杭州城市西湖区活力评价

1 前言

1.1 研究区简介

杭州是浙江省省会和经济、文化、科教中心，长江三角洲中心城市，重要的风景旅游城市，首批国家历史文化名城。杭州山水相依、湖城合璧，江、河、湖、海、溪五水共导，风景如画，堪称“人间天堂”。

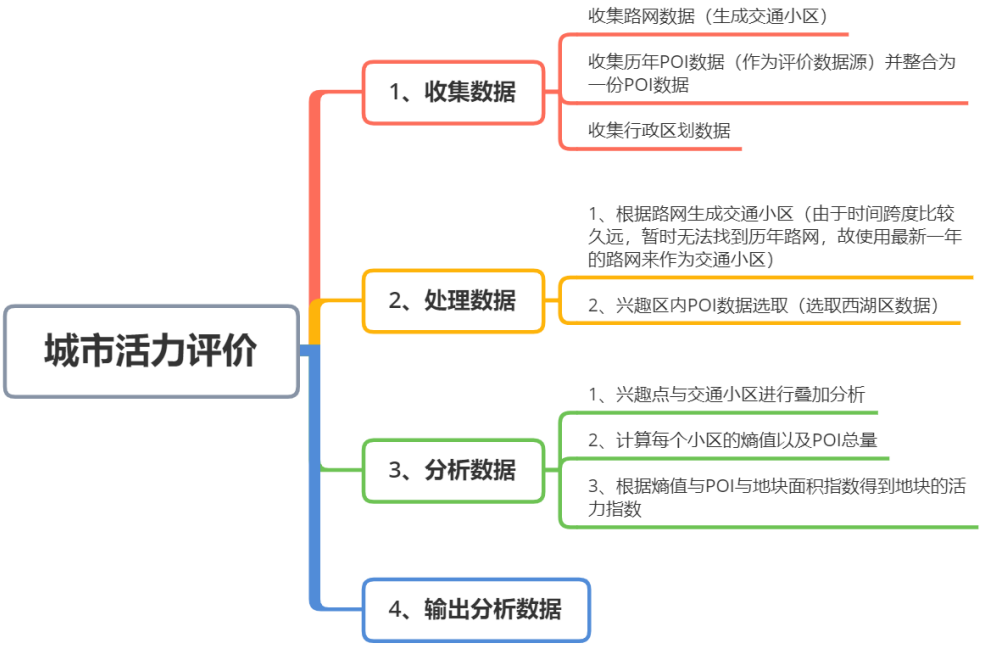
西湖区是浙江省委、省政府的所在地。辖区内有西湖景区、西溪国家湿地公园、西山国家森林公园等优质旅游资源，以及灵山风情小镇、兰里景区等特色景点，绿化覆盖率 40.6%、森林覆盖率 42.6%以上，拥有得天独厚的自然条件。

1.2 研究方法

如何评价城市空间的活力？有很多的专家从很多的维度进行了探索，有些专家从便捷性的角度出发，认为适宜居住的城市空间应该具有良好的可达性（在往期的博客大赛种也有很多相关的作品），也有些专家认为适宜的建设强度与建筑形态、足够的功能混合度是促生城市活力的空间形态基础。那么混合度从何而来，又应该怎么去量化呢？

杭州是历史文化名城，西湖区拥有得天独厚的自然条件。那么现在的西湖区是否有足够的活力？本文将采用 FME 平台，结合互联网上的资源，用数据的方式对西湖区各区块活力进行一个分析。

研究的思路如下图所示：



基于这样的思路，本文将开始对西湖区的活力展开分析。

2 收集数据

数据是分析的基础，本次的分析也不例外。按照思路，本文共需要三种数据，分别是：1、路网数据，2、POI 数据，3、行政区划数据。

在收集好各种数据之后，我将使用路网+行政区划数据，来构建交通小区（以下统称地块）；使用 POI 来分析城市的各地块的混合度；并给予此来评估地块的活力。

2.1 路网数据

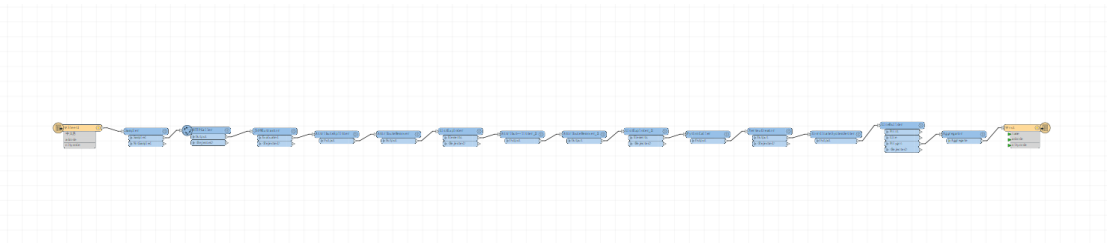
本文中使用到的路网数据，来自于 OSM。获取方式为从 OSM 平台直接下载。

2.2 历年 POI 数据

本次分析使用到的数据为 2012-2017 年 POI 数据，数据来源于互联网开放平台。获取方式为从互联网直接下载。

2.3 行政区划数据

来自于国内数据厂商：高德。获取方式为：使用 FME 模板下载。下载模板如下所示：



获取逻辑为：1、发请求；2、解析数据；3、纠偏；4、写出。由于数据获取不是本文重点，故不再赘述。

3 处理数据

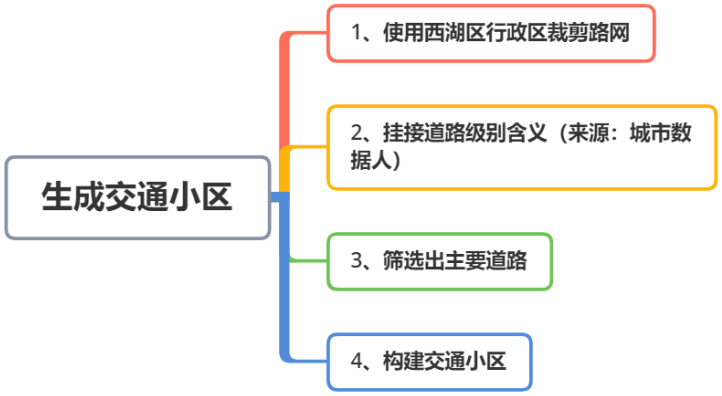
为了这次研究，我获取到了很多的数据，但这些数据都没办法直接使用。想要把数据应用到研究中，需要进行一些处理。在这一步骤主要完成两项数据处理：1、生成地块数据，2、筛选出研究区 POI 数据。那么一起来看下我是怎么处理这些数据的吧。

3.1 地块数据生成

地块（交通小区）是指由路网构成的不规则网格，地块是承担城市管理和城市规划社会经济功能的基本单元，也是城市区域的自然边界之一。

之所以采用地块而不是画格网来进行分区的原因是：与格网相比，地块内部具有更紧密的联系。一个地块由若干主干道围合而成，可能是一个小区、居民点或者商场，而格网不具有这种特性。因此后续的分析将基于地块进行。

