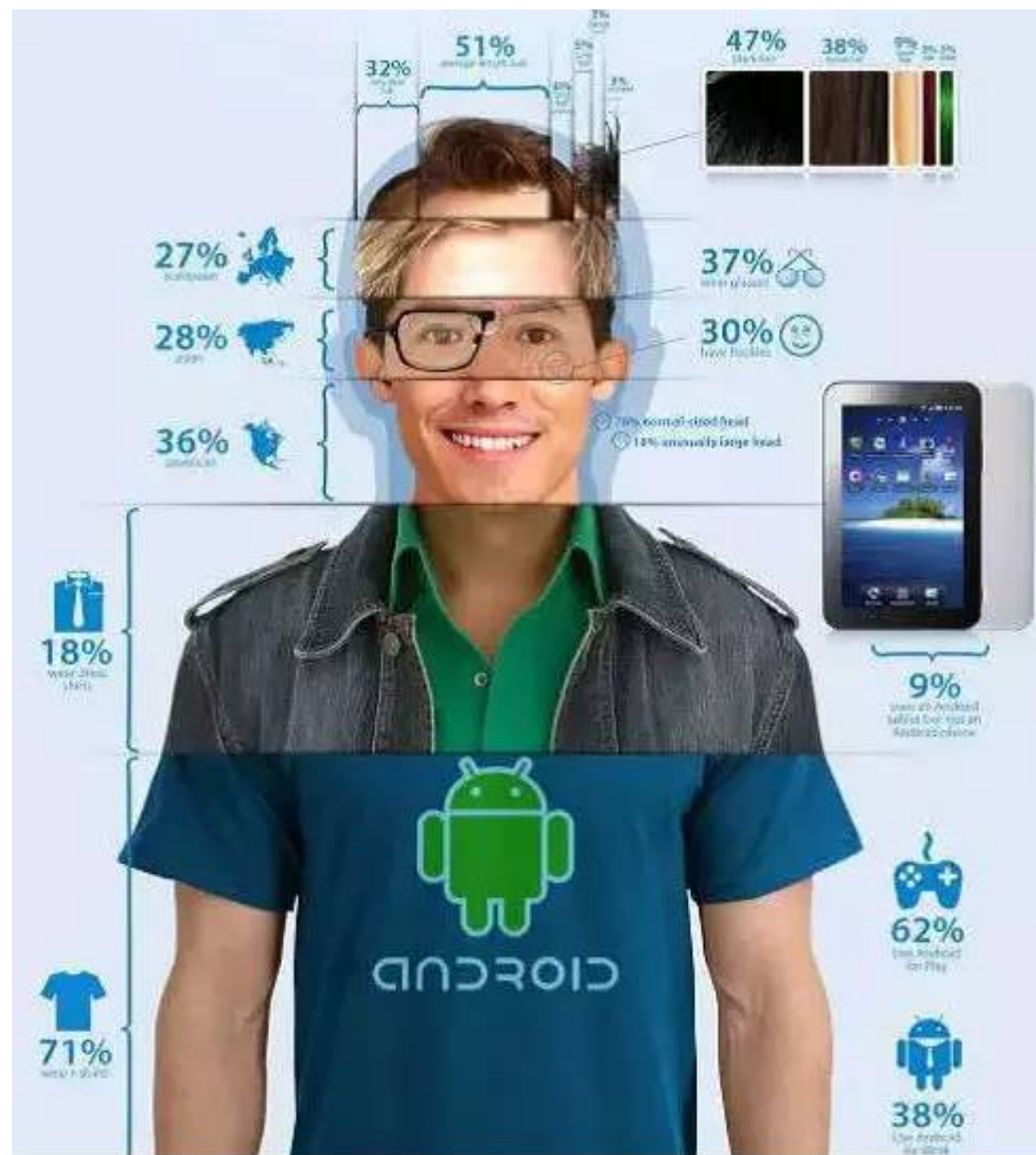
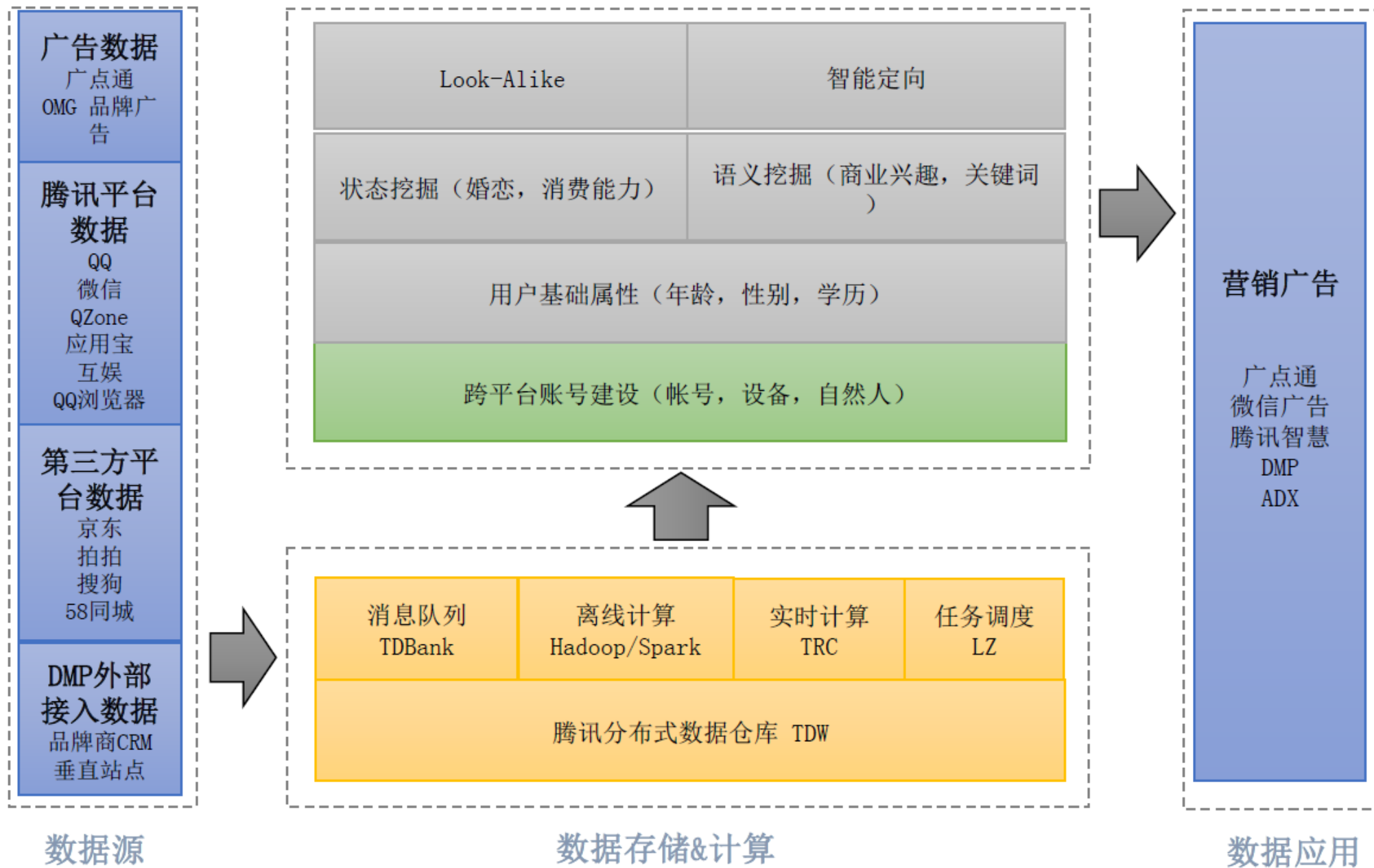
