

基于 Openpyxl 的高校期末成绩数据处理与分析方法研究

朱阁 胡硕娟

张家界航空工业职业技术学院 湖南 张家界 427000

摘要: 随着教育信息化的不断推进,如何高效、准确地处理高校期末成绩数据成为一个 important 问题。本文研究了基于 Python 和 Openpyxl 库的期末成绩数据处理方法,设计并实现了一个自动化的数据处理系统。通过实验验证,该系统显著提高了数据处理的效率和准确性,展示了在教育数据处理中的广泛应用潜力。

关键词: Openpyxl; 高校期末成绩; 数据处理; 自动化; Python

DOI:10.69979/3029-2735.24.3.032

1. 引言

1.1 研究背景

随着信息技术的飞速发展,教育信息化已经成为现代教育的重要趋势之一。在高校教育管理中,学生成绩数据的处理与管理一直是一个关键环节。传统的成绩管理方式主要依赖于人工录入和手动处理,不仅耗时费力,还容易出现错误。尤其在期末考试成绩汇总与分析过程中,手工操作的效率和准确性难以满足当前高效、精准的需求。近年来,自动化技术在教育管理中的应用逐渐增多,特别是利用编程语言和数据处理工具进行成绩数据的自动化处理,显著提升了工作效率和数据准确性。Python 作为一种高效、简洁且功能强大的编程语言,在数据处理领域表现尤为出色。其丰富的第三方库提供了多种数据处理解决方案,例如, Pandas、NumPy 和 Openpyxl,可以方便地进行 Excel 数据的读取、清洗、转换和分析^[1],其中 Openpyxl 库专门用于操作 Excel 文件,因其易用性和强大功能受到广泛关注。

1.2 研究目的

本研究旨在探索基于 Python 和 Openpyxl 库的高校期末成绩数据处理方法,通过自动化的方式实现成绩数据的读取、处理、计算和格式化,以提高数据处理的效率和准确性。同时,研究结果将为高校成绩管理提供一种高效、可靠的技术解决方案,推动教育数据处理的自动化发展。具体来说,本研究的目标包括:设计并实现一个自动化的成绩数据处理系统;提高成绩数据处理的效率和准确性;为其他教育数据处理应用提供参考和借鉴。

1.3 研究意义

本研究具有重要的理论和实践意义。从理论上讲,

本研究丰富了教育数据处理领域的技术手段,展示了 Python 和 Openpyxl 在实际应用中的潜力和优势。从实践上讲,研究结果可以直接应用于高校的成绩管理系统,解决实际工作中的数据处理难题,减轻教师的工作负担,提高工作效率和数据质量。此外,本研究还具有一定的推广价值。通过对成绩数据处理的研究,可以为其他类似的教育数据处理应用提供技术参考,推动教育管理的信息化和智能化发展。

2. 相关研究

2.1 数据处理与分析技术的现状

在教育管理中,数据处理与分析技术的应用越来越广泛。传统的成绩管理主要依赖手工操作和基本的电子表格工具,如 Microsoft Excel。然而,这些方法往往存在效率低、易出错的问题。随着大数据和人工智能技术的发展,教育数据处理技术逐渐向自动化、智能化方向转变。现有的成绩管理系统如学信网、教务管理系统等,虽然能够实现基本的数据处理功能,但其灵活性和可扩展性有限,难以满足复杂的数据分析需求方式^[2]。