生成地块的思路为：

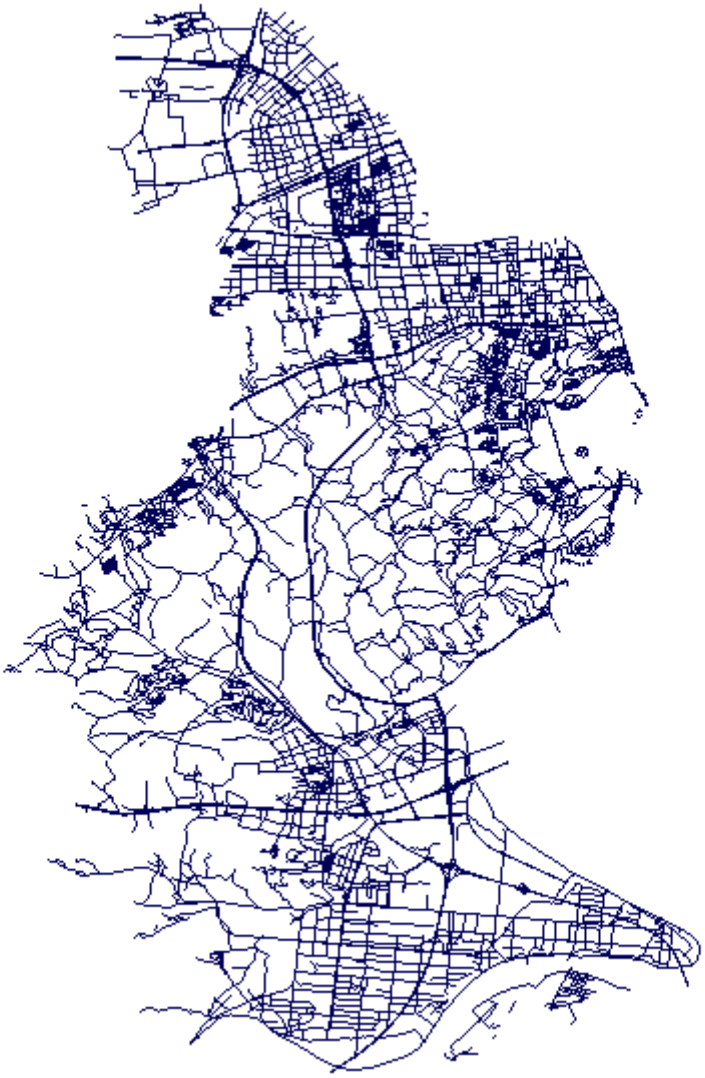


具体的处理流程为：

下面来介绍下处理流程中的几个关键步骤。

3.1.1 裁剪路网

从 OSM 下载的路网数据，并没有按照行政区划进行区分。也就是说：想要得到西湖区的路网数据，需要一些处理。我采用了 FME 来对数据做裁剪处理。使用 FME 的 Clipper 转换器，对 OSM 的路网进行裁剪后，得到杭州市西湖区的路网数据



3.1.2 挂接道路级别含义

OSM 道路数据有自己的一套分类体系，包含了主干道、高速路与各种小路。具体的如下图所示：

- tertiary（第三级道路）：城市支路，
- tertiary_link（第三级道路-连接）：匝道，机场集散车行道路（数据量极少，多是未知道路）
- residential（居住区道路）：居住区车行道路，
- unclassified（未分类道路）：居住区车行道路，滨水车行道路，机场机动车通道
- secondary（次要道路）：城市次要车行道路，机场外围车行道路，
- secondary_link（次要道路-连接）：城市次要车行道路立交、匝道（数据量少，部分零散分布未知类别）
- primary（主要道路）：城市主要车行道路
- primary_link（主要道路）：城市主要车行道路立交，城市主要车行道路匝道（数据量少，部分零散分布未知类别）
- motorway（高速公路）：高速公路，过江隧道
- motorway_link（高速公路-连接）：高速公路立交，匝道
- trunk（干道）：高架快速路，机场进站快速路，过江隧道，桥上快速路
- trunk_link（干道-连接）：立交，匝道，桥上引道，机场进站快速路，国道改道
- track（小路）：郊区、乡村、工矿区、田间、林间小路
- track_grade1（小路 级别1）：郊区、乡村、工矿区、田间、林间小路
- track_grade2（小路 级别2）：郊区、乡村、工矿区、田间、林间小路
- track_grade3（小路 级别3）：郊区、乡村、工矿区、田间、林间小路
- track_grade4（小路 级别4）：郊区、乡村、工矿区、田间、林间小路
- track_grade5（小路 级别5）：郊区、乡村、工矿区、田间、林间小路
- bridleway（马道）：体育场馆内部专用道路（数据量极少，零星道路在公园、居住区内部）
- living_street（生活街道）：居住区车行道路，公园车行道路
- path（小道）：公园车行道路，居住区车行道路（分布零碎，量少）
- service（服务性道路）：居住区车行道路，火车站集散车行道，公园车行道路，公共建筑集散车行道，公交枢纽入口车行道路，停车场入口车行道路
- footway（人行道）：滨水绿道，公园步行道，广场步行道，大学步行道路，人行道，火车站人行集散道路
- pedestrian（步行街道）：步行街，广场步行道路，公园步行道路，居住区步行道路
- steps（台阶踏步）：人行过街天桥台阶，广场台阶、公共建筑入口台阶，登山台阶
- cycleway（自行车道）：滨水绿道，非机动车道，公园自行车道
- unknown（未知道路）：滨水车行道路，校园广场车行道路，乡道（数据量少，比较难判别道路类型）

内容来源于城市数据人

英文部分为 OSM 数据中的分类信息，还好有人已经整理好了其分类对应的信息（上图）。但这样的内容没办法直接使用，借助 FME 的能力，我将上面的文本处理成了下图所示的表格：

	fclass	type	v	text_line_data
1	tertiary	第三级道路	城市支路，	tertiary（第三级道路）：城市支路，
2	tertiary_link	第三级道路-连接	匝道，机场集散车行道路（数据量极少，多是未知道路）	tertiary_link（第三级道路-连接）：匝
3	residential	居住区道路	居住区车行道路，	residential（居住区道路）：居住区车
4	unclassified	未分类道路	居住区车行道路，滨水车行道路，机场机动车通道	unclassified（未分类道路）：居住区车
5	secondary	次要道路	城市次要车行道路，机场外围车行道路，	secondary（次要道路）：城市次要车
6	secondary_link	次要道路-连接	城市次要车行道路立交、匝道（数据量少，部分零散分布未知类别）	secondary_link（次要道路-连接）：城
7	primary	主要道路	城市主要车行道路	primary（主要道路）：城市主要车行道
8	primary_link	主要道路-连接	城市主要车行道路立交，城市主要车行道路匝道（数据量少，部分零散分布未知类别）	primary_link（主要道路-连接）：城市
9	motorway	高速公路	高速公路，过江隧道	motorway（高速公路）：高速公路，过
10	motorway_link	高速公路-连接	高速公路立交，匝道	motorway_link（高速公路-连接）：高
11	trunk	干道	高架快速路，机场进站快速路，过江隧道，桥上快速路	trunk（干道）：高架快速路，机场进站
12	trunk_link	干道-连接	立交，匝道，桥上引道，机场进站快速路，国道改道	trunk_link（干道-连接）：立交，匝道，
13	track	小路	郊区、乡村、工矿区、田间、林间小路	track（小路）：郊区、乡村、工矿区、
14	track_grade1	小路 级别1	郊区、乡村、工矿区、田间、林间小路	track_grade1（小路 级别1）：郊区、
15	track_grade2	小路 级别2	郊区、乡村、工矿区、田间、林间小路	track_grade2（小路 级别2）：郊区、
16	track_grade3	小路 级别3	郊区、乡村、工矿区、田间、林间小路	track_grade3（小路 级别3）：郊区、
17	track_grade4	小路 级别4	郊区、乡村、工矿区、田间、林间小路	track_grade4（小路 级别4）：郊区、
18	track_grade5	小路 级别5	郊区、乡村、工矿区、田间、林间小路	track_grade5（小路 级别5）：郊区、
19	bridleway	马道	体育场馆内部专用道路（数据量极少，零星道路在公园、居住区内部）	bridleway（马道）：体育场馆内部专用
20	living_street	生活街道	居住区车行道路，公园车行道路	living_street（生活街道）：居住区车行
21	path	小道	公园车行道路，居住区车行道路（分布零碎，量少）	path（小道）：公园车行道路，居住区
22	service	服务性道路	居住区车行道路，火车站集散车行道，公园车行道路，公共建筑集散车行道，公交枢纽入口车行道路，停车场入口车行道路	service（服务性道路）：居住区车行道
23	footway	人行道	滨水绿道，公园步行道，广场步行道，大学步行道路，人行道，火车站人行集散道路	footway（人行道）：滨水绿道，公园步
24	pedestrian	步行街道	步行街，广场步行道路，公园步行道路，居住区步行道路	pedestrian（步行街道）：步行街，广
25	steps	台阶踏步	人行过街天桥台阶，广场台阶、公共建筑入口台阶，登山台阶	steps（台阶踏步）：人行过街天桥台阶
26	cycleway	自行车道	滨水绿道，非机动车道，公园自行车道	cycleway（自行车道）：滨水绿道，非
27	unknown	未知道路	滨水车行道路，校园广场车行道路，乡道（数据量少，比较难判别道路类型）	unknown（未知道路）：滨水车行道路

