Big Data



数据挖掘

广告数据

广点通 OMG 品牌广 告

腾讯平台 数据

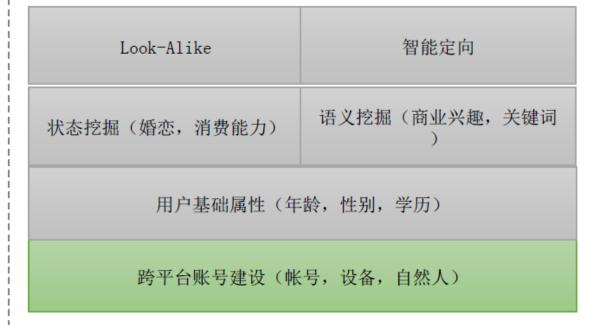
QQ 微信 QZone 应用宝 互娱 QQ浏览器

第三方平 台数据

京东 拍 搜狗 58同城

DMP外部 接入数据

品牌商CRM 垂直站点





消息队列 TDBank 离线计算 Hadoop/Spark 实时计算 TRC 任务调度 LZ

腾讯分布式数据仓库 TDW



营销广告

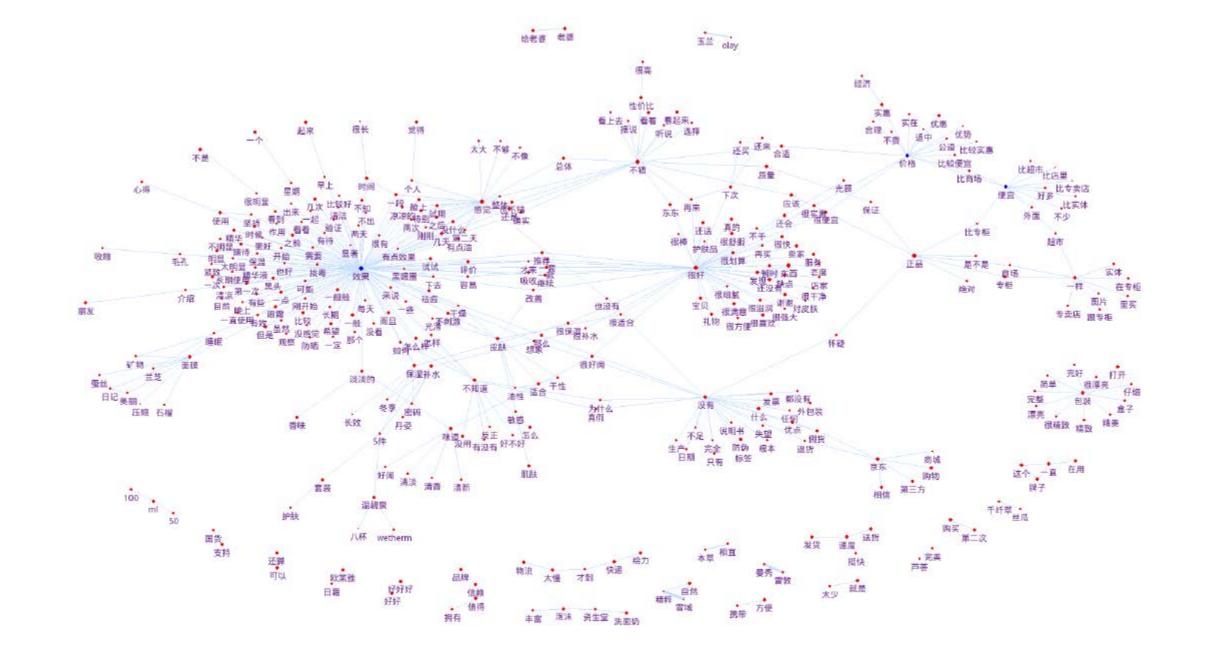
广点通 微信广告 腾讯智慧 DMP ADX



数据源

数据存储&计算

数据应用



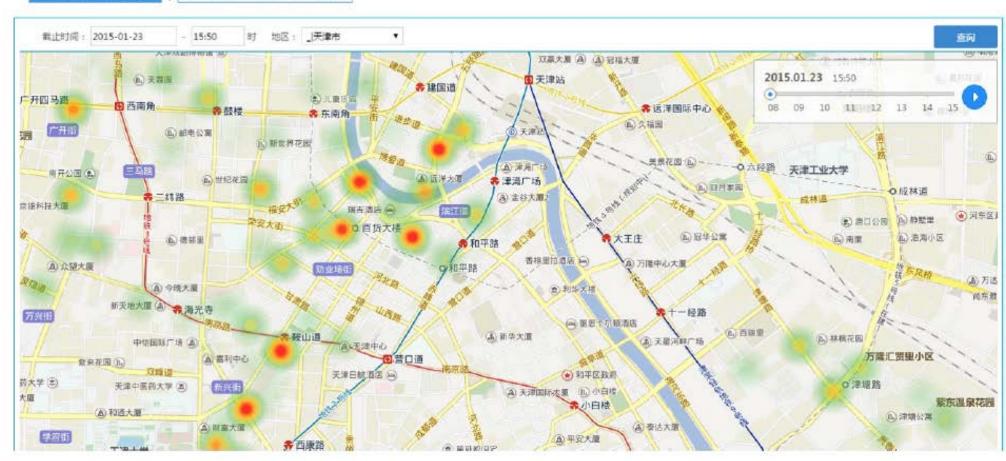
艺术	计算机与互联	羽 中小	小学教辅	男装服饰内衣		内衣	智能设备		1子教育	烹饪锅具		餐具	厨房电器		話电器	
少儿	大中专教材教辅	金融与投资	励志与成功		EUSES	服饰配件	数E 影音娱乐		紀 件 #85	,刀剪菜植		茶具/咖啡具 厨房配件	个护健康	家用电器 五金家装	大家电	
