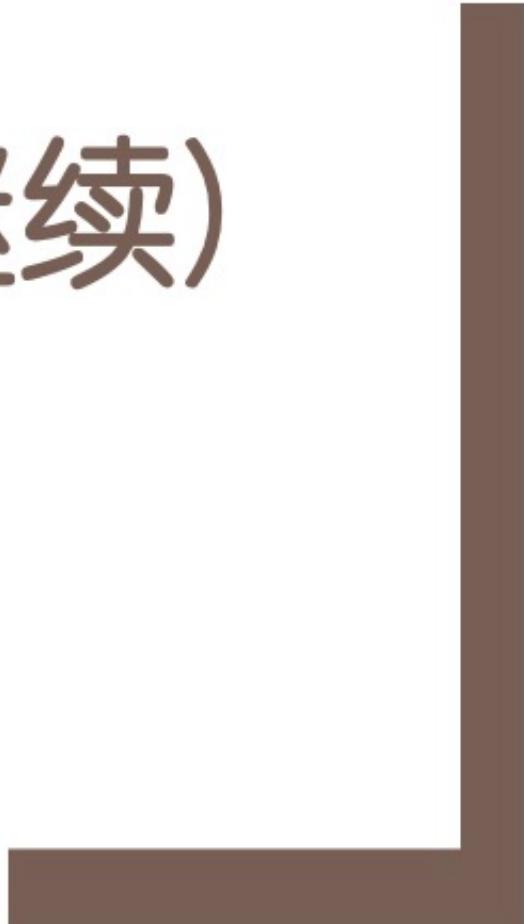


空间分析（继续）



表面分析：空间插值

空间插值 (spatial interpolation)：根据已知采样点的数值，预测或估算未采样区域的数值并创建出连续表面模型。

常用的空间插值方法

不规则三
角网

反距离权
重法

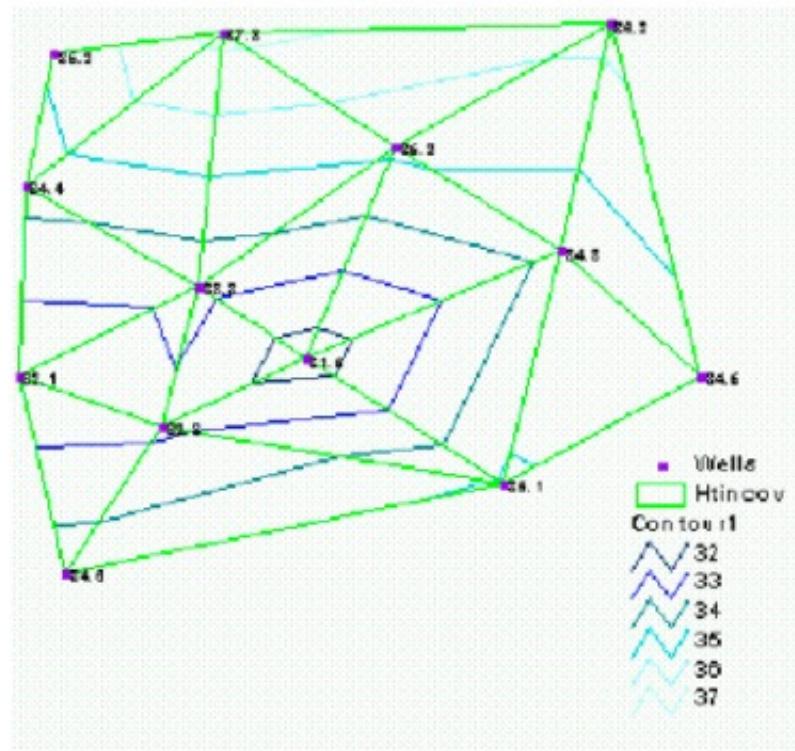
克里金法

趋势法

最小曲率
法

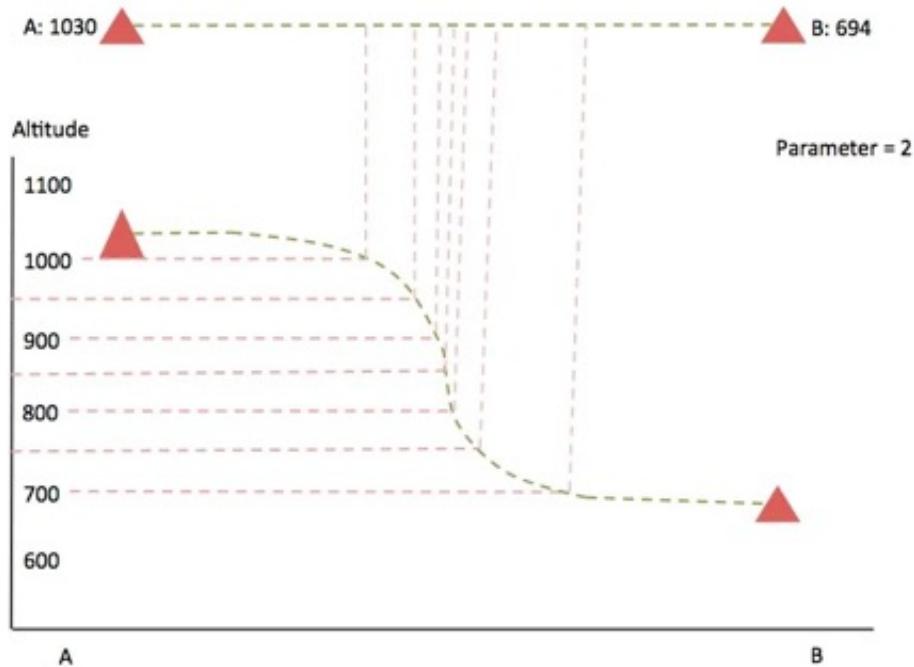
表面分析：空间插值

三角测量法（不规则三角网 Triangulated Irregular Network, TIN）：线性插值的方法，假设数据在空间内的变化呈线性趋势，通过在两个已知点之间创建线性模型来预测未知点的数值。



表面分析：空间插值

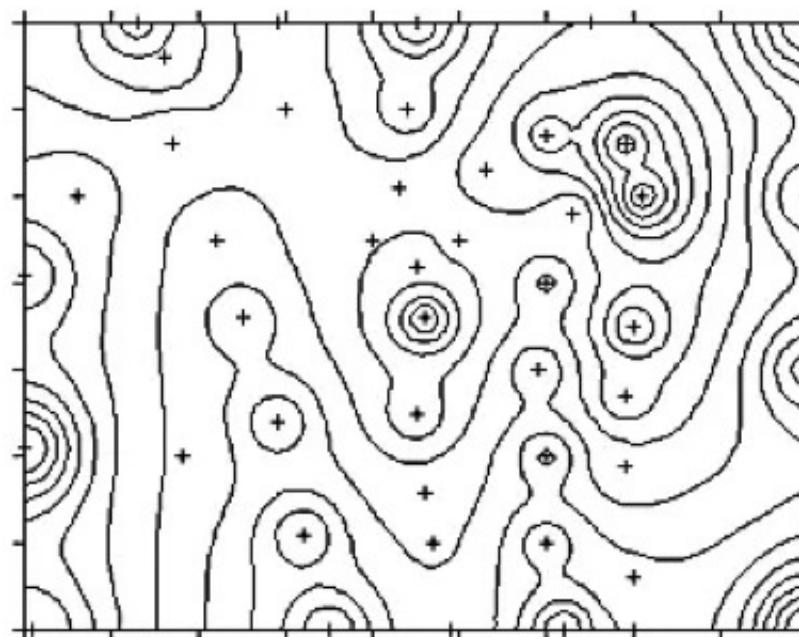
反距离权重法 (Inverse Distance Weighting - IDW)：已知点位的数值对其周围一定范围内施加了较大的影响（“近朱者赤、近墨者黑”）。符合Tobler的**地理学第一定律**：所有事物都是相关的，但离得近的事物之间的关联比较大。



Waldo Tobler

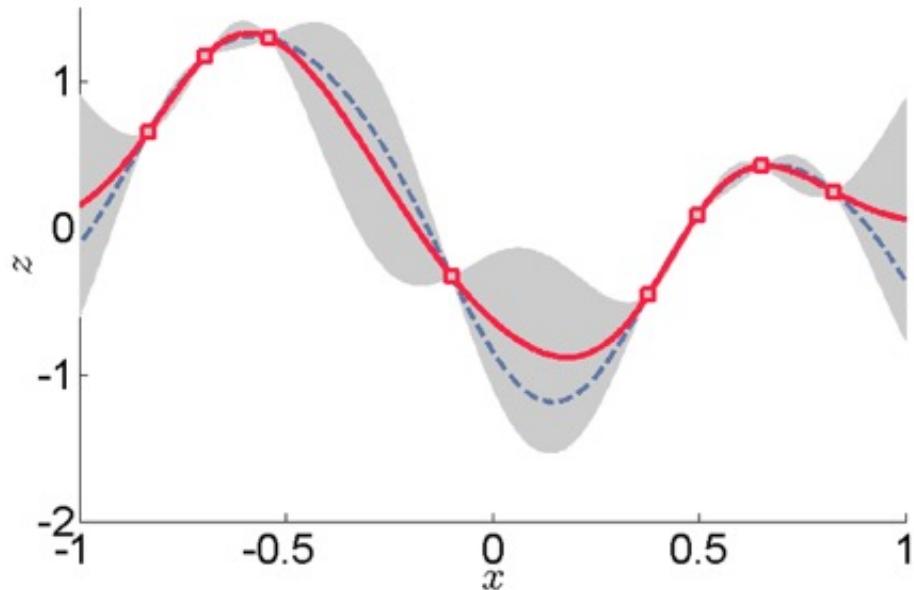
表面分析：空间插值

- 反距离权重法一般在数据点集中且均匀分布时使用。
- 特点：插值结果的等值线图中易出现“牛眼”形状。



表面分析：空间插值

克里金法（Kriging）：高斯过程主导的插值法，考虑指定范围内所有已知数据点的协方差，数学较复杂但误差相对较小。



克里金法的一维插值示意图

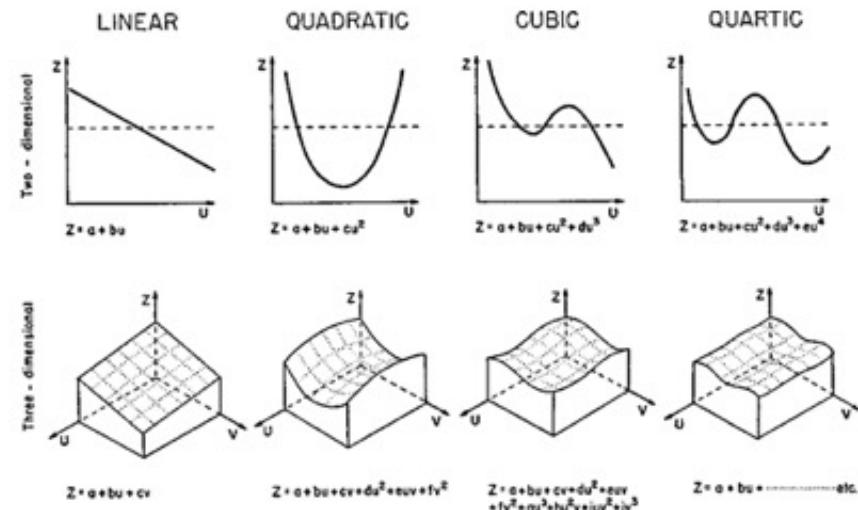
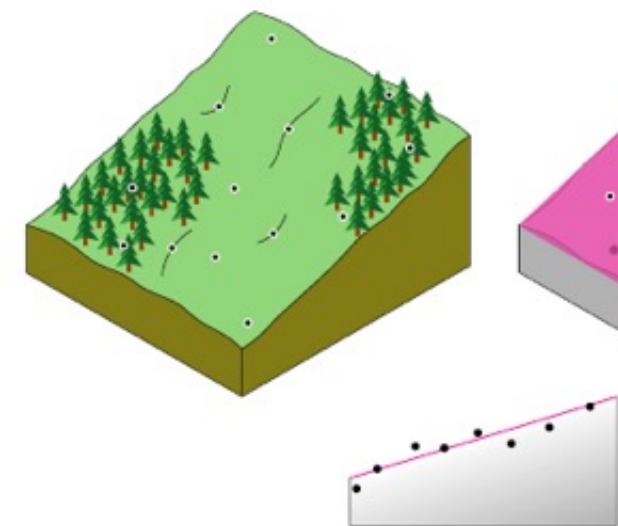


