**知彼知己——成绩评价与预测**

**姓名：郑茜 班级：18大数据1 学号：1810311130 成绩**

**0 引 言-问题的提出**

1. 采集18大数据1-3班上学期每个人各科成绩，学分绩点、成绩排名（大数据专业）

请根据成绩数据，对本专业同学的学习整体情况进行分析说明；

1. 请根据成绩数据，建立数学模型，全面、客观、合理的评价本专业同学的学习状况（自己成绩排名及分析）；
2. 试建立预测模型，预测本专业同学本学期的学习情况。（自己成绩排名预测及分析）；
3. 根据你的模型分析，试给学院写一篇短文，提出对本专业学生学习状况评价的建议和改进方案。

**1 模型与算法**

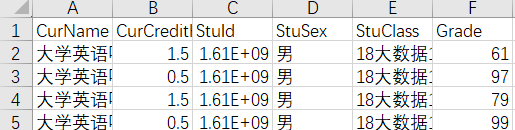
邻近算法，或者说K最近邻(kNN，k-NearestNeighbor)分类算法是[数据挖掘](https://baike.baidu.com/item/%E6%95%B0%E6%8D%AE%E6%8C%96%E6%8E%98/216477" \t "_blank)分类技术中最简单的方法之一。所谓K最近邻，就是k个最近的邻居的意思，说的是每个样本都可以用它最接近的k个邻近值来代表。

人工神经网络（Artificial Neural Networks，简写为ANNs）也简称为神经网络（NNs）或称作连接模型（Connection Model），它是一种模仿动物神经网络行为特征，进行分布式并行信息处理的算法数学模型。这种网络依靠系统的复杂程度，通过调整内部大量节点之间相互连接的关系，从而达到处理信息的目的。