2.2 Openpyxl 库的应用

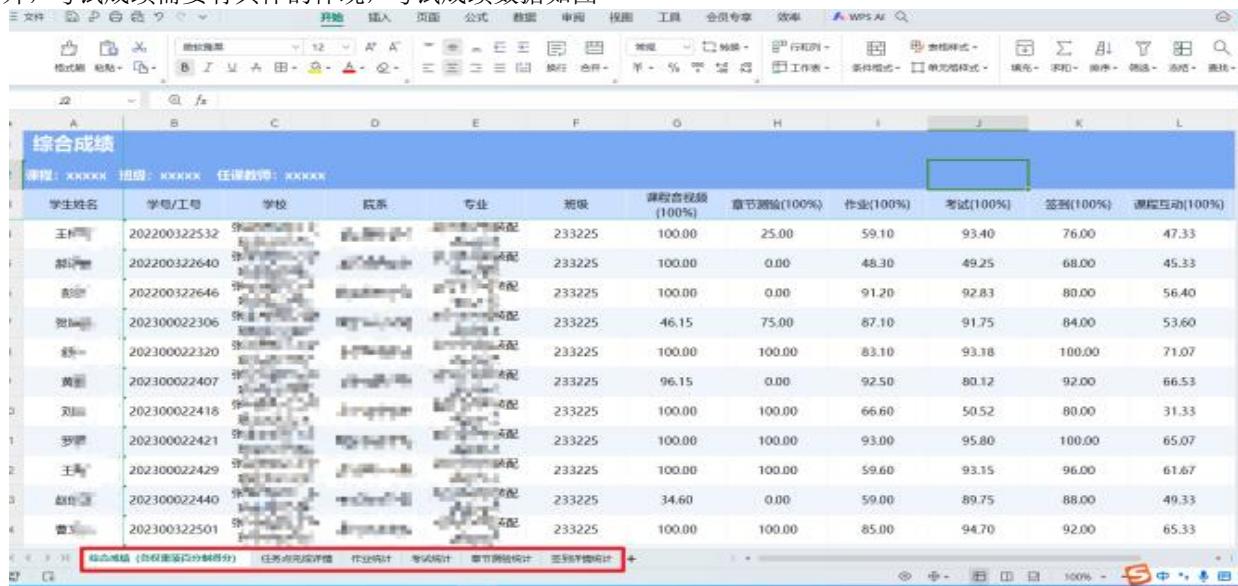
Openpyxl 是 Python 中一个强大的 Excel 文件处理库,广泛应用于各类数据处理场景。与传统的 Excel 操作方式相比,Openpyxl 能够高效地进行数据读取、写入、格式化等操作,极大地提高了数据处理的自动化程度。许多研究表明,Openpyxl 在教育、财务、统计等领域具有重要应用价值。例如,马宁与陈曦等(2020)在其研究中指出,Openpyxl 库可操作 Excel 数据表,这些库可以极大地节约时间,提升工作效率^[3]。

3. 研究方法

3.1 研究设计

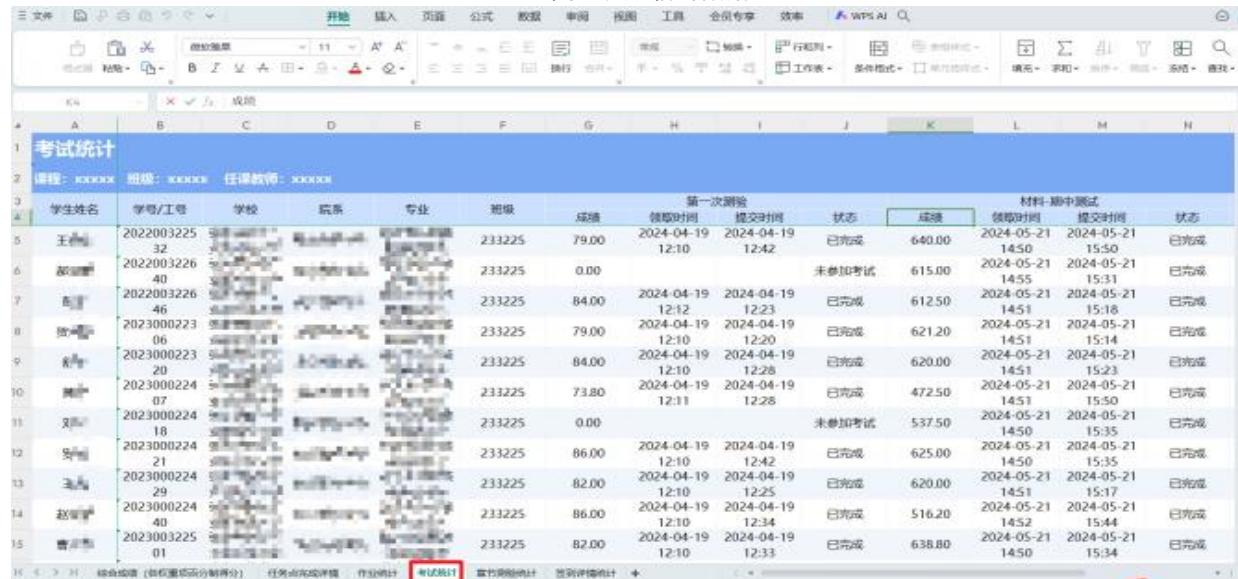
本研究为批量处理一批期末成绩数据,某一数据如图 1 所示,其余数据结构类似,该数据中各项权重为百分制,需要按要求修改为不同的比例,总权重为 100%。另外,考试成绩需要有具体的体现,考试成绩数据如图

