**实验一**

**学号： 220111028 姓名： 许辰涛**

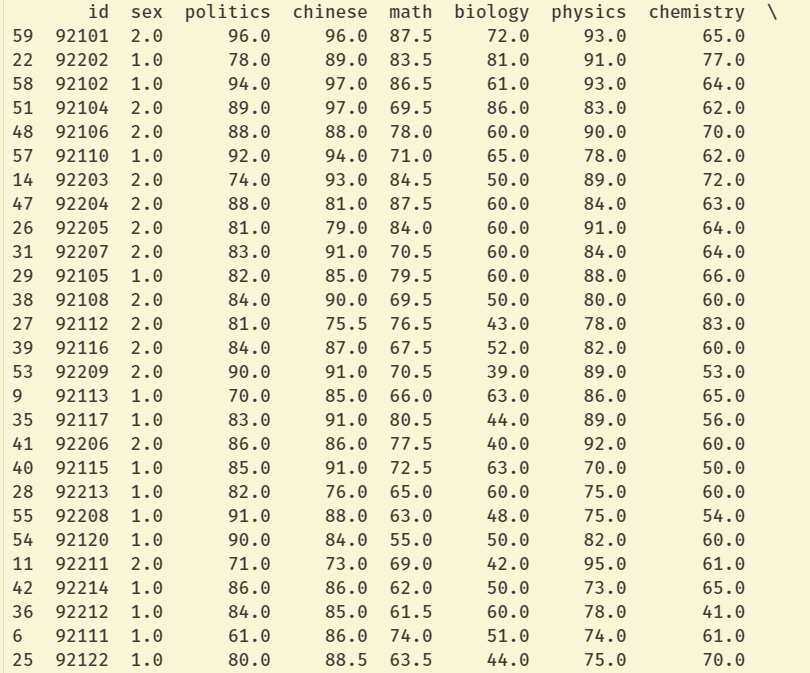
1. **运行截图**（配合截图结果，可以添加一些文字说明）

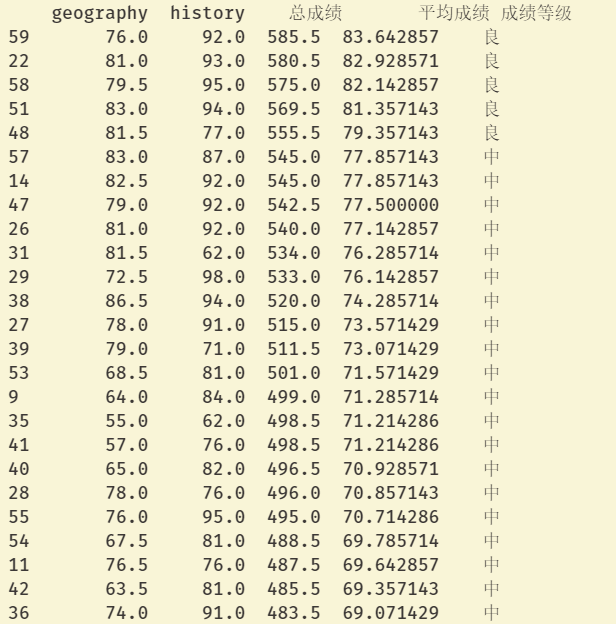
**合并两张表格**

**使用0填充NaN**

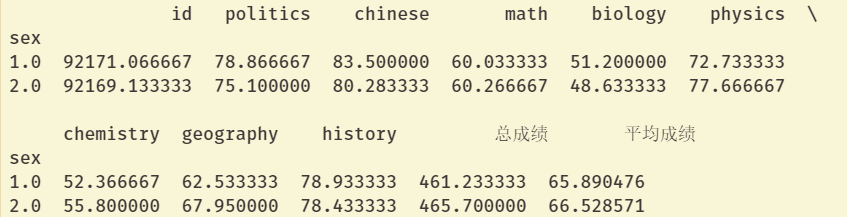
**计算总成绩与平均成绩**

**将数据按总成绩降序排序**

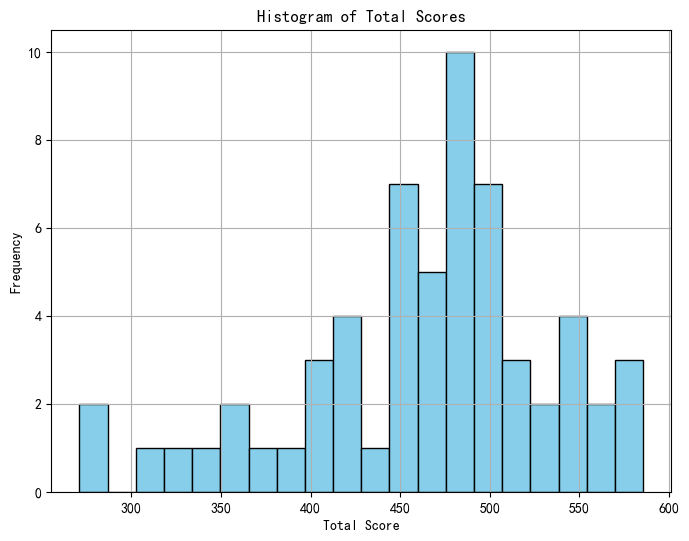




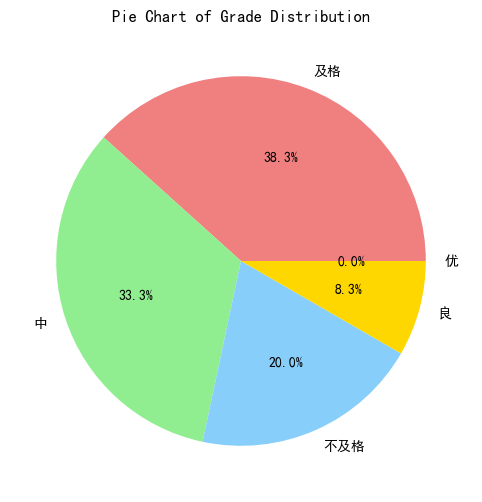
**按性别分别计算各门课程的成绩**



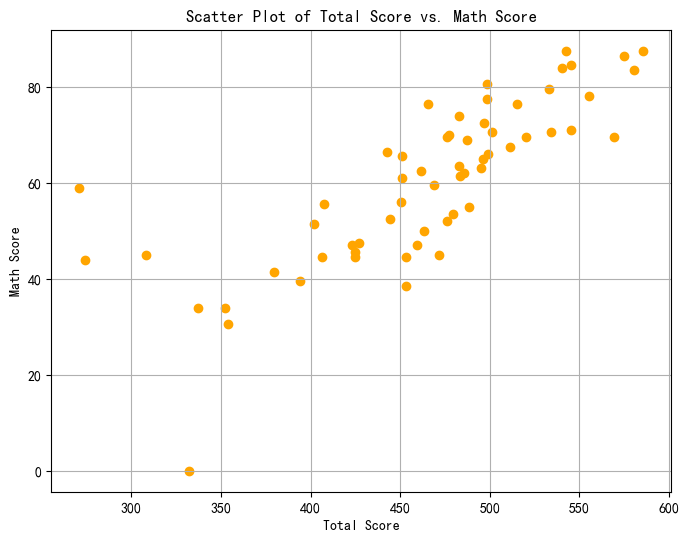
**绘制总成绩的直方图**

****

**绘制平均成绩的优、良、中、及格和不及格的饼图**

****

**提取数学（math）成绩，绘制总成绩和数学成绩的散点图**

****

1. **代码（文本式粘贴，重要代码处，需添加注释）**

**import pandas as pd**

**import matplotlib.pyplot as plt**

**plt.rcParams['font.sans-serif']=['SimHei'] #解决中文显示乱码问题**

**plt.rcParams['axes.unicode\_minus']=False**

**df1 = pd.read\_table("ReportCard1.txt")**

**df2 = pd.read\_table("ReportCard2.txt")**

**df = pd.merge(df1, df2, on='id')**

**# 用 0 填充缺考成绩为 NaN 的情况**

**df = df.fillna(0)**

**# 计算每个同学所有课程的总成绩和平均成绩**

**df['总成绩'] = df.iloc[:, 2:9].sum(axis=1)**

**df['平均成绩'] = df.iloc[:, 2:9].mean(axis=1)**

**# 将数据按总成绩的降序排序**

**df = df.sort\_values(by='总成绩', ascending=False)**

**# 按性别分别计算各门课程的平均成绩**