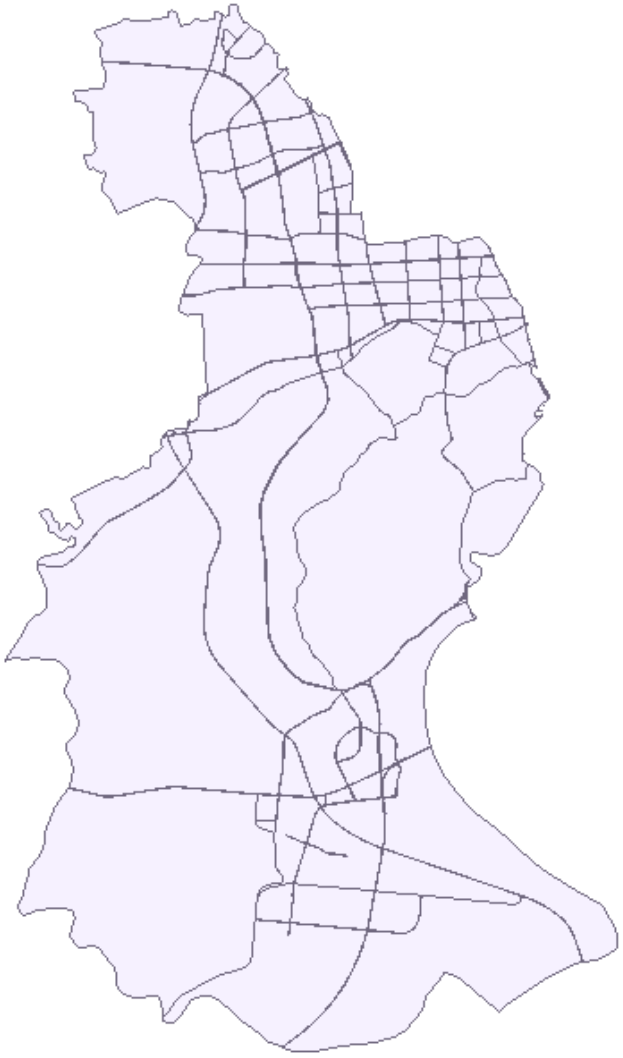
然后使用表格进行挂接，即可按照其中文类型进行道路级别的筛选。

3.1.3 筛选主要道路

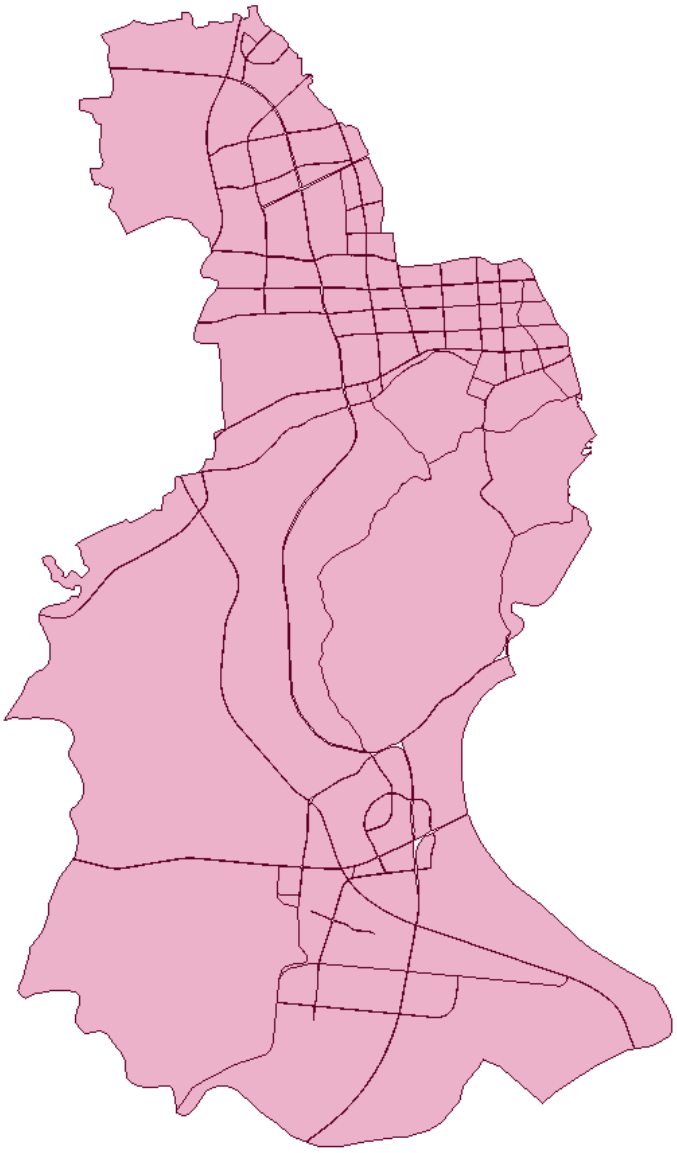
使用 FME 转换器，按照道路级别，筛选出主要道路、高速公路、干道、次要道路四种类型的数据。筛选后的路网如下图所示：

3.1.4 地块构建

使用路网与行政区进行拓扑分析，打断相交的面与线，再进行构面操作，最终形成了如下图所示的图层：

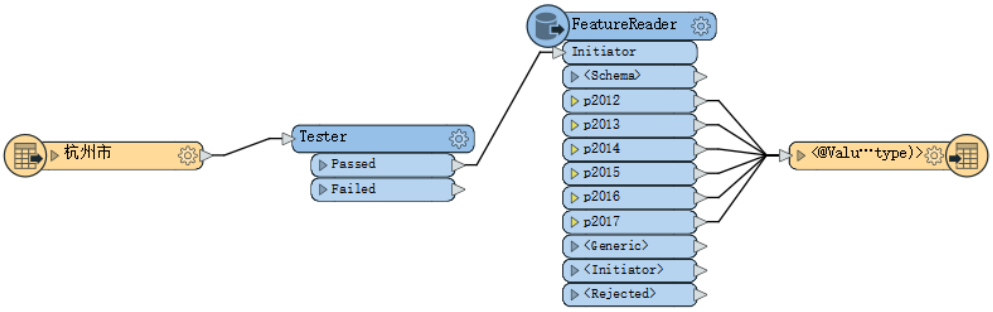


再通过面积与边长的比，去掉一些狭长面，输出后得到了最终的地块：



3.2 POI 筛选

POI 数据是各行各业使用比较多的数据。我获取到的 POI 数据也是全国数据，并没有分区。与地块数据的处理相比 POI 数据的筛选比较简单，用一个模板截图来说明吧。



如上图所示，我使用了 FeatureReader 结合 Where 字句参数来获取历年的数据。使用此方式来进行数据获取，无需读入全部数据，比较高效。

4 分析数据

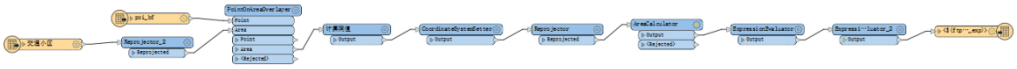
当数据准备好之后，就到了分析的阶段。

4.1 分析思路

我们首先将 POI 与地块进行叠加分析，然后再根据公式计算地块的熵值，再加上 POI 总量及面积指数，最终得到地块的活力指数。

分析的思路为：

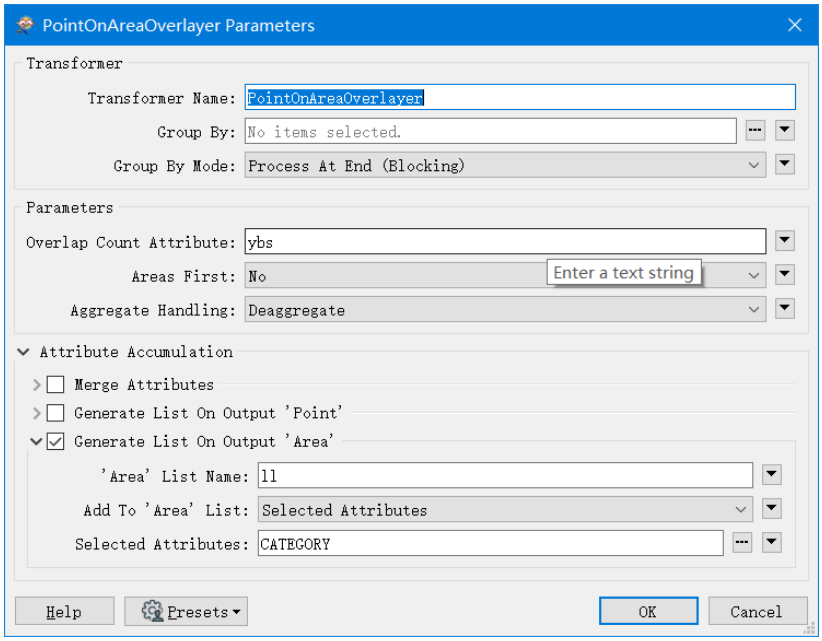
具体的分析流程为：



下面来介绍下分析流程中的几个关键步骤。

4.2 叠加分析

FME 有很多的转换器都可以完成空间数据的叠加分析，在本次分析中，我使用了 PointOnAreaOverlayer 转换器来完成 POI 点层与地块面层的叠加分析：



