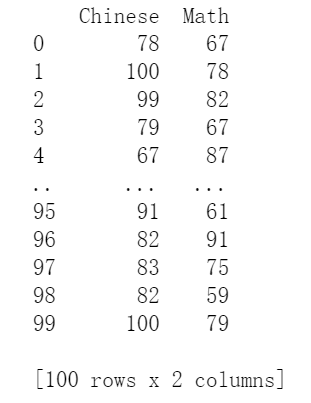
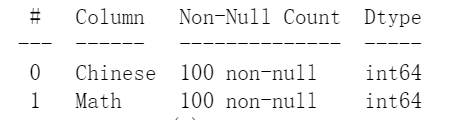
学生成绩数据分析

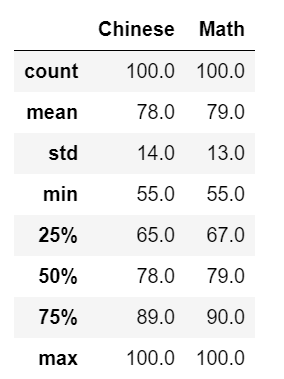
拿到的学生成绩有100条的数据。数据维度为2，两列类型都为整型。



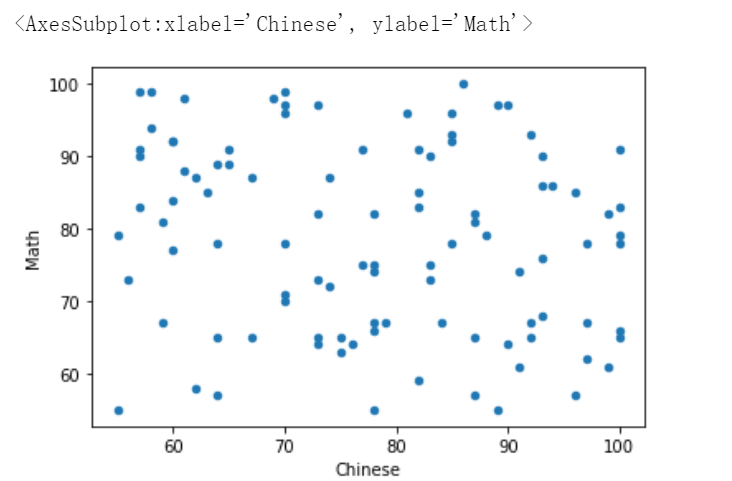
数据中没有缺失值：



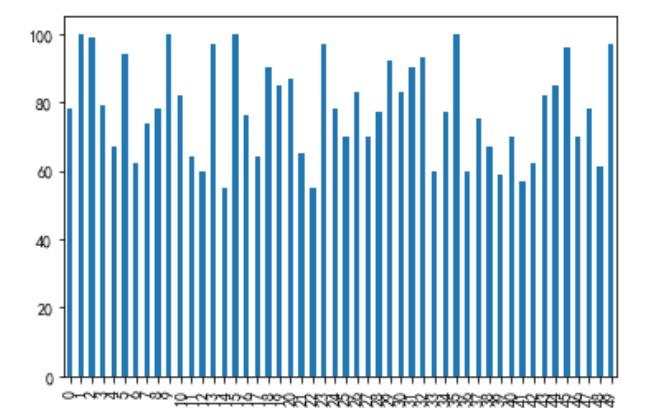
整个学生成绩数据共有100条有效值，语文成绩与数学成绩的平均分分别为78分与79分。数据的标准差为14与13。最高分与最低分均为100分、55分。



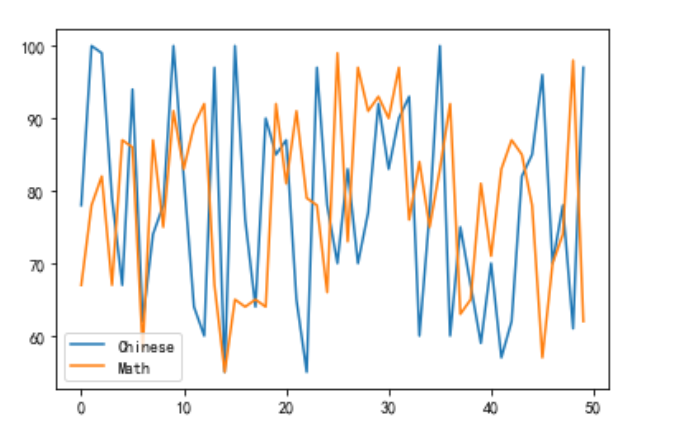
学生成绩数据的散点图较为分散，部分学生比较偏科，单一科目接近满分而另一科目处于及格边缘。



前五十位学生的语文成绩柱状图：



前五十位学生成绩的折线图：



由此可见数据间的距离还是相对较大的，可通过对数据进行归一化处理。

归一化：１）把数据变成(0,1)或者(1,1)之间的小数。主要是为了数据处理方便提出来的，把数据映射到0～1范围之内处理，更加便捷快速。２）把有量纲表达式变成无量纲表达式，便于不同单位或量级的指标能够进行比较和加权。归一化是一种简化计算的方式，即将有量纲的表达式，经过变换，化为无量纲的表达式，成为纯量。

通过归一化处理后最优解的寻优过程明显会变得平缓，更容易正确的收敛到最优解。