Danie Krige
南非金山大学教授

表面分析：空间插值

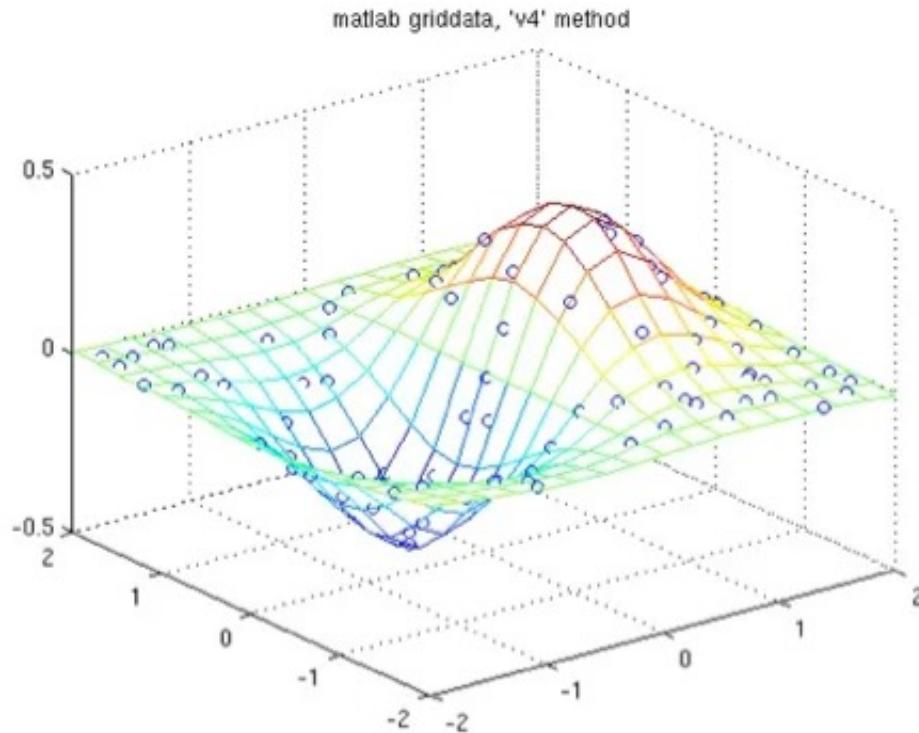
趋势法（trend）：尝试用较简单的表面去大致拟合已知数据点，**不要求**所有的已知数据点都被囊括在表面里，但要让所有已知数据点到表面距离的**均方根**达到最小值。

- 均方根：每个点到表面的距离先平方，再取平均数，再开方。
- 通常适用于探索性数据分析。



表面分析：空间插值

最小曲率法（minimum curvature）：要求表面经过所有的已知数据点，并让表面的总曲率达到最小（用有弹性的『橡胶膜』去拟合所有数据点）。



地统计学

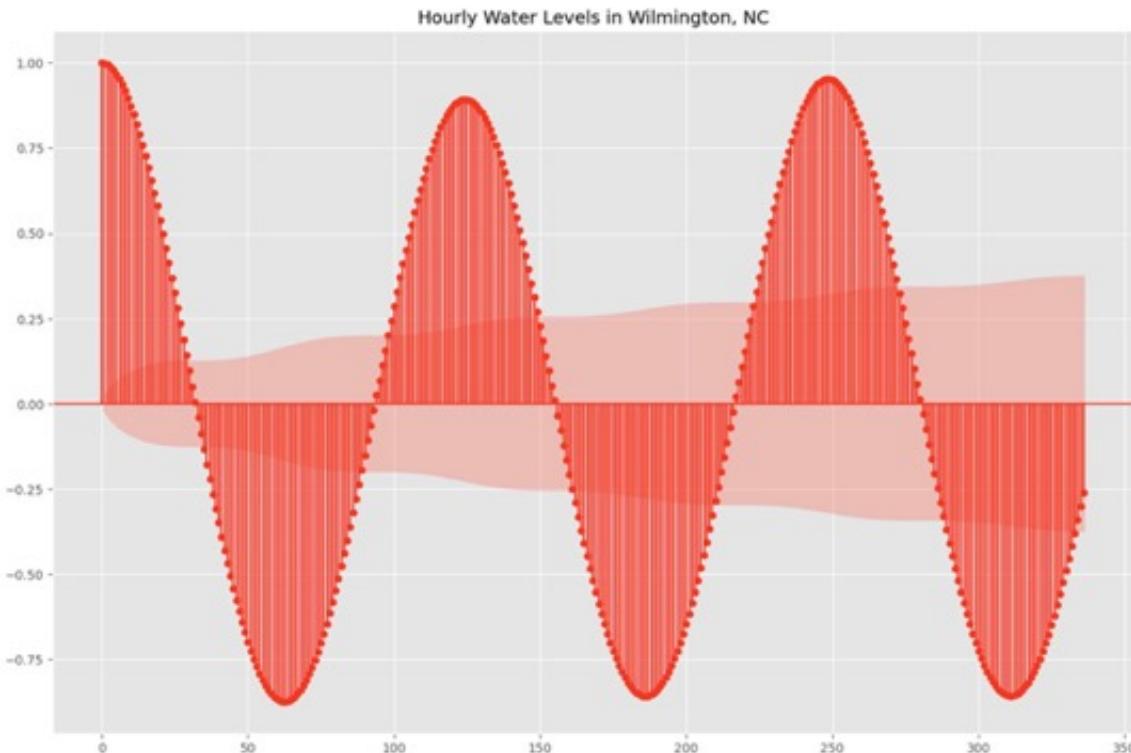
空间统计/地统计学 (geostatistics) : 借助统计学方法理解和量化空间模式。

常用工具:

- 空间自相关 (spatial autocorrelation)
- 热点分析 (hot spot analysis)
- 地理加权回归 (geographically weighted regression – GWR)

空间自相关

自相关 (autocorrelation)：描述一个序列中各元素之间的关联性。在时间序列分析中，自相关被用来衡量当前时刻的数值与其过去时刻数值之间的关系，从而得知**时间序列**中的信号是否有**周期性**或出现**重复模式**。

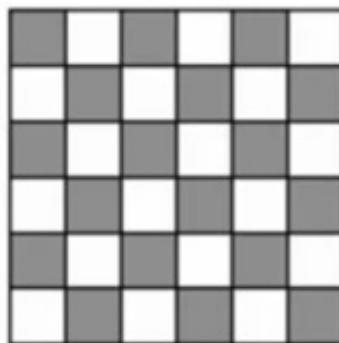


空间自相关

空间自相关：地理空间内相邻位置上特征值或属性值之间的相互关联性。

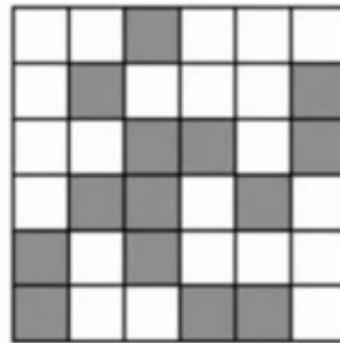
- 地理特征的分布模式：分散式（dispersed，负相关）、随机式（random，不相关）、聚团式（clustered，正相关）

Dispersed



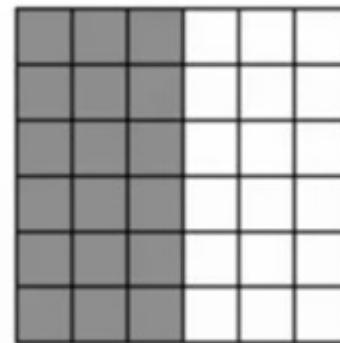
Negative spatial
autocorrelation

Random



No spatial
autocorrelation

Clustered

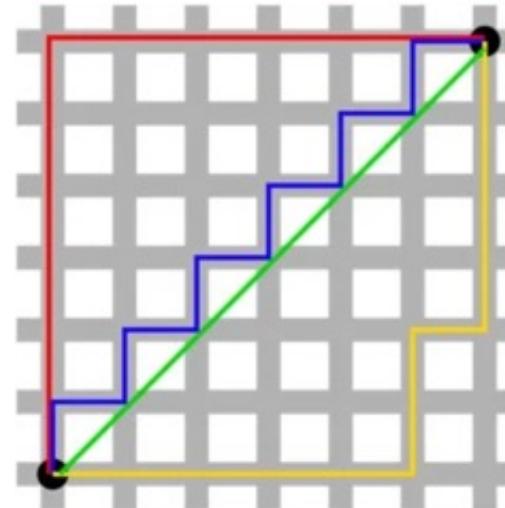
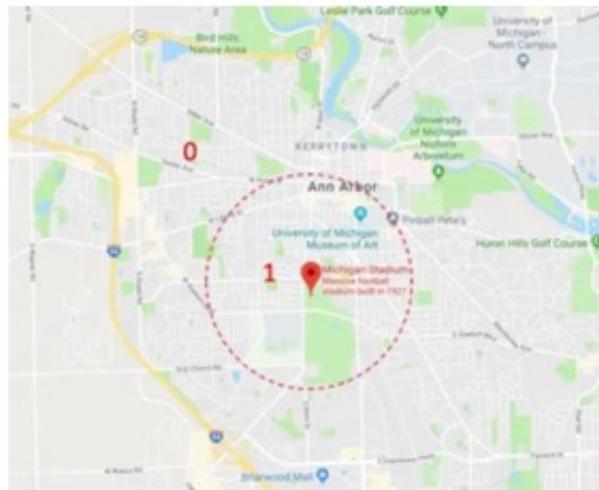


Positive spatial
autocorrelation

空间自相关

邻域（neighbors）的定义：点状数据可采用自然距离。在某个地理要素周围，位于某个**距离阈值**以内的其他要素为**邻域**，**阈值以外的要素为非邻域**。