**2 成绩分析与预测**

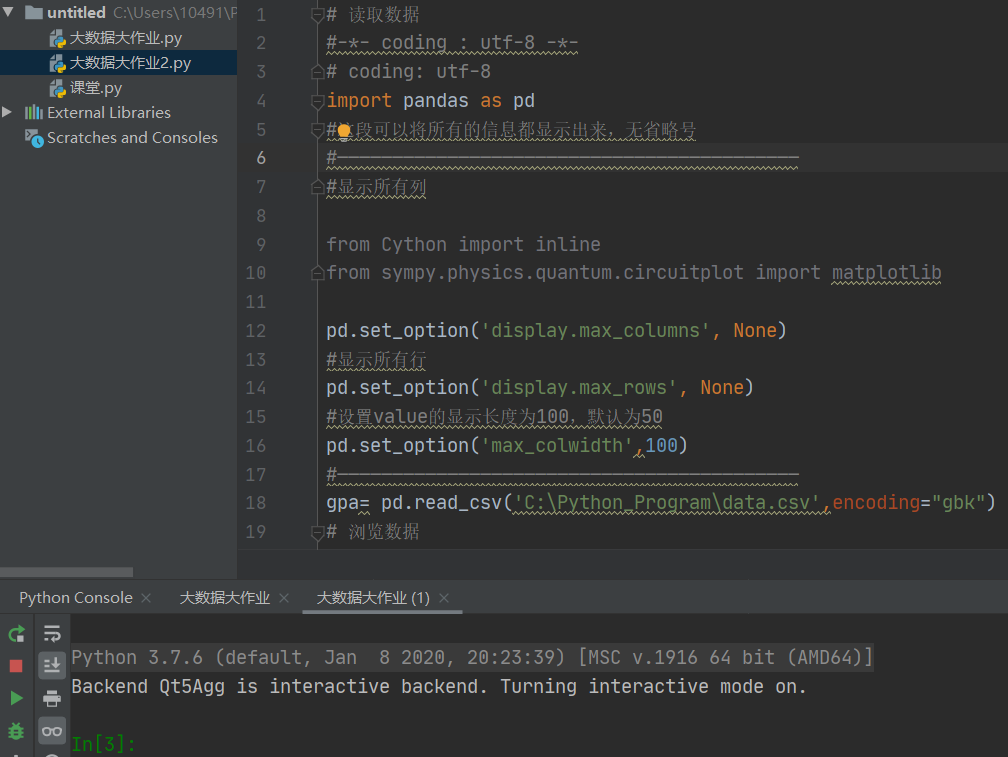
**整体分析**

1. 首先将得到的成绩表进行简单的处理，将对本次分析无用的信息删除，得到以下表

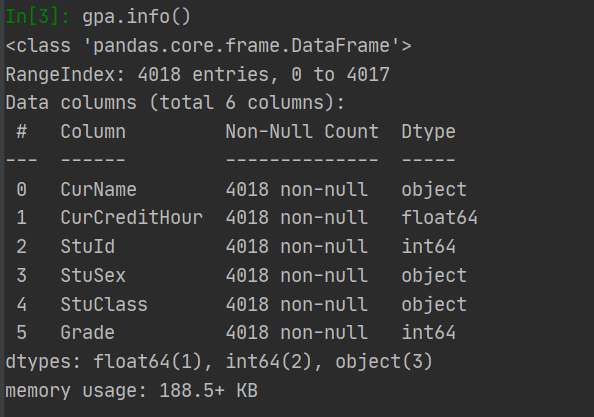


这里用StuId代替姓名的作用

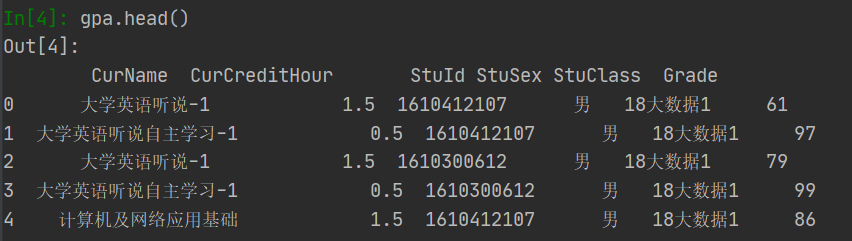
1. 首先导入数据



