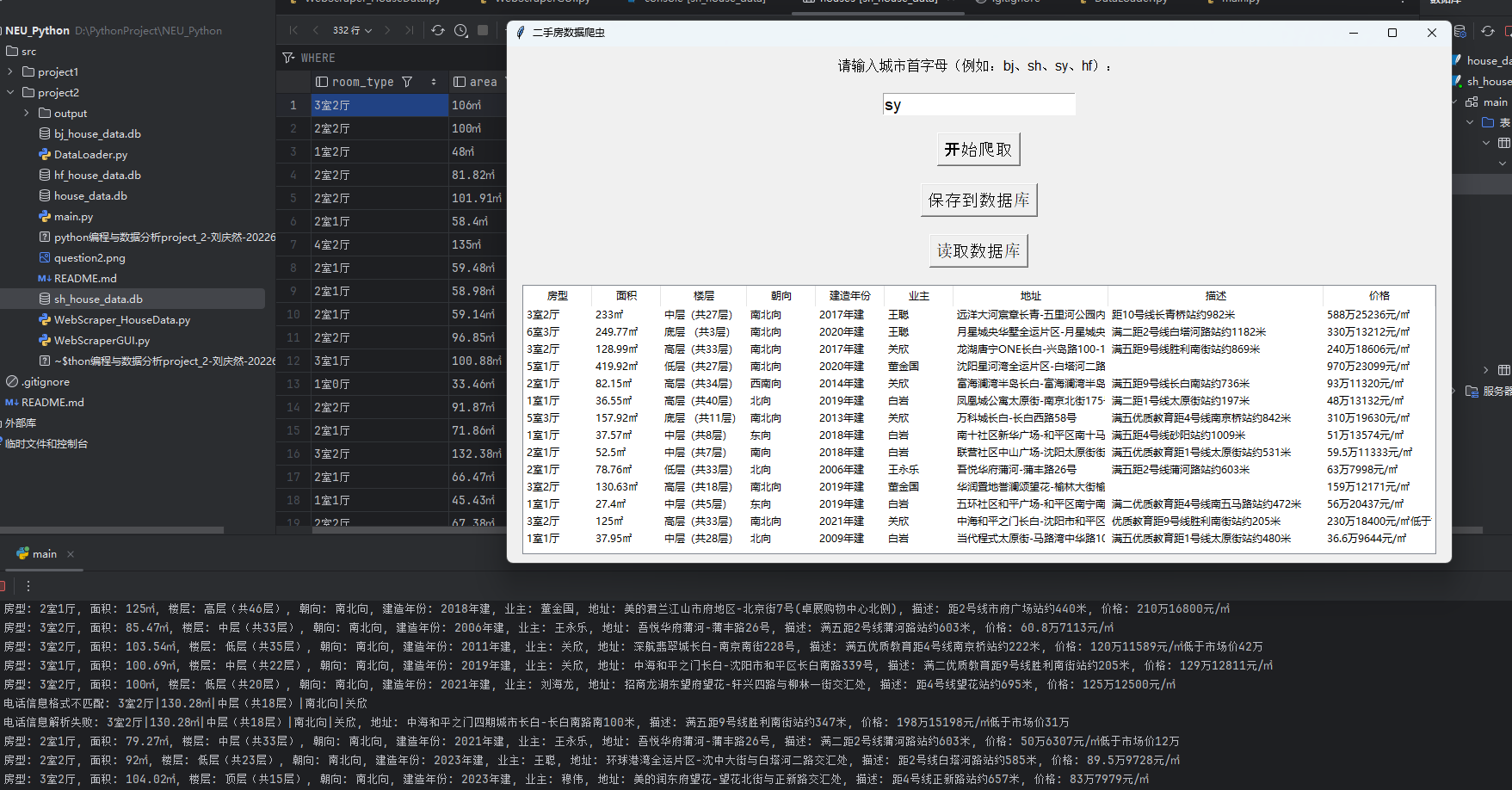
默认储存在当前目录下，

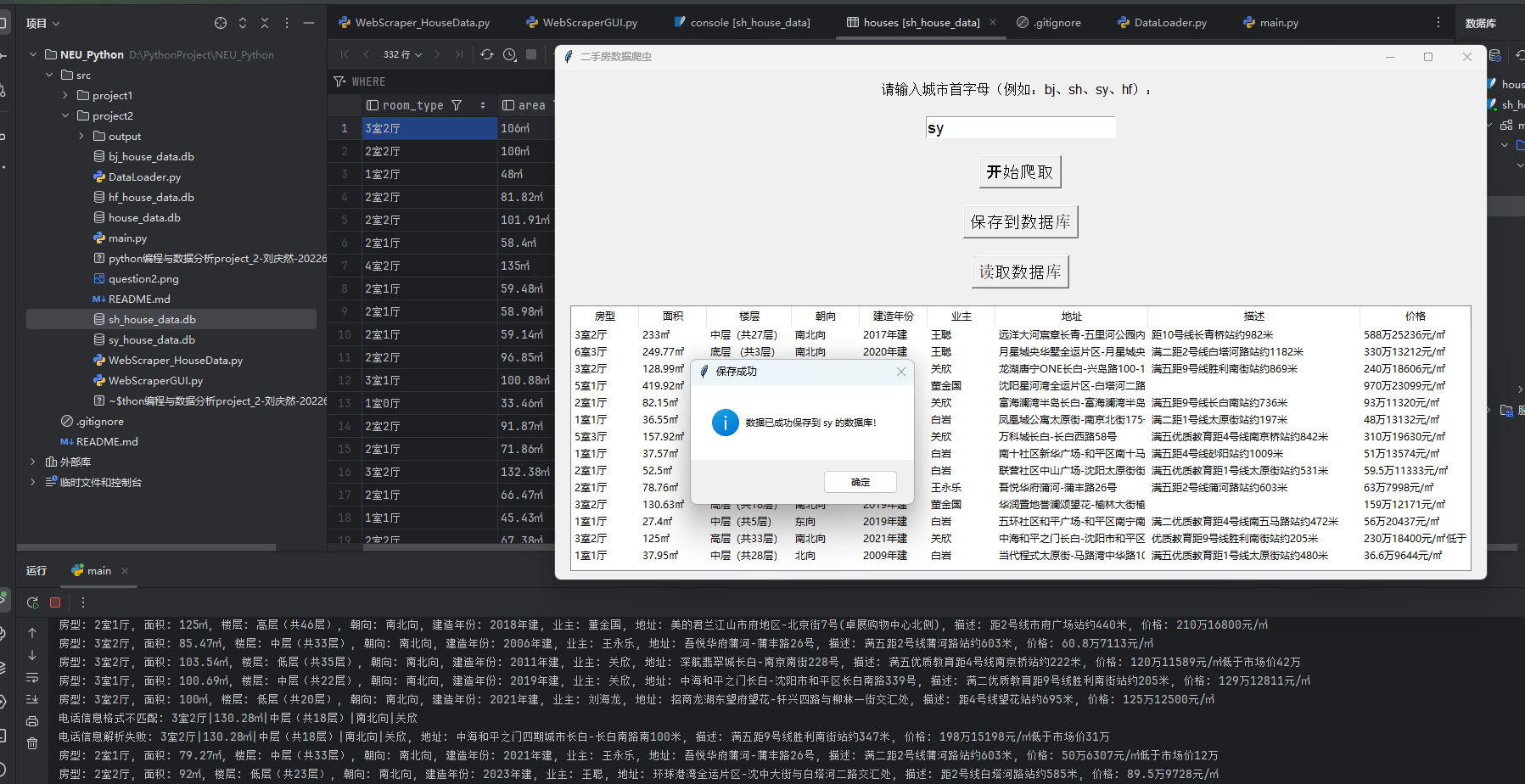


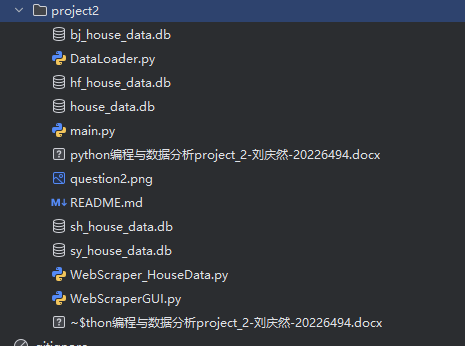
可直接双击打开查看：



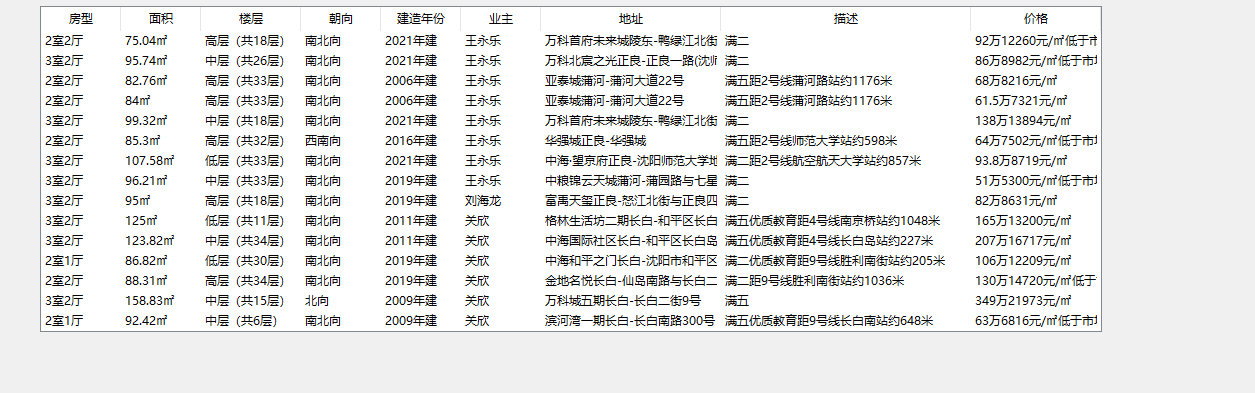






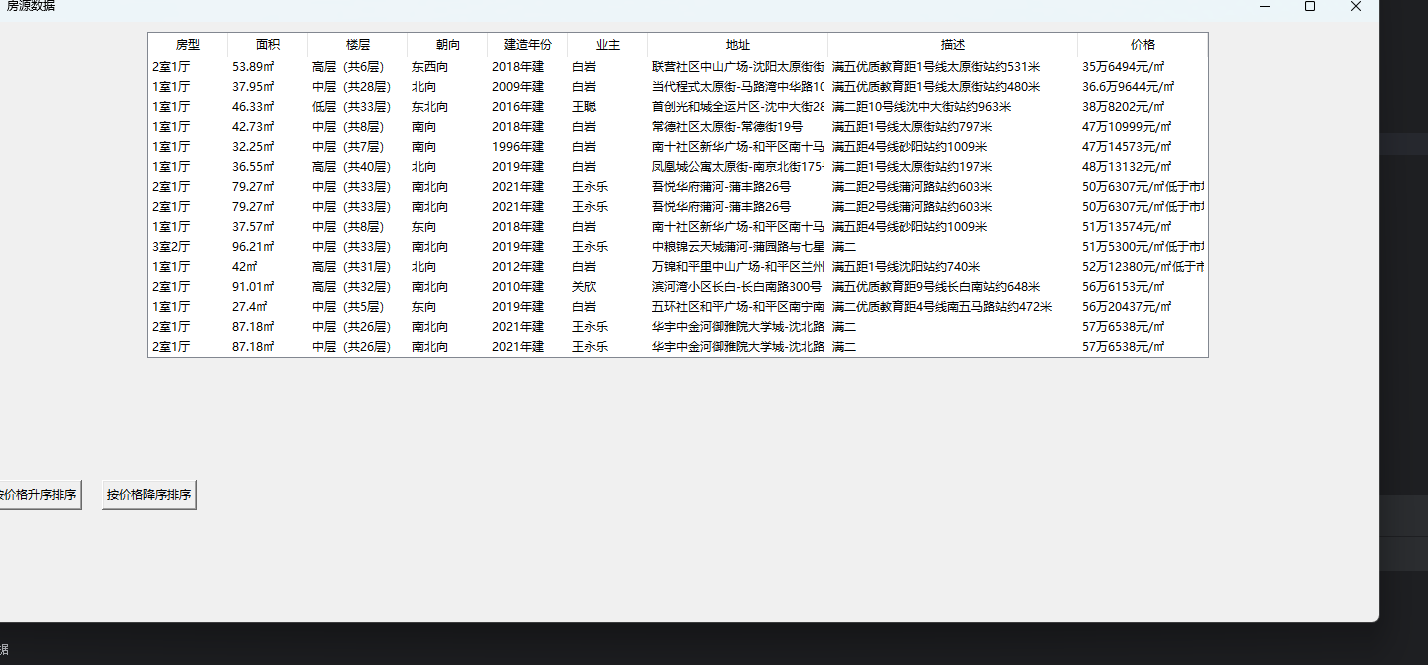