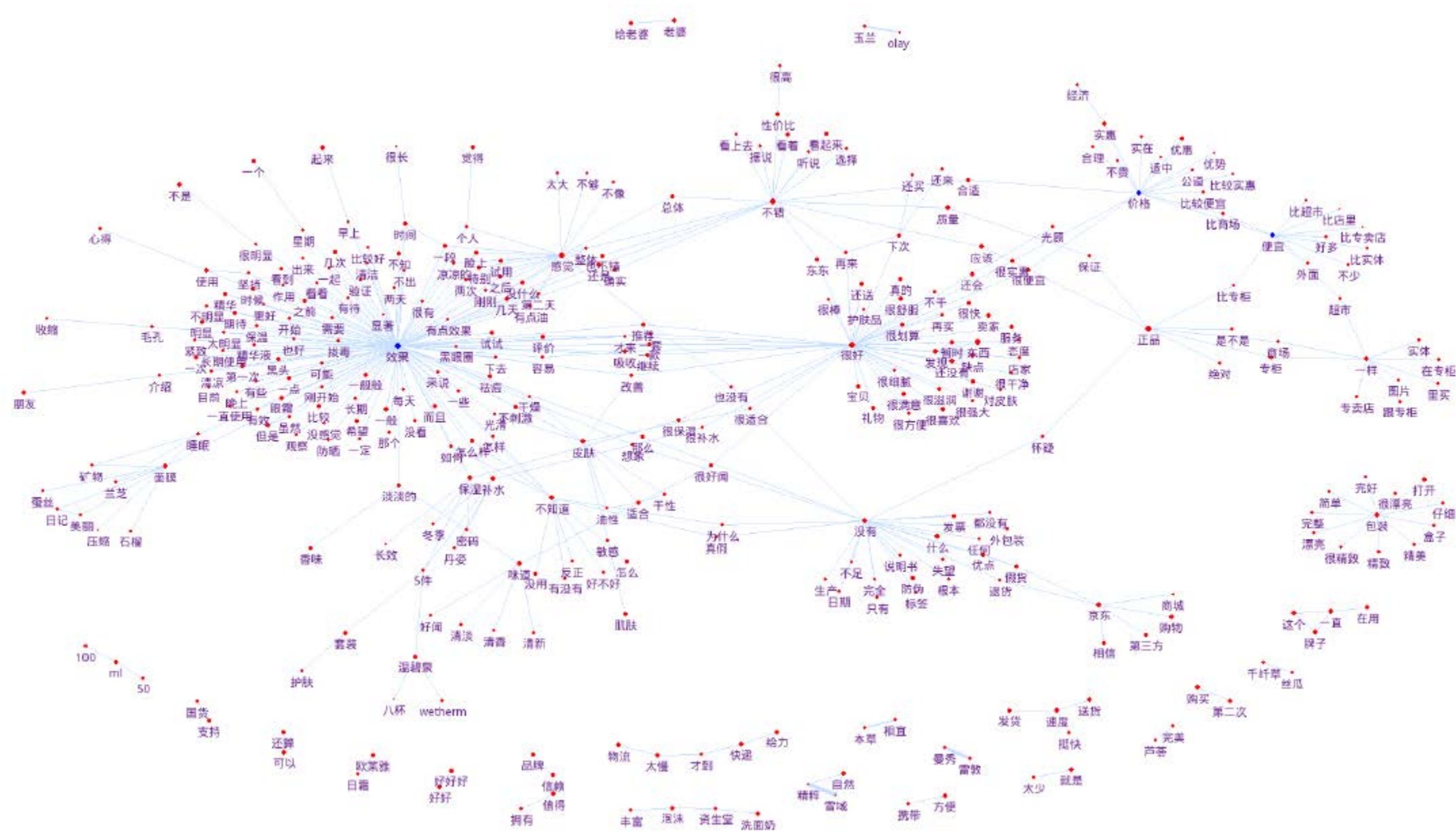


