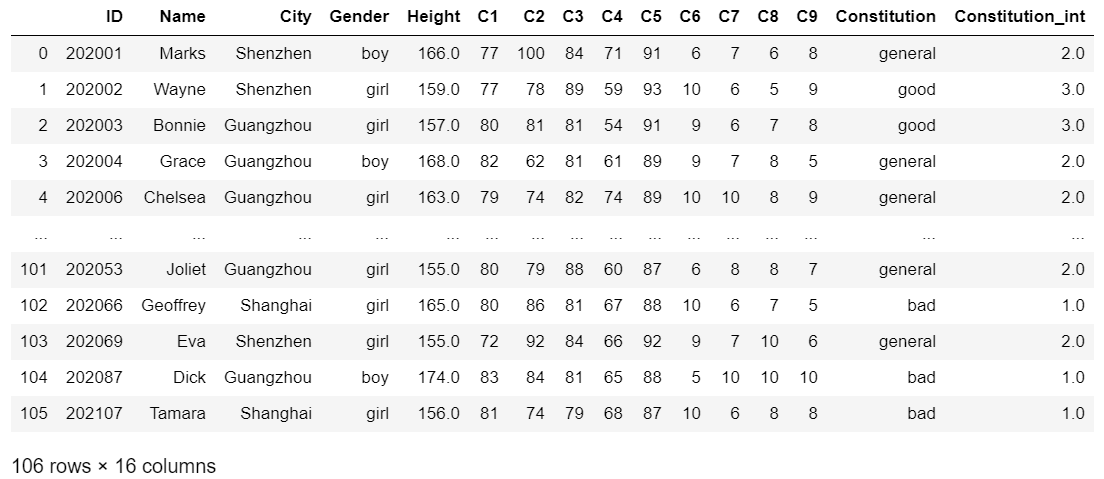
实验二 《数据统计和可视化》

### 实验题目

基于实验一中清洗后的数据练习统计和视化操作，100个同学（样本），每个同学有11门课程的成绩（11维的向量）；那么构成了一个100x11的数据矩阵。以你擅长的语言C/C++/Java/Python/Matlab，编程计算：

1. 请以课程1成绩为x轴，体能成绩为y轴，画出散点图。
2. 以5分为间隔，画出课程1的成绩直方图。
3. 对每门成绩进行z-score归一化，得到归一化的数据矩阵。
4. 计算协相关矩阵，并画出混淆矩阵。
5. 根据协相关矩阵，找到距离每个样本最近的三个样本，得到100x3的矩阵（每一行为对应三个样本的ID）输出到txt文件中，以\t,\n间隔。

