**用python进行数据分析**

**1、有关库**

numpy：基于矩阵的数学计算库。

pandas：基于表格的统计分析库，它建立在NumPy之上

scipy：科学计算库，支持高阶抽象和复杂模型。Scipy库依赖于NumPy，它提供便捷和快速的N维向量数组操作。

matplotlib：流行的数据可视化库。

requests：HTTP库，用于网络访问。

jeba：中文分词工具库。

Wordcloud：词云展示第三方库

beautiful soup：HTML和XML解析器

pip install numpy

pip install pandas

pip uninstall 库名

pip show  库名

pip list

**2、进行数据分析**

1、分析过程：导入库，打开文件，增加列（字段）并赋值，显示内容，写入文件

**2、编程实现 （按行求和）**

coding=utf-8

import pandas as pd #导入所需的库并设置别名

import numpy as np #导入所需的库并设置别名

df = pd.read\_csv('grade.csv', encoding='utf-8', header=0) 打开CSV文件

print(df) #显示CSV文件的内容

df["总分"] = df.sum(axis=1) #增加总分列，并按行计算出总和赋值给该列

print(df[["姓名","总分"]]) #显示姓名、总分列的内容

df.to\_csv('sum\_grade.csv', encoding='utf-8',float\_format='%.3f') #将更新内容的CSV文件存储为新的CSV文件 命名为sum\_grade.csv