基础知识：2019.12.16–2020.

一. MOOC–大数据与城市规划(2019秋)

学校	清华大学
老师	龙瀛（清华大学 建筑学院）
体验频率	1小时/天
体验行为	看视频+笔记+kaggle实战
体验中的评估和反馈	* 线上的问答题 * 线下和本校建筑规划设计学院同学参加比赛和交流
可能的认知偏见	* 行业术语理解 *

1.1 课程概论

课程大纲

以这样的研究方向为指导，整个课程可以分为以下几个章节：

章节	大致内容
概述篇（第1–3章）	
技术篇（第4–8章）	
数据篇（第9–11章）	
应用篇（第12–14章）	
展望篇（第15章）	

预计的收获

- 课程老师的预计

☐ 数据：提供的案例地区一整套的城市空间数据集
☐ 方法：基本的数据抓取、分析和可视化
☐ 思维：利用新数据、新技术认识城市和规划设计城市

- **我的预计（实时拓展）**

☐ 数据：这套的城市空间数据集涉及的城市维度和获取难易的评估
☐ 方法：数据抓取、分析和可视化在这个场景中可能遇到的问题和解决方案，现有科研成果或产品的解决方案
☐ 思维：大数据或人工智能算法在城市规划上的科研和**商业路径**、战略思维

1.2 概述篇（第1–3章）

概念

城市

- 城市规划学科同定义（经济学、地理学、社会学等等），城市规划学科中的思考维度：
  - 行政领域，市辖区/市区 = 市域–县
  - 实地领域，城镇化用地
  - 功能领域，与实际功能上的关联，比如人口、就业

城市群：比如长三角、珠三角
城市变化

- 全天候在线化
  - 传感设备的应用和普及–多元的线上线下数据

研究机构与项目	项目内容
芝加哥城市运算和数据中心“物联城市”（Array of Things）	通过搭建城市传感器网络，为居民、城市管理者和科学家提供认识、分析和改造城市的数据基础。
哥伦比亚大学 智慧城市研究中心	构建了智慧社区的传感器系统，提出营造社区安全性的方案。
麻省理工学院（MIT）市民数据设计实验室	通过开发新型传感器，测量公共空间中的人群感知与行为。
清华同衡技术创新中心团队“CITYGRID城市数据传感器”	CITYGRID支援的测量指数多元、精细，可结合路灯、站牌等处悬挂安装，从街道层面获取实时且准确的人口、交通、或是环境状况等流动数据，以支持更永续、更智慧的城市决策。

- 小型化
  - 室内外公共空间的新元素（路边KTV、录音亭、自动贩卖机、自主按摩椅等）
- 居家化
  - 人流、物流流向的该变（比如外卖）
  - 居住空间的混合使用（上门修手机、美甲等）
- 个性化
  - 以体验为目的的小众需求
- 智能化
  - 取代高危险性、重复性的工作（比如无人商店，阿里的广告设计机器人“鲁班”）
- 算法化
  - 算法与人们的行为选择和背后的城市空间
- 共享化
  - “闲置”的重新利用，服务半径的改变
- 连锁化
  - 大者恒大，强者恒强，单体经营到并入商业综合体
- 自然化
  - 自然环境的保留（城市绿廊，慢行系统，绿地公园，登山路道，农家乐等）

城市数据

- 传统城市数据
  - 特点：受行政区域限制，主要基于空间属性
  - 包括六类：遥感测绘数据，统计数据（统计年鉴数据），调查数据（普查），知识数据，规划成果数据（总规控规专项规划），业务数据（规划院行政管理。委办局等）
  - 主要问题：获取成本，及时性，精确度，数据质量（类别，可否验证，定量OR定性）
- 新数据
  - 特点：精度高，覆盖面广，更新快
  - 分类
    - 数据来源：政府数据（信息公布平台），开放组织数据（百度地图开放平台，open street map，sightsmap），企业数据（需要一定的相关性分析），社交数据（筛选使用的工作量大），智慧设施数据（传感器等）
    - 数据环境：建成环境数据（多维度多尺寸），行为活动数据（人类电子足迹）
    - 其他：数据时空分辨率（时间–空间四象限），数据几何形态（点线面），空间关联（位置数据，联系数据），动静状态
- 典型城市数据
  - 手机信令数据
    - 用户与发射基站、微站之间的联系数据
    - 始终带有时间、位置和话单等信息

        - 空间分辨率：多为基站，时间分辨率：精确到秒
    - 全球定位系统数据（GPS）
    - 点评及签到数据（大众点评，微博签到）
    - POI数据（兴趣点）
    - 公交智能卡刷卡数据
        - 空间分辨率：站点，时间分辨率：精确到秒
        - 特点：连续性好，信息全面，动态更新
    - 地图数据
    - 住房数据（搜房网，安居客）
    - 夜光影像数据
        - 卫星数据
        - 评估社会经济收入、城市化和光污染等
    - 位置数据（Flicker照片）
    - 街景数据
    - 腾讯宜出行数据（关注公众号，只能移动端 查询）
    - 人口热力图（百度）
    - 网上消费数据
    - 智慧足迹（中国联通，smartsteps）
    - 谷歌地球引擎（影像，地球，气候及水文，人口数据）