### 实验数据

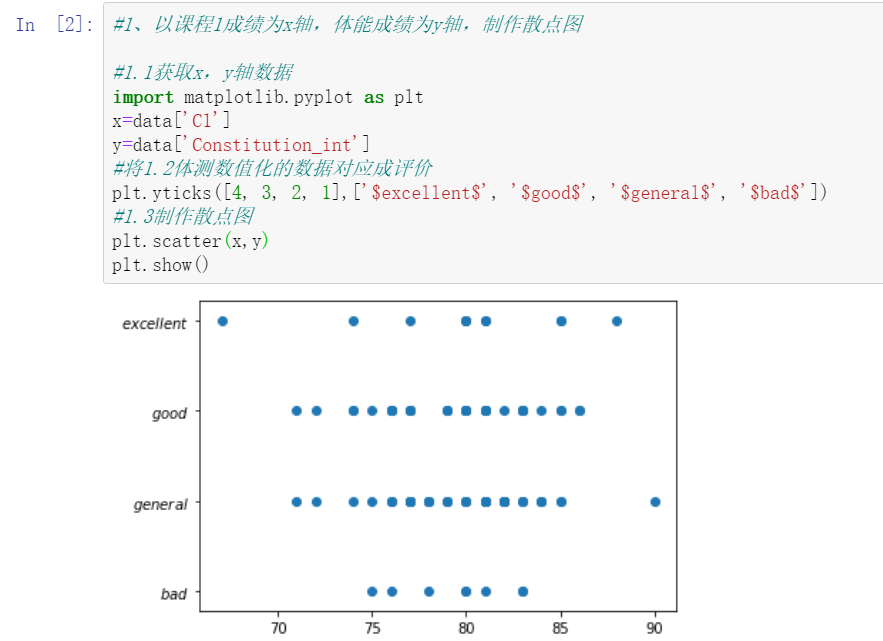


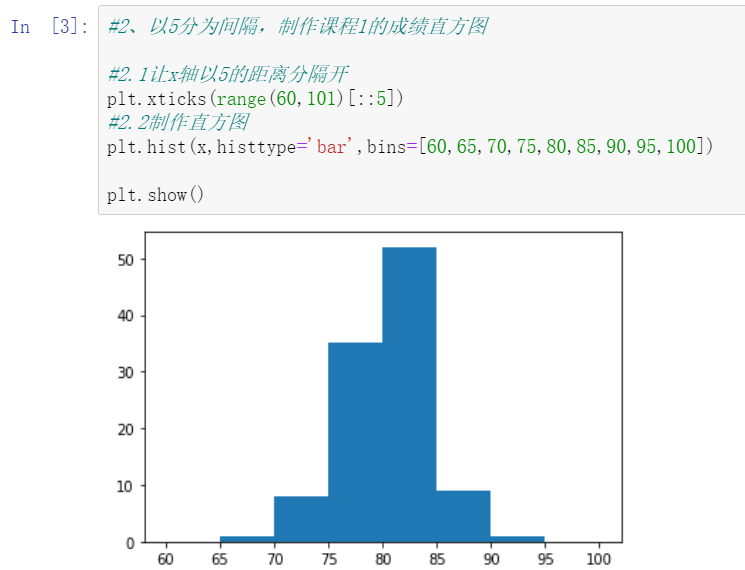
### 实验环境

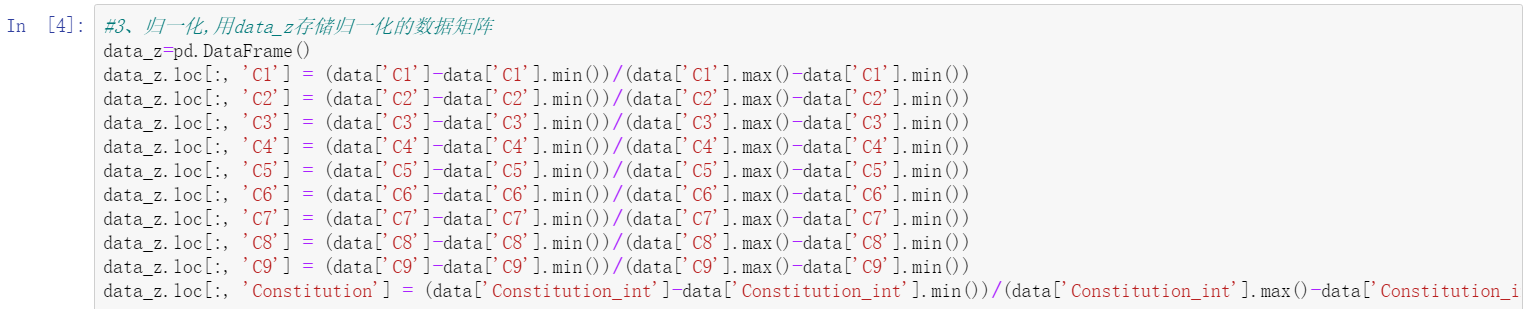
Windows 10、Jupyter

