

额外加分项课题

数据结构与算法

Outline

- **任务：**课题四选一做即可
- **满分：**5分
- **基本完成：**3分（报告基本完整，代码可以直接跑通并画图且注释清晰 / 输出json文件）
- **截止时间：**2018/05/31 23:59
- **提交方式：**见课程网站
- **课题1-3数据链接：** https://pan.baidu.com/s/1nRBusv0g5n_FiYMe58Ev7Q 密码:aho3
- **推荐地图绘图方法：** Python folium库、百度地图API、MapBox
- **推荐Voronoi图绘制方法：** Python scipy.spatial.Voronoi()

课题

- 全天人口分布变化分析
- 不同类型信令数据分布分析
- 高峰期人口流动分析
- 最短路径应用

课题1: 全天人口分布变化

- **数据:** 一天的信令数据（抽样5w用户）、基站位置分布数据
- **任务:** 画出全天每个小时的人口分布热力图或Voronoi图等，并根据图描述合肥市当天的人口流动规律。
- **提交:**
 - 将24张**人口分布图**以小时数命名，保存在pic/文件夹下，如pic/voronoi9.jpg。
 - 提交**分析报告**一份，需要包括但不限于数据描述（python版本、需要的包以及其他用到的语言和版本）、数据处理过程（如何定义某时刻的人口分布，数据预处理等）、当天人口流动规律分析并赴几张关键图，报告提交pdf，如学号_姓名.pdf。
 - 提交带注释的**代码**，放在src/文件夹下。
 - 所有结果打包成ZIP格式，命名：学号_姓名.zip。压缩后文件小于10MB。

Dataset

- 信令数据: 1天5万用户

random_users_signals20170607.csv

- cell_id (基站ID)
- dates (日期时间)
- service_type (服务类型)
- user_id (用户ID)
- web (4G, 3G, 2G)

	cell_id	dates	service_type	user_id	web
0	110334208	2017-06-07 10:28:59	LTE_ATTACH	/hzAXS5MaEuvCLz5rSYdNw==	4G
1	110112563	2017-06-07 10:28:58	LTE_ATTACH	ffWmYB9PWpm7iL7PIiUolg==	4G
2	111050293	2017-06-07 10:28:59	LTE_ATTACH	9saso6/7c5qP02csiPNaMA==	4G
3	110331443	2017-06-07 10:28:59	LTE_ATTACH	t2CyXmTY2xLJ0RQh3DTlHw==	4G
4	110419508	2017-06-07 10:28:59	LTE_ATTACH	0DwTTThk1aYJxszg7NQLyw==	4G

	cell_id	lat	lon	radius
0	110468915	31.875043	117.381348	300
1	110468914	31.880910	117.382230	300
2	232408085	31.850418	117.273734	300
3	232408086	31.855779	117.270840	300
4	110976259	31.885605	117.355936	300