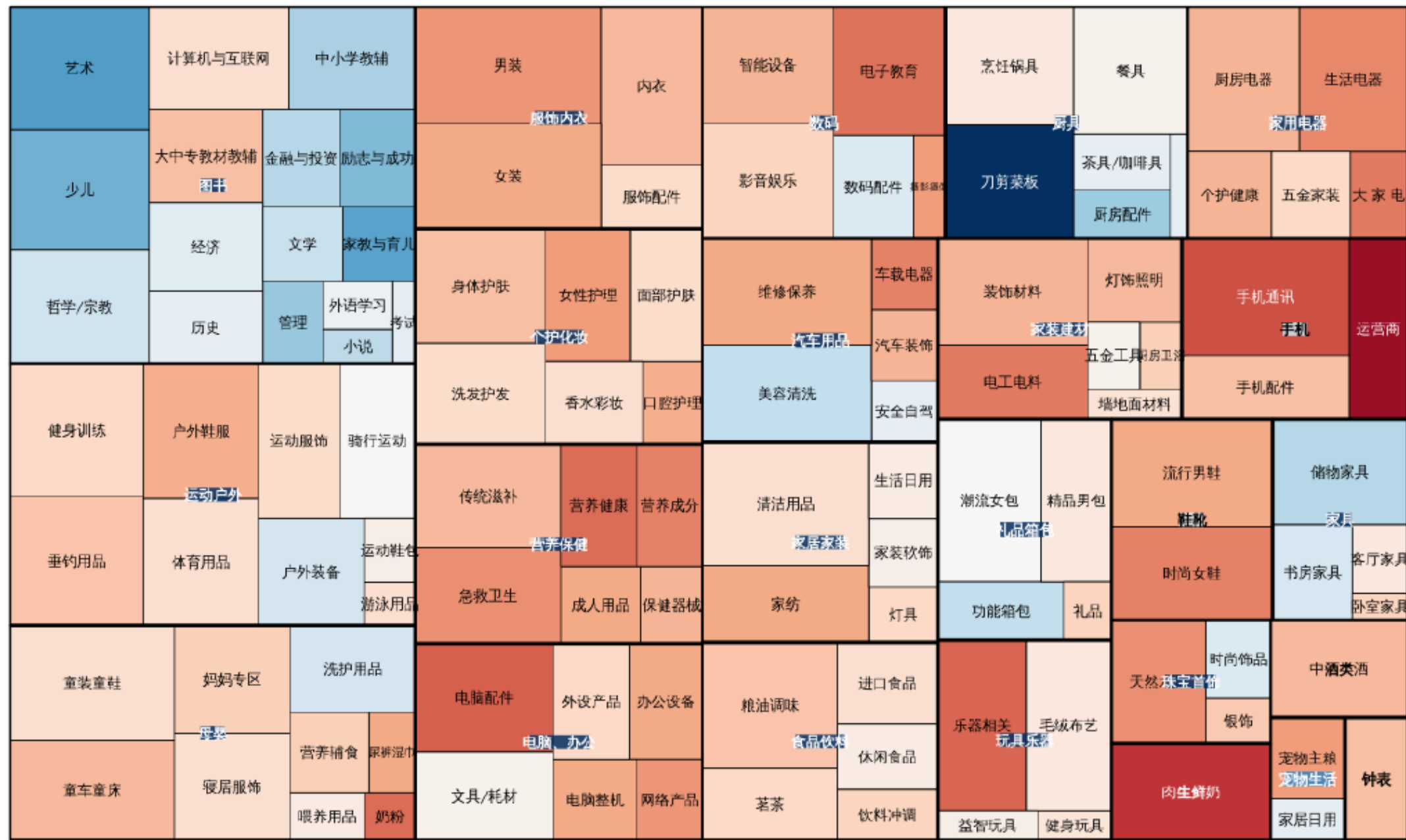
Big Data



数据挖掘



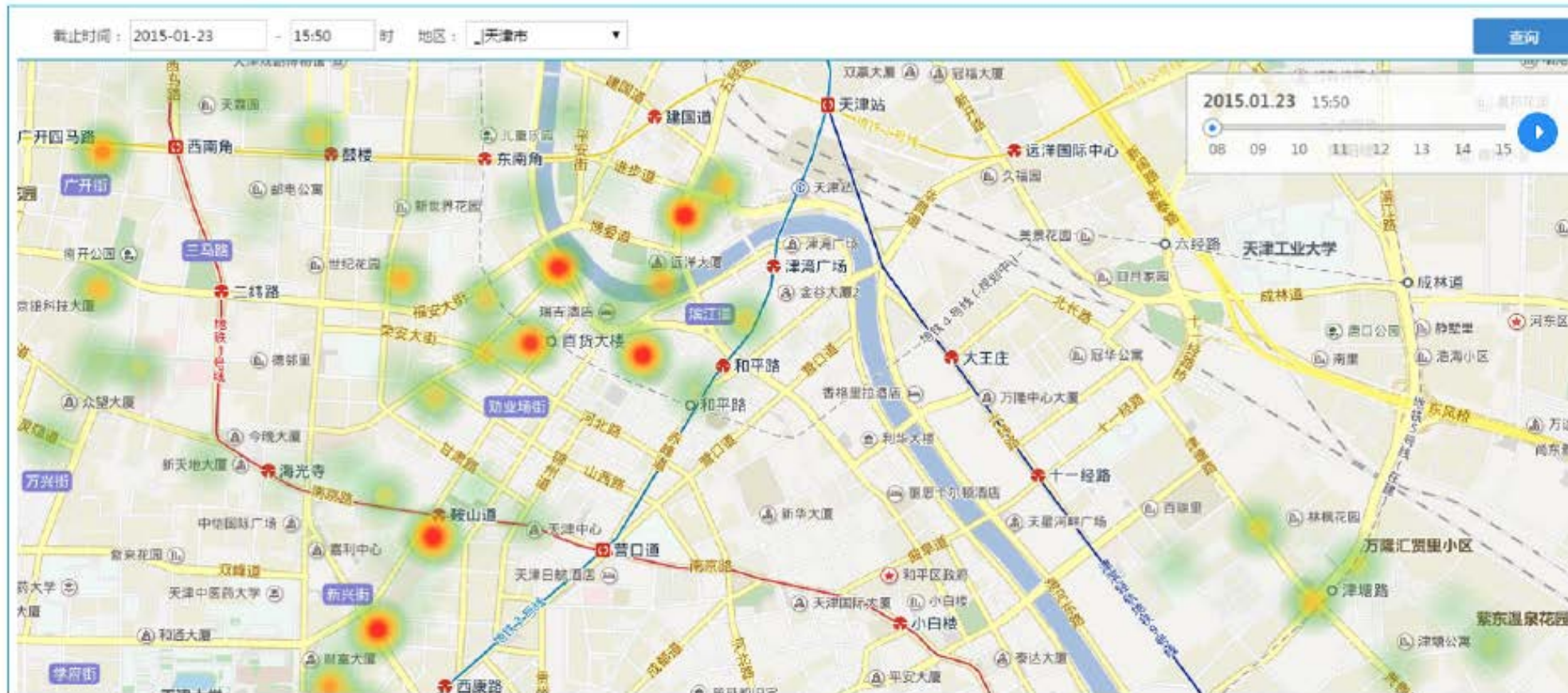






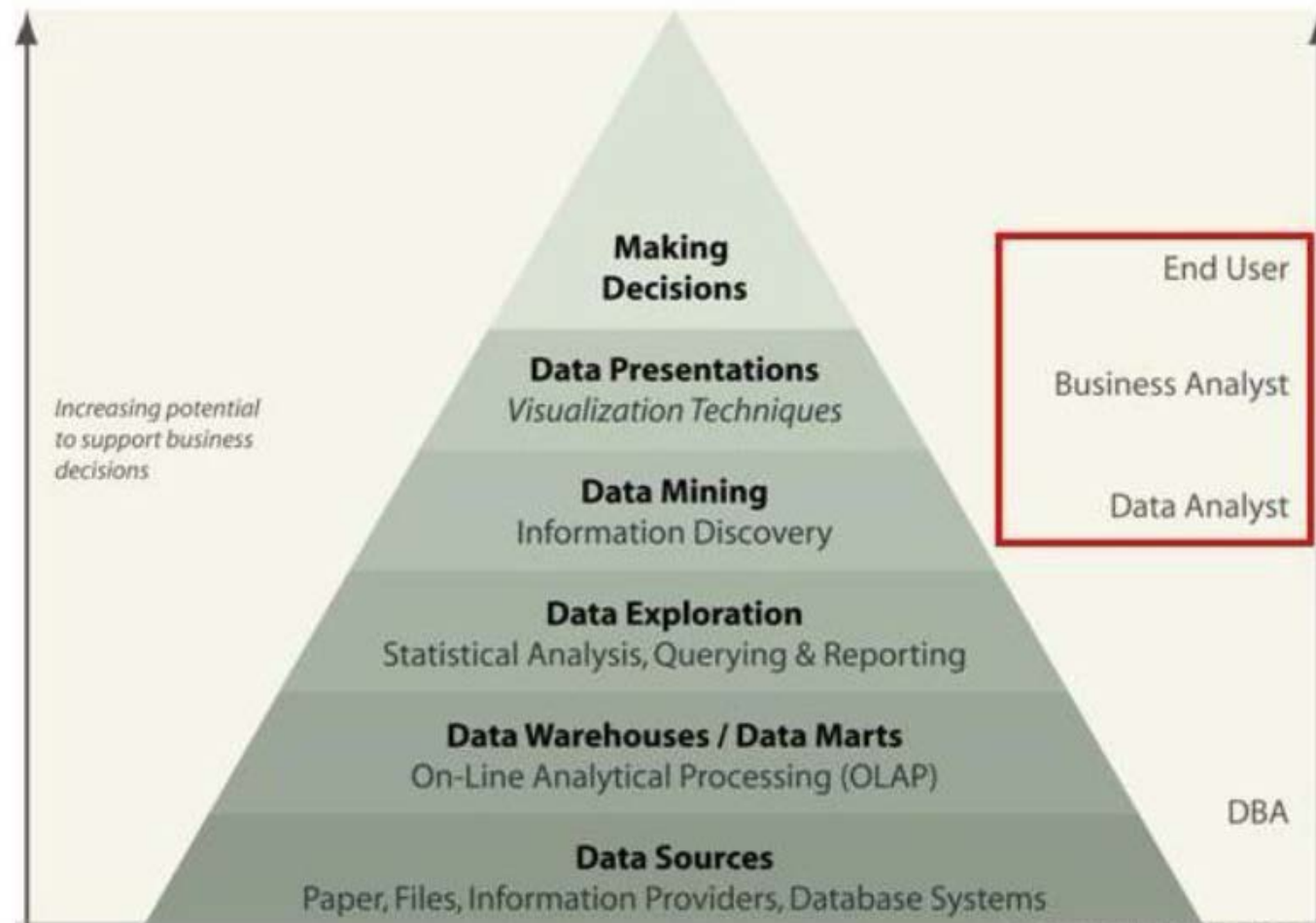