### 实验代码和结果截图



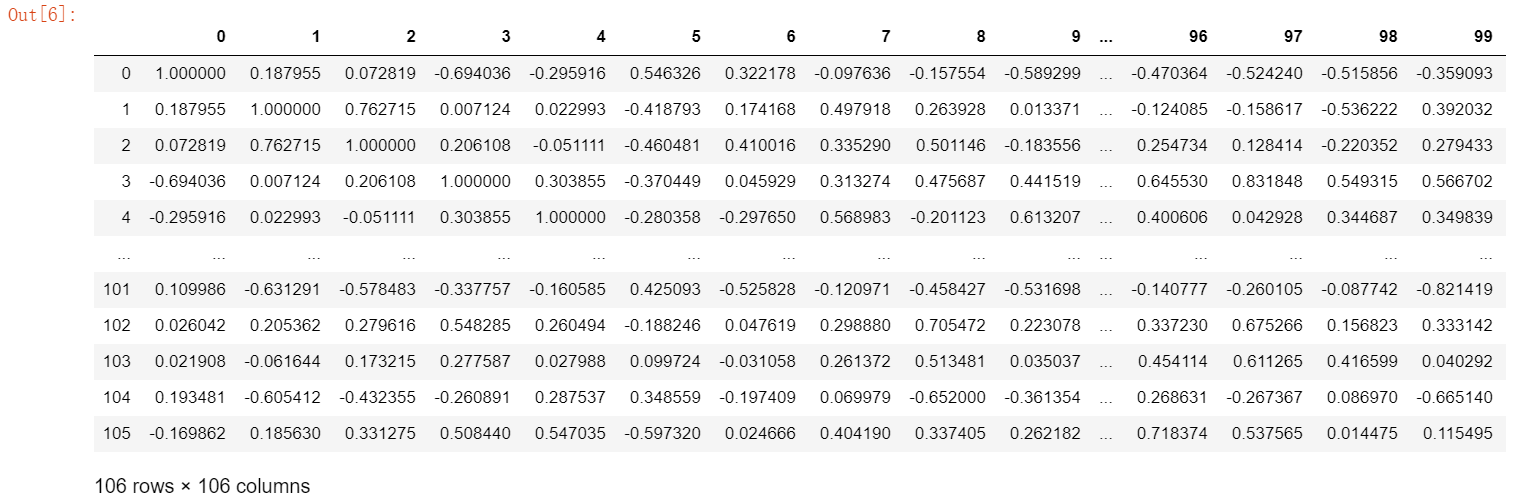


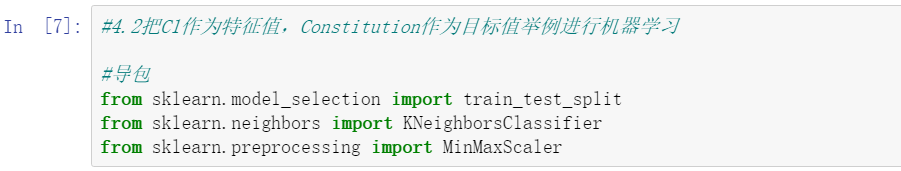


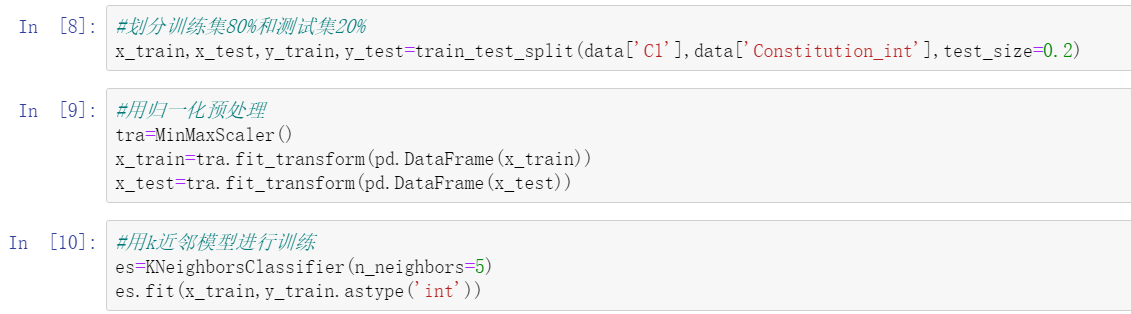




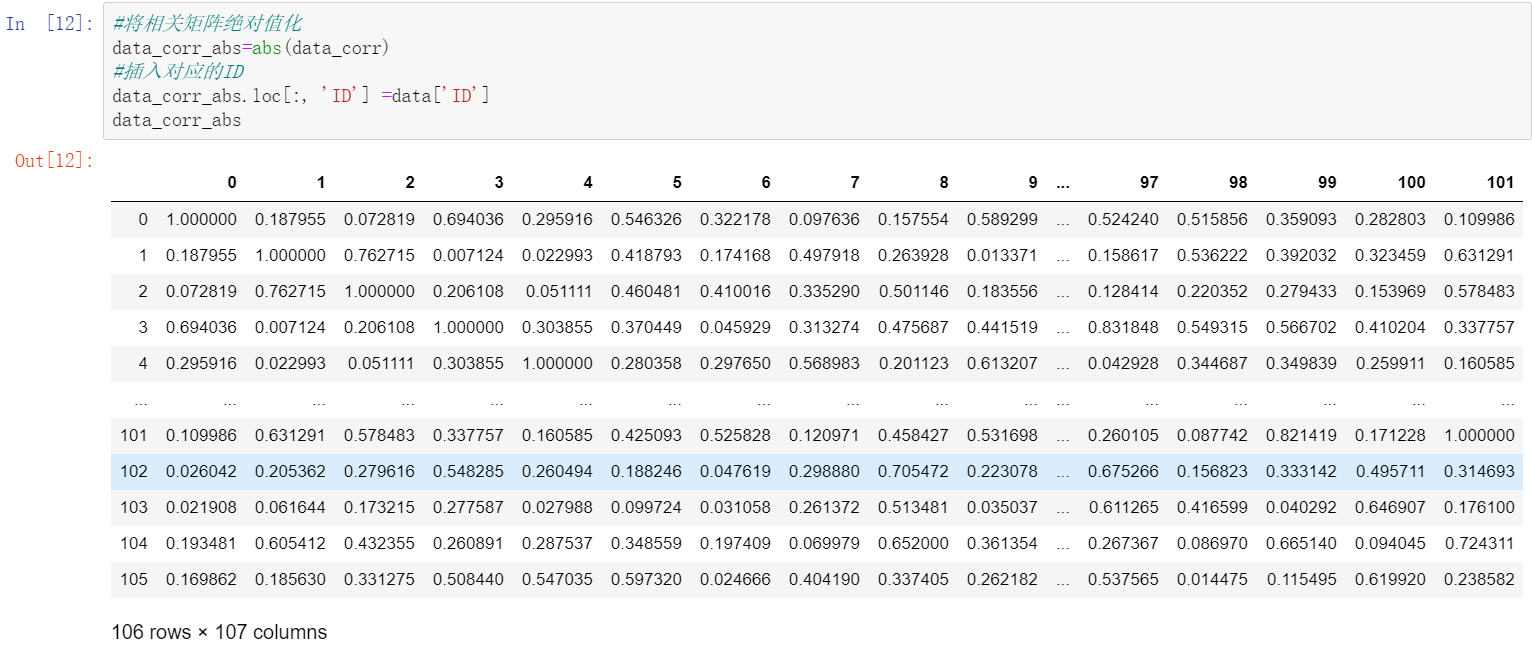