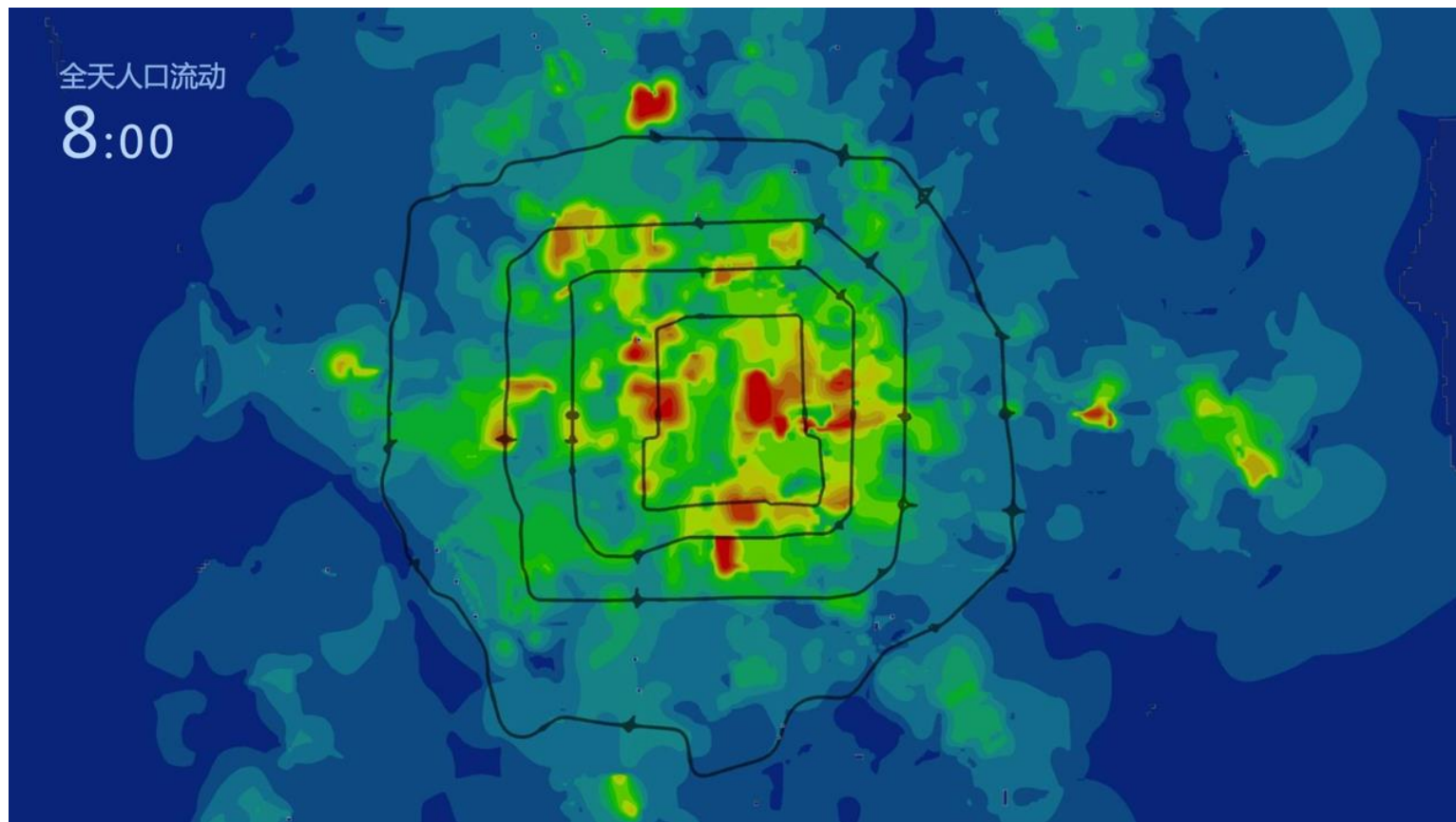
- 基站分布数据:

cells_baidu_hf.csv

- cell_id (基站ID)
- lon (经度)
- lat (纬度)
- radius (辐射半径)