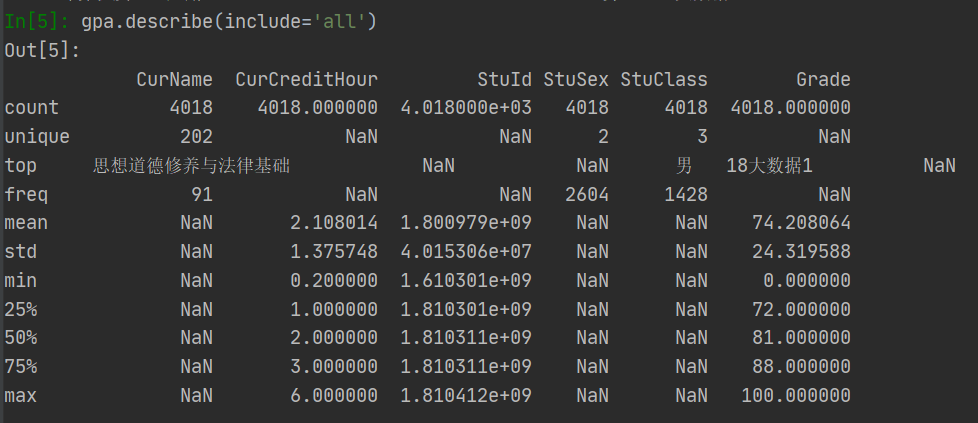
1. 浏览一下数据集的基本信息



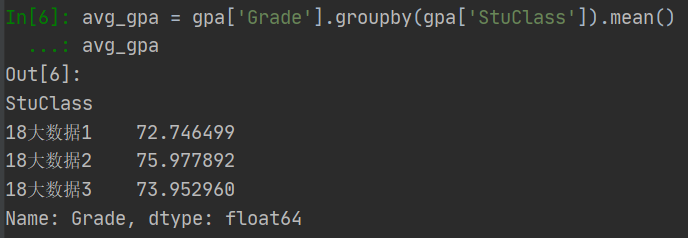
1. 查看前五行



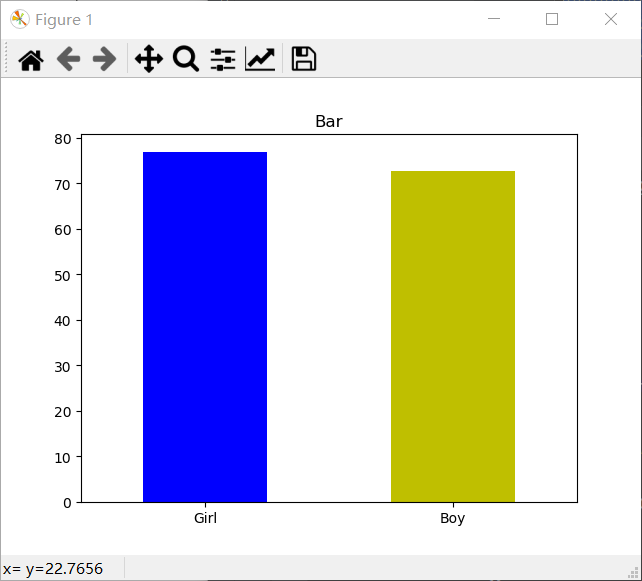
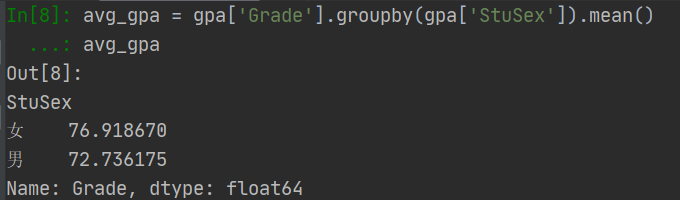
1. 基本统计量