## 3.读取数据库数据

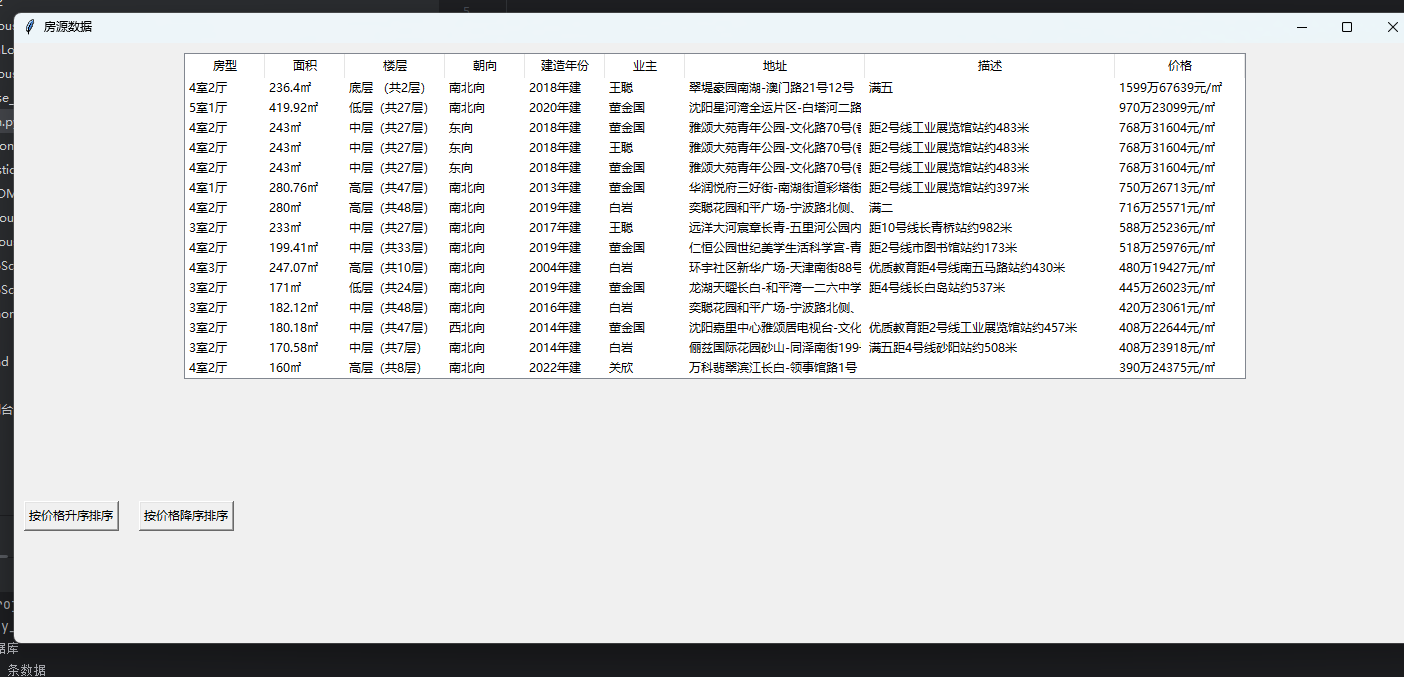


### 复用project1中的代码进行排序

#### 价格升序

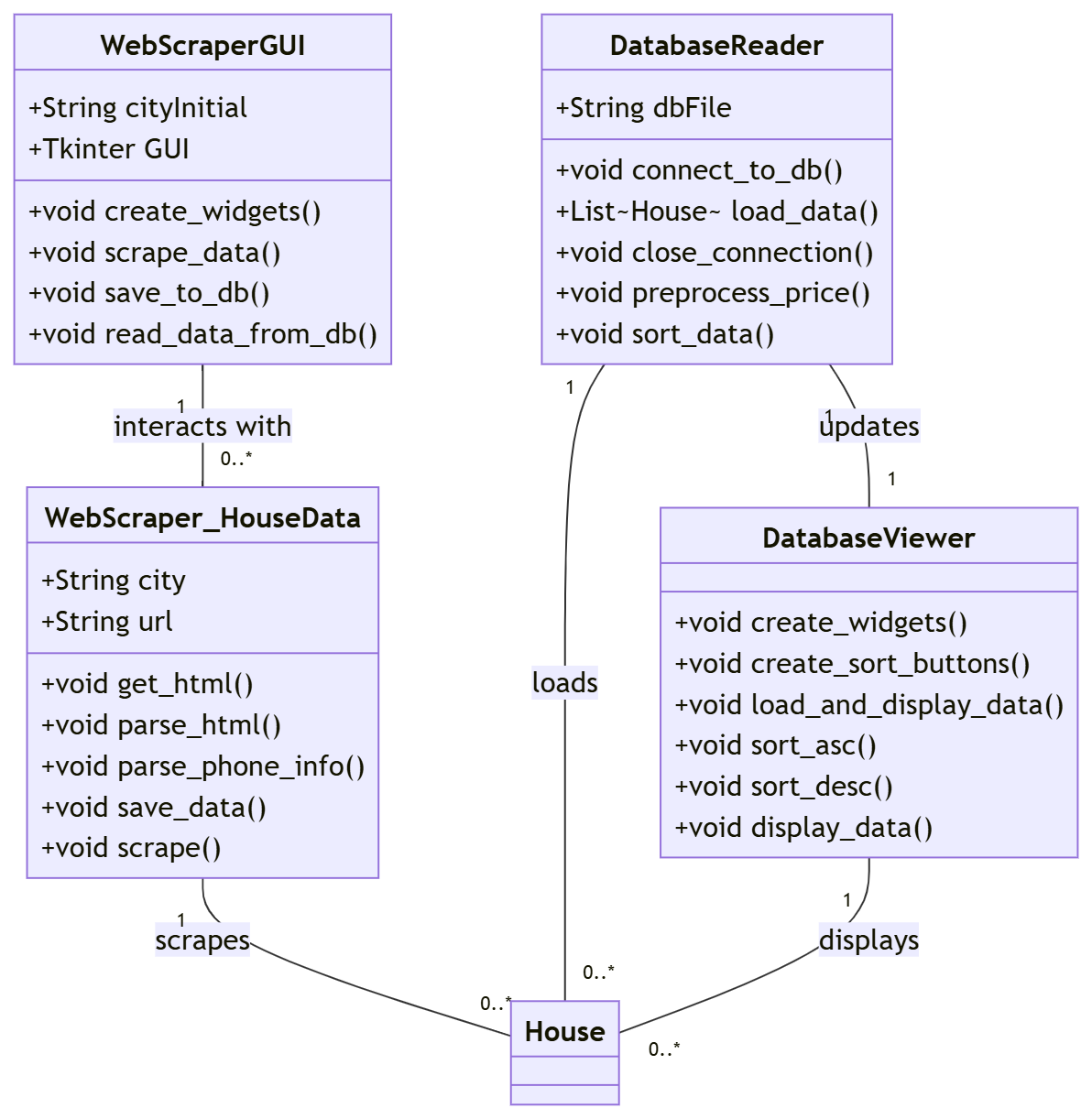


#### 价格降序



此页面上并没有完全显示，鼠标滚轮上下滚动即可查看更多数据

# 四、源码解读



**说明：**

* **WebScraper\_HouseData**：负责爬取二手房数据，解析并保存数据。