链接:https://pan.baidu.com/s/1nRBusv0g5n_FiYMe58Ev7Q
密码:aho3

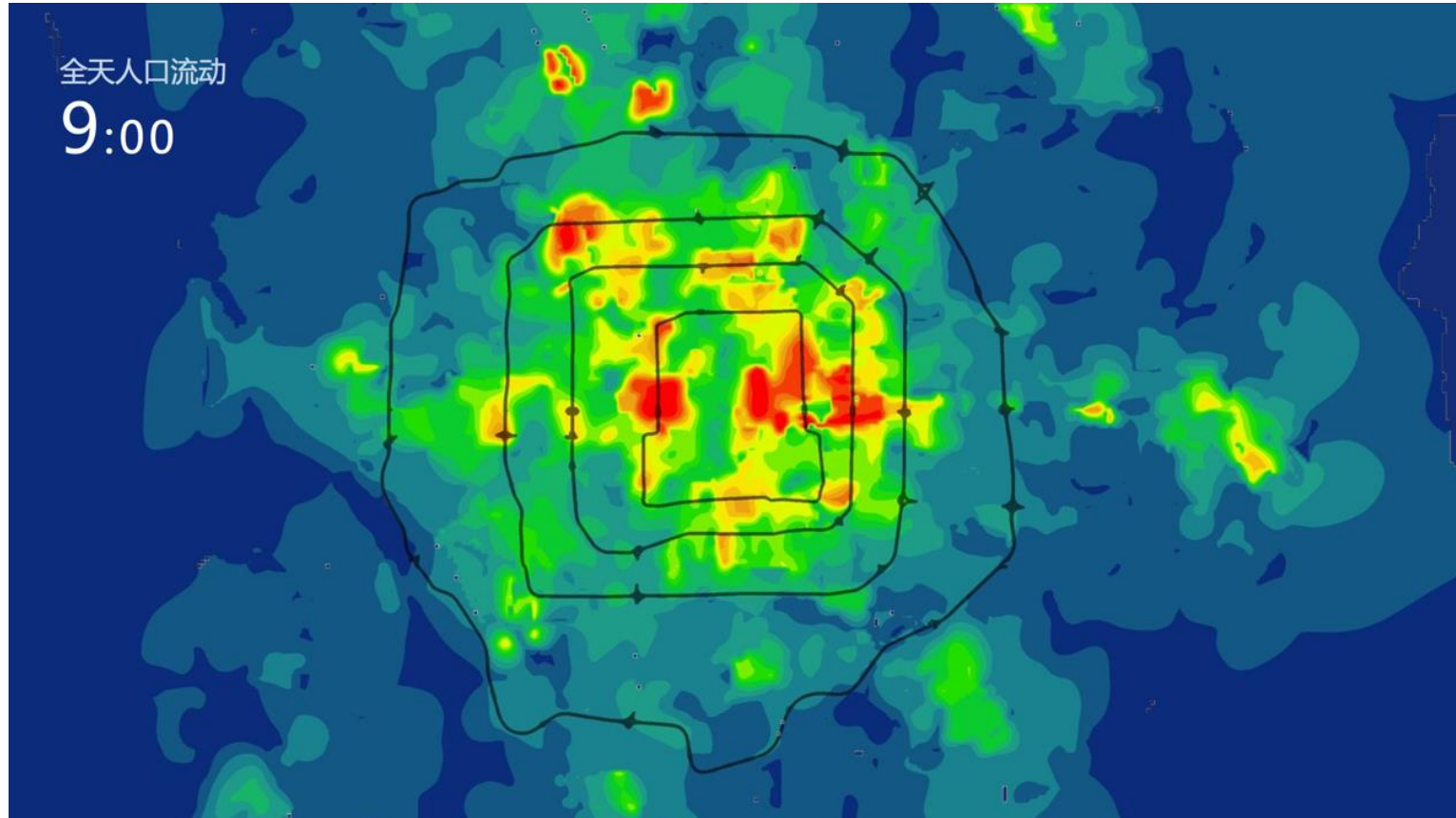
示例图（热力图）



图例

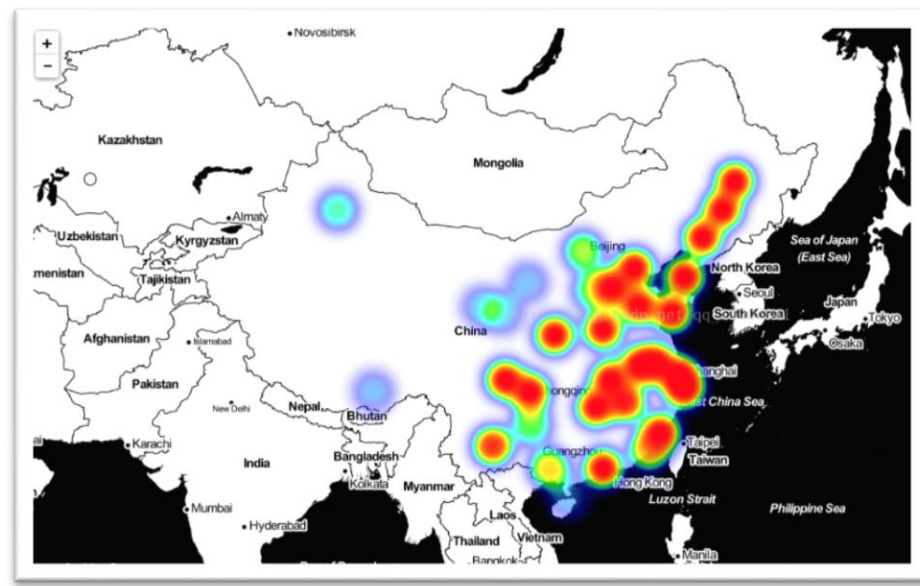
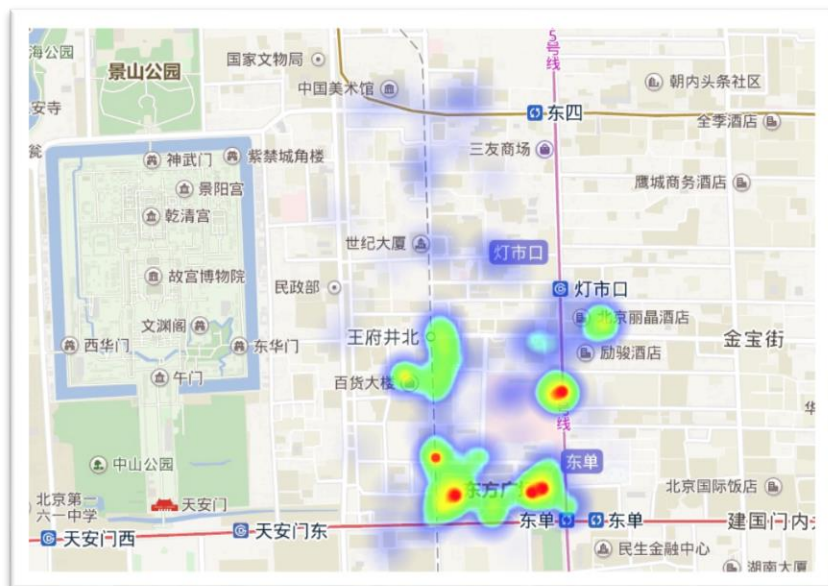


示例图（热力图）



热力图

- 推荐：百度地图API / Python folium库
- 参考：http://developer.baidu.com/map/jsdemo.htm#c1_15
- 参考：https://blog.csdn.net/qq_14906811/article/details/74906275
- 百度地图需要先申请秘钥AK：<http://lbsyun.baidu.com/apiconsole/key#>



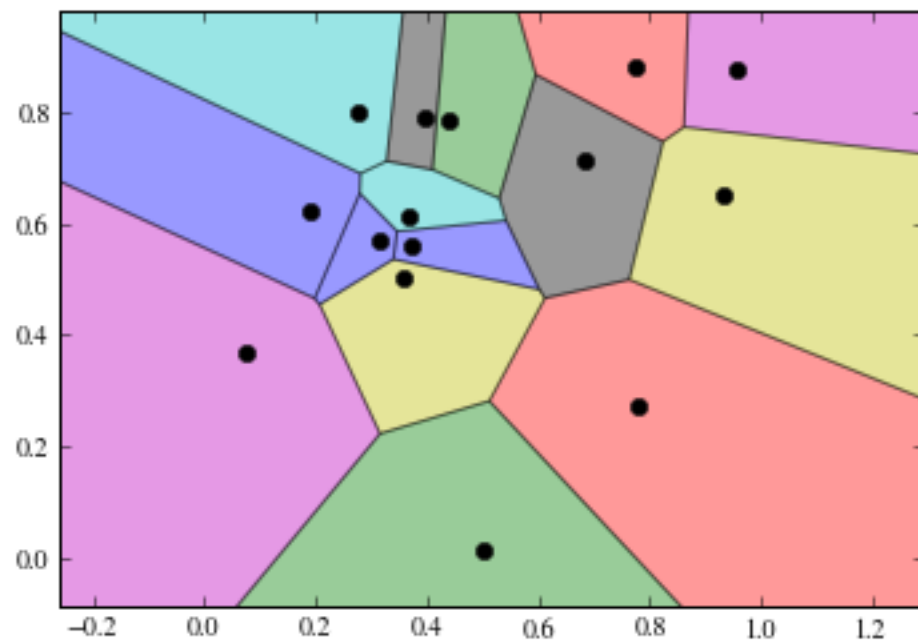

```

<!DOCTYPE html>
<html>
<head>
    <meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
    <meta name="viewport" content="initial-scale=1.0, user-scalable=no" />
    <style type="text/css">
        body, html, #allmap {width: 100%;height: 100%;overflow: hidden;margin:0;font-family:"微软雅黑";}
    </style>
    <script type="text/javascript"
src="http://api.map.baidu.com/api?v=2.0&ak=您的密钥"></script>
    <title>地图展示</title>
</head>
<body>
    <div id="allmap"></div>
</body>
</html>
<script type="text/javascript">
    // 百度地图API功能
    var map = new BMap.Map("allmap");    // 创建Map实例
    map.centerAndZoom(new BMap.Point(116.404, 39.915), 11);    // 初始化地图,设置中心点坐标和地图级别
    //添加地图类型控件
    map.addControl(new BMap.MapTypeControl({
        mapTypes:[
            BMAP_NORMAL_MAP,
            BMAP_HYBRID_MAP
        ]}));
    map.setCurrentCity("北京");    // 设置地图显示的城市 此项是必须设置的
    map.enableScrollWheelZoom(true);    //开启鼠标滚轮缩放
</script>