- 自然距离：欧几里得距离（根据坐标计算所得） vs 曼哈顿距离（两点在**正东西**和**正南北**方向上的距离之和）

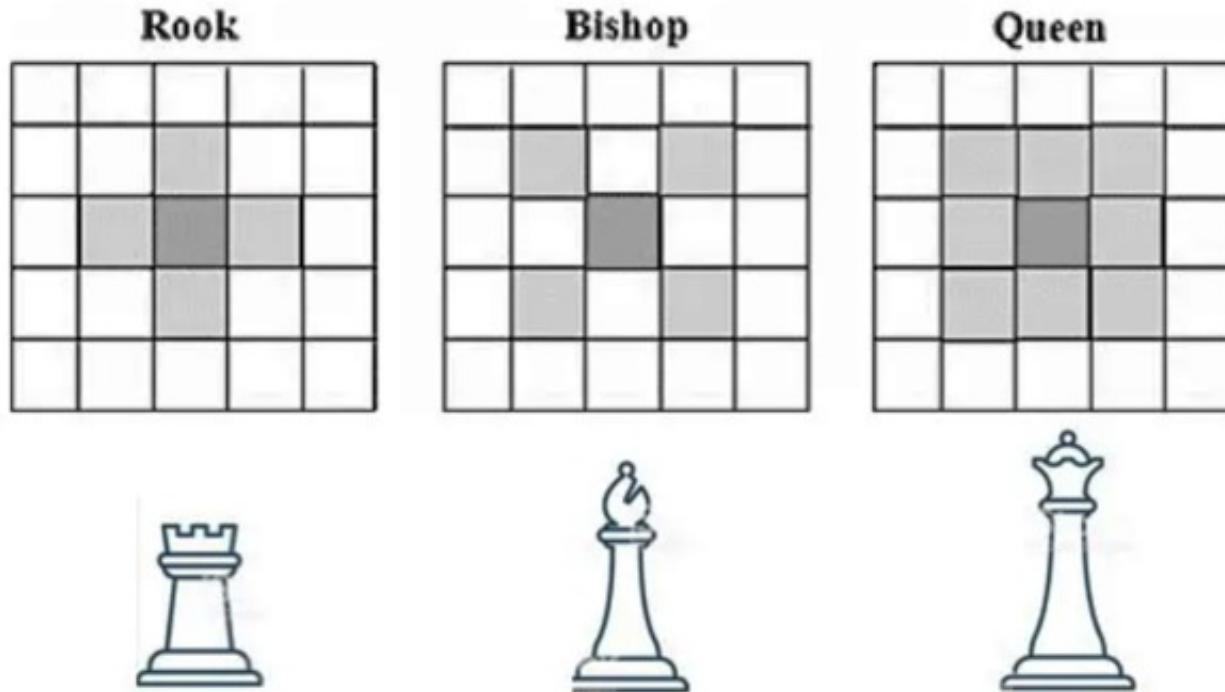


欧几里得距离：绿
曼哈顿距离：红黄蓝



空间自相关

多边形的邻域：空间邻接矩阵（spatial contiguity matrix）



Rook: 共享边的多边形为邻域

Bishop: 不共享边但共享端点的多边形为邻域

Queen: 共享边或端点的多边形为邻域

空间自相关

空间自相关参数

- Moran's I: 相邻数据点和平均值之间关系的比较
- Geary's C: 相邻数据点自身数据之间的比较

对比	Moran's I	Geary's C
考察内容	数据值在空间上的聚类程度	相邻数据点之间的差异
参照基准	所有数据点的均值	
取值范围	[-1,1]	[0,2]
缺点	必须知道均值	不考虑样本量和空间权重，稳健性较低
重要取值	聚团式分布	1
	随机分布	0
	分散式分布	-1

举例：

A: 50只 

B: 55只 

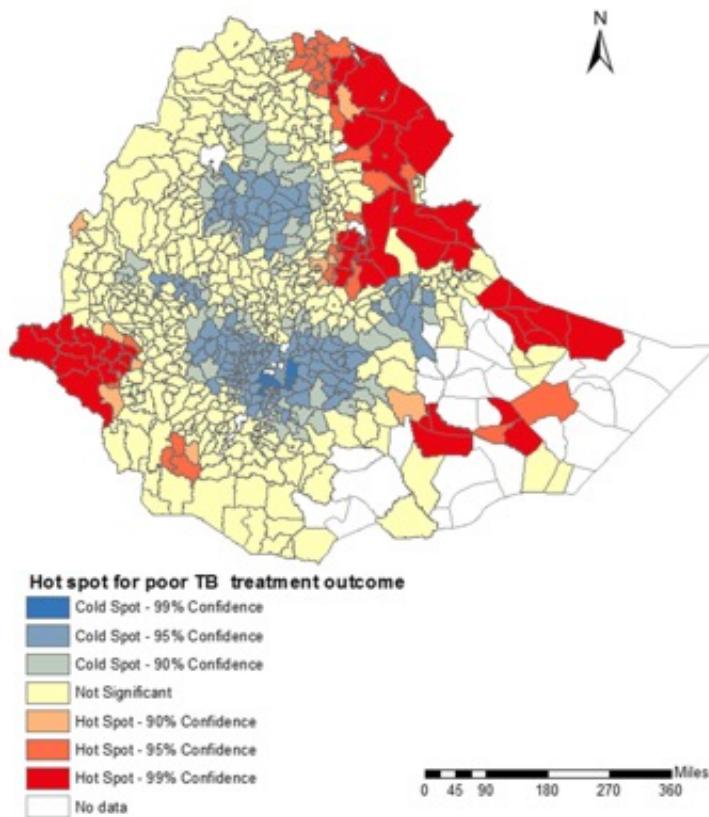
C: 1只 

全岛平均：30

热点分析

Getis-Ord G 检验：发现全局中的高/低值数据分布趋势。

Getis-Ord Gi* 检验：根据邻域发现局部的高/低值数据聚集区，即热点（hot spots）或冷点（cold spots）。

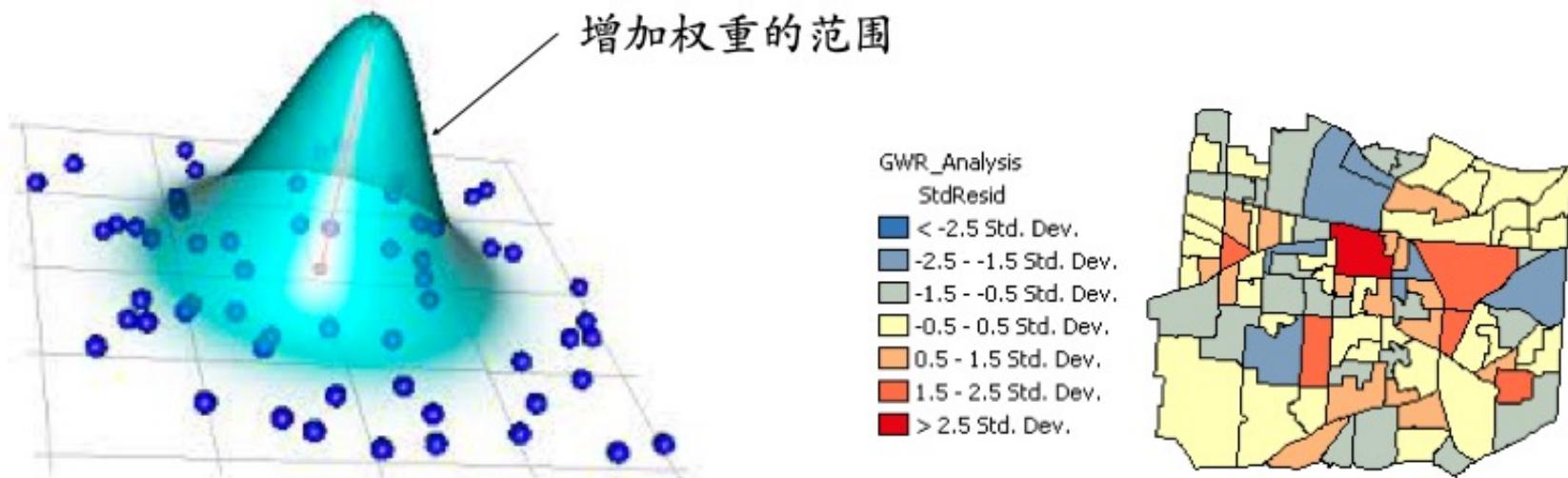


埃塞俄比亚结核病医疗不足程度的热点，越红的地方代表结核病的医疗水平越差。

地理加权回归

地理加权回归（GWR）：允许模型的参数随空间位置而发生改变，从而更好地描述和预测复杂动态的空间趋势。

每一个观察点都会进行一次单独的回归计算，且每次计算都会给邻域更高的权重，以此能从局部估计出各个位置的自变量对于因变量影响力大小及方向。最后得到各数据点的局部残差。



Pop Quiz

回答正确得2分
回答错误得1分
缺席不得分

1. 以下哪种图表常被用于描述类别之间的数值流向关系?

- A. 旭日图 | B. 桑基图 | C. 峰峦图 | D. 树形图

2. 以下哪一项不属于矢量数据格式?

- A. GeoJSON | B. Shapefile | C. GeoTIFF | D. KML

3. 以下哪一项通常采用栅格数据格式?

- A. 遥感影像 | B. 河道水系 | C. 政区边界 | D. 公交站点

4. 以下哪个操作将多个输入图层结合成同一图层?

- A. 交集 | B. 擦除 | C. 溶解 | D. 融合

5. 期末考试在哪一天?