2 所示,其中只需要保留成绩信息,并且最终还需计算综合成绩,并进行格式化处理等操作。最终处理结果如图 3 所示。



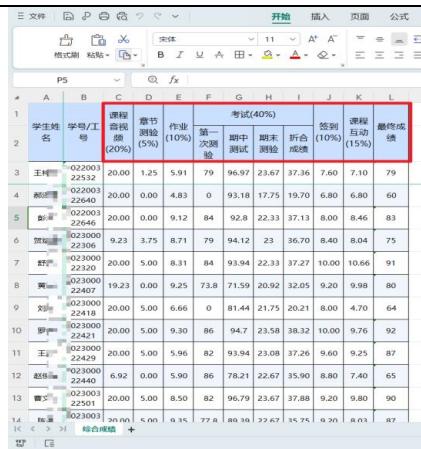
综合成绩											
课程: XXXXX 班级: XXXXX 任课教师: XXXXX											
学生姓名	学号/工号	学校	院系	专业	班级	期末音视频 (100%)	章节测验(100%)	作业(100%)	考试(100%)	签到(100%)	课堂互动(100%)
王强	202200322532	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	100.00	25.00	59.10	93.40	76.00	47.33
赵雷	202200322640	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	100.00	0.00	48.30	49.25	68.00	45.33
张伟	202200322646	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	100.00	0.00	91.20	92.83	80.00	56.40
刘伟	202300022306	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	46.15	75.00	87.10	91.75	84.00	53.60
孙伟	202300022320	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	100.00	100.00	83.10	93.18	100.00	71.07
黄伟	202300022407	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	96.15	0.00	92.50	80.12	92.00	66.53
刘伟	202300022418	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	100.00	100.00	66.60	50.52	80.00	31.33
罗伟	202300022421	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	100.00	100.00	93.00	95.80	100.00	65.07
王伟	202300022429	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	100.00	100.00	59.60	93.15	96.00	61.67
赵伟	202300022440	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	34.60	0.00	59.00	89.75	88.00	49.33
曹伟	202300322501	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	100.00	100.00	85.00	94.70	92.00	65.33

图 1 处理前的数据



考试统计											
课程: XXXXX 班级: XXXXX 任课教师: XXXXX											
学生姓名	学号/工号	学校	院系	专业	班级	成绩	第一次测验	状态	成绩	材料-期中测试	状态
王伟	2022003225	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	79.00	2024-04-19 12:10	2024-04-19 12:42	已完成	640.00	2024-05-21 14:50
赵雷	2022003226	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	0.00			未参加考试	615.00	2024-05-21 14:55
张伟	2022003226	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	84.00	2024-04-19 12:12	2024-04-19 12:23	已完成	612.50	2024-05-21 14:51
刘伟	2023000223	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	79.00	2024-04-19 12:10	2024-04-19 12:20	已完成	621.20	2024-05-21 14:51
孙伟	2023000223	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	84.00	2024-04-19 12:10	2024-04-19 12:28	已完成	620.00	2024-05-21 14:51
罗伟	2023000224	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	73.80	2024-04-19 12:11	2024-04-19 12:28	已完成	472.50	2024-05-21 14:51
王伟	2023000224	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	0.00			未参加考试	537.50	2024-05-21 14:50
刘伟	2023000224	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	86.00	2024-04-19 12:10	2024-04-19 12:42	已完成	625.00	2024-05-21 14:50
孙伟	2023000224	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	82.00	2024-04-19 12:10	2024-04-19 12:25	已完成	620.00	2024-05-21 14:51
赵伟	2023000224	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	86.00	2024-04-19 12:10	2024-04-19 12:34	已完成	516.20	2024-05-21 14:52
曹伟	2023003225	华中科技大学	电气与电子工程学院	电气工程及其自动化	233225	82.00	2024-04-19 12:10	2024-04-19 12:33	已完成	638.80	2024-05-21 14:50