```

示例图（Voronoi图）

- Voronoi图，又叫泰森多边形或Dirichlet图，它是由一组由连接两邻点直线的垂直平分线组成的连续多边形组成。N个在平面上有区别的点，按照最邻近原则划分平面；每个点与它的最邻近区域相关联。
- 参考：<http://nbviewer.jupyter.org/gist/pv/8037100>
- 颜色：由人口数量控制颜色深浅



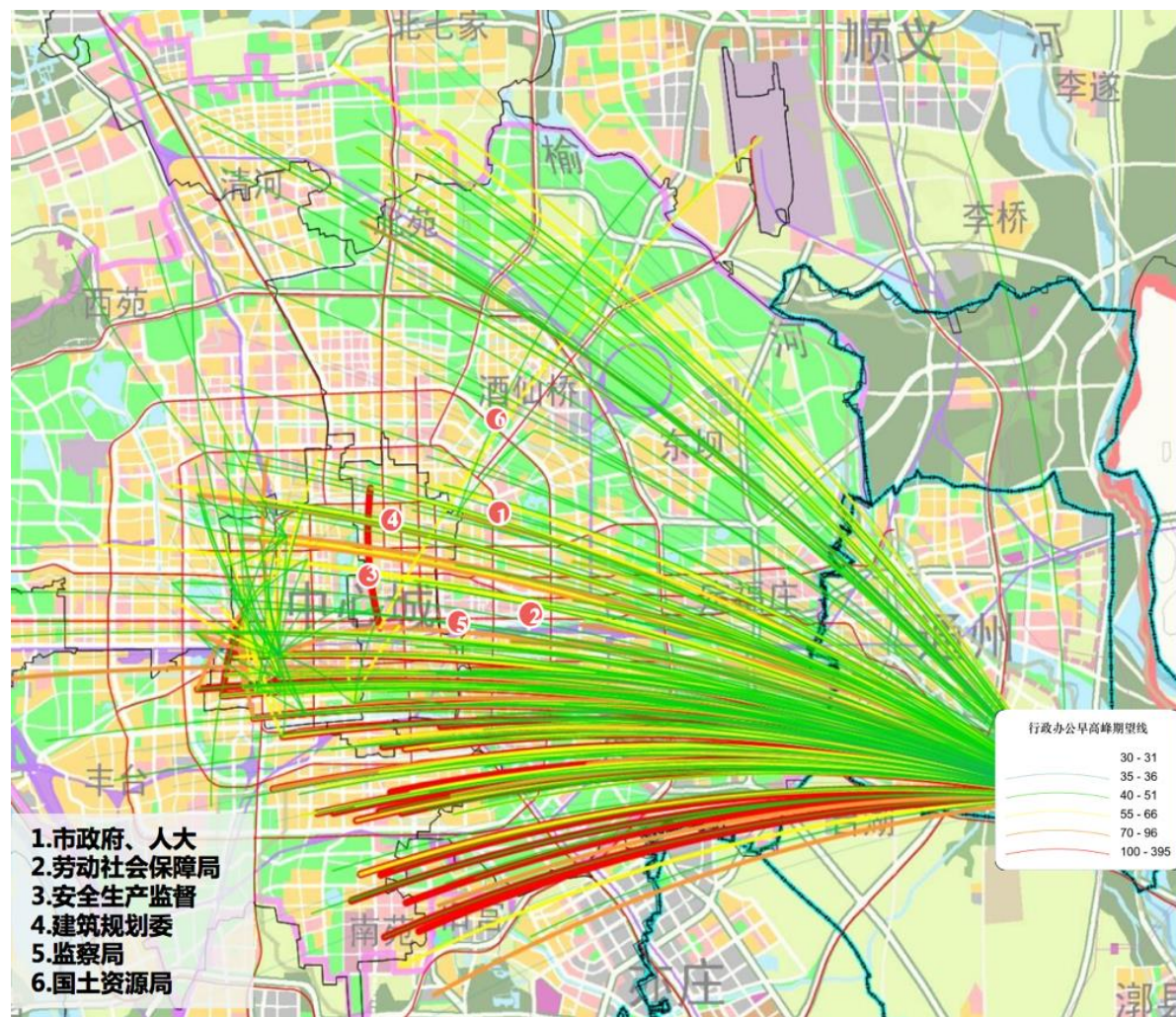
课题2:不同类型信令数据分布分析

- **数据：**一天的信令数据（抽样5w用户）、基站位置分布数据
- **任务：**画出全天8点、12点、18点不同类型信令数据的分布，如热力图，并根据图分析不同类型的信令数据在不同的时间段在分布上的差异。
- **提交：**
 - 将不同类型信令数据的分布图保存在pic/文件夹下。
 - 提交**分析报告**一份，需要包括但不限于数据描述（python版本、需要的包以及其他用到的语言和版本）、数据处理过程（如数据预处理等）、当天人口流动规律分析并赴几张关键图，报告提交pdf，如学号_姓名.pdf。
 - 提交带注释的**代码**，放在src/文件夹下。
 - 所有结果打包成ZIP格式，命名：学号_姓名.zip。

课题3:高峰期人口流动分析

- **数据：**一天的信令数据（抽样5w用户）、基站位置分布数据