- A. 1月3日 | B. 1月4日 | C. 1月5日 | D. 1月6日



地图设计

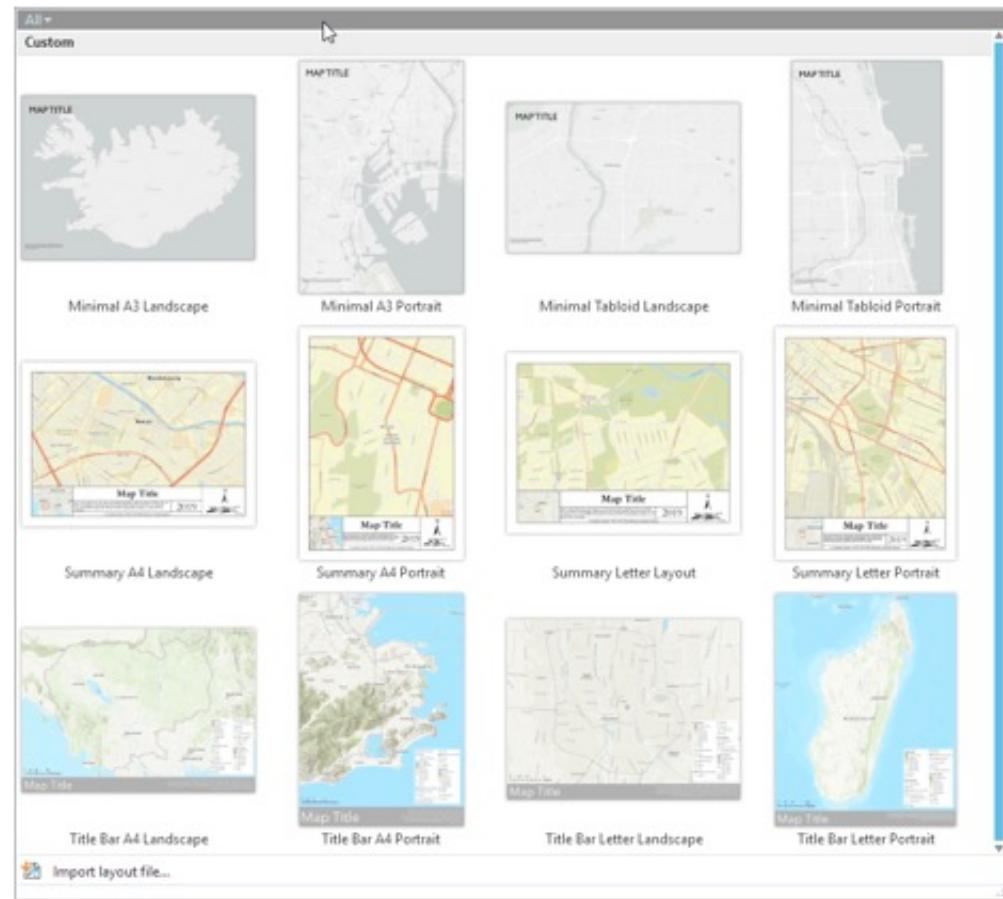


地图设计

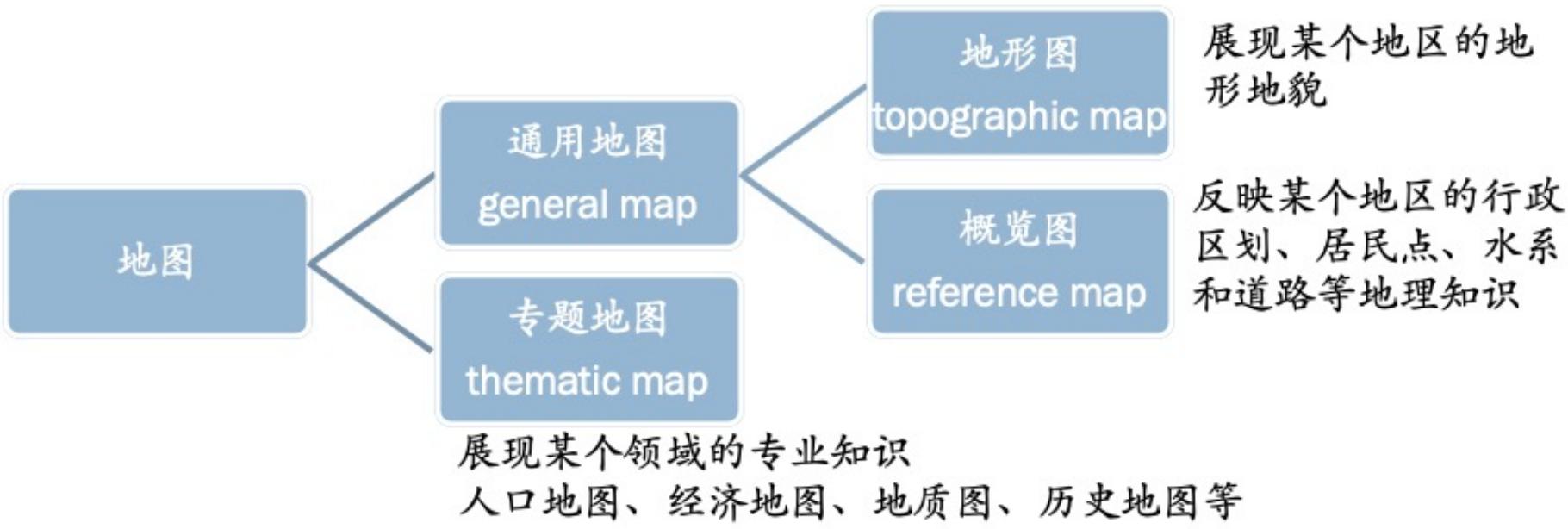
地图既是一种信息载体，也是一种艺术作品。

制作优秀的地图需要注意的内容：

- 用途
- 布局
- 投影
- 底图
- 标注和配色

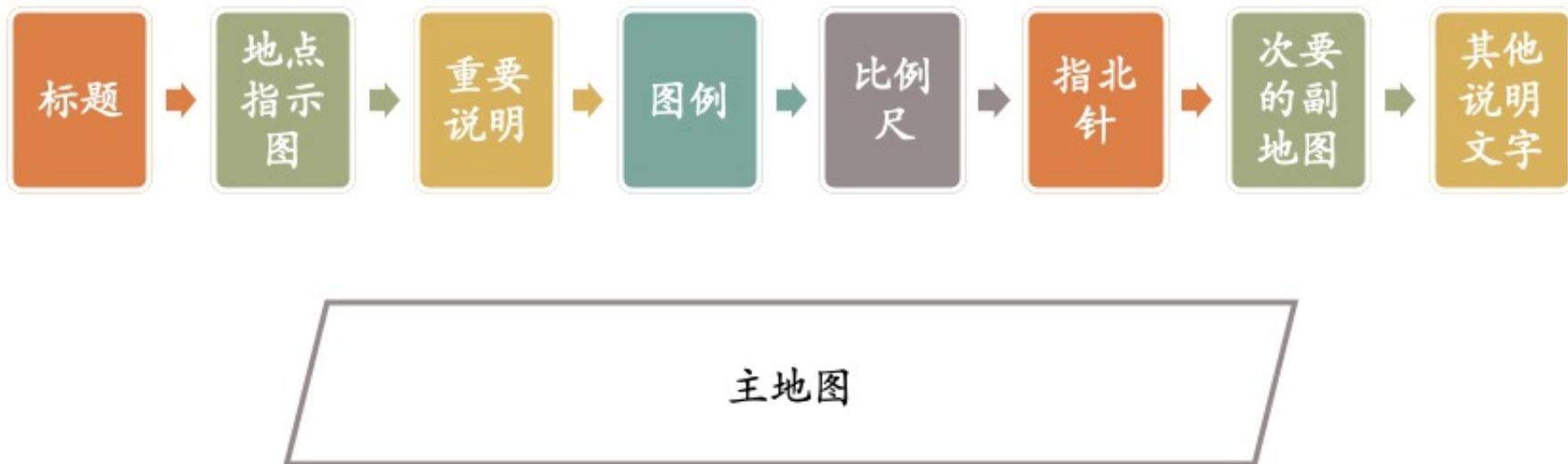


地图设计：用途



地图设计：布局

原则：重要的信息放在视觉突出的位置，不能被遮挡。



地图设计：布局

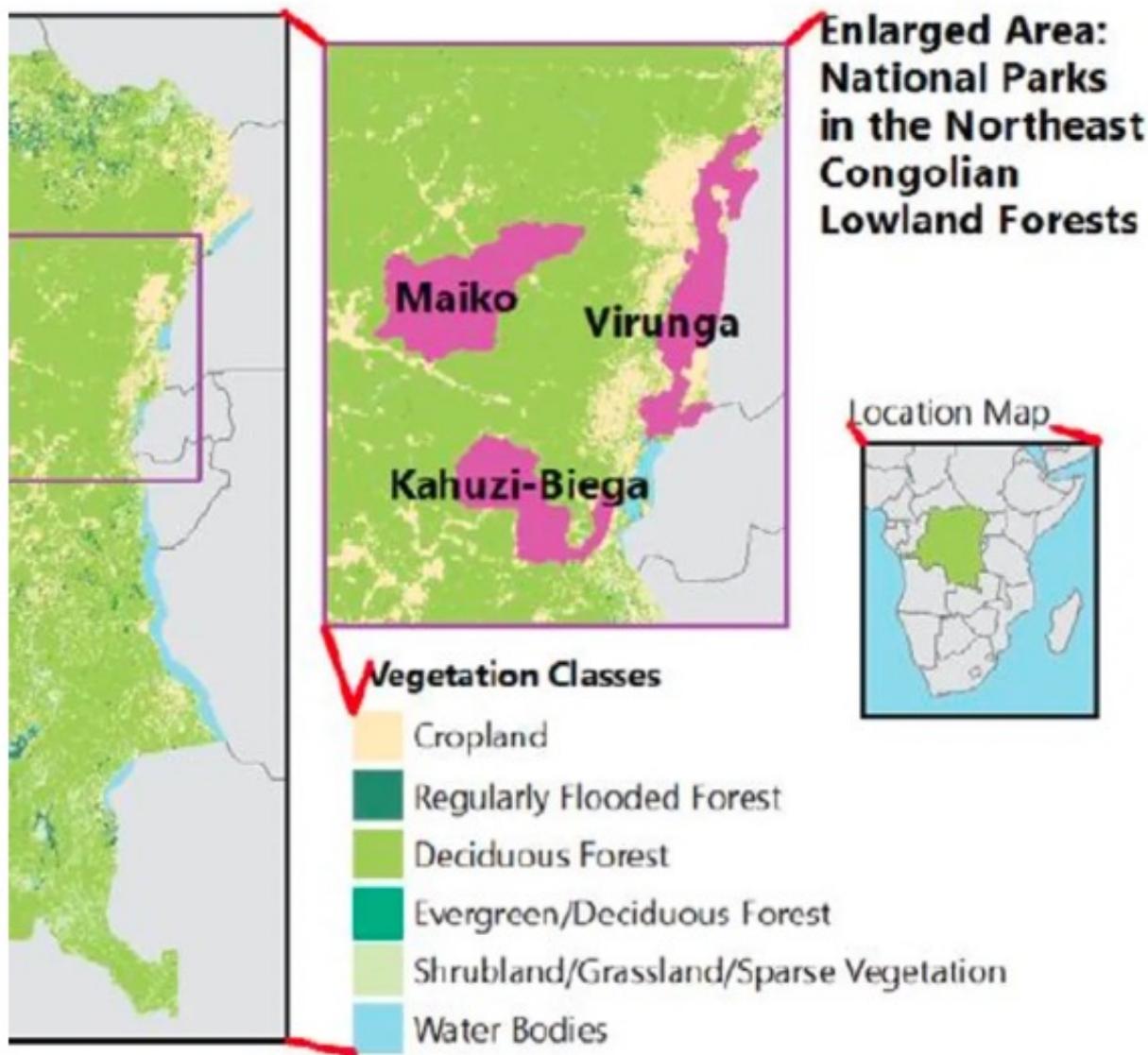
不要刻意用地图元素填满空白区域。空白区域尽量做到平衡。



Transportation and Land Use
Prince George's County, Maryland

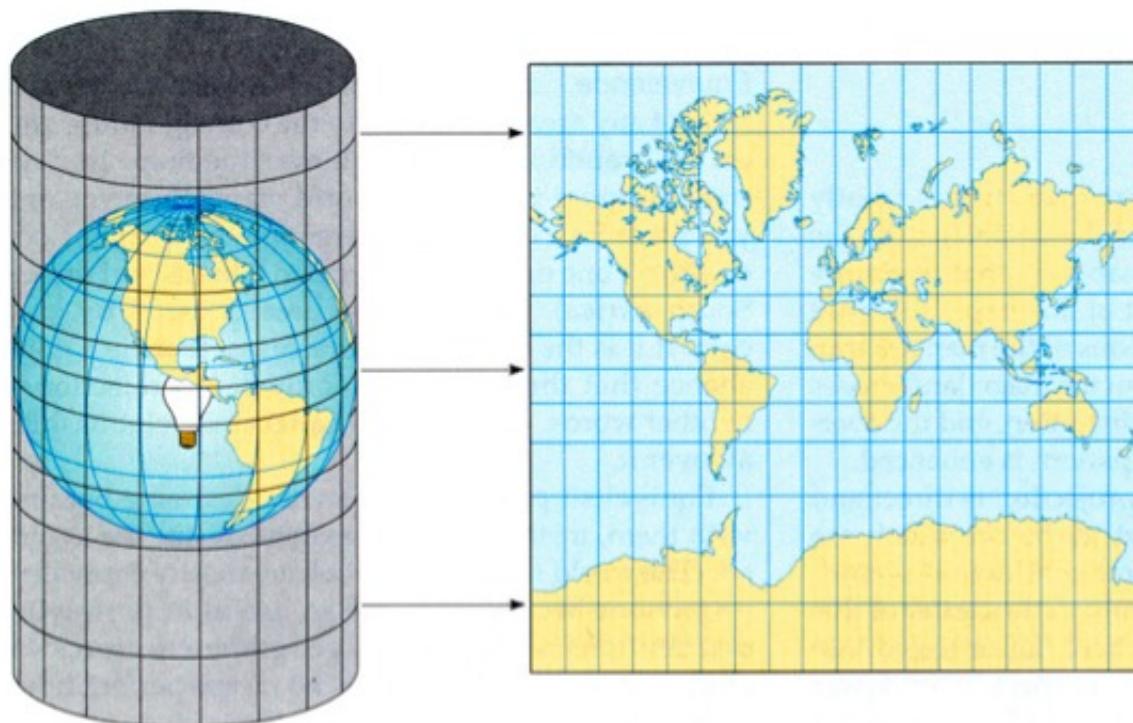


地图设计：布局



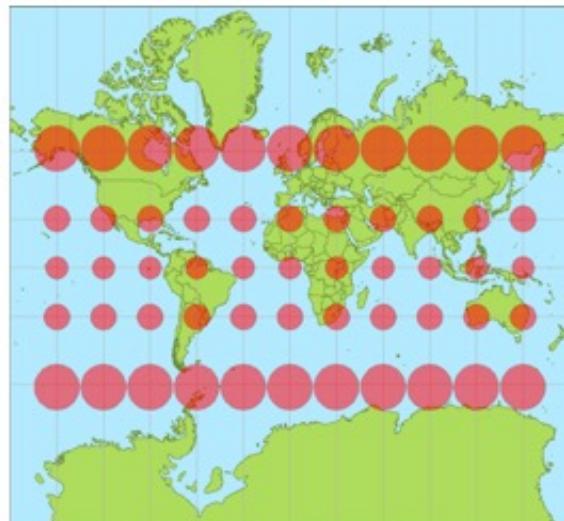
地图设计：投影

地图投影 (map projection) : 将三维的地球表面展开到二维平面的方法。在投影过程中，球面会有一定程度的变形，轮廓形状和面积比例无法兼顾。

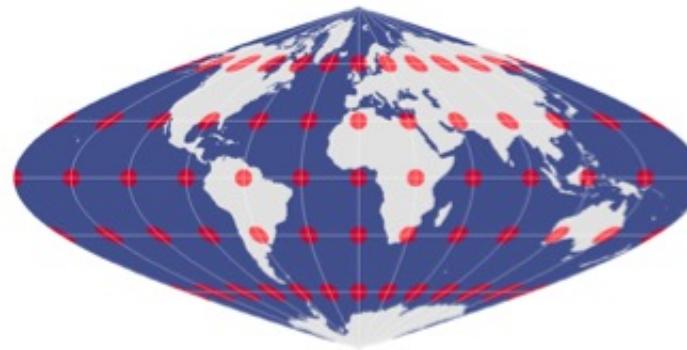


地图设计：投影

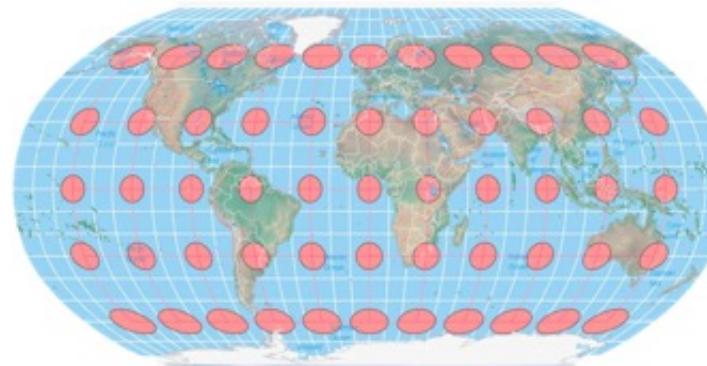
- 等角投影 (conformal projection) : 牺牲面积大小，保证形状不扭曲。