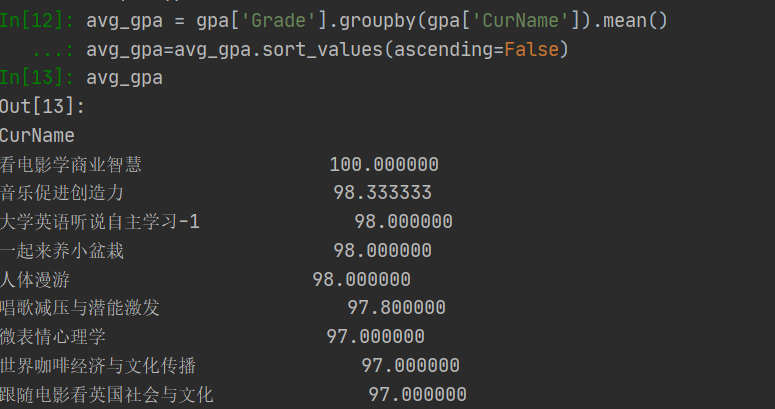
1. 我们首先查看一下三个班的平均值



1. 男女的平均值

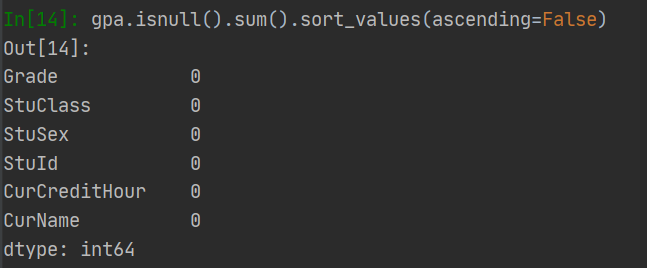


1. 查看以课程名分组的平均成绩



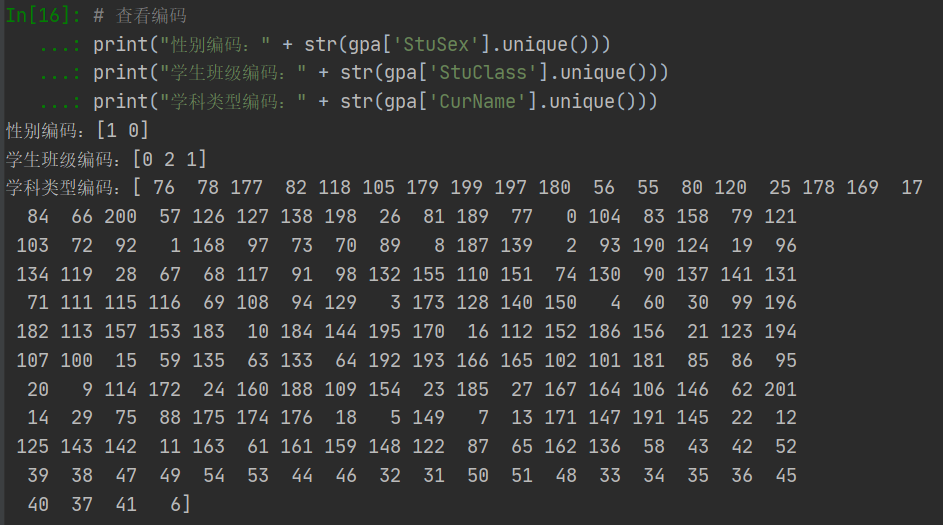
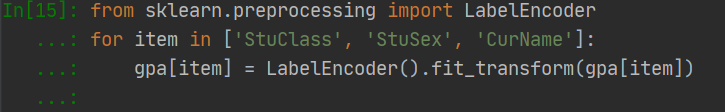
1. 数据预处理