**avg\_score\_by\_gender = df.groupby('sex').mean()**

**# 分组计算平均成绩的优、良、中、及格和不及格**

**bins = [0, 59, 69, 79, 89, 100]**

**labels = ['不及格', '及格', '中', '良', '优']**

**df['成绩等级'] = pd.cut(df['平均成绩'], bins=bins, labels=labels)**

**grade\_counts = df['成绩等级'].value\_counts()**

**# 绘制总成绩的直方图**

**plt.figure(figsize=(8, 6))**

**plt.hist(df['总成绩'], bins=20, color='skyblue', edgecolor='black')**

**plt.title('Histogram of Total Scores')**

**plt.xlabel('Total Score')**

**plt.ylabel('Frequency')**

**plt.grid(True)**

**plt.show()**

**# 绘制平均成绩的优、良、中、及格和不及格的饼图**

**plt.figure(figsize=(8, 6))**

**plt.pie(grade\_counts, labels=grade\_counts.index, autopct='%1.1f%%', colors=['lightcoral', 'lightgreen', 'lightskyblue', 'gold', 'lightgrey'])**

**plt.title('Pie Chart of Grade Distribution')**

**plt.show()**

**# 提取数学（math）成绩，绘制总成绩和数学成绩的散点图**

**plt.figure(figsize=(8, 6))**

**plt.scatter(df['总成绩'], df['math'], color='orange')**

**plt.title('Scatter Plot of Total Score vs. Math Score')**

**plt.xlabel('Total Score')**

**plt.ylabel('Math Score')**

**plt.grid(True)**

**plt.show()**