

第八章课后作业

1. 读取外部 csv 格式的数据文件 GDP.csv, 去掉存在缺失数据的那些行, 再把剩余的数据以 csv 格式存储到相同的文件夹中的 GDP_new.csv 文件中。
2. 读取外部 csv 格式的数据文件 GDP_new.csv, 以 Year 为自变量, CPI 为因变量画图, 线设置为实线, 点为实心点, 颜色设置为绿色, 在纵坐标 y=1 处画水平线。添加坐标轴标签, 横坐标标签为“Year”, 纵坐标标签为“CPI”, 图形标题为“Figure”, 并利用 legend 添加图例介绍。
3. 读取外部 csv 格式的数据文件 GDP_new.csv, 提取前 30 行数据, 以 Kapital 为自变量, GDP 为因变量, 画散点图, 散点大小设置为 15, 形状设置为“*”, 颜色设置为红色。添加坐标轴标签, 横坐标标签为“Kapital”, 纵坐标标签为“GDP”, 图形标题为“The scatter figure of GDP and Kapital”, 横坐标范围设置为 0~2000, 纵坐标范围设置为 0~8000。
4. 读取外部 csv 格式的数据文件 GDP_new.csv, 提取 KR, 分别绘制 KR 的频数直方图 (bar 的个数为 6, 颜色设置为绿色, X 轴标签为“KR”, 图形标签为“frequency”)、频率分布直方图 (bar 的个数为 6, 颜色设置为红色, X 轴标签为“KR”, 图形标签为“PDF”) 和累计分布直方图 (bar 的个数为 6, 颜色设置为粉色, X 轴标签为“KR”, 图形标签为“CDF”)。
5. 已知某班同学的身高数据, 分布在 150-160, 160-170, 170-180 和 180 学生人数分别为 5, 28, 35, 4, 绘制有阴影的饼图, 每一块离开中心距离分别为 0, 0.1, 0, 0.1, 颜色分别设置为黄色、绿色、蓝色和粉色, 每一块饼图外侧显示的标签分别为“150-160, 160-170, 170-180, 180”, 饼图半径设置为 0.8, 从 y 轴正方向顺时针绘制, 饼图内百分比设置为小数点之后两位。
6. 读取外部 csv 格式的数据文件 GDP_new.csv, 提取 CPI, 绘制带有凹槽的箱形图, 显示均值并且用线表示, 箱体垂直摆放, 内部颜色设置为黄色边框为绿色, 图形大小设置为宽 4 高 6, 箱子下面标签为“CPI”, 整个图形的标题设为“The boxplot of CPI”。
7. 假设某项针对男女大学生购买饮用水爱好的调查结果如下表:

	男	女
碳酸饮料	6	9
绿茶	7	4
矿泉水	6	4
其他	2	6
果汁	1	5
总计	22	28

- (1) 画出男生饮用水情况的垂直条形图, 条形宽度为 0.5, 条形内部设置为绿色、边框为黄色, 条形底部加上每种饮料的标签, 添加图形标题“男性购买饮用水情

况的调查结果”

(2)画出女生饮用水情况的水平条形图，条形高度为 0.5，条形内部设置为红色、边框为蓝色，条形旁边加上每种饮料的标签，添加图形标题“女性购买饮用水情况的调查结果”。

(3)画出男生和女生饮用水情况的分组垂直条形图(水平分组条形图)，条形宽度为 0.3，男生的条形图设为蓝色、女生的条形图设为绿色，条形底部(左边)加上每种饮料的标签，y 轴(x 轴)标签设为“购买量”，图形标题设为“购买饮用水情况的调查结果”，并在图形中添加图例介绍。

(4) 将男生数据放在底层，女生数据放在男生数据的上方绘制叠加的垂直条形图(柱状图)，要求底层为蓝色，上层为绿色，并使用反斜杠填充上层柱，条形宽度为 0.5，条形底部加上每种饮料的标签，y 轴标签设为“购买量”，图形标题设为“购买饮用水情况的调查结果”，并在图形中添加图例介绍。

(5) 将男生数据放在左端，女生数据放在男生数据的右方绘制叠加的水平条形图(柱状图)，要求左边为蓝色，右边为绿色，并使用“*”填充右边层柱，条形高度为 0.5，条形左边加上每种饮料的标签，x 轴标签设为“购买量”，图形标题设为“购买饮用水情况的调查结果”，并在图形中添加图例介绍。