- 等 (面) 积投影 (equal area projection) : 允许轮廓形状发生扭曲，保证面积大小的准确性。
- 折中投影 (compromise projection) : 根据需求，在面积和形状的扭曲程度之间找到一个合适的平衡点。
 - 等距投影 (equidistant projection) : 折中投影的特殊类型，让沿特定方向的距离 (长度) 不改变 (但不能保证任意两点之间的距离不变)。



等角投影：陆地轮廓不扭曲，但图中的圆圈本应一样大。



等积投影：面积比例不改变，但陆地轮廓被扭曲 (圆形不再是圆形)。

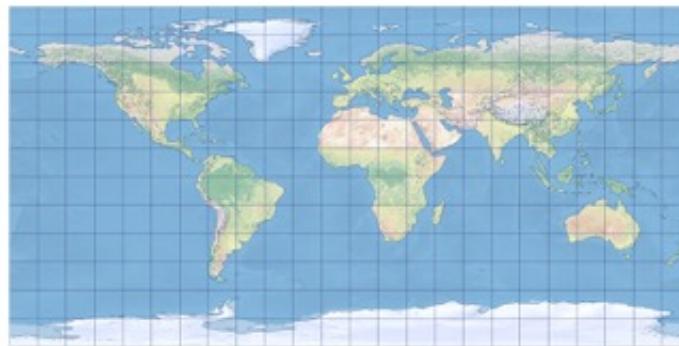


折中投影：面积和形状都改变，但总体幅度较小。

地图设计：投影



麦卡托（墨卡托）投影



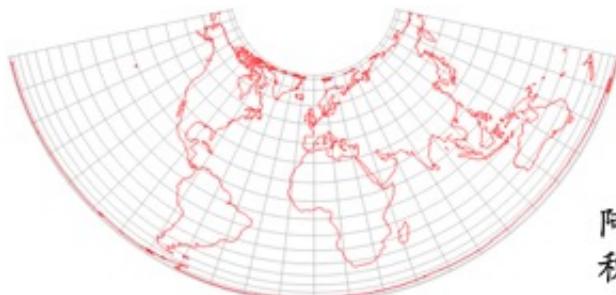
简易圆柱投影（伪投影）



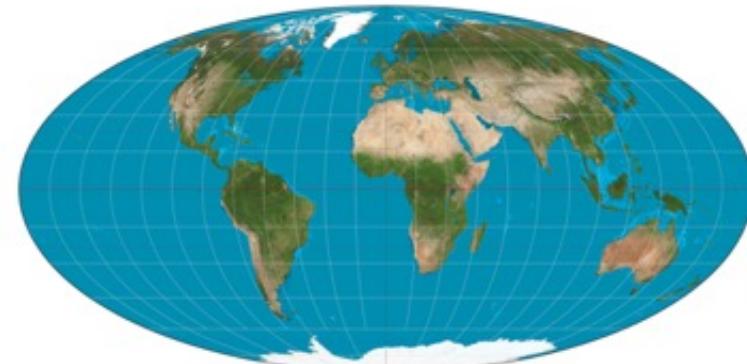
兰伯特等角圆锥投影



罗宾逊投影



阿尔伯斯等积圆锥投影



莫尔韦德投影

地图设计：投影

- 常见的中国地图投影：

- 等积斜切方位投影 (oblique azimuthal equal area projection)，是兰伯特等积方位角投影的变形。投影中心坐标是北纬 30° ，中央经线为东经 105° ，中央纬线为北纬 30° 。
- 兰伯特等角圆锥投影 (Lambert conformal conic projection)，两条标准纬线分别是北纬 25° 和 47° ，中央经线东经 102° 或者 110° （比如谭其骧版历史地图）。