图 2 考试成绩数据



学生成绩名	学号/工号	课程成绩(0%)	平时成绩(5%)	作业(10%)	考试(40%)			签到(10%)	互动(15%)	最终成绩
					基础题	初中题	高中题			
王华	0220001	22.53	20.00	1.25	5.91	79	96.97	23.67	37.36	7.60
张华	0220002	20.00	0.00	4.83	0	93.18	17.75	19.70	6.80	6.80
李华	0220003	22.64	20.00	0.00	9.12	84	92.8	22.33	37.13	8.00
王华	0220004	22.64	20.00	0.00	9.12	84	92.8	22.33	37.13	8.00
张华	0220005	22.30	9.23	3.75	8.71	79	94.12	23	36.70	8.40
李华	0220006	22.30	20.00	5.00	8.31	84	93.94	22.33	37.27	10.00
王华	0220007	22.32	19.23	0.00	9.25	73.8	71.59	20.92	32.05	9.20
张华	0220008	22.40	20.00	5.00	6.66	0	81.44	21.75	20.21	8.00
李华	0220009	22.40	20.00	5.00	9.30	86	94.7	23.58	38.32	10.00
王华	0220010	22.40	20.00	5.00	9.30	86	93.94	23.08	37.26	9.60
张华	0220011	22.40	20.00	5.00	5.96	82	78.21	22.67	35.90	8.80
李华	0220012	22.40	6.92	0.00	5.90	86	96.79	23.67	37.88	7.40
王华	0220013	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220014	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220015	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220016	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220017	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220018	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220019	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220020	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220021	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220022	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220023	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220024	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220025	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220026	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220027	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220028	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220029	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220030	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220031	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220032	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220033	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220034	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220035	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220036	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220037	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220038	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220039	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220040	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220041	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220042	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220043	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220044	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220045	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220046	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220047	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220048	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220049	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220050	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220051	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220052	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220053	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220054	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220055	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220056	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220057	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220058	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220059	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220060	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220061	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220062	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220063	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220064	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220065	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220066	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220067	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220068	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220069	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220070	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220071	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220072	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220073	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220074	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220075	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220076	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220077	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220078	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220079	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220080	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220081	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220082	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220083	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220084	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220085	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220086	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220087	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220088	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220089	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220090	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220091	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220092	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220093	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
王华	0220094	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
张华	0220095	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20
李华	0220096	22.40	20.00	5.00	8.50	82	96.79	23.67	37.88	9.20

```
def unmerge_cell(ws):
    """
    查找合并单元格，并拆分
    """
    m_list = ws.merged_cells # 合并单元格的位置信息
    cr = []
    for m_area in m_list:
        # 合并单元格的起始行坐标、终止行坐标。。。
        r1, r2, c1, c2 = m_area.min_row,
        m_area.max_row, m_area.min_col, m_area.max_col
        # 横向合并单元格的位置信息提取
        if c2 - c1 > 0 or r2 - r1 > 0:
            cr.append((r1, r2, c1, c2))
            print('符合条件%s' % str(m_area))
    for r in cr:
        ws.unmerge_cells(start_row=r[0], end_row=r[1],
        start_column=r[2], end_column=r[3])
```

图 6 查找合并单元格并拆分程序

4.2.3 数据处理

在预处理后的数据基础上再次进行成绩数据的处理。这里主要进行的操作有：1、在考试统计表单中获取多次考试的成绩，写入到对应的考试成绩详情列中。2、将课程音视频、章节测验等项目的权重由 100% 按要求转化为特定的权重值。3、对考试成绩中非百分制的成绩需进行换算为百分制，4、计算最终的综合成绩并将综合成绩列的值进行四舍五入处理。部分代码如图 7 所示：

```
# 处理考试统计表单
ws2 = wb['考试统计']
col_list = []
for j in range(1, ws2.max_column + 1):
    if ws2.cell(4, j).value == '成绩':
        col_list.append(j)
# 复制考试数据到成绩详情表单
ws1.cell(1, 6).value = ws2.cell(1, col_list[0]).value
ws1.cell(1, 7).value = ws2.cell(1, col_list[1]).value
ws1.cell(1, 8).value = ws2.cell(1, col_list[2]).value
for r in range(2, ws1.max_row + 1):
    # 成绩总分为100的情况
    ws1.cell(r, 6).value = round(float(ws2.cell(r, col_list[0]).value)/1,2)
    # 成绩总分不为100的情况，如：250分、300分
    ws1.cell(r, 7).value = round(float(ws2.cell(r, col_list[1]).value)/2.5,2)
    ws1.cell(r, 8).value = round(float(ws2.cell(r, col_list[2]).value)/3.0,2)
# 将综合成绩列的值进行四舍五入
print(round(float(ws1.cell(1, 12).value)))
for i in range(3, ws1.max_row + 1):
    value_cj=float(ws1.cell(i, 12).value)
    ws1.cell(i, 12).value=str(round(value_cj))
```