在叠加时，将叠加次数存储在 ybs 字段中，并且为面要素启用列表参数来存储面上的所有点的类别。

4.2.1 混合度计算

在计算混合度之前，先来了解一下什么是熵，什么是信息熵？

4.2.1.1 什么是熵？

熵，最初是用来描述“能量退化”的物理状态的物质状态参数。但是在很长一段时间内，大家虽然一直都在使用熵，却并没有认清熵的本质。一直到统计物理、信息论等一系列科学理论的发展，熵的本质才逐渐被解释清楚，即：熵的本质是一个系统“内在的混乱程度”。

国内也有专家学者尝试借助“信息熵”的概念来度量城市土地利用的多样性和混合度。

4.2.1.2 什么是信息熵

信息是个很抽象的概念。人们常常说信息很多，或者信息较少，但却很难说清楚信息到底有多少。比如一本五十万字的中文书到底有多少信息量。

直到 1948 年，香农提出了“信息熵”的概念，才解决了对信息的量化度量问题。信息论之父克劳德·艾尔伍德·香农第一次用数学语言阐明了概率与信息冗余度的关系。

信息熵计算公式：

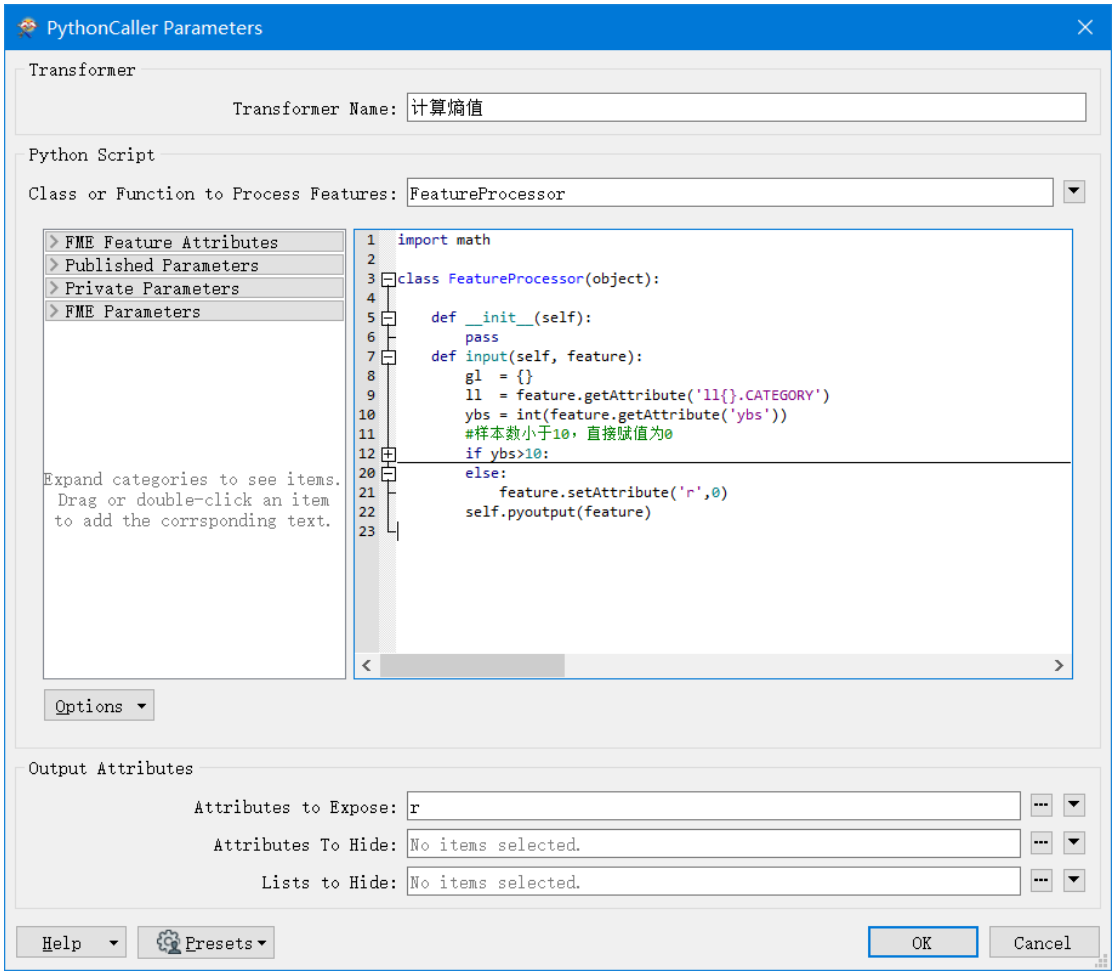
$$H(U) = -p_i \sum_{i=1}^{n} \log p_i$$

其中 p_i 为事件 i 发生的概率。

4.2.1.3 混合度计算

按照熵的计算公式（$H(U) = -p_i \sum_{i=1}^{n} \log p_i$）我们来计算一下每块地的熵值。我们每类 POI 可以看成是同一类型的事件，再把叠加分析后的列表中 POI 数据进行分析，获取到每类 POI 所占比例，然后再根据公式计算，即可得到每个地块的熵值（混合程度）。POI 本身有自己的分类体系，我使用的西湖区历年 POI，涉及到大类 23 类，中类 178 类，小类 533 类。为了更好的体现出科学性，在后续的分析中，我使用最精细的小类来计算熵值。