区域整体用户流量热力图

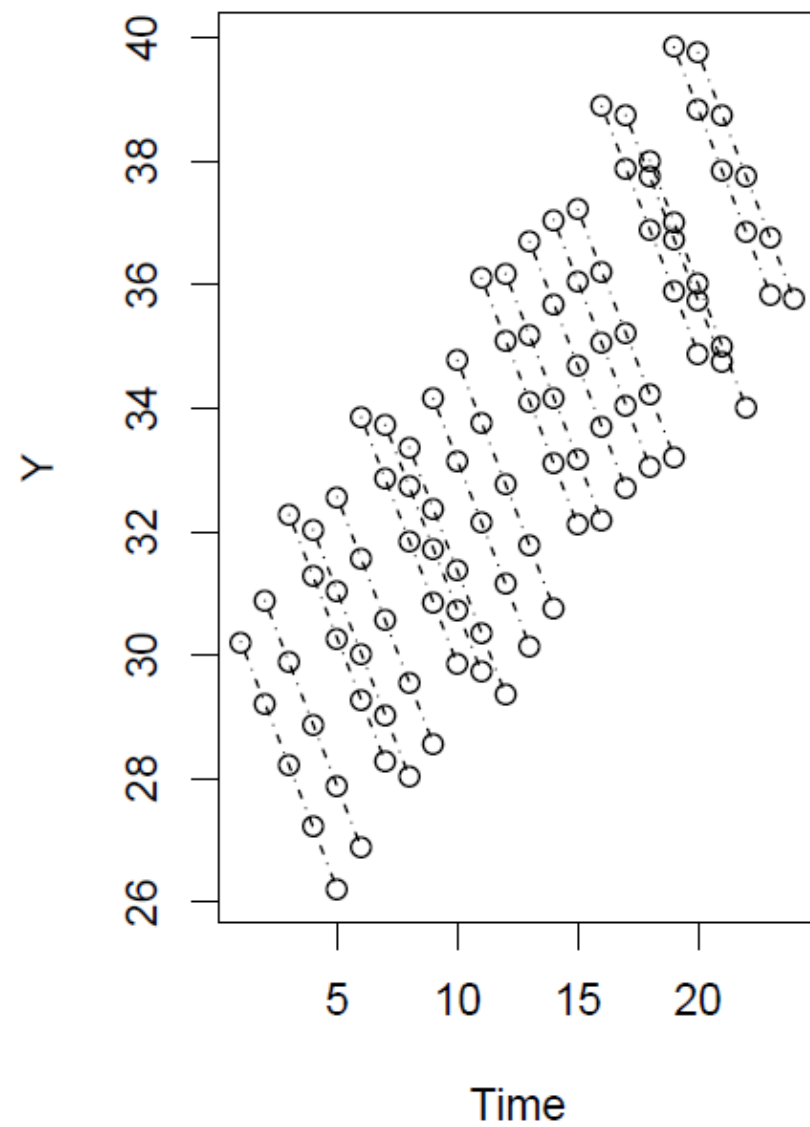
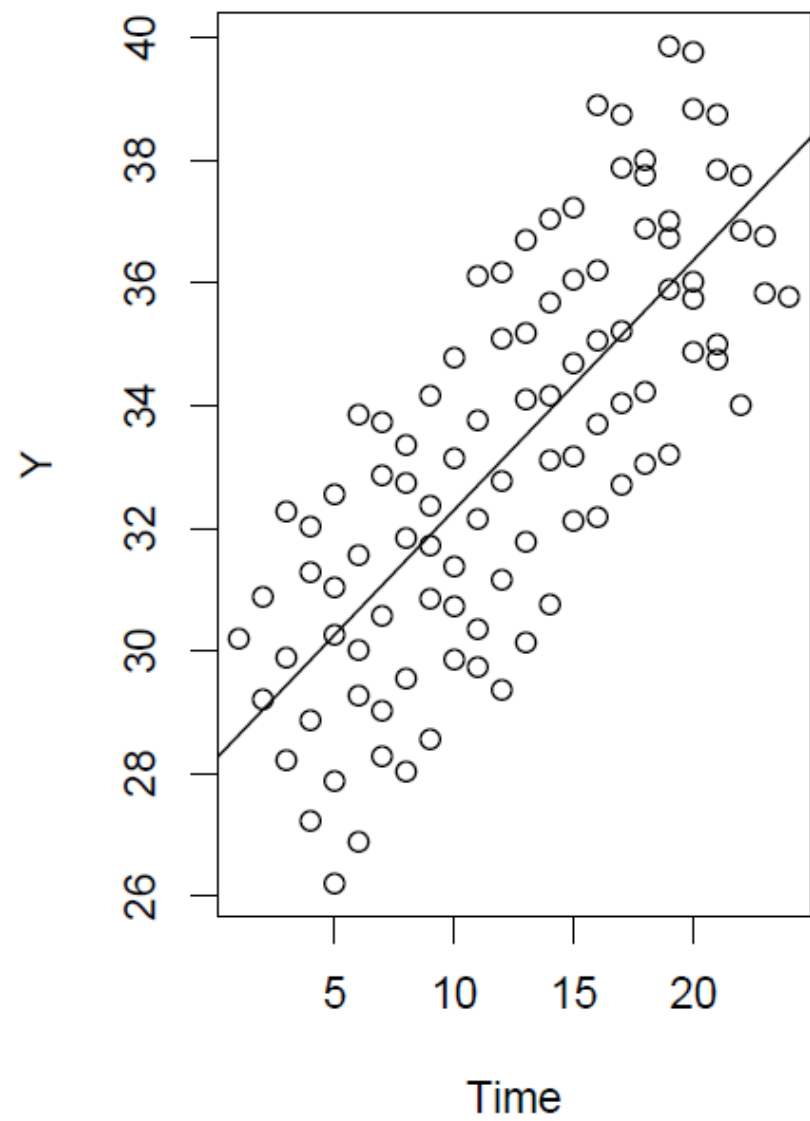
区域整体用户（全部）流量热力图