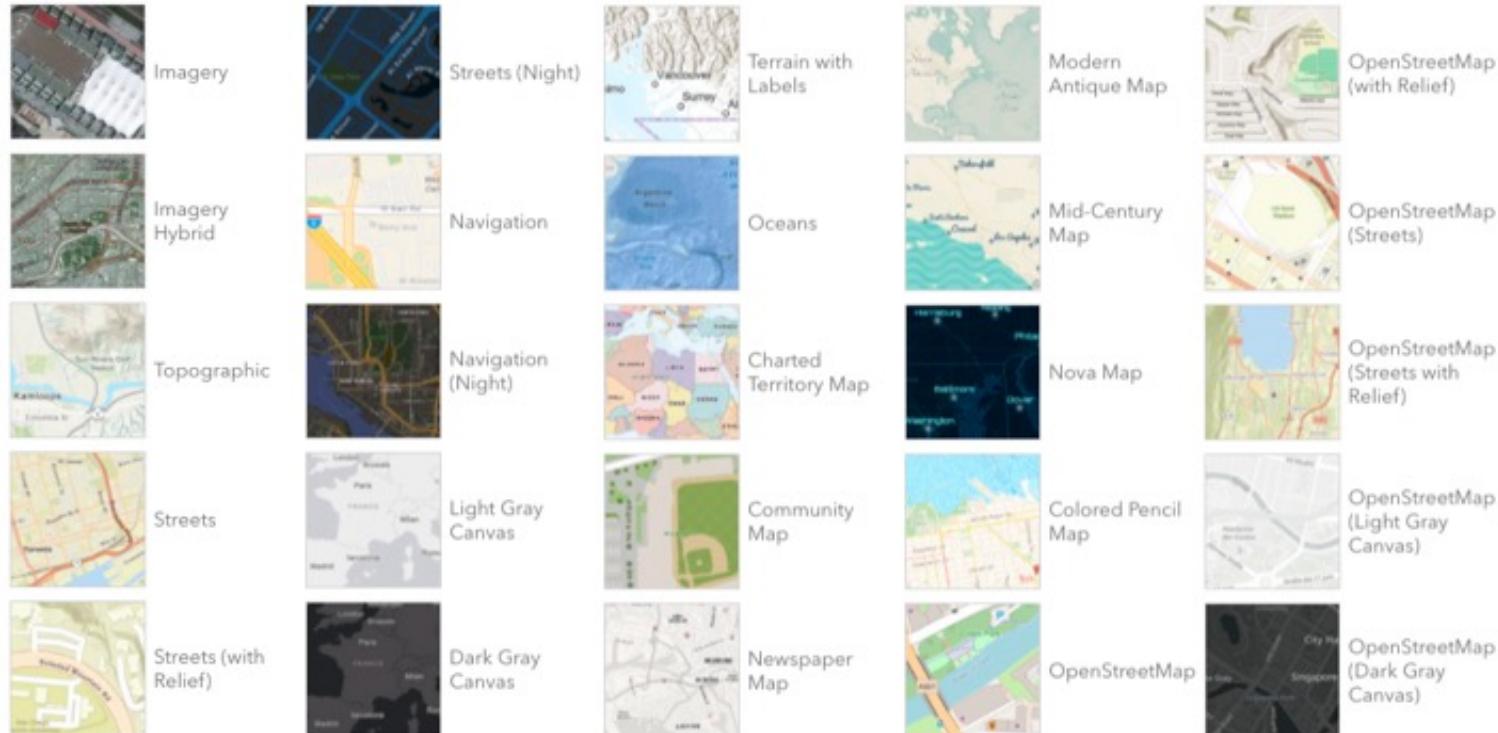
等积斜切方位投影



兰伯特等角圆锥投影

地图设计：底图

- 地形类：高程、坡度阴影、地表弯曲度等
- 遥感图像：用地类型、地表覆盖物等
- 矢量图：行政区边界、河流、湖泊、城镇、道路等



Esri的底图集

制图综合



制图综合

制图综合 (map generalization) : 在不同比例尺的地图上精确、有效地表示重要地理信息的方法，在保持重要信息准确性的同时，使地图清晰易读。

				
Initial map (1 :1000)	1 : 1582	1 : 2550	1 : 4100	1 : 27176

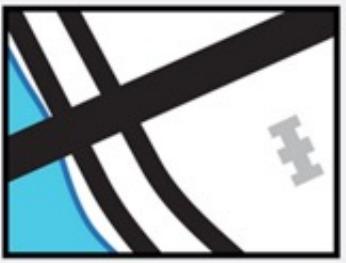
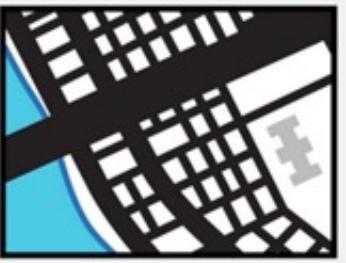
制图综合



轮廓简化 (Simplified) : 简化复杂的线性要素和多边形轮廓。



融合 (Fused) : 把靠得很近且性质相同、相关性大的几个地理要素合并为一个。



筛选 (Omitted) : 删 除不重要的地理要素，只保留重要的要素。



错位 (Displaced) : 在对地理精度要求不高时 (通常为示意图)，适当让某些地理要素发生位移，以突出显示或避免要素过于密集。

制图综合

矢量数据
简化算法

隔点法

垂距法

分裂法

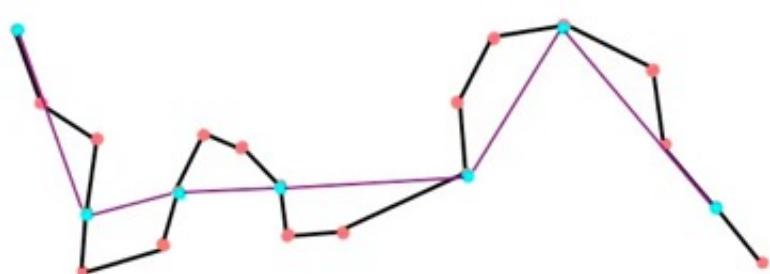
制图综合

隔点法：每隔n个点选取一个保留，删除没选中的点。

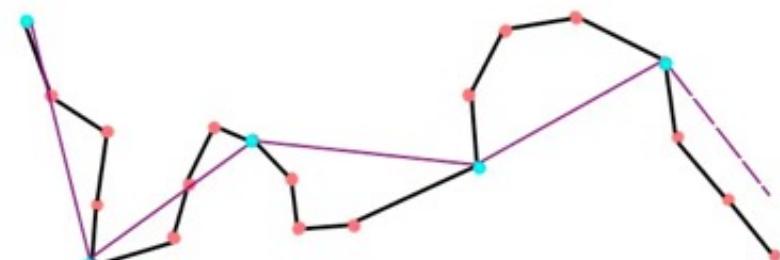


初始曲线

优点：算法和操作简单
缺点：有特色或重要的点可能
会被漏掉（比如海湾或半岛）



每隔两个点留一个

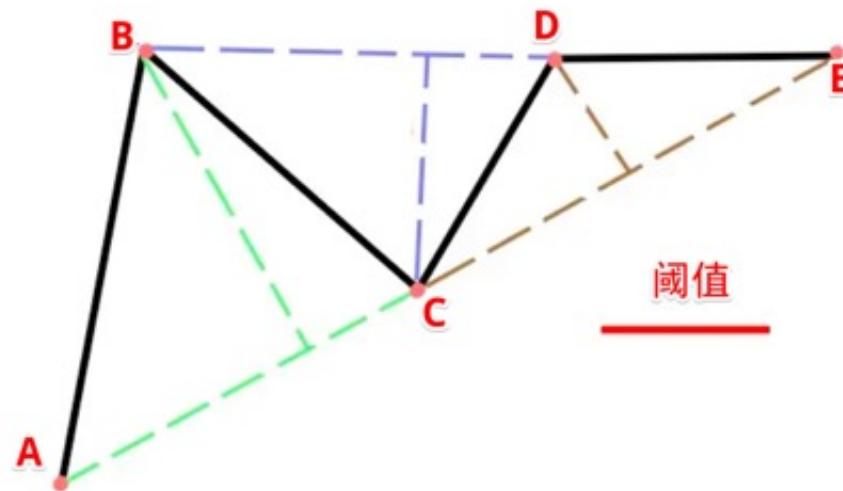


每隔三个点留一个

制图综合

垂距法：从局部出发，按照长度阈值（参数）依次筛选端点

1. 从左到右简化曲线（端点依次按ABCD等编号）
2. 设置长度阈值(参数)
3. 选中ABC三个点，比较B到AC距离和阈值的长短
4. 如果B到AC的距离大于阈值则保留B，反之则删除B
5. 如果B被保留，移动到BCD，比较C到BD的距离和阈值
6. 如果B被删除，则比较C到AD的距离和阈值
7. 除起点和终点外，其他隔点都需遍历一次

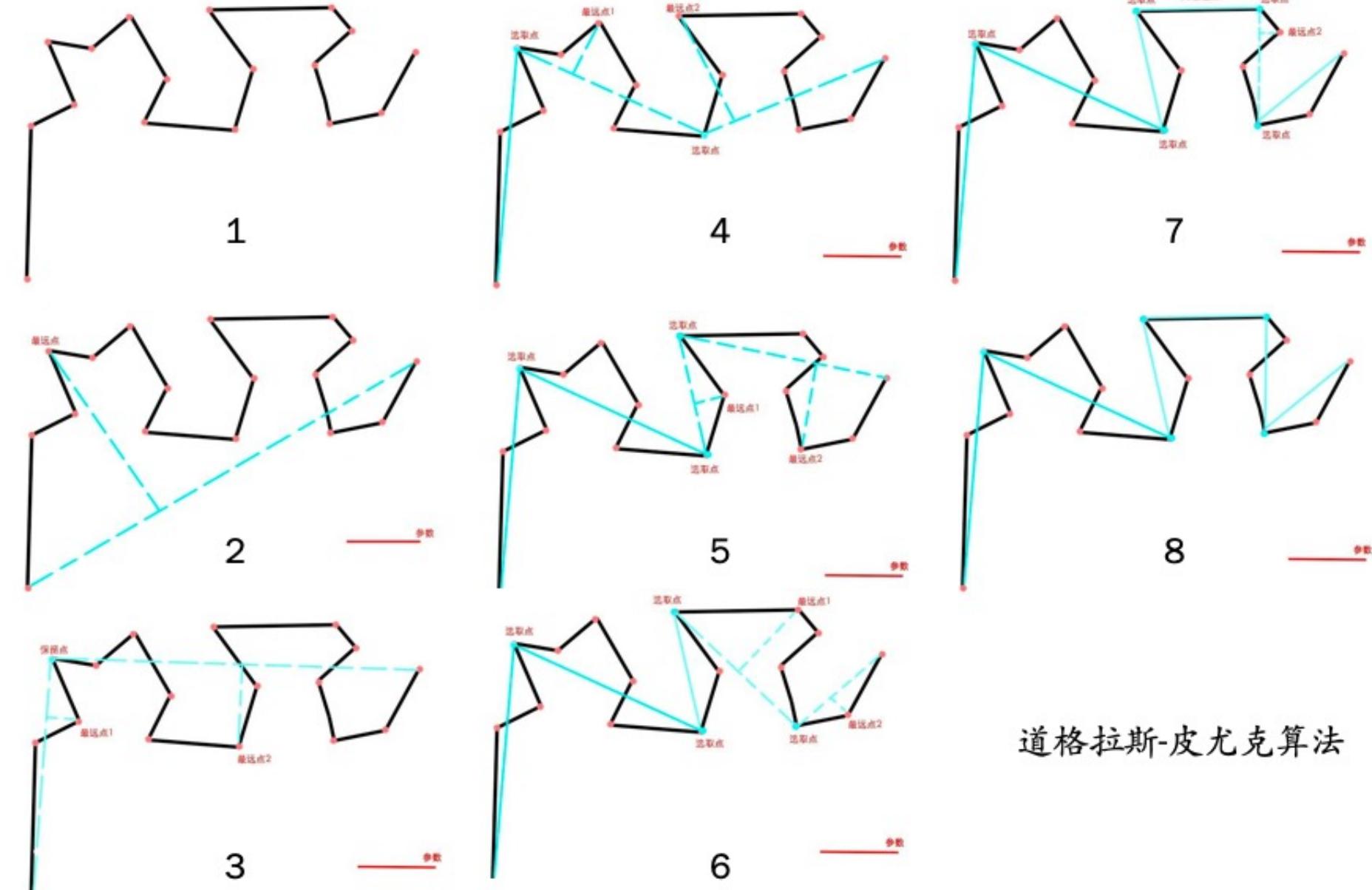


制图综合

分裂法通常为道格拉斯-皮尤克算法 (Douglas-Peucker Algorithm)：根据阈值从宏观角度决定每个端点的保留与否。

1. 定义长度阈值(参数)
2. 连接首尾两个点，在其余端点里寻找距离该连线最远的点
3. 如果该端点到连线的距离小于阈值，则删除首尾间沿途各点，算法结束
4. 如果该端点到连线的距离大于阈值，则该点保留，并用该点将要素一分为二，在其两侧进行递归操作

制图综合



道格拉斯-皮尤克算法

同时展示时空数据



时空数据

时间戳

基线技术

图片序列

时空立方
体STC

3D渲染

时间戳

- 时间戳：一系列带有日期和时间信息的事件。在地图上，我们可以使用特定的图形和符号来表示时间的变化。

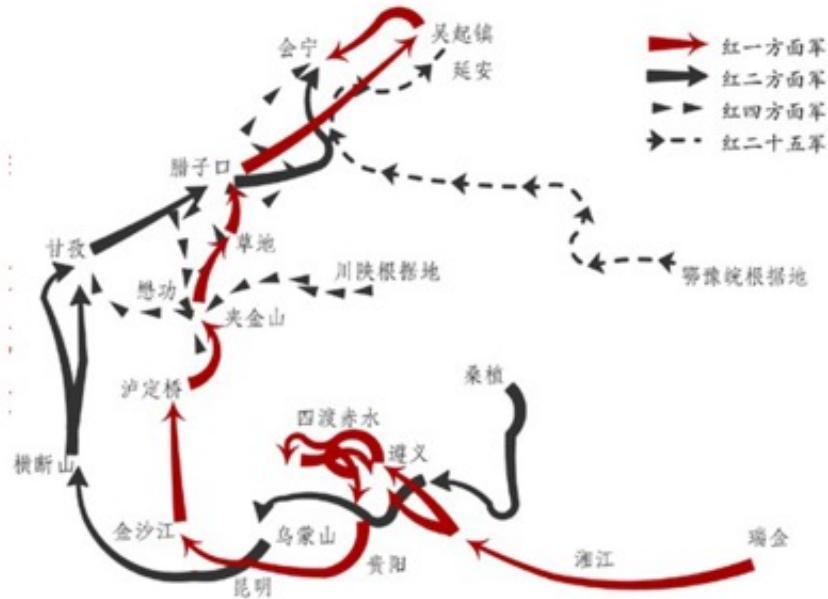
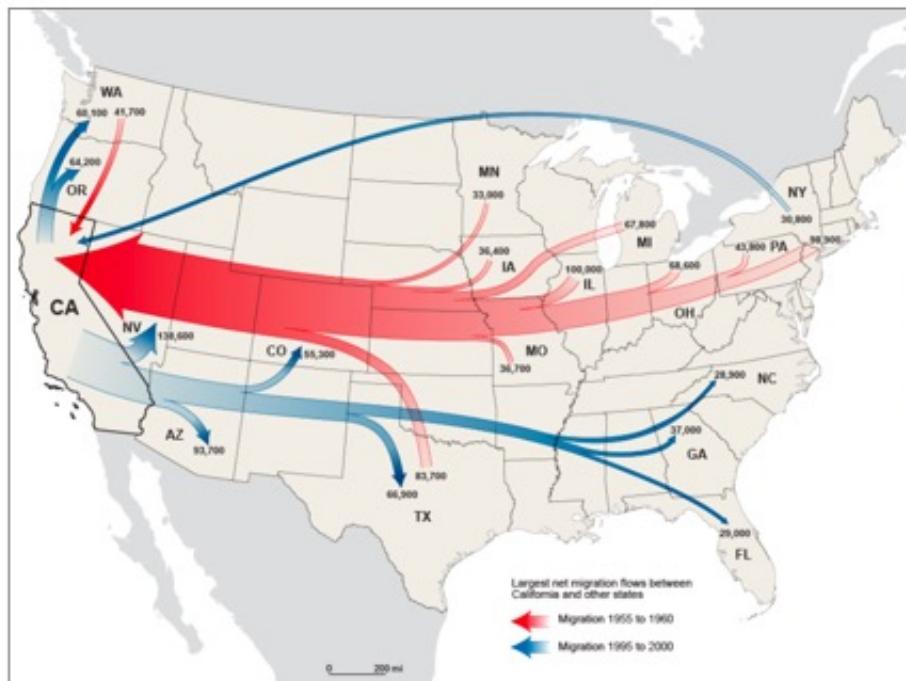


基线技术

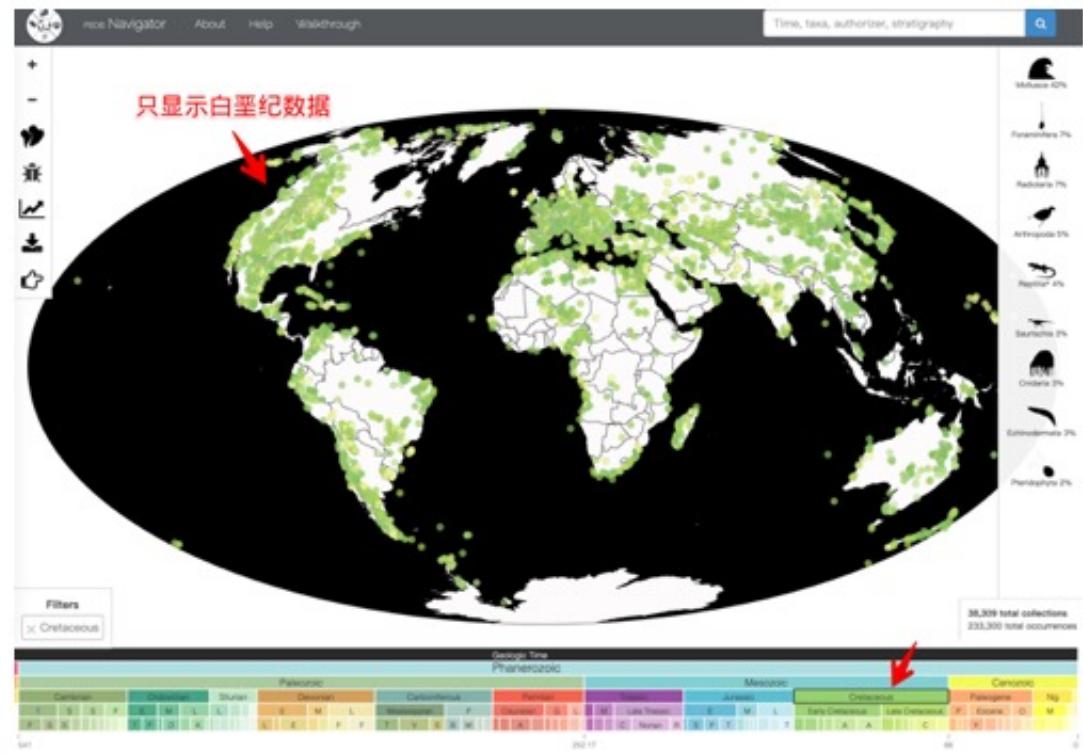
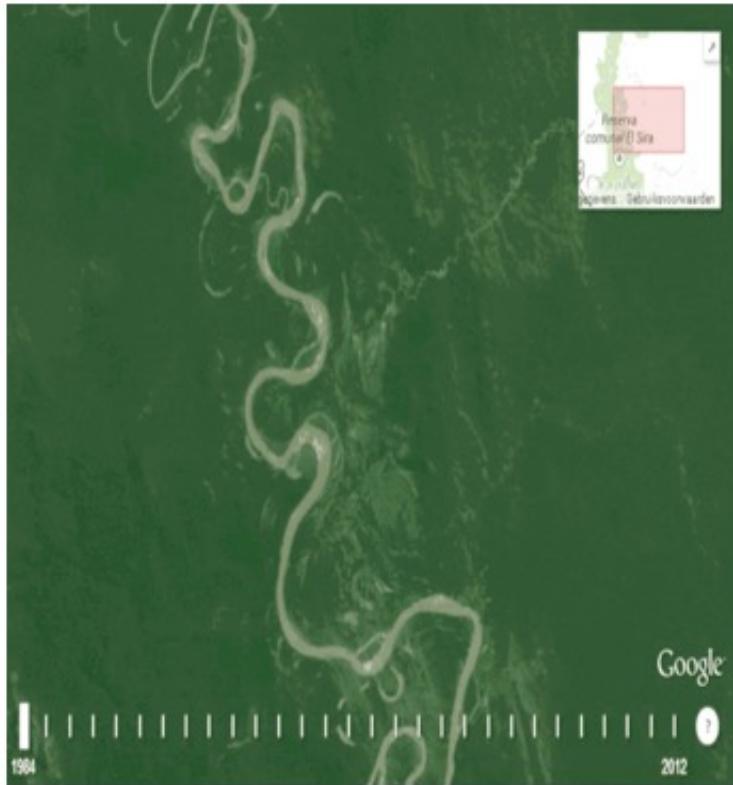
- 基线技术：用箭头等线状图形展示随时间变化的空间数据。

Net Migration Between California and Other States: 1955-1960 and 1995-2000

March 7, 2013



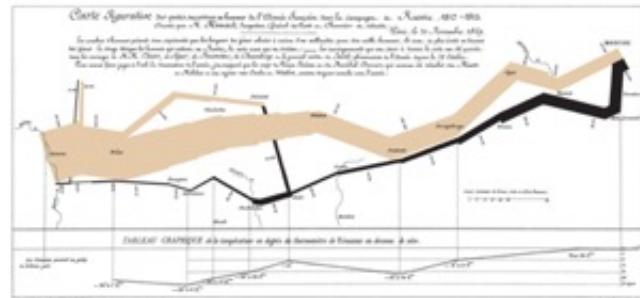
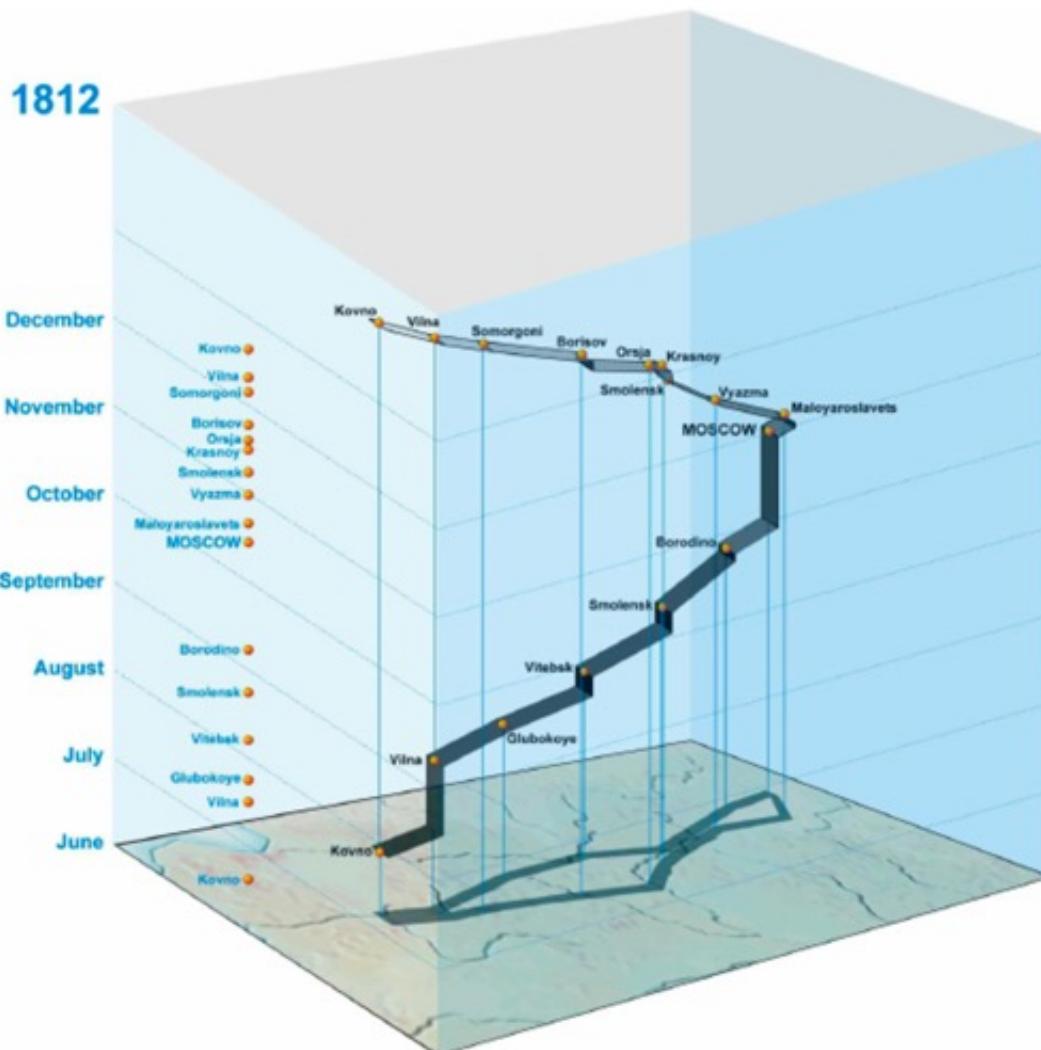
图片序列