包含方法如 get\_html()，parse\_html()，save\_data() 等。
* **WebScraperGUI**：负责图形用户界面的构建。允许用户输入城市首字母，启动数据爬取，并将数据保存到数据库中，或从数据库中读取数据。
* **DatabaseReader**：负责与数据库的交互，加载数据，处理房价数据（如预处理房价），并对数据进行排序。
* **DatabaseViewer**：负责在 GUI 界面上显示数据库中的房源数据，并允许用户对数据进行排序。

**关系：**

* **WebScraper\_HouseData** 与 **House**：WebScraper\_HouseData 类从网站爬取数据并生成多个 House 实例。
* **WebScraperGUI** 与 **WebScraper\_HouseData**：GUI 与爬虫模块交互，启动数据爬取过程。
* **DatabaseReader** 与 **House**：DatabaseReader 从数据库中读取房源数据。
* **DatabaseViewer** 与 **House**：DatabaseViewer 显示房源数据，支持排序。
* **DatabaseReader** 与 **DatabaseViewer**：DatabaseReader 更新数据库内容，并通过 **DatabaseViewer** 显示给用户。

源码及README.md均在压缩包中，并且均包含了详细注释，在此就不再cv一下，感觉没有什么意义，还显得文档冗余，故文档只保留了一些必要说明和功能展示。