哲学/宗教	经济	文学	家教与育儿	身体护肤	女性护理	里 面部护肤	维修保养		车载电器	装饰材料	4	灯饰照明	手机	通讯		
	历史	管理	小说	洗发护发	担化 放 香水彩	妆 口腔护理	单数往往		汽车装饰	电工电	延問認 4	五金工具房工場地面材料	手机	配件	运营商	
健身训练	户外鞋服 运动户外——	运动服饰	骑行运动	#46 2 21431					安全自驾生活日用	潮流女包	特品男	菏	行男鞋	储物家具		
垂钓用品	体育用品	户外装备	运动鞋包	- 1-g	营养(ii	建康 营养成分	清洁用品 区居家装		家装软饰	礼品箱·包			首女鞋 书房家具		安厅安日	
			游泳用品	急救卫生	成人用	用品 保健器械	家纺		灯具	功能箱包 礼		LA	_		具家室径	
童装童鞋	妈妈专区	洪	护用品	电脑配件	外设产	品办公设备	粮油调味	进	口食品	乐器相关	毛绒布		財尚佈品 天然兼 宝首 都 银饰		中酒类酒	
老	寝居服饰		前食 尿辨湿巾	电 <u>服</u> 文具/耗材	电脑整机		食品校	体	闲食品	乐器相关 玩具	d.		al Alkary	宠物主粮 宠物生活	钟表	
童车童床	-TX767JK11	喂养用	品 奶粉			4机 网络产品	茗茶	飲	料冲调	益智玩具	健身功		肉生鲜奶		Ħ	