- **要求：**分析当天某区域早高峰、晚高峰人口流入流出。区域自选，可以是居住区、办公区，需要有明显的流入流出趋势。
- **提交：**
 - 提交**分析报告**，需要包括但不限实验环境（python版本、需要的包以及其他用到的语言和版本）、数据描述、数据处理过程（如选取的分析区域，早晚高峰的定义，流入流出重点的判断方法等）、人口流动分析并附图，报告提交pdf，如学号_姓名.pdf。
 - 提交带注释的**代码**，放在src/文件夹下。
 - 所有结果打包成ZIP格式，命名：学号_姓名.zip。

示例图



课题3:最短路径应用

- **数据：**不提供，从OpenStreetMap上下载北京市路网数据（OSM无法直接下载某个城市的数据，用矩形框框个大概即可）。
- **要求：**1. 用合适的数据格式来存储路网数据。2. 实现给定任意两条路段的最短路径算法，并考虑算法优化方法。3. 地图可视化。
- **提交：**
 - 提交实验报告，需要包括但不限实验环境（python版本、需要的包以及其他用到的语言和版本）、数据来源（来源、城市选取的矩形范围）、算法描述和实验（具体算法和改进，复杂度分析，实验可以是北京大学东门到奥林匹克公园站的最短路径），报告提交pdf，如学号_姓名.pdf。
 - 提交带注释的代码，放在src/文件夹下，实现功能包括但不限于给定一个经纬度找到最近的路段、给定两个路段的最短路径算法、最短路径的可视化。
 - 所有结果打包成ZIP格式，命名：学号_姓名.zip。