PBDB: 古生物数据库

时空立方体 (STC)

时空立方体 (space-time cube) : 采用三维空间，使用两个水平轴 (x和y) 来表示地理坐标，而垂直轴 (z) 用来表示时间。



3D渲染和交互式可视化

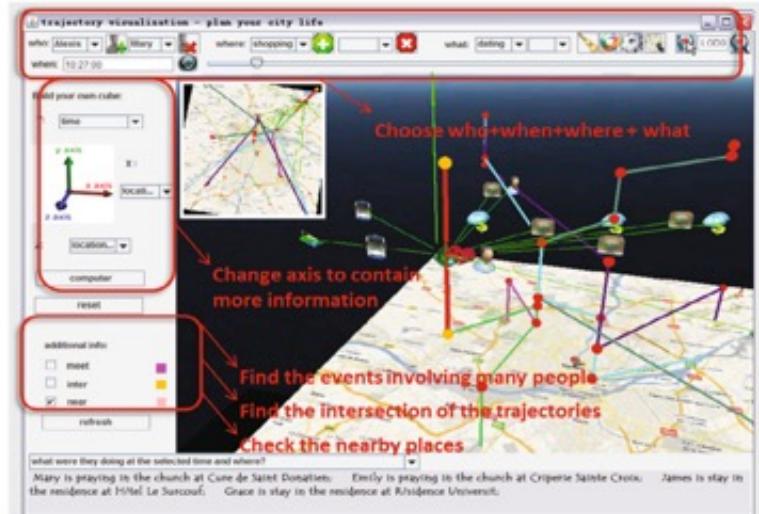
设计交互式时空可视化应考虑的问题

用户是否能在没有任何预先指导的情况下理解坐标轴含义?

用户是否能解释他们在界面上的基本操作生成结果的含义?

用户从视图中能够获取多少信息?

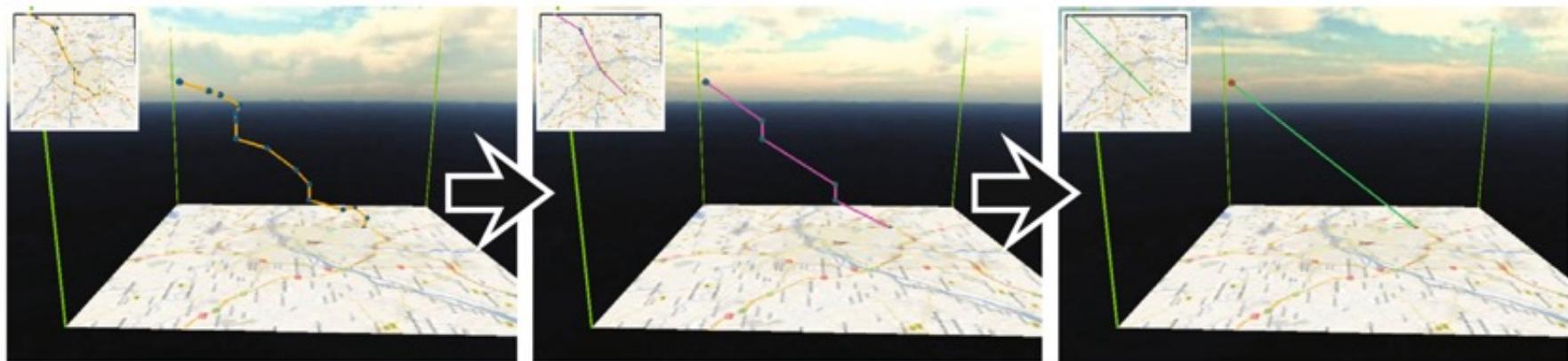
是否需要引导用户进行某些特定的交互操作?



3D渲染和交互式可视化

交互式聚合 (Interactive Aggregation)

- 手动过滤无关的数据 (关闭图层)
- 根据需求调整显示层级



3D渲染和交互式可视化

不规则时间切片 (Irregular Profiles)

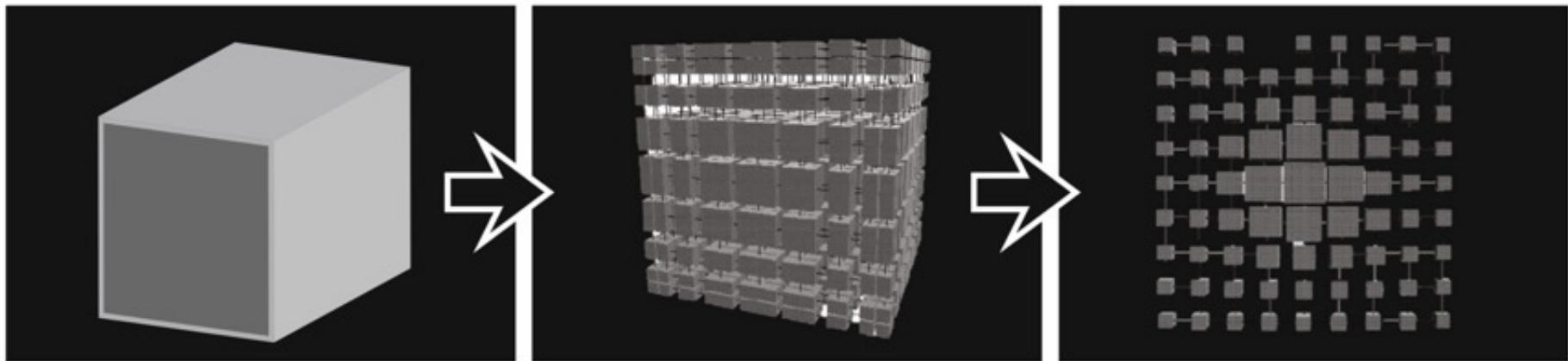
- 允许用户自定义时间截面，跨时空选择所需数据。



3D渲染和交互式可视化

3D扭曲 (3D Distortion)

- 把整个时空立方体分解成可以单独查看和处理的小块，让用户能把精力集中在项目所需的内容上。

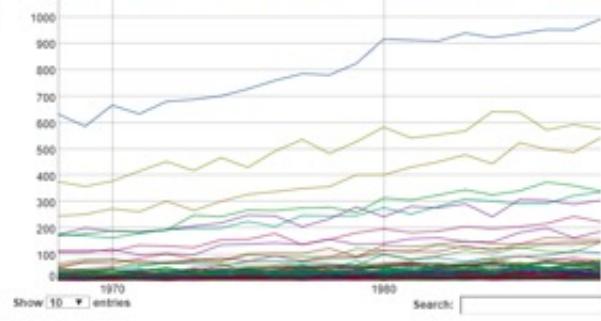
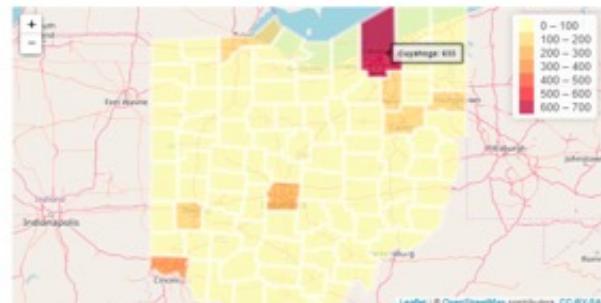
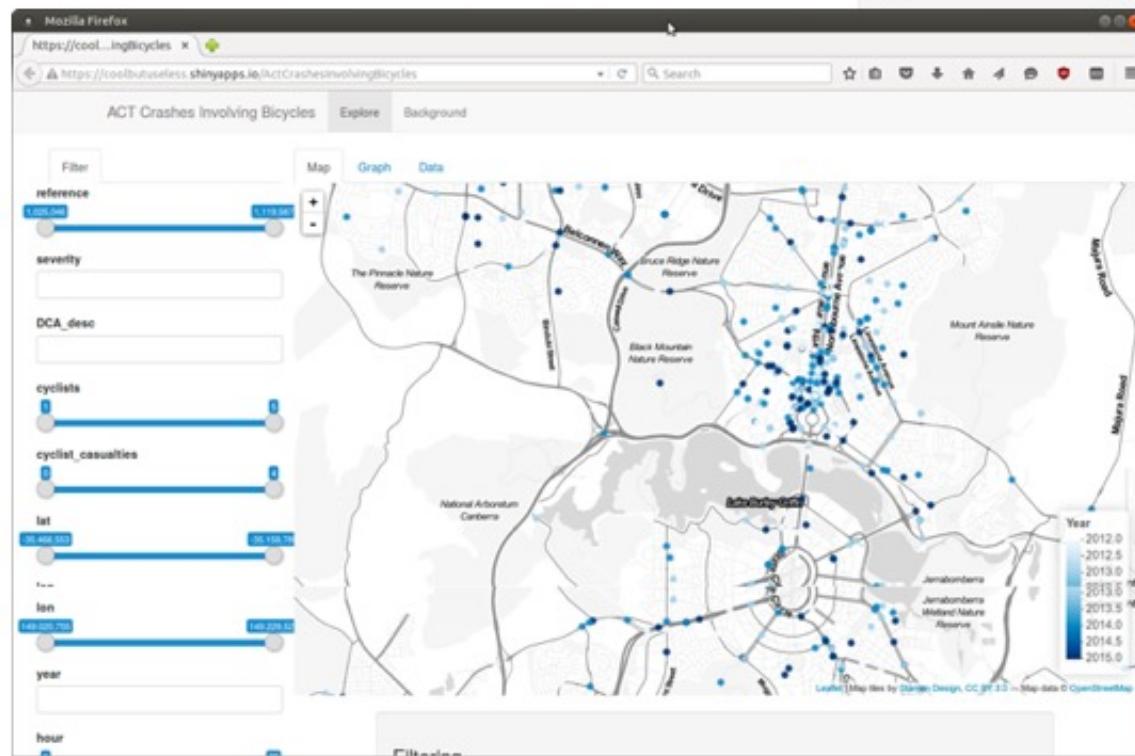


R中的交互式可视化

Shiny是R语言（及Python）中的一个包，用于建立交互式Web应用程序。Shiny擅于创建能够实时反映数据变化且具有高度交互性的网页端可视化界面。



Shiny



<https://shiny.posit.co/>

期末考试：1月5日下午2点

课程号: SR043268 课程名: 数据分析及可视化 考试时间: 2024-01-05 14:00--16:00	教室	监考教师	考试总人数	课序号	课程班别名
综合楼303 输出考试名条 普通考试桌帖名条			30	1	SR043268