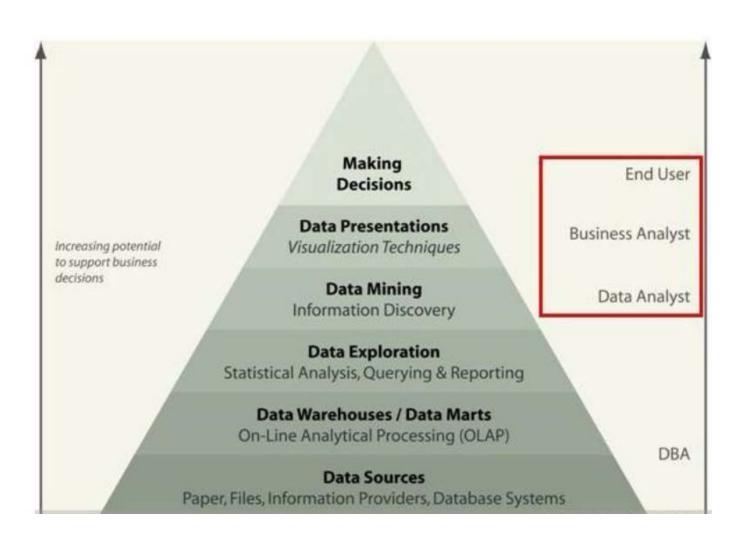
异常预警

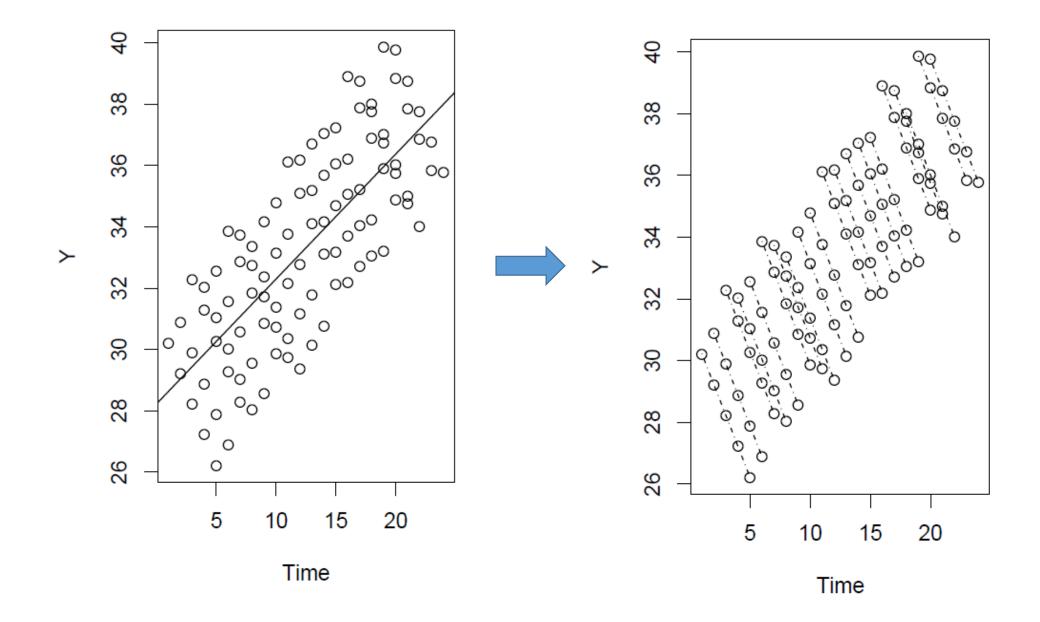
区域整体用户流量热力图

区域整体用户(全部)流量热力图

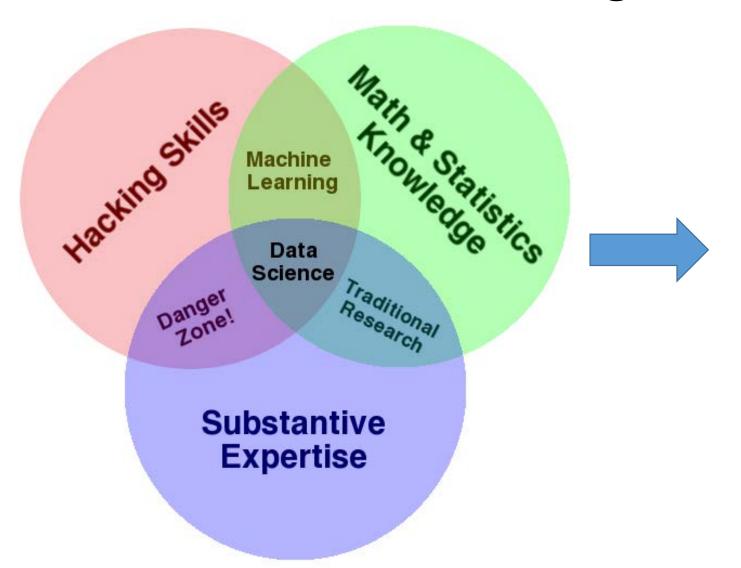


Data Mining 数据分析多层模型介绍



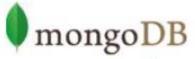


The Data Science Venn Diagram



- 1. 研究数据科学,一方面需要 如极客那般刻苦钻研,一方 面需要像统计学家那样拥有 完美的理论。
- 2. 数据科学家也不仅仅是统计学家——后者只关注如何完成一个理论的证明或构建出一个完美的模型,很少有人会使用R语言将数据文件读入系统,从而进行后续的分析。
- 3. 数据科学是一门关于数据的 工程,它需要同时具备理论 基础和工程经验,需要掌握 各种工具的用法。