在实际的计算中，我使用了 Python 来进行熵值的计算，由于 FME 对 Python 的友好支持，使得这个计算非常的简单。



需要注意的是，在取对数的计算中，我使用了每个地块内的 POI 总量做底来计算地块的混合度。

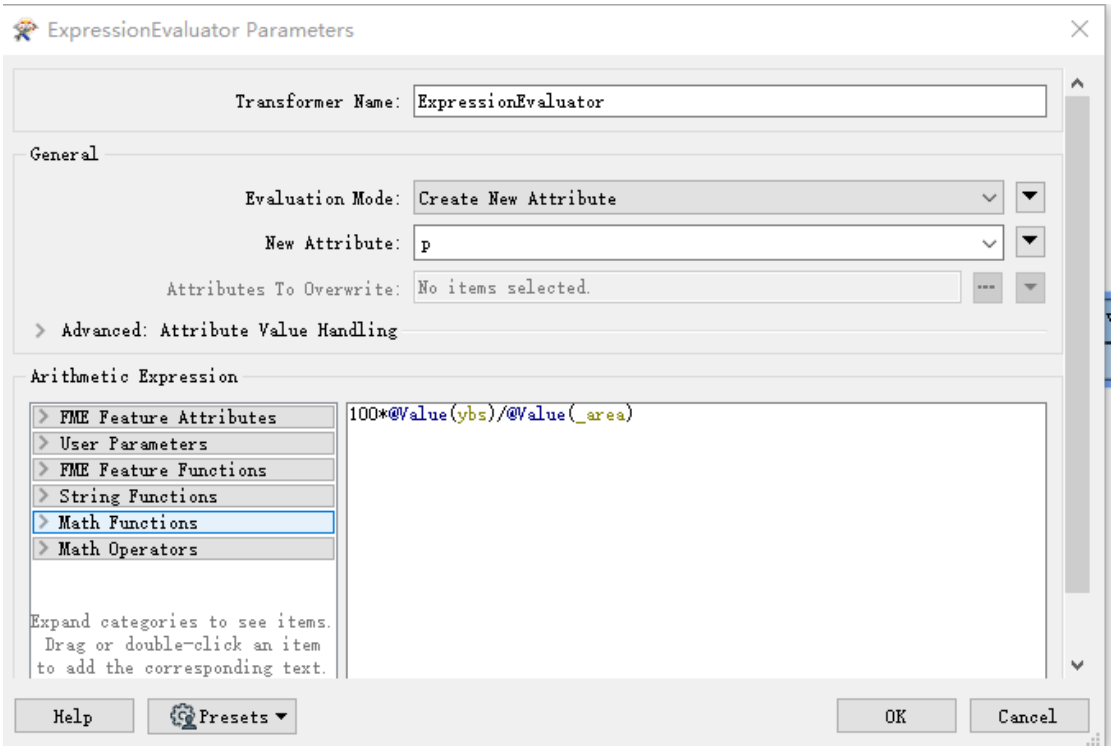
4.2.2 计算 POI 总量及面积指数

在设计这个指数的时候，充分考虑到与总量相关的面积，我设计了如下公式来计算该指数：

$$P = k * m/a$$

其中 P 为指数，k 为常数，m 为地块内 POI 数量，a 为地块面积（单位为平方米）。

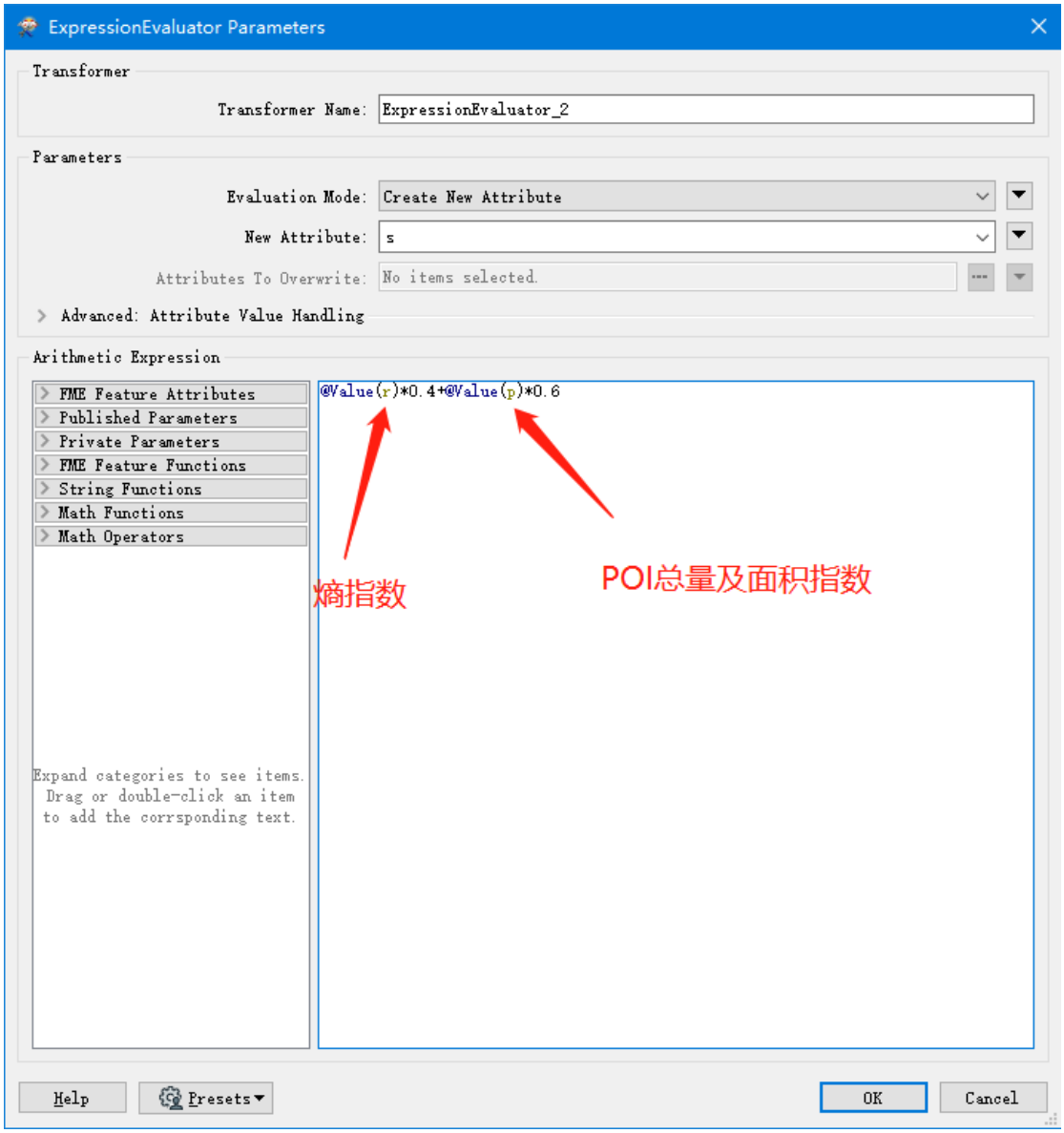
与熵值的计算相比，借助于 FME 对四则运算的友好支持，P 值的计算要更简单一些：



计算方式如上图所示，上图中 k 值取 100。这个值是依经验取的，科学性，还有待验证。

4.2.3 计算活力指数

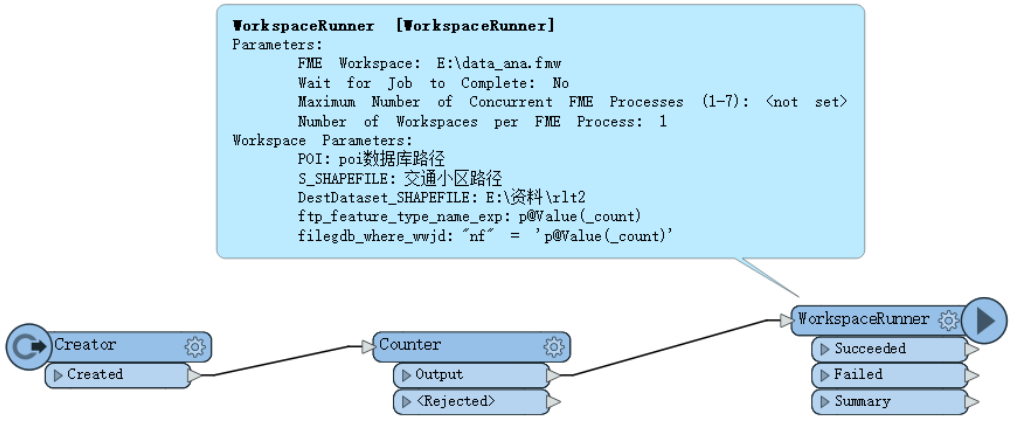
单纯的依靠熵值，并不能很好的体现出地块的活力。POI 数量本身也是一个能够反应城市活力的很重要的一项指标，所以在最终结果中我加入了 POI 总量与地块面积指数来体现地块活力指数。计算得到的最终数值，来作为地块活力的评价依据。



这里的四六开，也是根据经验取的，给 POI 总量及面积指数较大的权重，是为了凸显其重要性。但这种做法的科学性还有待验证。

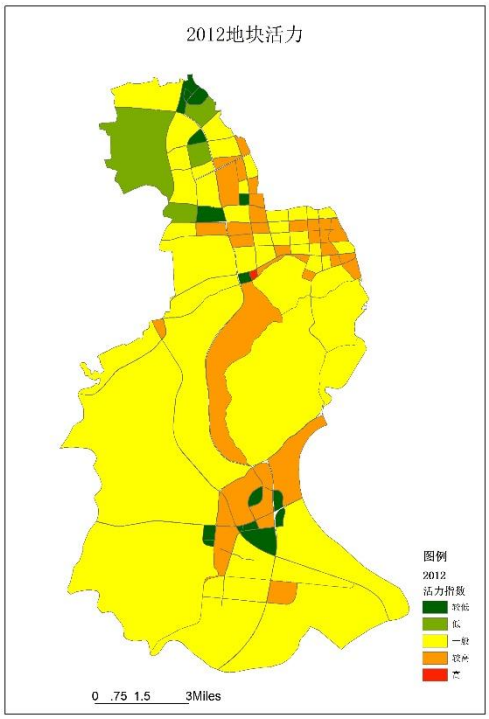
4.2.4 批量计算历年地块活力指数

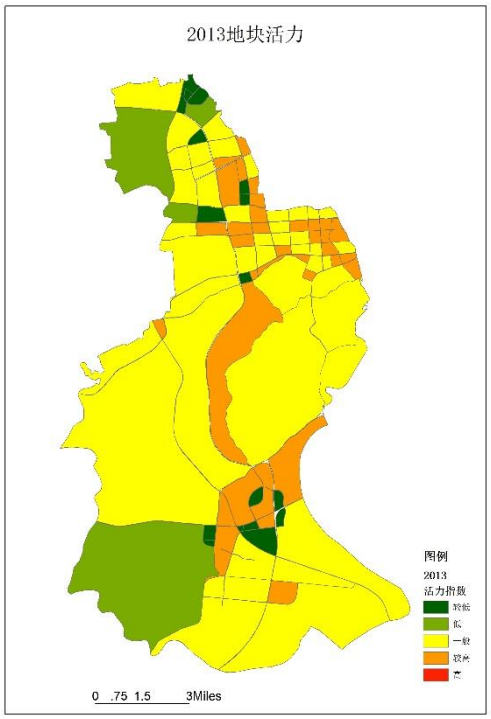
由于使用的数据比较多，在调试好上述功能模板之后，在功能模板的基础上，直接使用批处理方式进行历年地块活力指数的计算