统计数据中的缺失值：

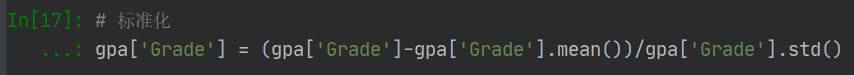


无缺失值，所以不用进行缺失值插补

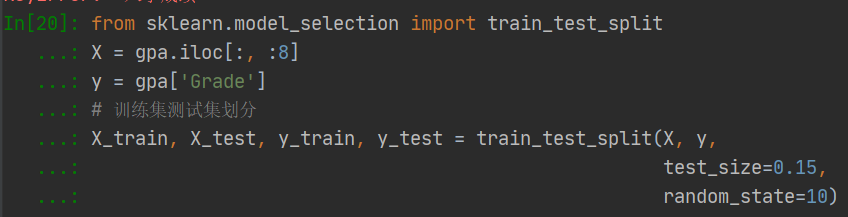
将班级，性别和课程名进行特征编码：



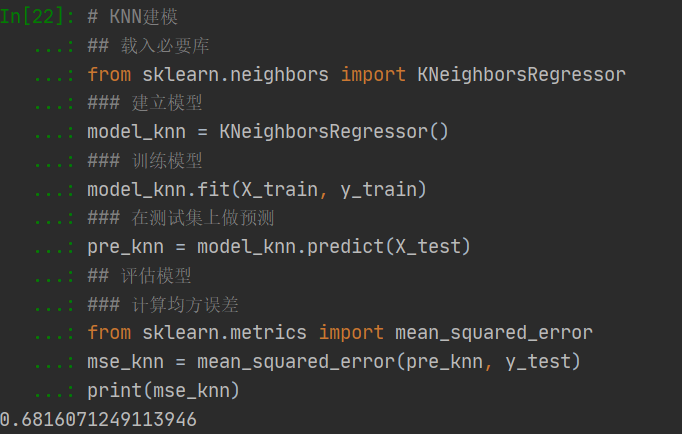
对入学成绩这个特征进行标准化：



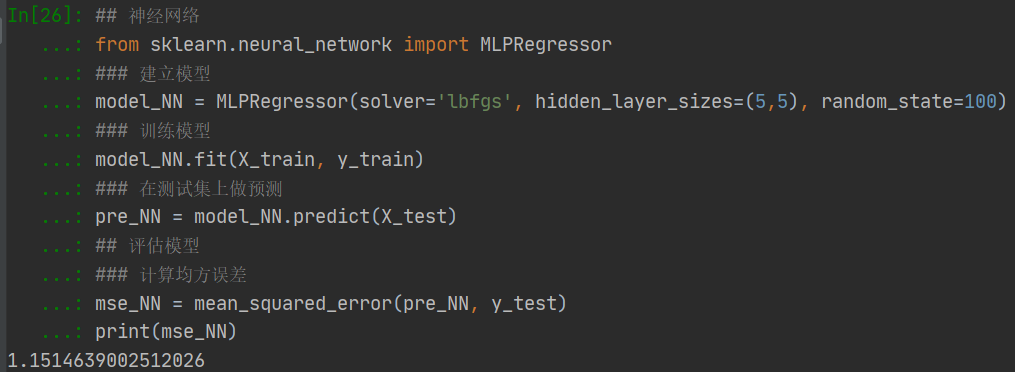
数据划分：



KNN模型：

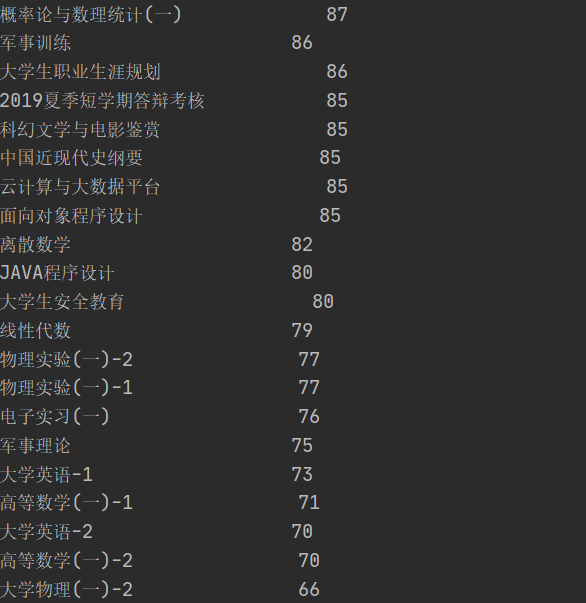
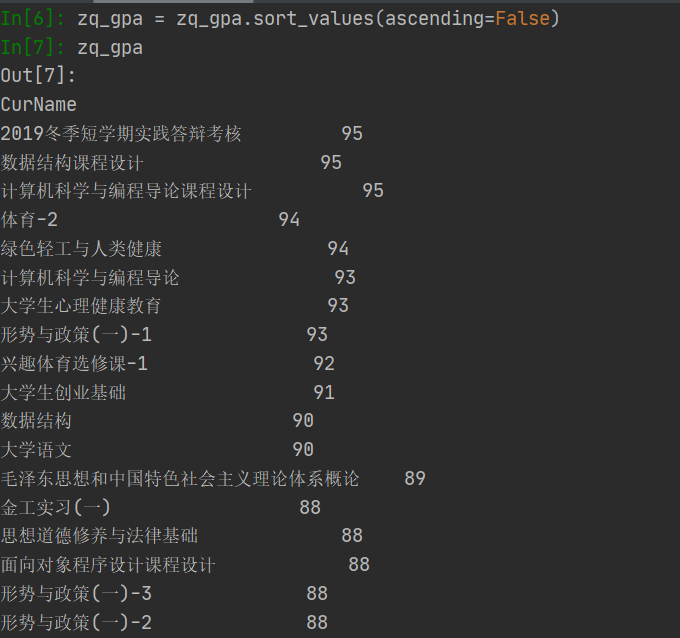