http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram





















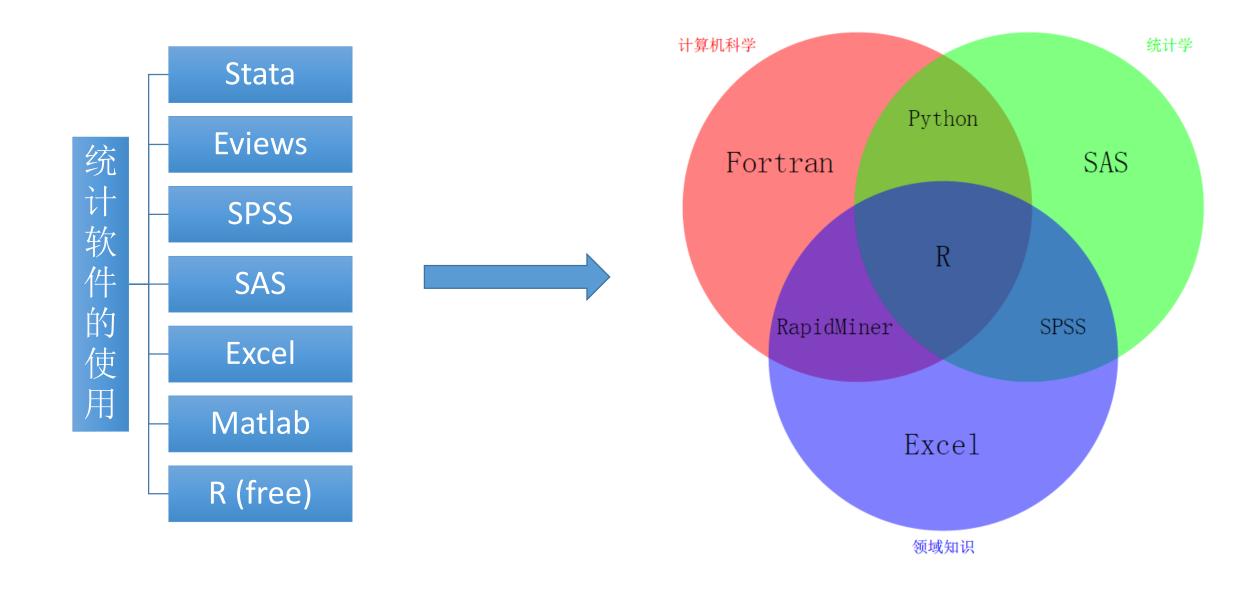












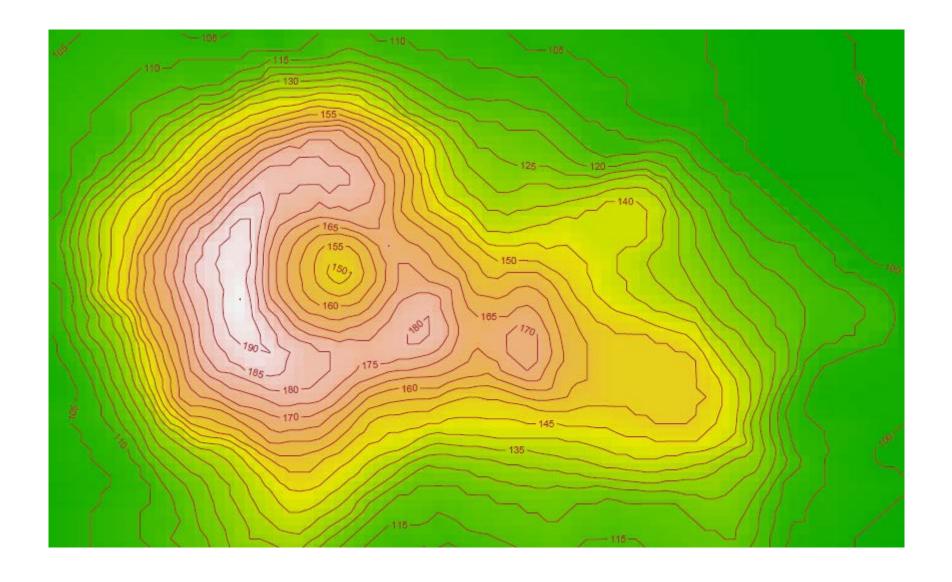
Data Visualization



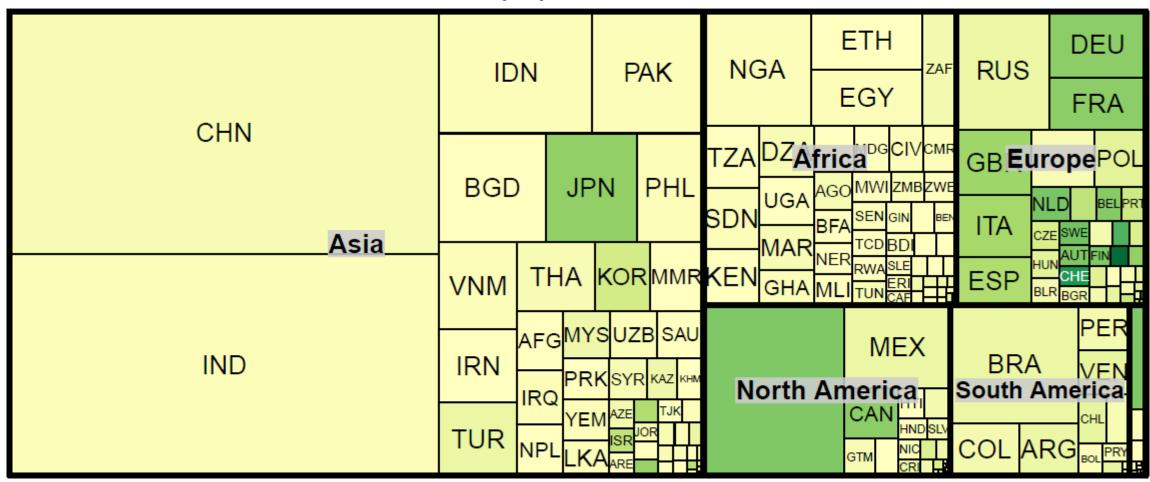
http://echarts.baidu.com/index.html



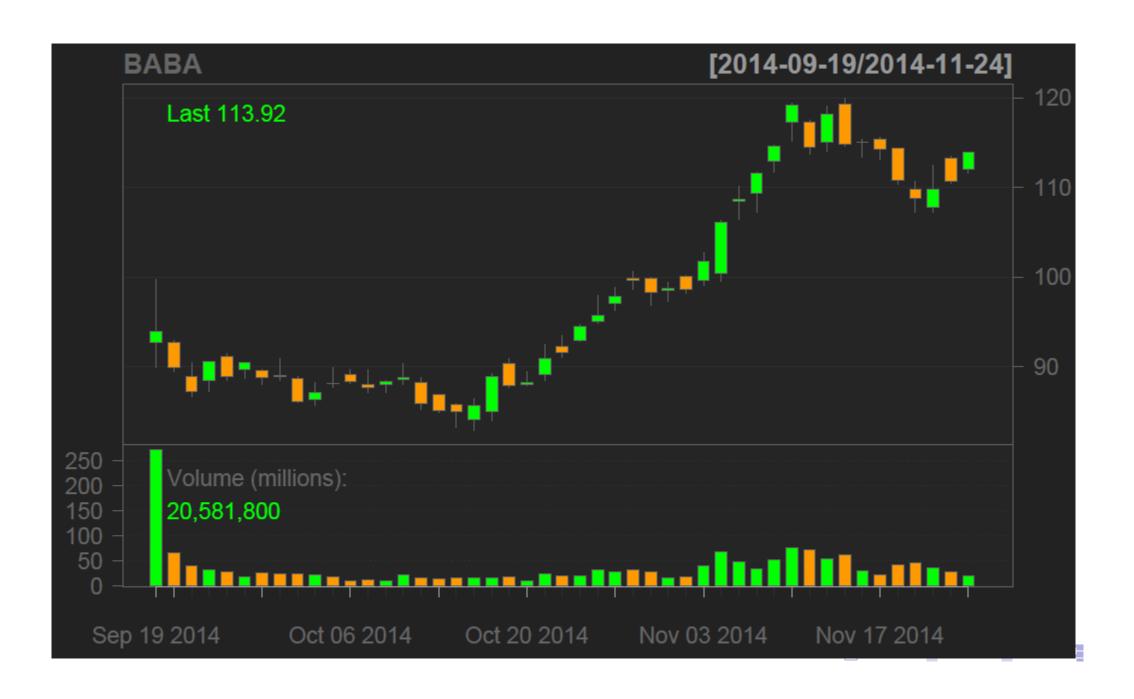