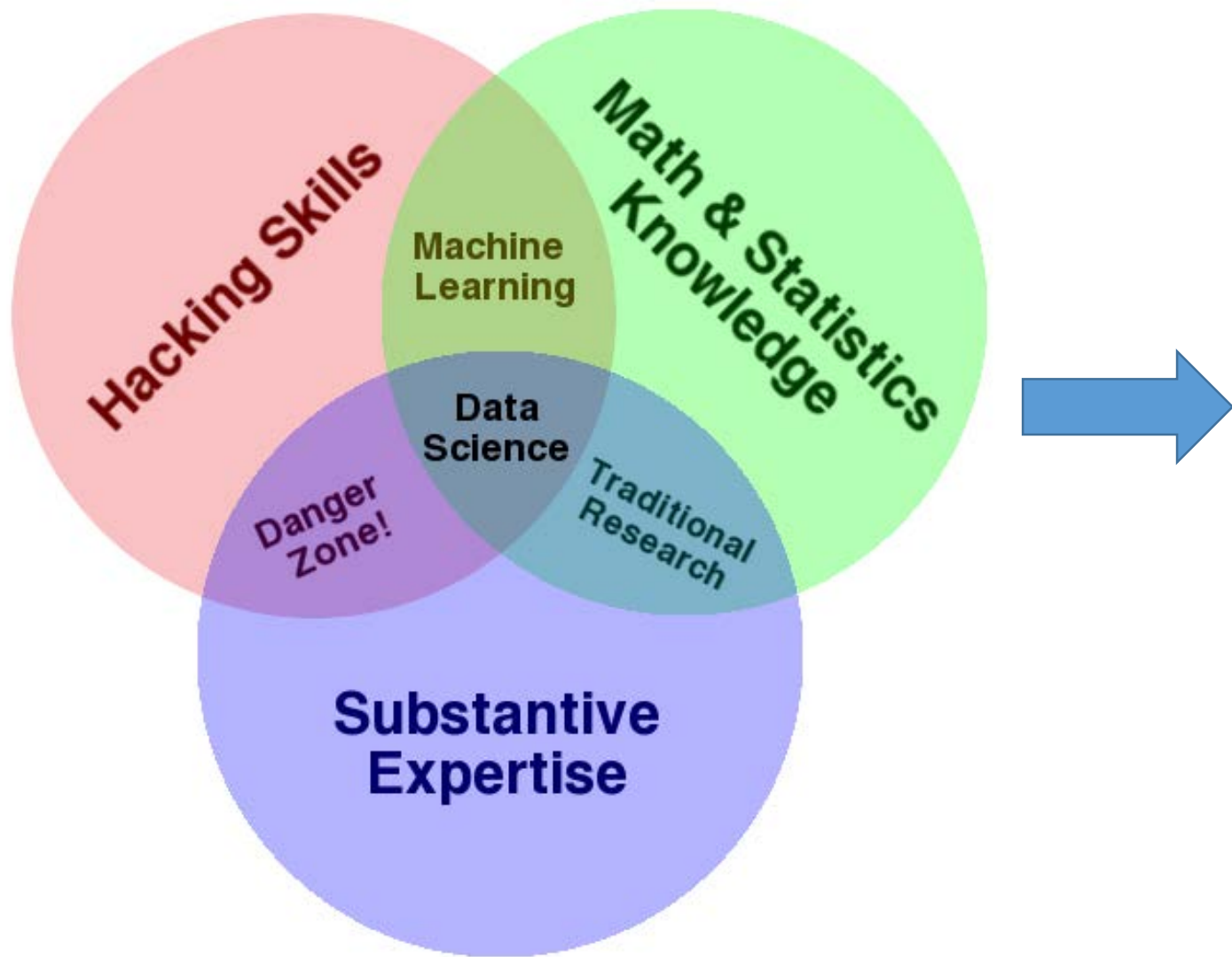
Data Mining

数据分析多层模型介绍

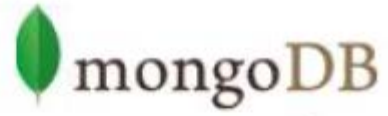


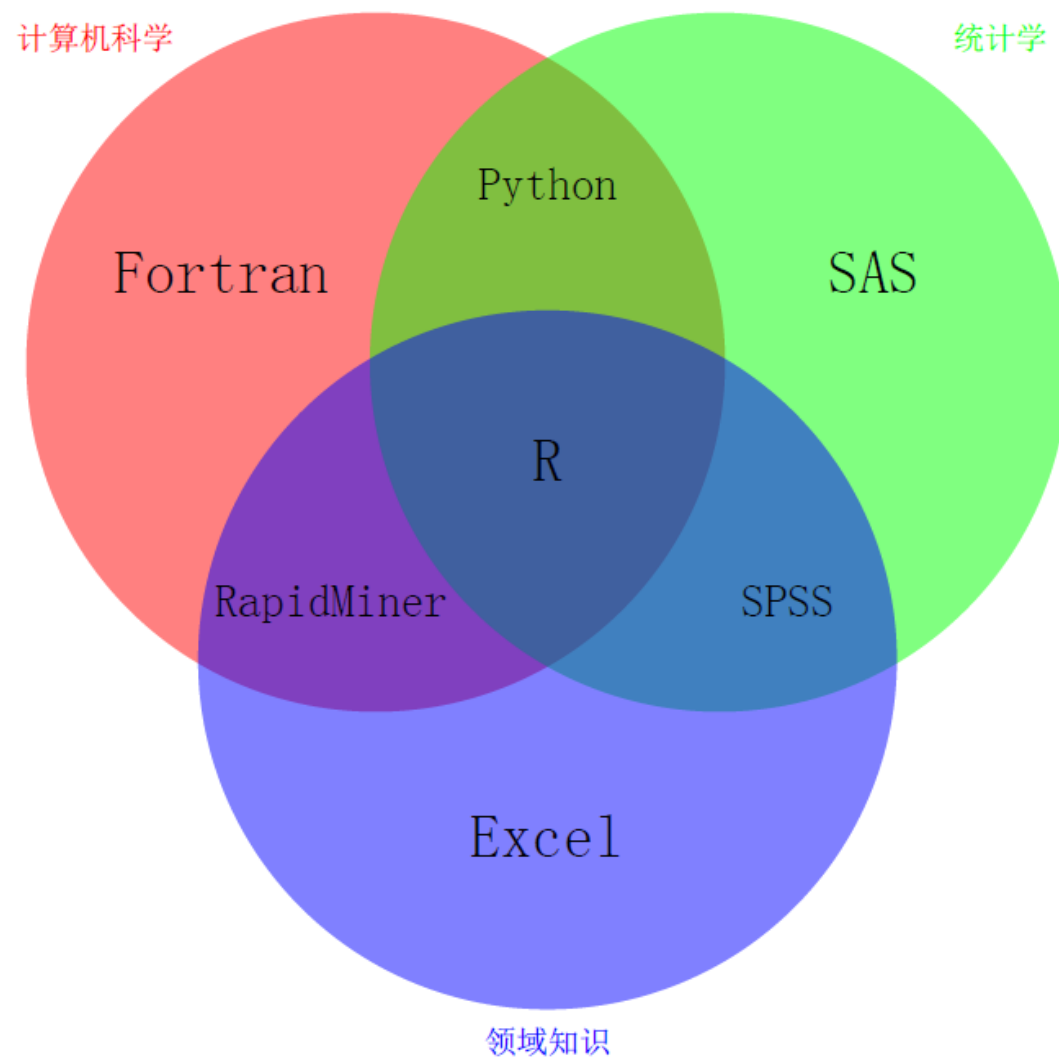


The Data Science Venn Diagram



