# 大数据可视化技术考试题目2

Pandas数据分析与可视化

步骤与要求：

1. 基于students\_score\_1.csv 和students\_score\_2.csv；
2. 将两张宽表转换为长表；
3. 将两张长表合并；
4. 对各科目数据分别进行归一化，使各科目最高分为1，最低分为0，写出归一化公式；
5. 利用Pandas自身可视化API和计算分析API完成下列任务；
6. 算出平均成绩最好的3名同学和他们的成绩，要求不能有缺考和挂科；
7. 画出各科目成绩分布的箱型图；
8. 画出所有成绩分布的直方图；
9. 基于分析选出各课程的课代表人选；
10. 结合归一化之前的数据，用统计和可视化方法分析出最容易考的3门课，和最不容易考的3们课。

------------------------------------分割线------------------------------

上交材料要求：

1. 打印版纸质的实验报告；首页、目录单面打印，报告正文双面打印；
2. 实验报告文档电子版，pdf或word文件均可；目录名：学号姓名
3. 实验代码和数据电子版，.py和.ipynb均可，实验代码与所需数据放在一个子目录中，这个子目录放置在实验报告电子档的目录中；
4. 自评报告纸质版：一张打印，按照评分要求对自己上交的材料进行评分，分析优势和扣分原因；

------------------------------------完结线------------------------------