

# “用 R 包 tidyverse 掌握数据” — 教学大纲

授课人: 李响

网址: 网易云课堂待定

Email: [miao.cai@outlook.com](mailto:miao.cai@outlook.com)

软件包: R, RStudio, tidyverse

**课程目标:** 能够熟练使用 `dplyr` 和 `ggplot2` 进行、数据清洗、合并、分组操作、探索性数据分析和绘图等。

**课程简介:** R 是近期崛起的广泛用于学术研究和数据科学的语言，开源免费并且功能强大，值得各个阶段的高校本科生、研究生、老师以及研究人员进行学习。本课程主要介绍用 `dplyr` 和 `ggplot2` 对原始数据进行管理、清洗、合并、以及探索性数据分析的常用工具和最佳代码习惯。

**参考书籍:**

- Grolemund, Garrett & Hadley Wickham (2018). “[R for data science](#)”. [中文版购买链接](#).
- Hadley Wickham (2016). [ggplot2: elegant graphics for data analysis](#). [中文版购买链接](#).

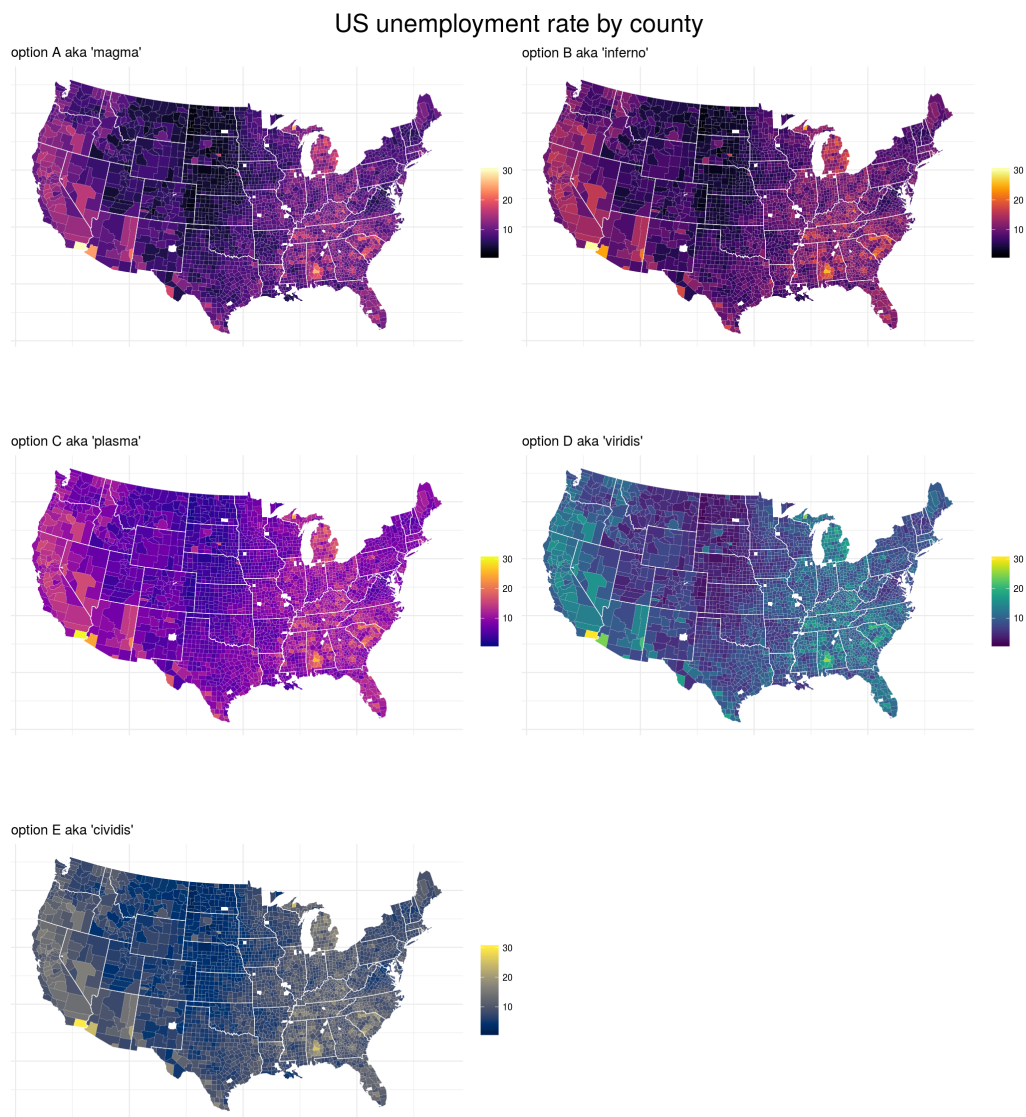


图 1: ggplot2 绘制的地图举例

## 课程安排:

序号	课时名称	课时内容
<b>dplyr 数据管理和清洗部分</b>		
1	R 引入	安装 R 与 RStudio, 设置项目文件, 安装 tidyverse, data.table, fst 等包
2	读写数据	读写 csv, Excel, fst, SQL 等数据文件
3	dplyr (1)	简介 dplyr, 介绍常用函数 select, filter, mutate, case_when 等
4	dplyr (2)	dplyr 分组操作 (group_by, summarize)
5	dplyr (3)	用 dplyr 进行数据聚合 → 连接多个表 (left_join 等)
6	tidyr	用 tidyr 进行长宽数据转换、数据清洗
7	案例实战 (1)	实际数据项目实战
8	案例实战 (2)	实际数据项目实战
9	案例实战 (3)	实际数据项目实战
<b>ggplot2 绘图部分</b>		
1	ggplot2 引入	安装 ggplot2, 介绍其基本语法
2	散点图和折线图	散点图、折线图、LOESS 曲线等绘制
3	条形图	条形图绘制 (单组或多组)
4	箱型图	箱型图、小提琴图绘制 (单组或多组)
5	直方图	直方图和核函数绘制
6	地图	sf, tmap 等包引入和地图
7	散点地图	散点图与地图相结合
8	热力图	热力图
9	颜色自定义	介绍 viridis, colorbrewer 等包
10	主题自定义	介绍 ggthemes 等包
11	图标自定义	自定义图标、标题、图例等元素
12	案例实战 (1)	实际数据项目实战
13	案例实战 (2)	实际数据项目实战
14	案例实战 (3)	实际数据项目实战