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法。





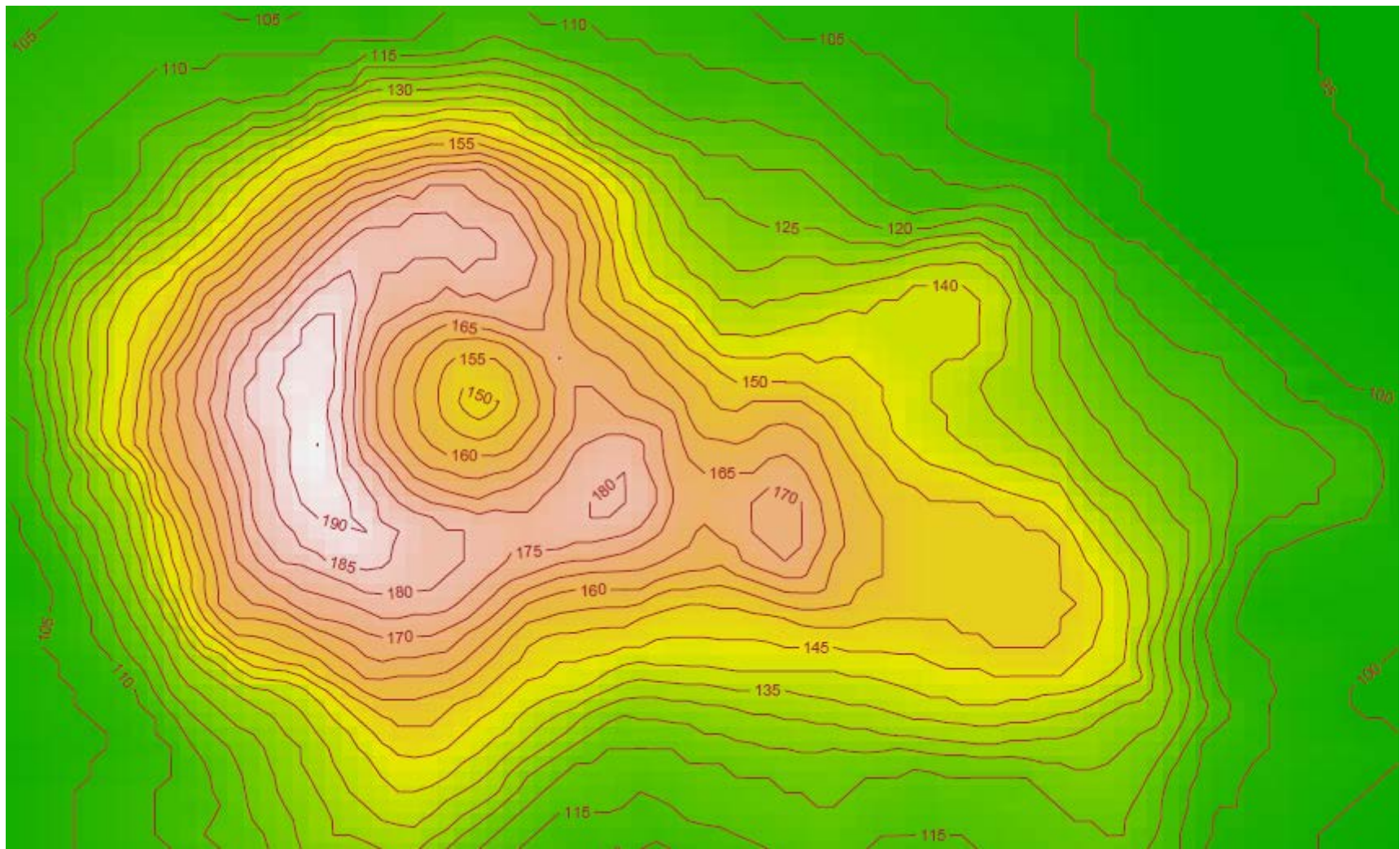
Data Visualization



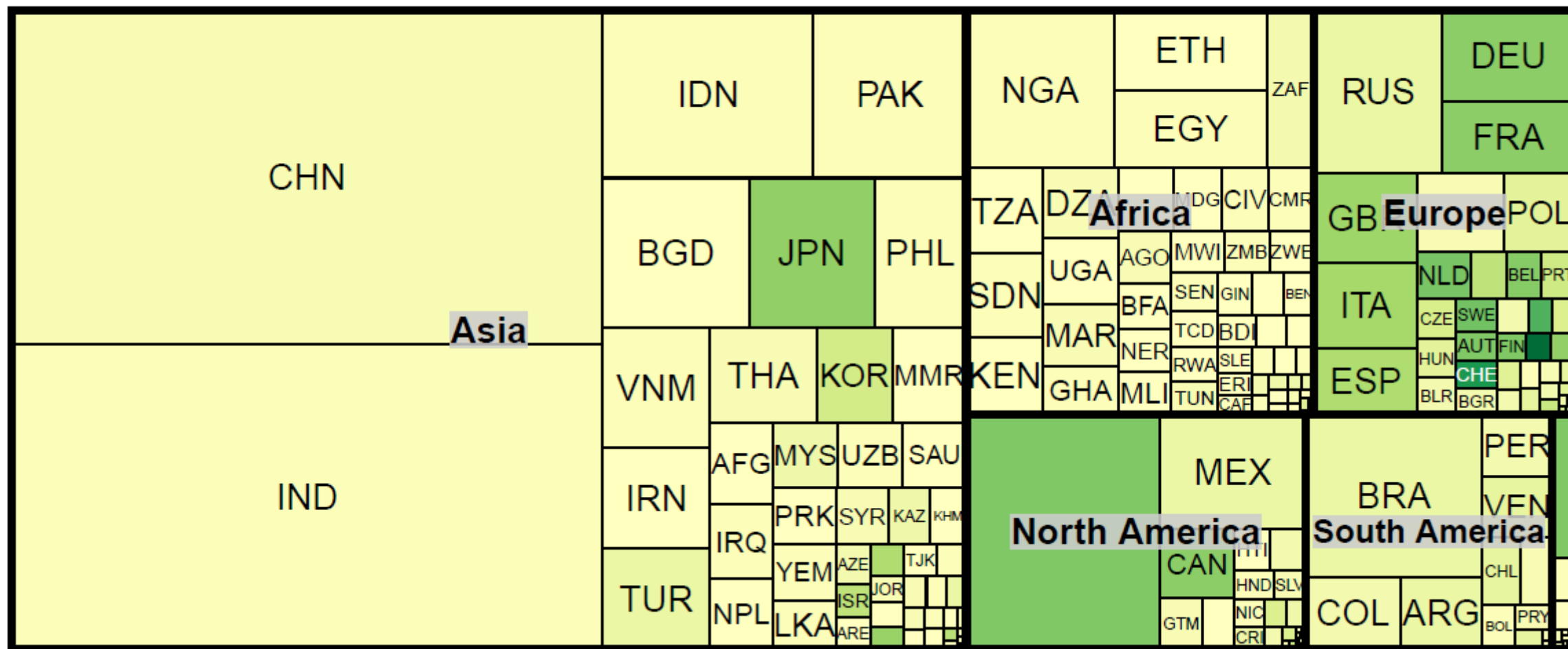
- <http://echarts.baidu.com/index.html>



- <http://vis.pku.edu.cn/wiki/doku.php>