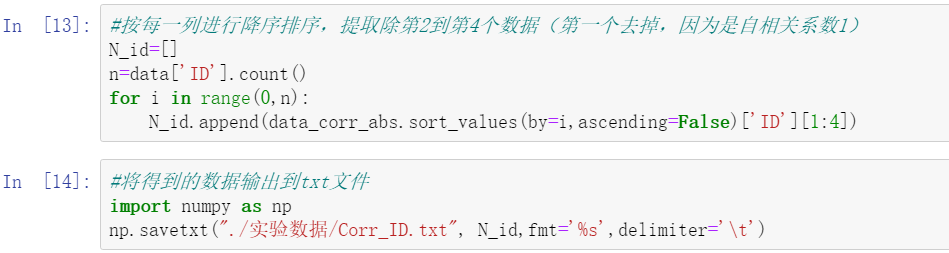


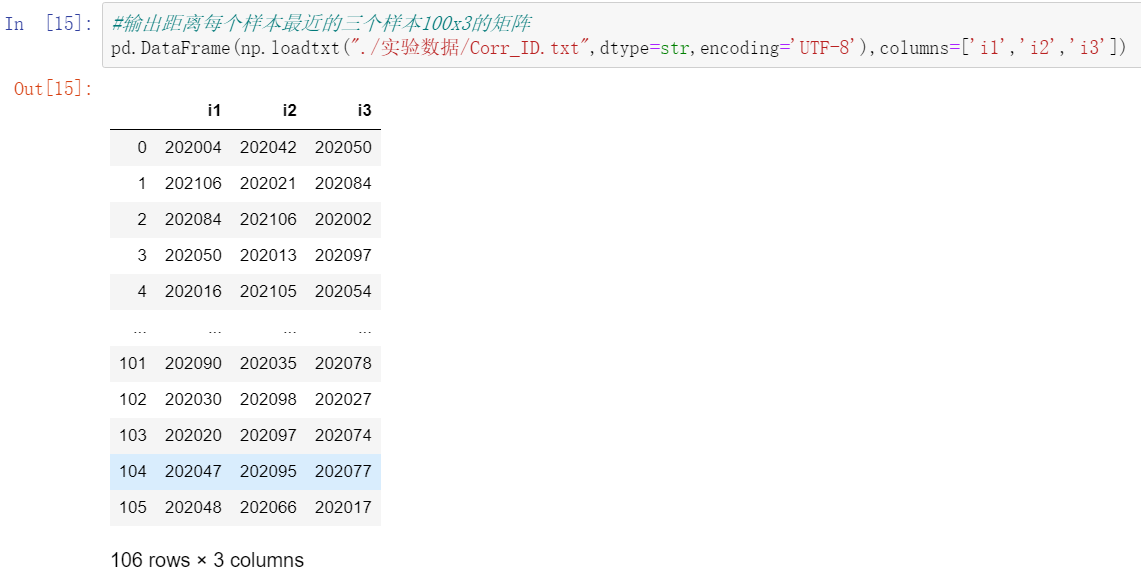












### 实验心得

通过该实验，我对使用数据画图更加熟练；用数据进行简单的机器学习，实现预测，用来比较与现实值的差别，画出混淆矩阵；最后，最重要的是我对数据的读取和使用有了更深刻的理解。