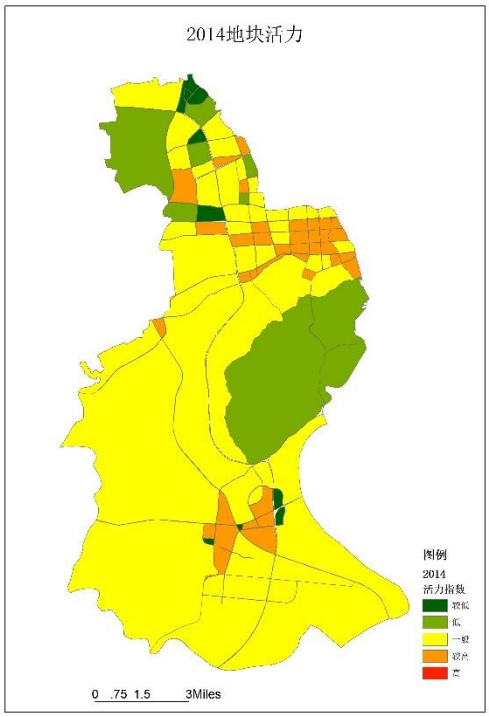


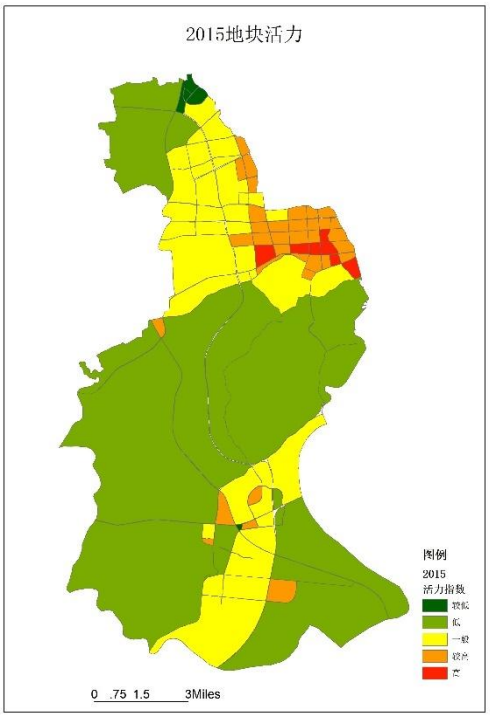
4.3 分析结果

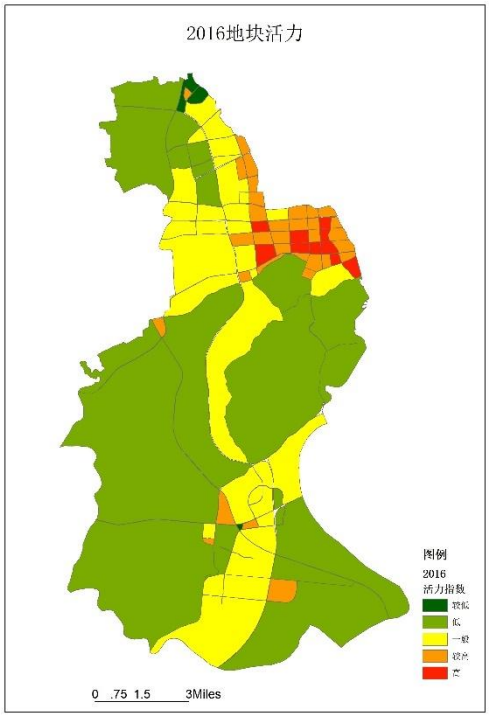
经过上述计算，最终得到了 2012 年至 2017 年的地块活力图。

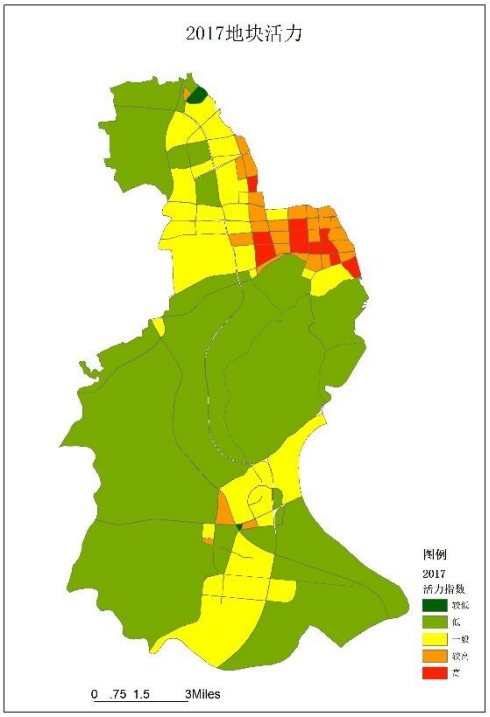








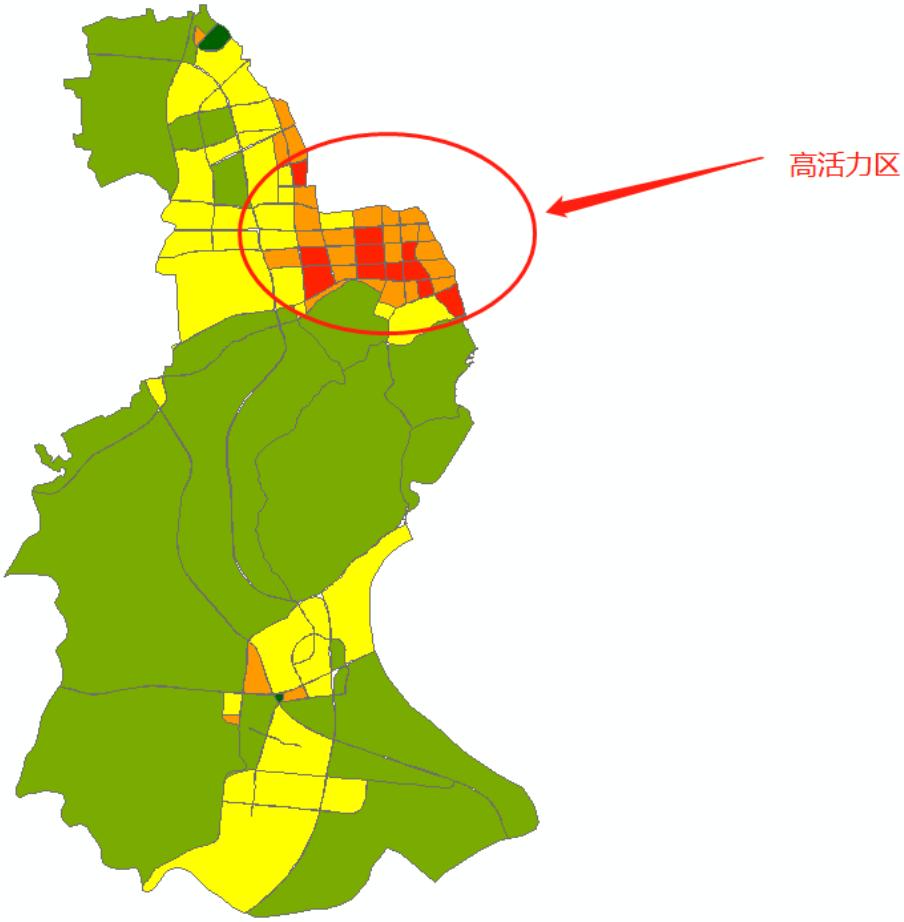




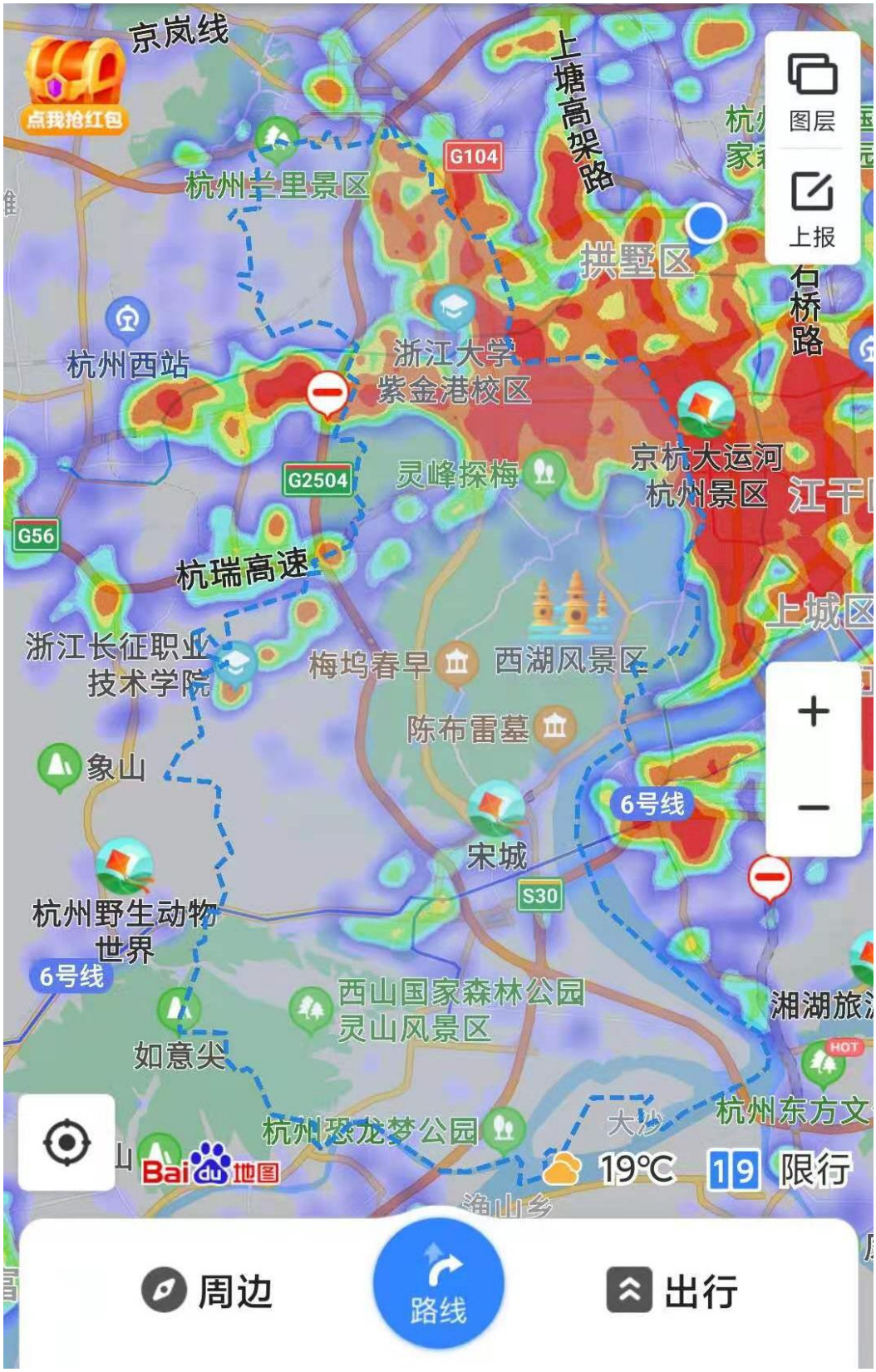
从上图我们看到，2012 年地块活力还属于一般级别（黄色）的部分数据，到了 2017 年却变成了绿色。其实这不是说这些地块的活力降低了，而是这些地块这些年变化不大，而同区的其他地块发展的太快。放在同一尺度下来看，就好像是这些地块的活力降低了一样。