http://vis.pku.edu.cn/wiki/doku.php

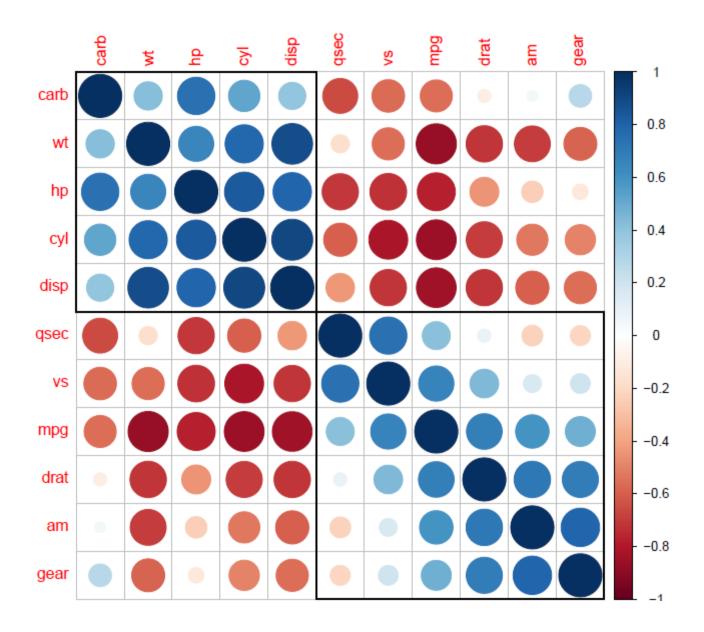


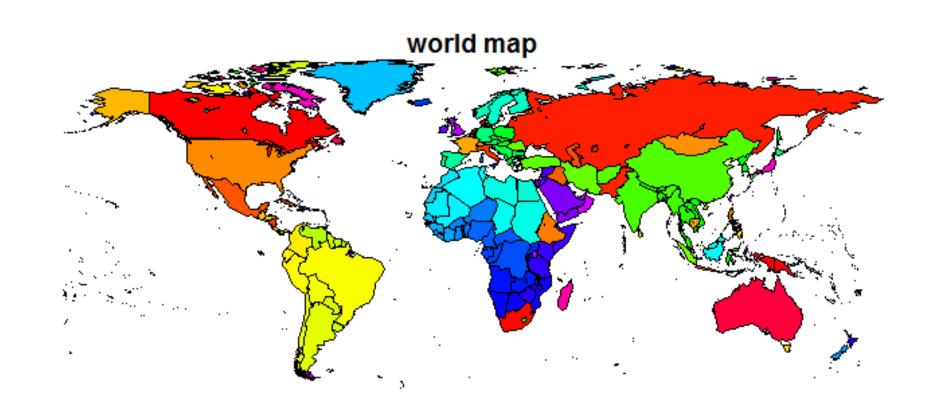
population

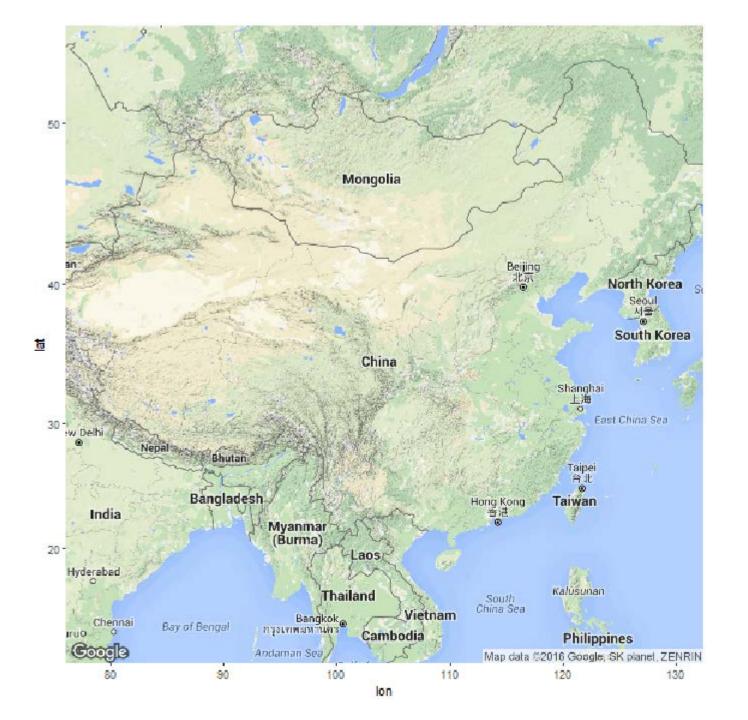


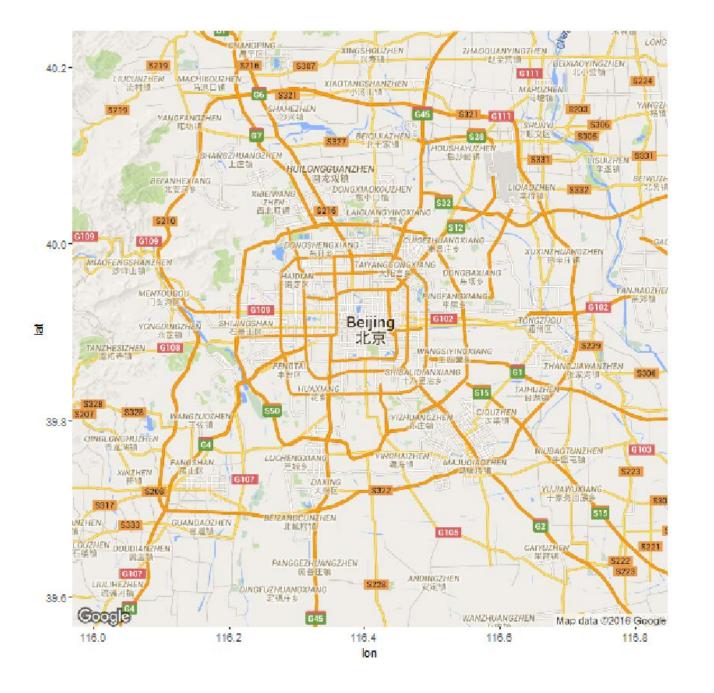












```
library(R2SWF)
if (capabilities("cairo")) {
    olddir = setwd(tempdir())
    svg("Rplot%03d.svg", onefile = FALSE)
    set.seed(123)
    x = rnorm(5)
    y = rnorm(5)
    for (i in 1:100) {
         plot(x <- x + 0.1 * rnorm(5), y <- y + 0.1 * rnorm(5), xlim = c(-3, y + 0.1 + y + 0.1) = c(-3, y + 0.1)
          3), ylim = c(-3, 3), col = "steelblue", pch = 16, cex = 2, xlab = "x",
             vlab = "v")
    }
    dev.off()
    output = svg2swf(sprintf("Rplot%03d.svg", 1:100), interval = 0.1)
    swf2html(output)
    setwd(olddir)
```

```
library(animation)
output = dev2swf({
    par(mar = c(3, 3, 1, 1.5), mgp = c(1.5, 0.5, 0))
    kmeans.ani()
}, output = "test.swf")
swf2html(output)
```

- Rstudio
- shiny (http://shiny.rstudio.com/tutorial/)
- knitr