1.3 技术篇（第4–8章）

概念

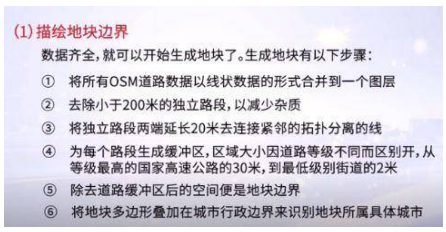
- 空间分析（侧重于空间数据的整合）
    - 数据维度的选择

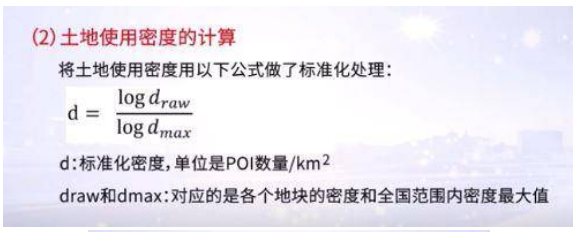
	对应的数据维度	分析方法
点	兴趣点（POI）	核密度，网格聚合
线	街道	路段预处理、评价指标、分类
面	地块	识别地块类型

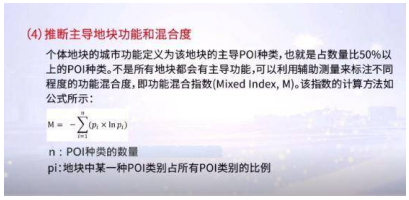
    - POI数据：ArcMap
    - 街道数据
        - 路段预处理（合并道路–ArctoolBox, Merge，细化道路–ArctoolBox, Thin，拓扑处理）
        - 量化指标

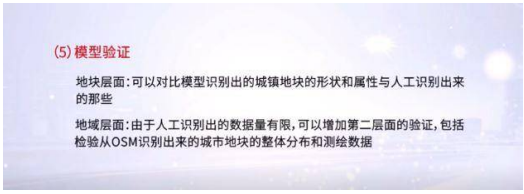
	数据类型	可采用的数据
外在特征	人口密度	人口普查数据
		手机信令
		基于位置的服务（LBS，互联网公司提供）
	城市活力	经济活力（经济普查、居民出行调查中的居民家庭调查、大众点评）
		社会活力（位置微博、街景、大众点评）
自身特征	城市功能	功能密度，多样性和中心性（兴趣点、用地现状图）
	物理特征	街道长度，地面铺装，是否非隔离、行道植被（街景）
	界面特征	建筑立面连续度，橱窗比（建筑、街景）
	交通特征	道路等级，限速，车流量（居民出行调查，出租车轨迹和城市基础地理信息系统GIS）
环境特征	区位特征（城市基础地理信息系统GIS）	所处功能分区
		是否处于城镇建设用地内
		与城市中心，城市次中心，商业综合体的距离
	城市设计	周边街坊机理（街道交叉口，用地现状图·）
	开发强度（建筑）	
	可达性	地铁站、公交站点与线路数量（城市基础地理信息系统GIS）
	控制变量（？）	

    - 街道分类
        - 标准：不同时段的人类活动，功能密度等级，功能多样性等级，周边城市设计情况，步行性等级
        - 根据街道周边最大地块的占比定义（50%以上和0–50%）
- 地块数据
    - 较难获得
    - 地块的自动识别和分类（AICP）
        - 数据需求：中国城市行政边界，中国的开放截图（OSM），POI数据（可从新浪微博整合），夜间灯光数据
        - 操作（1–5）