4.4 总体评价

经过对历年的地块活力图进行观察，发现西湖区活力最高的地块集中在如下图所示的区域：



虽然偶有变化，但上图中标出的区域基本上在历年都处于比较高的活力水平。而这个分析结果，跟百度热力图基本吻合，说明这种评价方式还是可行的。

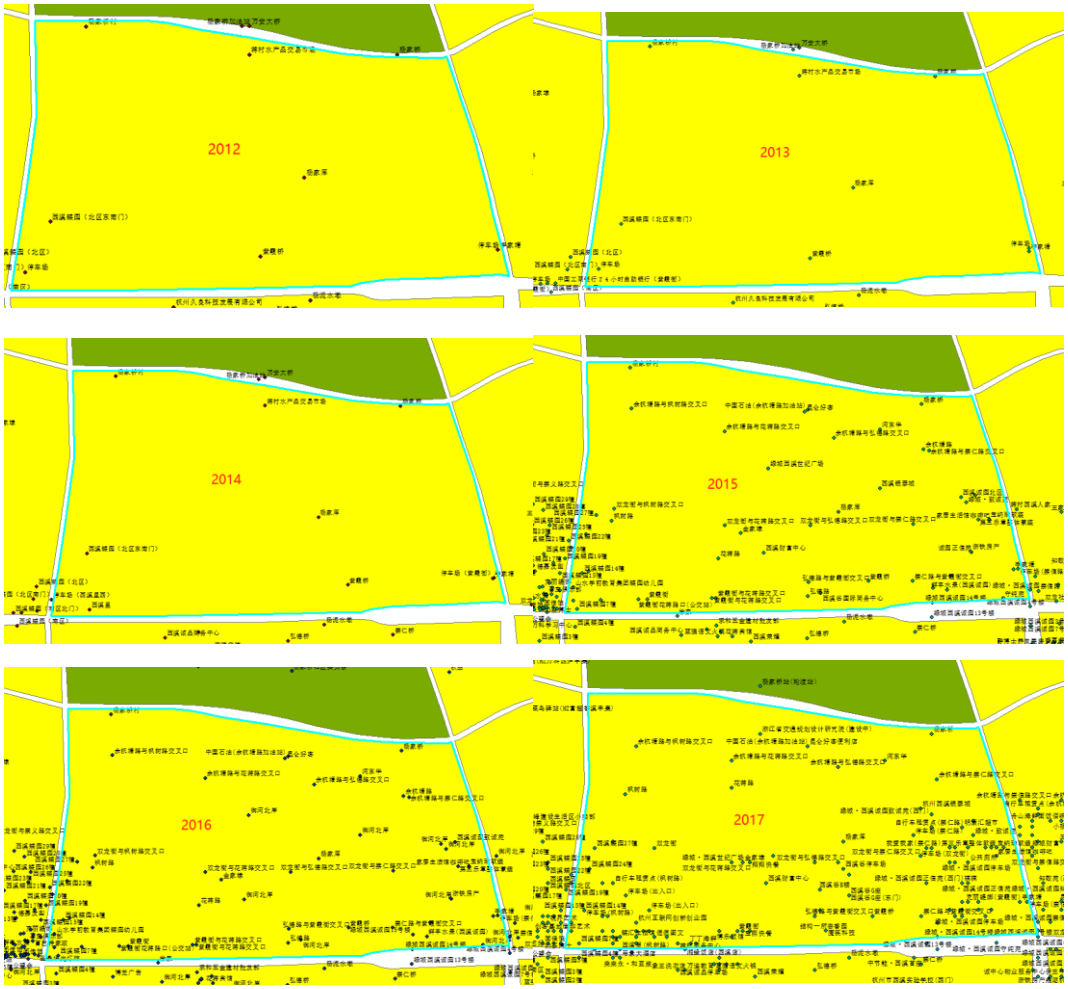

京岚线
点我抢红包
上塘高架路
图层
上报
杭州兰里景区
G104
拱墅区
杭州西站
浙江大学
紫金港校区
京杭大运河
杭州景区
灵峰探梅
G2504
G56
杭瑞高速
浙江长征职业
技术学院
梅坞春早
西湖风景区
陈布雷墓
象山
6号线
宋城
S30
杭州野生动物
世界
6号线
西山国家森林公园
灵山风景区
如意尖
杭州恐龙梦公园
19℃
19
限行
周边
路线
出行

4.5 分地块评价

看完整体，再来分地块看一下。

本次选取如下图所示的这个地块作为例子

再来查看一下该地块历年 POI 的情况



可以看到，从 2015 年开始，该地块的 POI 数量开始显著增加。那么在 2015 年发生了什么呢？从互联网上的公开资料中，我找到了如下图所示的资料：

杭州日报 HANGZHOU DAILY | 都市快报 | 每日商报 | 都市周报 | 城报 METRO | 萧山日报 | 余杭晨报 | 富阳日报

搜索 高级搜索 版面概览 | 版面导航 | 标题导航 返回主页

上一篇 下一篇 字体：放大 缩小 默认

西溪银泰城2017年建设完工

2015-05-15

简讯

近日，“杭州，进入西溪时间”西溪板块发展论坛暨西溪银泰城价值研讨会成功举行。作为银泰置地今年最重要一子，西溪银泰城落户西溪，正是看中了区域内优越的景观资源、交通配套、潜在人气，尤其是未来发展利好。

总建筑面积约30万方的西溪银泰城，涵盖云概念购物中心、4座国际办公楼写字楼以及约3万方的国际风情步行街。未来将整合银泰商业资源，引入国际一线品牌主力店、大型精品生活超市、巨幕影院、银泰百货等。与以往不同的是，西溪银泰城将创新采用互联网思维进行规划，以O2O模式重新思考传统经济，打造“网络+实体经济”一体化的全新商业模式。西溪银泰城将于2017年全面建设完工。（项晶晶）

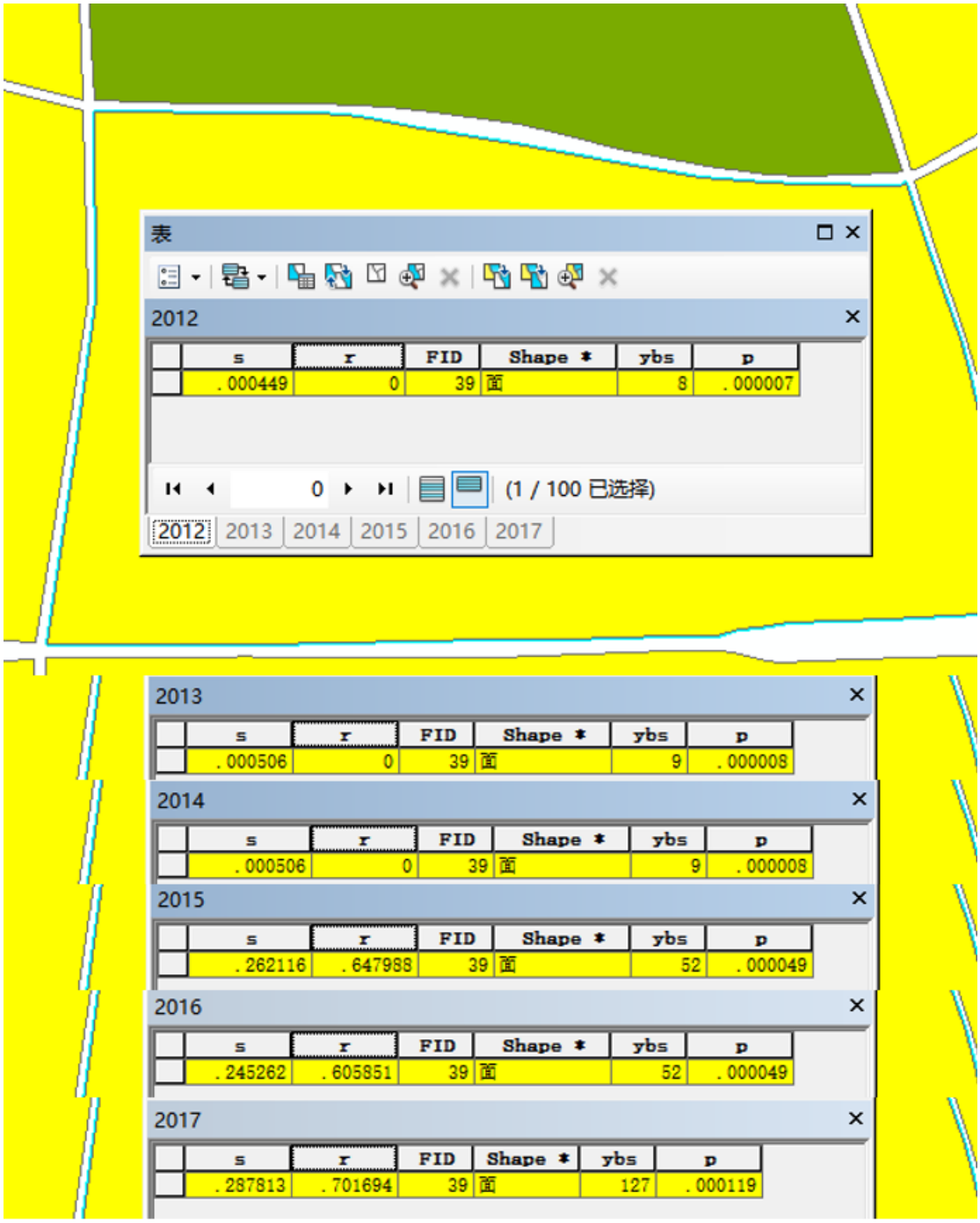
交通晓时代样板房亮相

交通晓时代72变样板房最近亮相，主题营销中心开放，活动当天500多人到场。项目由拥有11年高端品质物业营造经验的国企杭州交通置业有限公司开发建设，正对103万平方米七堡站地铁上盖综合体，自带电影主题商业街；区域周边中小学校聚集。户型是交通·晓时代的最大看点，59平方米稀缺两房、82平方米创新三房、115平方米全能四房，都以超高性价比成就主城置业梦想。（陈夏）

随园嘉树7月推特色养老机构

母亲节当天，万科·随园嘉树随园会举办了一场精彩的花样母亲节专题活动。包括业主、媒体以及随园嘉树的意向客户等共约400位嘉宾参与，现场欢声笑语、熙熙融融。随园嘉树作为万科致力于中国养老地产的高端范本，力求为长者带来美好的生活体验。还将在今年7月推出万科首个以护理服务为特色的养老机构，打造独立的护理院，独创24小时管家式服务。（刘志访）

2015 年，该地块迎来了以西溪银泰为代表的一批商业、办公、住宅等类型的 POI（一般 POI（商业/地产）在项目开始的时候就会出现在互联网地图上，并不是完工了才会出现）。这些 POI 的到来，反应在地块活力指数上如下图所示：



我们可以从上图看到，从 2015 年开始，其地块活力值发生了数量级的增长，并且总体趋势一致在上升。这与现实情况也是一致的，现在该地块已经变成了集高端住宅、商业、娱乐为一体的高活力地块。

5 总结

5.1 研究存在的不足

由于作者水平及精力限制，本次研究还存在很多的不足。总结起来主要有以下几条：

1、 缺少历年路网及 POI 数据。研究中使用的路网数据是最新的；POI 数据也只有 2012 年至 2017 年，2017 年至今的部分，是缺少的；

2、 在计算地块混合度的时候，其实每种类型的 POI 对地块的混合度影响并不相同。这个很好理解：一所医院跟一个小吃店相比，显然医院对地块的混合度影响更大，但在本次的分析中并没有考虑权重；

3、 用 POI 总量与面积指数结合地块混合度（熵值）来评价地块活力，还处于探索阶段，并没有得到充分的验证；

4、 在评价阶段验证的时候使用了现在的百度热力图，来验证 2017 年的评价结果，在逻辑上不是很严谨。

以上是列出的几条本次研究存在明显不足的地方，由于作者水平限制，本文难免会存在其他不足，敬请各位读者斧正。

5.2 研究的意义

本文从收集数据，到处理分析数据再到得出结论，针对地块活力的量化给出了一套完整比较完整的可行的方案。但各位读者也能看到，

评价本身，受制于数据、模型的限制，还存在一些不足，还有很大的改进空间。虽有诸多的不足，但这个评价的结果跟现实还是比较吻合的，这又说明这种方式是可行的。通过本文的探索，希望能给各位带来一些启发。

本文采用的主要数据处理工具是 FME，通过这次研究，充分体现出了 FME 强大又灵活的能力，在数据处理的各个阶段，都能使用 FME 来进行数据的处理与分析。做数据分析用 FME，真香！