population



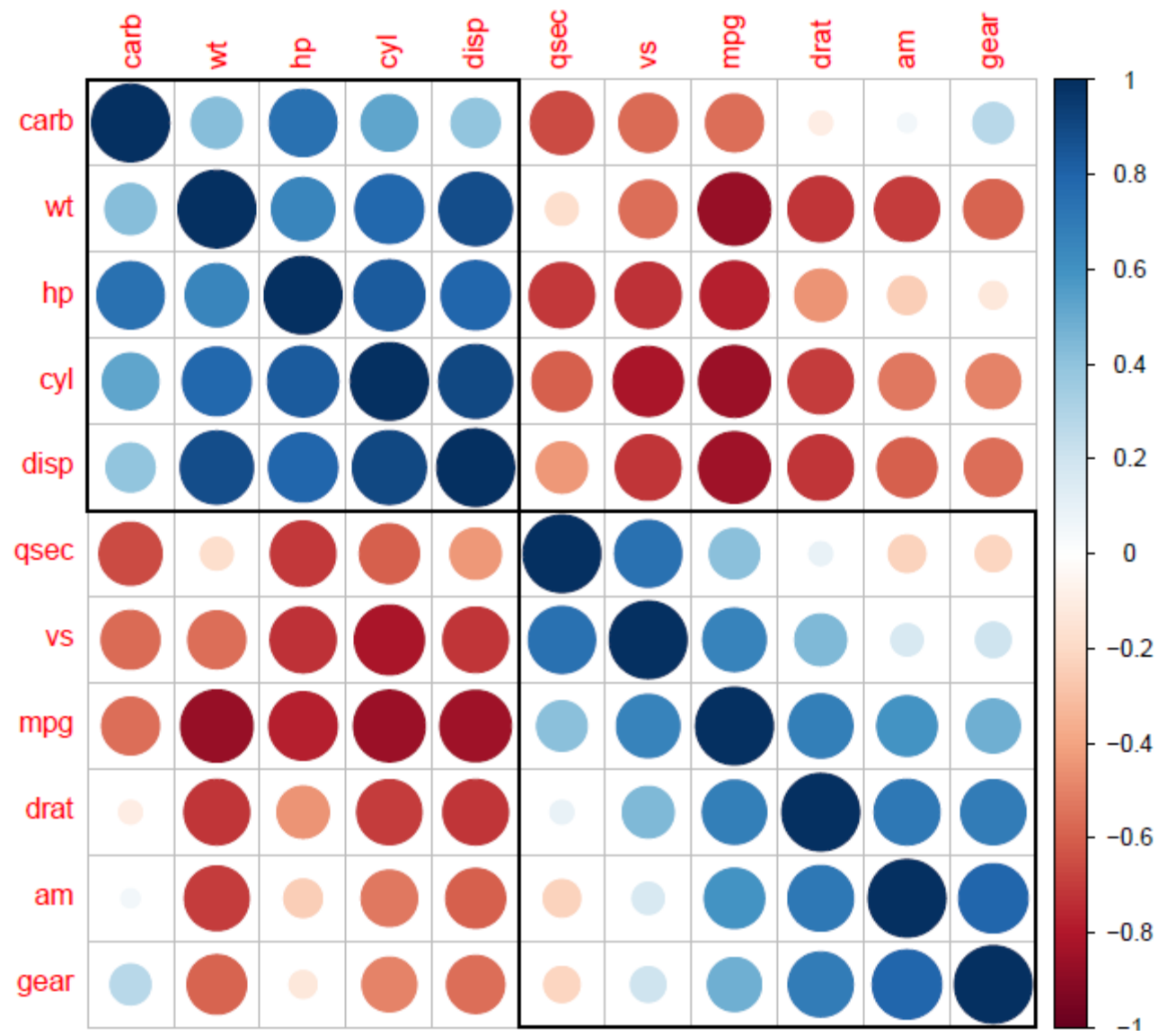


BABA

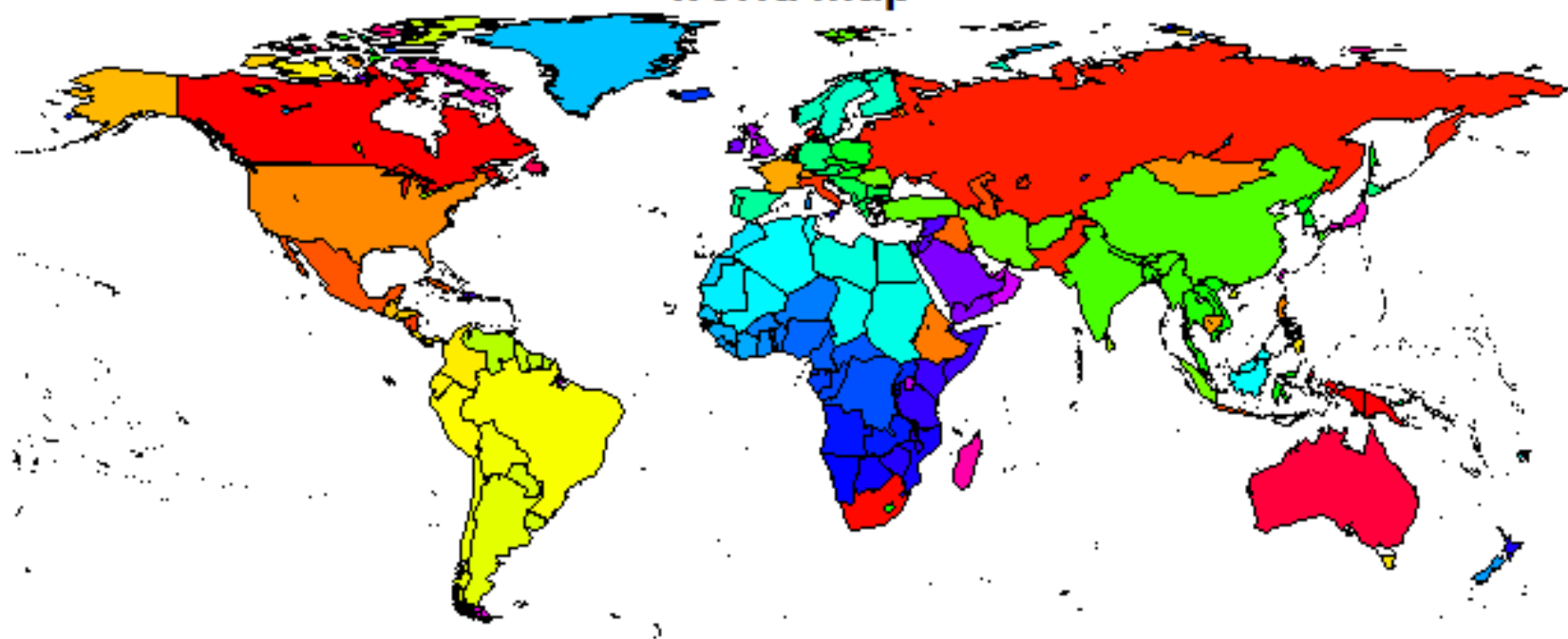
[2014-09-19/2014-11-24]

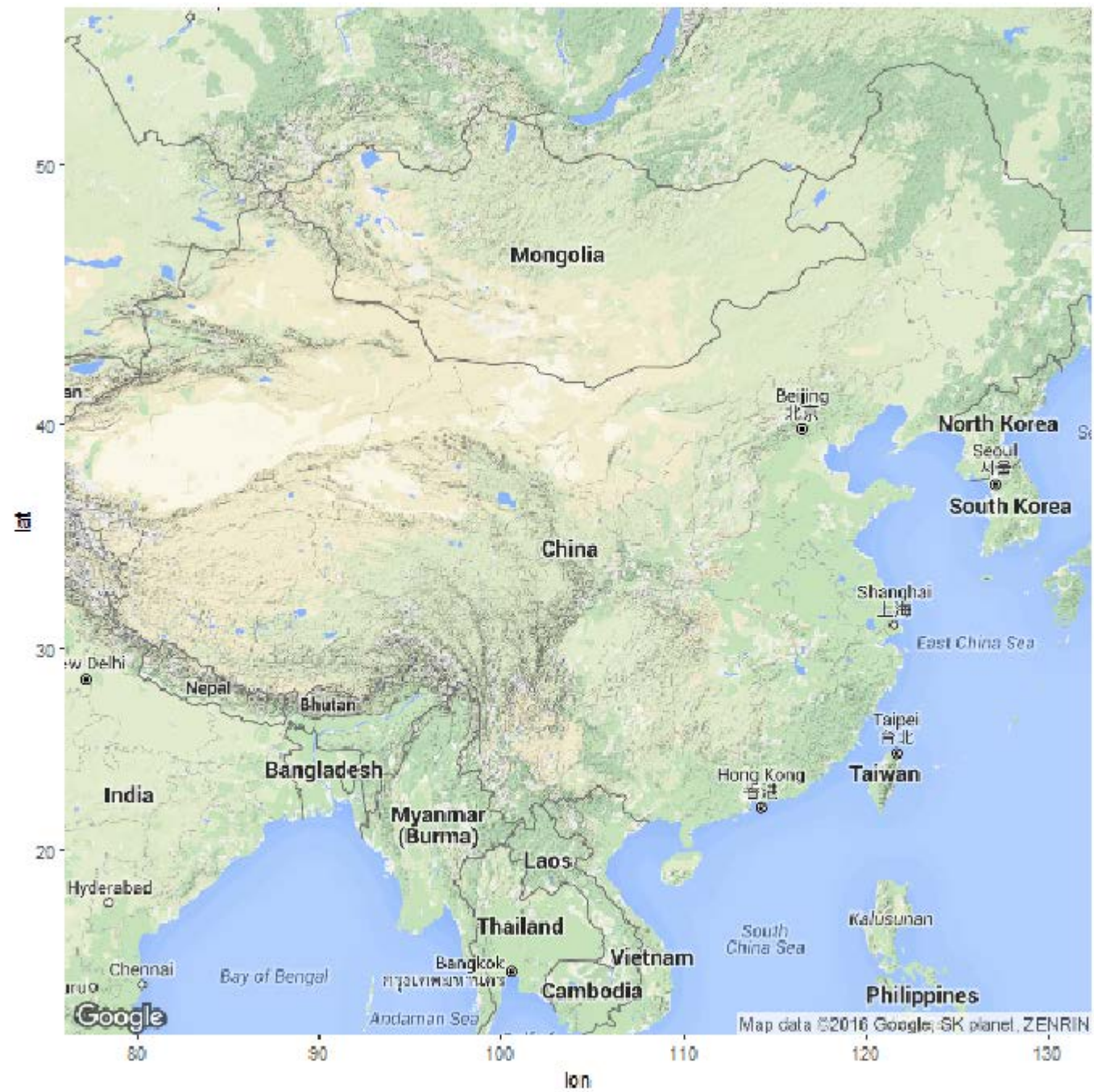
Last 113.92





world map







```
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,
      3), ylim = c(-3, 3), col = "steelblue", pch = 16, cex = 2, xlab = "x",
      ylab = "y")
  }
  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