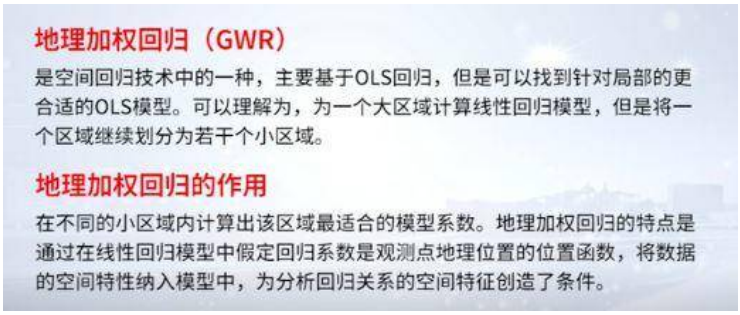






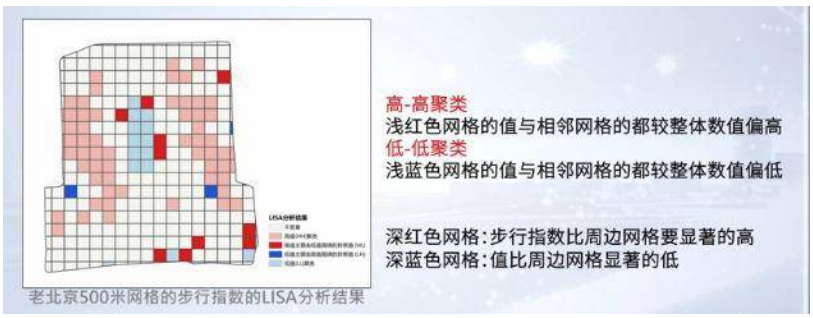


- 统计分析（侧重于深层次的规律挖掘）
  - 空间地理加权回归
    - 概念

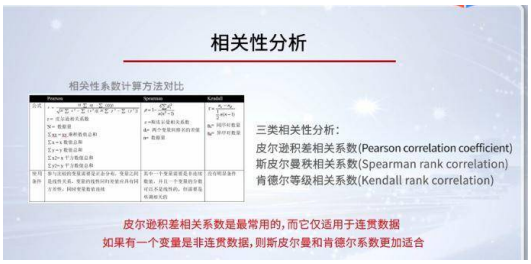


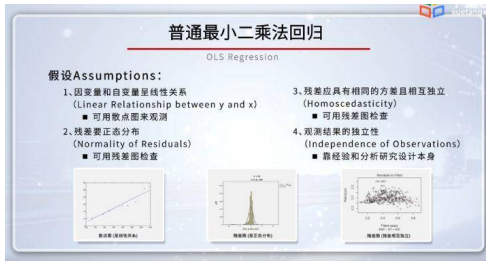
    - 前置条件：线性回归和残差分析以保证基本假设的合理性，300以上的样本量
    - 每个网格对应一套模型系数
  - 空间自相关（SA）
    - 概念

    - 全局莫兰指数绝对值越大，自相关越强烈（正-正相关，样本数值越接近，更加聚合；反之负相关，样本数值相差越大）；随机排列后的可靠性指标Z-score越大，零假设越可以被推翻
    - 一般使用曼哈顿距离
  - 聚类和异常值分析（LISA）



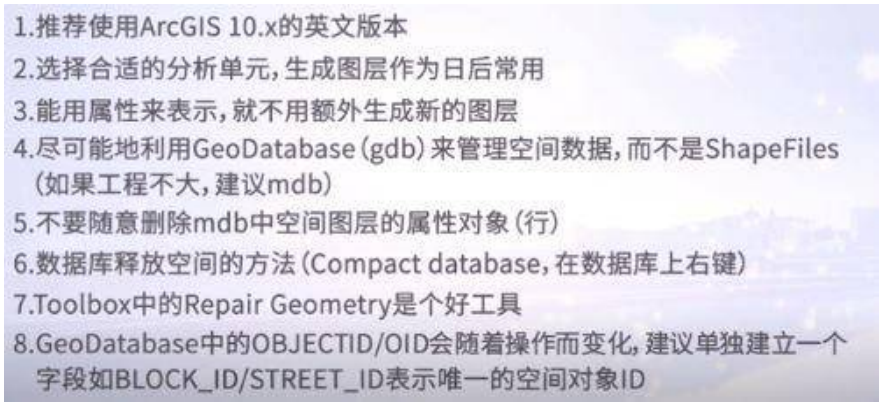
  - 线性回归，R2相似情况下，参数越少越好，变量显著值越接近0越好
  - 一些数学





工具+案例

- 工具
  - ArcGIS、ArcMap、ArcTool Box



  - SPSS
- 案例
  - 城市空间大数据获取
    - 问卷：问卷星，microsoft form
    - 定制/购买：猪八戒，idata api
    - 免费渠道：BCL，国匠城，Geohey，国家地球系统科学数据中心
  - 结构化网页数据获取
    - 路径

  - API数据获取
    - 接口查找：？
  - 抓包工具获取数据
    - 手机：fiddler
    - 目的是获取源数据的url，必要时对字段解码
  - 影像数据获取
    - LSV（LocaSpace Viewer）
    - 遥感集市云平台(?)
    - 地理空间数据云
    - Google earth engine（全球尺度，长时间序列）

        - United States Geological Survey
    - 数据清理
        - 坐标系转换问题

高德地图API、腾讯地图API	GCJ-02坐标（火星坐标）
百度API	BD-09坐标
搜狗API	搜狗坐标
Google Earth	GPS坐标

经纬度位置，单位换算，最大放大倍数后选点
工具：Geosharp，BCL坐标系转换软件，API接口调用

        - 正/逆地理编码

工具：Geosharp，API接口调用

三.