神经网络：

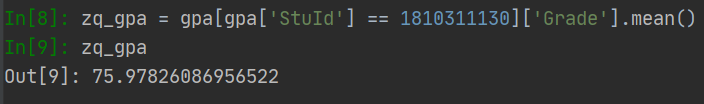


**个人成绩分析**

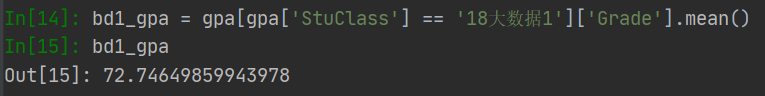
先查看自己的各科成绩：



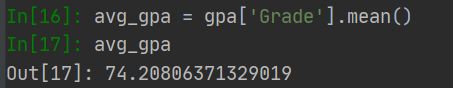
再来查看自己的总体平均分：



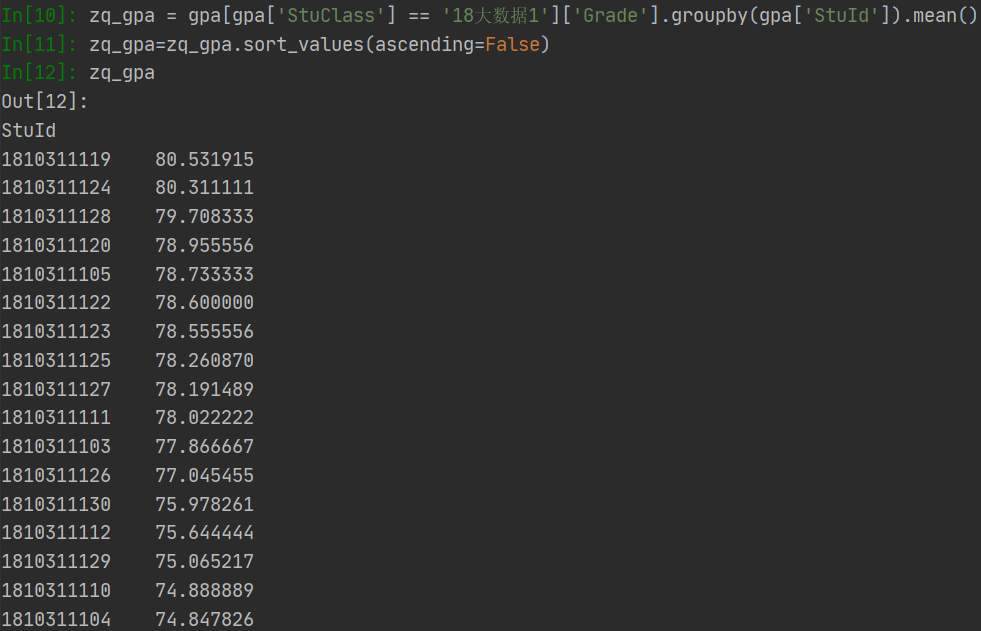
班级平均分：



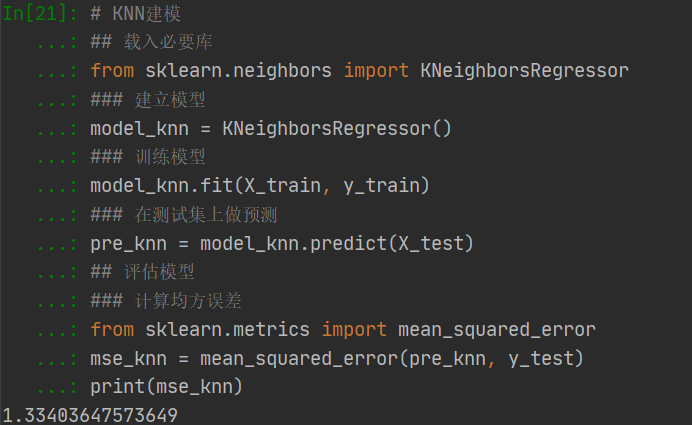
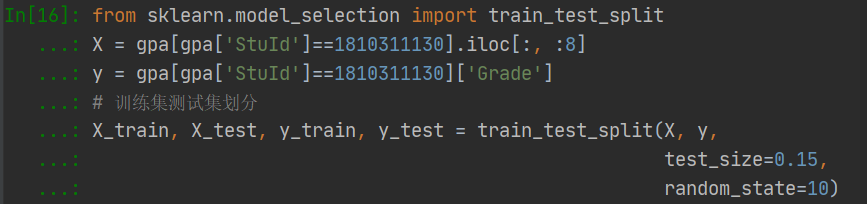
年级平均分：



