## 运行环境

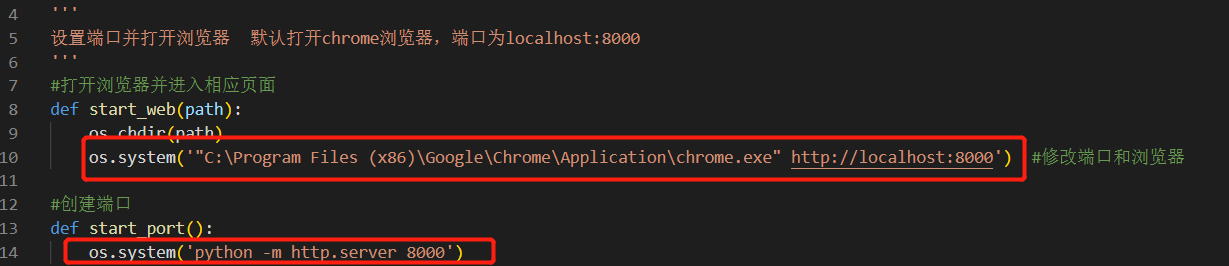
程序的运行环境为python 3.7。

## 运行方法

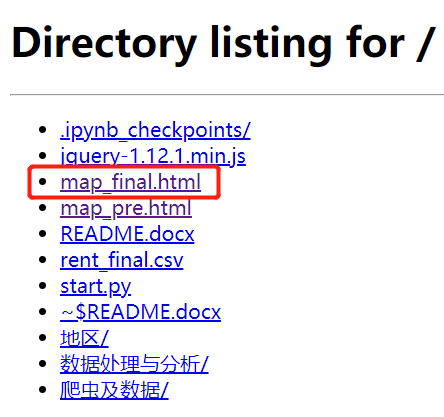
演示视频见：

<视频-初版功能演示.wmv>，<视频-终版功能①.wmv>，<视频-终版功能④.wmv>，<视频-终版功能⑤.wmv>

①运行当前文件下的start.py文件。默认设置当前文件夹为localhost:8000端口，并默认以Chrome浏览器打开localhost:8000页面。如需修改端口和浏览器，在start.py中相应位置修改。



②成功打开浏览器后，打开map\_final.html文件，会显示地图页面。



③在地图页面右上角选择工作地点、通勤方式、价格、来源等信息后，导入rent\_final.csv文件。



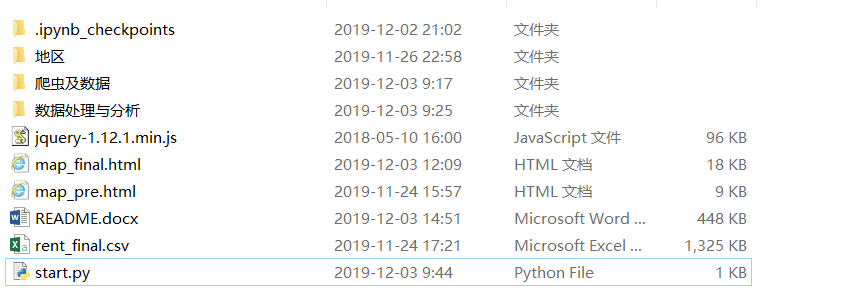
④地图出现的带数字六边形，为聚合的房源点，数字代表聚合数量，不同颜色代表不同来源（58，链家，赶集）的房源。

蓝色圆形为单个房源地点，点击蓝色圆形，会显示该房源的来源、价格、类型信息，并提供相应页面跳转连接，以及工作地到此处的通勤方式。



## 项目各文件介绍

本项目的文件夹结构如下：



①爬虫及数据

该文件夹下包含三个爬虫文件，分别爬取58、链家和赶集的数据，并分别生成rent\_58.csv、rent\_gj.csv、rent\_lj.csv文件。

**rent\_合并.csv**由rent\_58和rent\_lj合并成，包含四个字段”title”，”location”， ”price”，”url”。由于rent\_gj.csv字段不同，并未在此处合并。



②数据处理与分析

**analysis**为数据分析文件，该文件对rent\_source\_合并.csv数据进行了描述性统计、分组统计等分析。

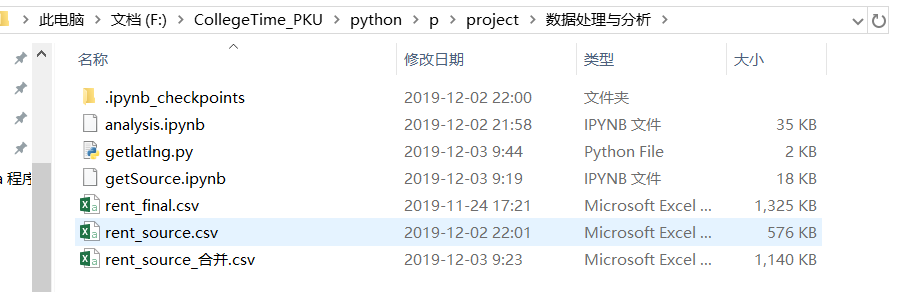
**getLatlng.py**为rent\_fianl.csv文件的lat（纬度）和lng（经度）列获取数据，生成的数据仍命名为rent\_fianl.csv。

**getSource**为爬虫及数据文件夹下的rent\_合并.csv文件新增source列，其中来源为58的房源source列值为“58”，来源为链家的房源source值为“lj”，生成文件为rent\_source.csv。

**rent\_source.csv**包含五个字段”title”，”location”，”price”，”url”，”source”。包含58和链家数据。

**rent\_source\_合并.csv**为rent\_source.csv和rent\_gj.csv文件合并成，包含五个字段”title”，”location”，”price”，”url”，”source”。包含58、链家和赶集数据。

**rent\_fianl.csv**文件由rent\_source\_合并.csv添加lng、lat列得到。lng和lat数据由getLatlng.py获取。其中，极个别数据获取不到经纬度信息，该部分数据已经手动剔除。该文件包含七个字段”title”，”location”，”price”，”url”，”source”，“lng”，“lat”。包含58、链家和赶集数据。



③map\_fianl.html

为最终版的页面文件，由map\_pre.html经过优化和改进得到，即运行时展示的页面。

④map\_pre.html

为初版的页面文件，实现了部分功能，但存在缺陷。若想运行此页面，在浏览器打开该文件即可。

⑤rent\_fianl.csv

由数据处理与分析文件夹下的同名文件拷贝而来，为最终需导入的数据文件。