

## Lab 6: R 数据可视化综合练习

Lab 6 为 R 可视化的综合性练习，要求撰写一份开放式的数据可视化报告。要成为一名优秀的数据科学家，你必须既是一名“侦探”，又是一个“讲解员”。作为“侦探”，你需要善于提出问题、寻找数据，并且用数据科学的技能（包括数据预处理和可视化技术）从数据中发现有用的信息和知识。作为“讲解员”，你需要利用图表和报告的形式，将信息和知识展示给其他人。

在这次综合性练习里，你需要用一系列的图表和相应文字来展现一个你自选的主题及其分析结果。报告可以是说明性质（介绍一个知识），也可以是议论性质（阐释一个观点）。具体的要求如下。

### 1. 主题

在自然科学（物理、化学、天文学、地球科学、生物、生态学、环境科学等）、社会科学（商业、历史、经济学、心理学、教育学、城市规划、大众传媒、流行病学、考古学等）、工程技术（计算机科学、电子工程、石油化工、电气工程、航空航天、船舶、汽车、交通、土木工程、医学、材料科学等）、文体艺术（体育赛事、音乐、美术、戏曲、影视、博物馆、电竞等）、地理区域（某个国家、省份、城市、街区、公园、旅游景点、河流、山脉、湖泊等）等领域，自选话题作为本次报告的主题。

主题应围绕某个核心内容/观点展开。威斯康辛大学数据可视化课的部分学生选题如下（仅供参考）：

- 美国的龙卷风灾害评估
- 美国古生物类博物馆的藏品概览
- 底特律各社区的犯罪率分布
- NOAA 数据中的气候变化现实
- 美国医学机构的癌症类科研经费变化
- 伊特鲁里亚文明的考古遗迹
- 威斯康辛州观鸟指南
- 英超联赛的票价和球队成绩
- 气温变化与战争频率
- 世界各国国旗的颜色分布（沿海国家的国旗是否有更多的蓝色）
- 雅哈拉湖群的水体化学成分

### 2. 数据

围绕自选的主题，自主寻找相关数据。可以是现成的数据集（比如来自论文、新闻、教材、各种官网、开源数据集等），也可以自主收集和统计。如果使用现成的数据集，需注明来源；如果是自主收集和统计，需简述过程。

### 3. 图表

报告中需包括 **5 组图表**（不可多、不可少）。多画幅图表（图表中包含 A、B、C 等不同部分）按一张图表计算。其中，需要至少包含一张地图和一张时间序列图。所有图表需要使用 R 语言的 `ggplot2` 包进行绘制，绘图的代码储存为 `.R` 格式，作为附件提交。

### 4. 报告文档

报告的格式不限，字数不限（但不宜过长）。所有的图表及其相应的文字部分需紧密围绕一个主题展开。与图片配套的文字部分至关重要（占一半的分数），它不仅要解释图表本身，还应展开相应的介绍或讨论。文字部分写作风格不限，但需通顺流畅，符合逻辑，且和图表密切对应。

学习通上有报告样本可以作为参考标准。

### 提交内容

在学习通的相应位置以附件形式提交报告和附加内容（绘图用的 R 语言文本和数据）。报告以 **PDF** 格式提交，文档请命名为“Lab6+姓名+学号”，比如“Lab6 张三 1010101010.pdf”。绘图用的 R 语言代码请命名为“Lab6+姓名+学号.R”提交，比如“Lab6 张三 1010101010.R”。

截止时间：2024 年 1 月 3 日北京时间 23:59。