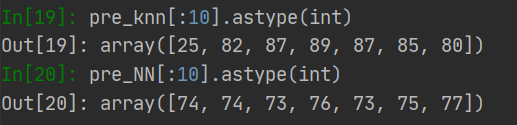
在班级中的大致排名：



预测个人成绩：



预测值：

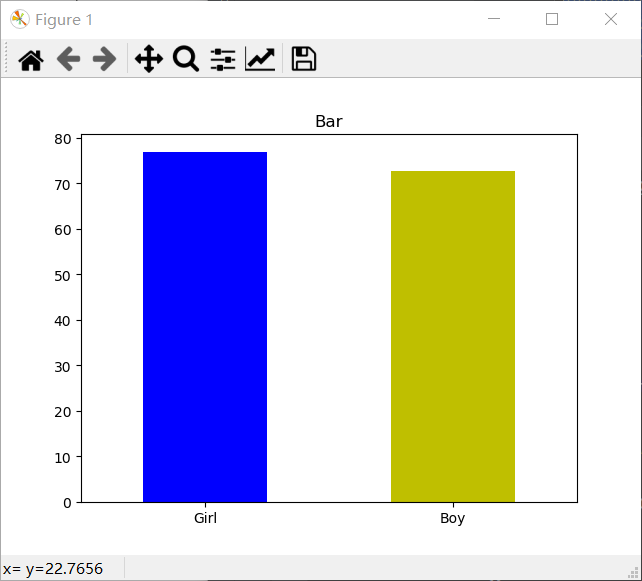


**3 结果分析**

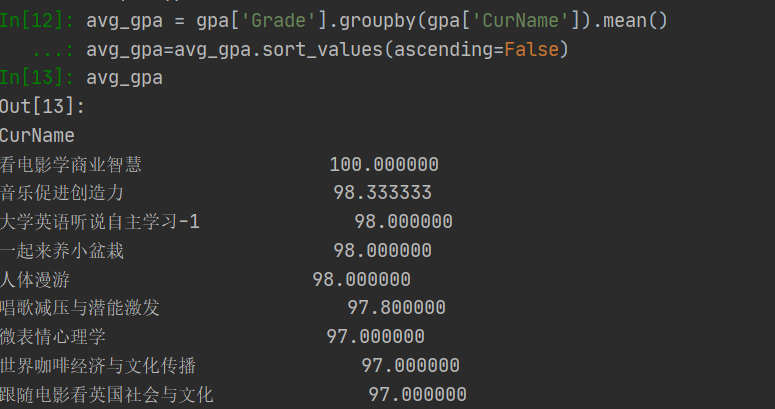
1. 从整体来看大数据2班的成绩较好，而一班的平均分比二班的低了三分的差距，但是这次求平均值是直接来求，并不是求几点那样来求，可能会存在一定的差异，但是整体来说大数据三个班的成绩都差不多，有着三分的差异。



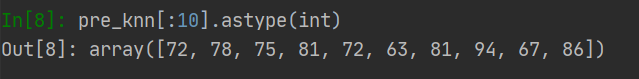
1. 从男女性别来看，女生的平均成绩是要比男生高的，也希望男生继续努力

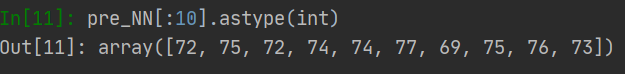


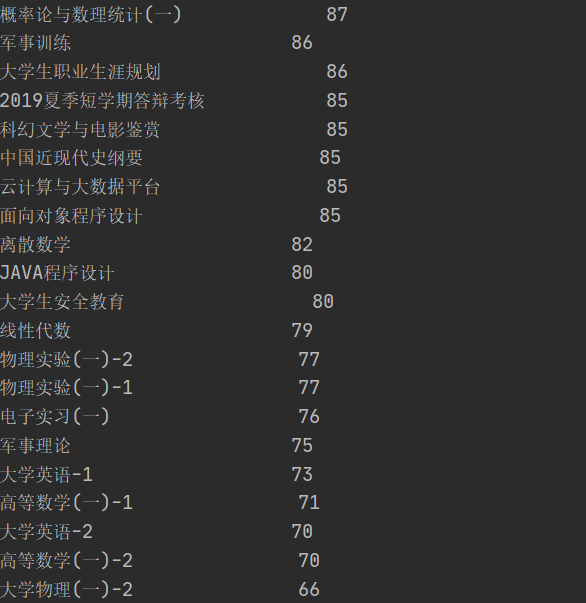
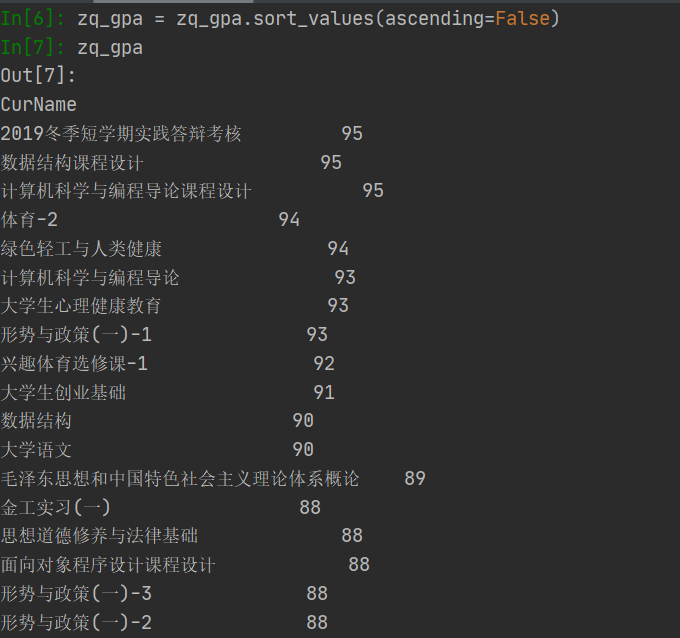
1. 从全部课程的平均分表明，最高的大多都是选修课，或是形势与政策和课设