OpenStreetMap

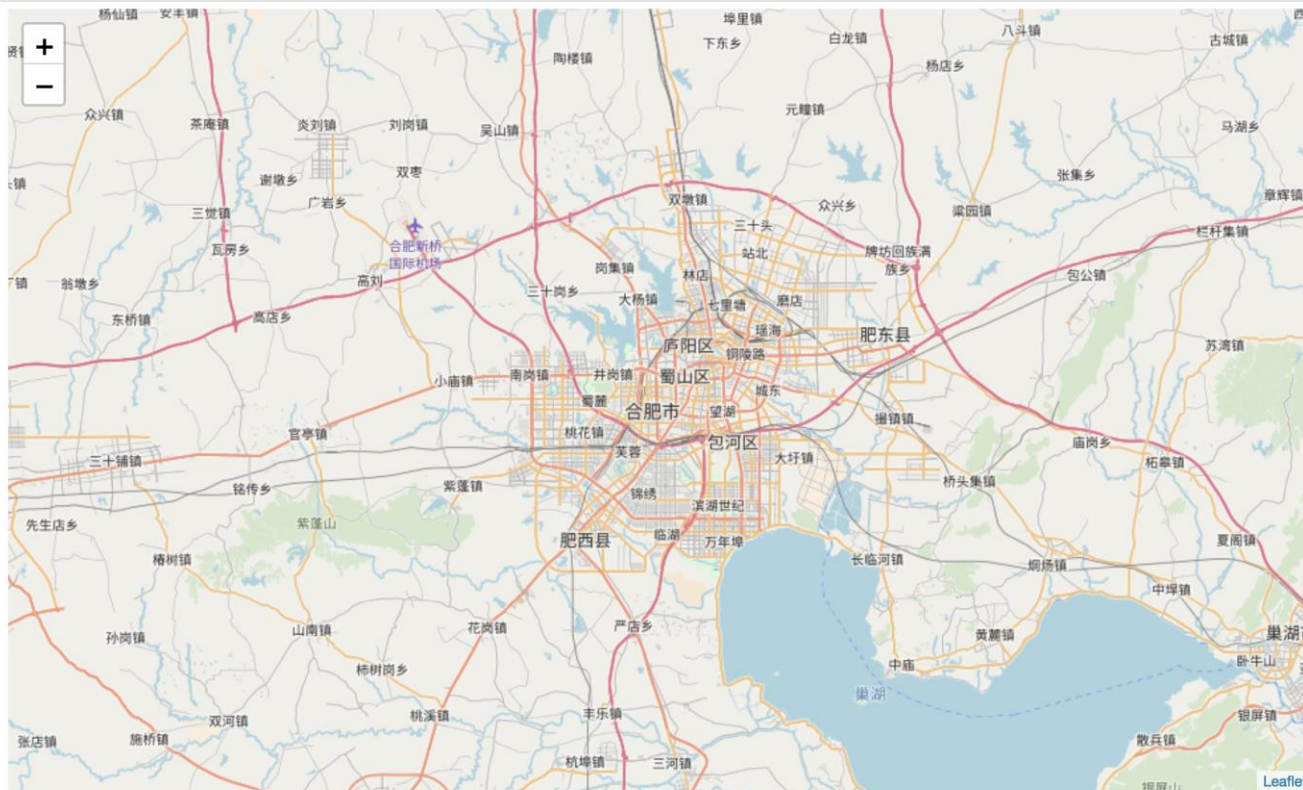
- 是一个开放街道地图
- 官网: <https://www.openstreetmap.org/>
- JOSM下载地图: <https://josm.openstreetmap.de/wiki/Download>
- 选中道路->搜索道路类型->另存为Json文件

Python folium库

```
In [1]: import folium  
map_osm = folium.Map(location=[31.83, 117.23])
```

```
In [2]: map_osm
```

Out[2]:



Python folium库

```
In [3]: folium.Map(location=[31.83, 117.23],
                        tiles='Stamen Toner',
                        zoom_start=13)
```

Out[3]:

