

Lec1: 作业题目

1. 文件UN1.txt(安装ahr3包后直接使用attach(UN1))包含了全球193个地方, 大多为联合国成员国, 其各自2001年的人均国内生产总值(PPgdp, 单位: 美元), 以及2000年女性的生育率(单位: 1000人)和国家(地区)名称. 我们的目标是研究生育率在给定PPgdp 下分布的情况.
 - (a) 确定预测变量和响应变量;
 - (b) 以Fertility为纵轴, PPgdp为横轴, 绘制散点图. 简单线性回归拟合此数据看起来合适吗?
 - (c) 以 $\log(\text{Fertility})$ 为纵轴, $\log(\text{PPgdp})$ 为横轴, 绘制散点图. 简单线性回归拟合此数据看起来合适吗?
2. **Water run-off in the Sierras** 能否利用以往数据预测南加州未来的水供应量? 影响水供应量的一个因素是径流, 如果径流可以被预测, 则工程师, 决策者等就可以更效率的制定其工作. 文件water.txt收集了Sierra Nevada山脉6个地点43年的降水量(标记为APMAM, APSAB, APSLAKE, OPBPC, OPRC, and OPSLAKE), 以及靠近加州Bishop的一个地点的径流量. 利用此数据绘制两两变量之间的散点图, 从这些散点图中可以得到什么信息?