1. KNN模型预测出来的值：

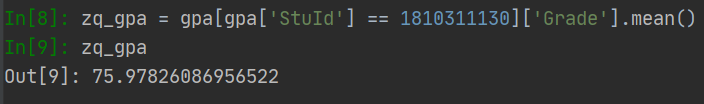


1. 神经网络预测出来的值：
2. 个人成绩分析
3. 先查看自己的各科成绩：

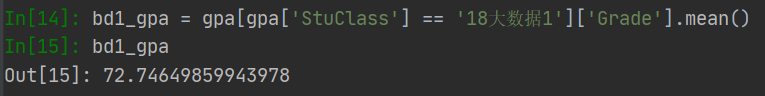


可以知道自己专业课成绩还是可观的，但是一些基础课程，比如高数物理来说还是太低了。

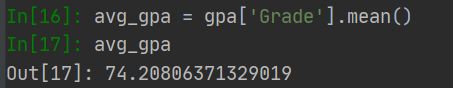
（2）再来查看自己的总体平均分：



班级平均分：

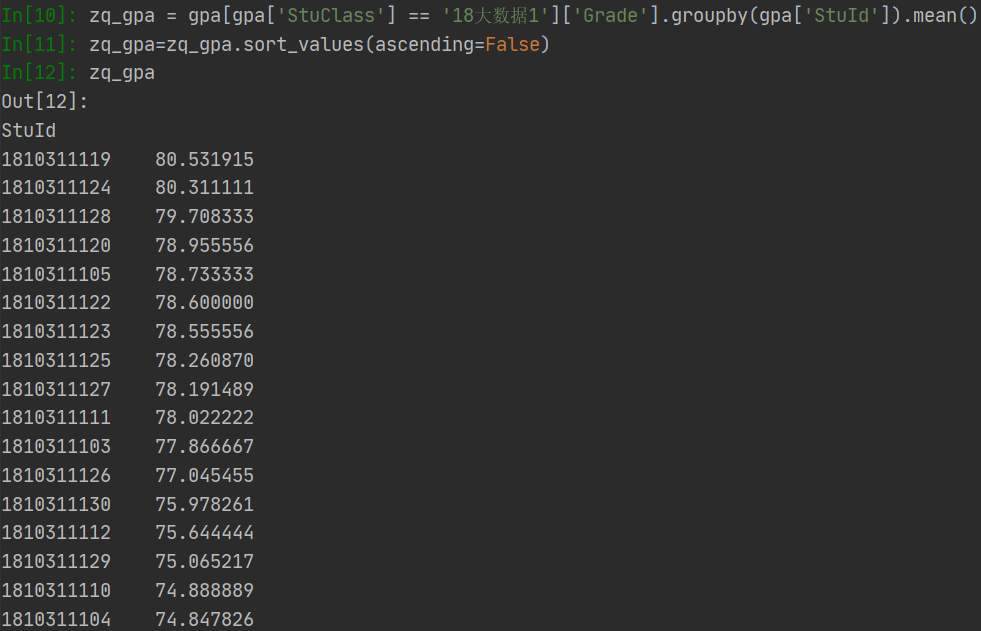


年级平均分：



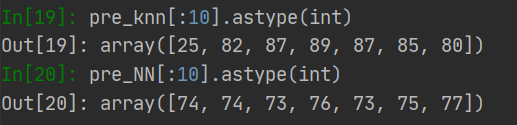
这里可以看出来在平均分上都高于班级和年级的平均分

（2）在班级中的大致排名：



在班级中排名13名，处于中等水平，还需继续努力

（3）预测未来成绩值：



**4 给学院建议**

从分析结果来看学院可以适当地给每个成绩比较靠后的学生进行面谈。

对于一些平均分比较低的基础课程如高数大物等可以进行一些策略提高同学们的分数。

对于专业课而言也要努力提高同学们的积极性，将专业课成绩提升。

**参考文献**

<https://blog.csdn.net/qq_35892623/article/details/82941473>

<https://www.jb51.net/article/140728.htm>

<https://blog.csdn.net/weixin_40283816/article/details/83242083>

<https://blog.csdn.net/weixin_43694361/article/details/99414231?utm_medium=distribute.pc_relevant.none-task-blog-BlogCommendFromBaidu-1&depth_1-utm_source=distribute.pc_relevant.none-task-blog-BlogCommendFromBaidu-1>