图 7 部分成绩处理程序

4.2.4 数据格式化与保存

成绩数据处理完成后，需对其进行格式化处理，包括表头和单元格中字体的样式，大小，对齐方式，单元格行和列的宽度。借助 openpyxl 库的 styles 模块即可很好的实现。最好将格式化好的文件进行重命名并保存。部分代码如图 8 所示：

```
from openpyxl.styles import Border, Side, Alignment, Font,
PatternFill
# 设置样式
border1 = Border(left=Side(style='thin'), right=Side(
style='thin'), top=Side(style='thin'), bottom=Side(
style='thin'))
align = Alignment(horizontal='center', vertical='center')
font1 = Font(bold=True, color= 'FFFFFF')
font2 = Font(bold=False, color= '000000')
fill_1 = PatternFill(fill_type='solid', fgColor= '000000')
fill_2 = PatternFill(fill_type='solid', fgColor= 'DDDDDD')
# 数据格式化
for i in range(1, ws.max_row + 1):
    for j in range(1, ws.max_column + 1):
        ws.cell(i, j).border = border1
        ws.cell(i, j).alignment = align
        if i == 1 or i == 2:
            ws.cell(i, j).font = font1
            ws.cell(i, j).fill = fill_1
        else:
            ws.cell(i, j).font = font2
            if i % 2 == 0:
                ws.cell(i, j).fill = fill_2
# 保存处理后的文件
new_path = os.path.join(path, '处理后的成绩文件.xlsx')
wb.save(new_path)
```

图 8 数据格式化与保存程序

5. 结果与分析

在完成代码实现后，进行测试以验证数据处理的准确性。将处理结果与手工处理结果进行比较，确保数据处理的正确性和一致性。实验结果表明，基于 Python 和 OpenPyXL 的期末成绩数据处理系统显著提高了数据处理的效率和准确性。在测试过程中，系统能够快速、准确地读取和处理大量成绩数据，并生成符合要求、格式规范的结果文件。具体结果分析如下：

1. 数据读取与处理效率：系统能够在短时间内读取并处理大量 Excel 文件，显著减少了人工操作的时间。
2. 数据准确性：系统通过自动化处理，避免了人工录入和计算错误，提高了数据的准确性。
3. 数据格式规范性：处理后的成绩文件格式规范，便于后续的查阅和分析。
4. 可扩展性强：代码可扩展性强，易于根据实际需求进行修改和扩展。

6. 结论与展望

本研究通过设计并实现基于 Python 和 OpenPyXL 的期末成绩数据处理系统，验证了自动化技术在教育数据处理中的应用潜力。研究结果表明，自动化数据处理系统显著提高了数据处理的效率和准确性，解决了传统手工操作的诸多问题，为高校成绩管理提供了一种高效、可靠的技术解决方案。

未来的研究可以进一步优化系统功能，增加可视化模块，实现更加智能化的教育数据处理和分析。此外，可以探索将该系统应用于其他教育管理领域，如学生考勤数据处理、教学质量评估等，以推动教育管理的信息化和智能化发展。

参考文献

- [1] 刘易. Python 实现 Excel 文档转换到 Word 文档的自动化方法 [J]. 电脑编程技巧与维护, 2024, 467(05): 47-49. DOI: 10.16184/j.cnki.comprg.2024.05.017.
- [2] 杜江舟. 高校教务管理系统信息化建设研究 [J]. 科教文汇, 2023, 597(21): 24-26. DOI: 10.16871/j.cnki.kjwh.2023.21.004.
- [3] 马宁, 陈曦, 张李铭. 基于 Selenium 与 Openpyxl 的 Web 脚本自动化设计研究 [J]. 电脑知识与技术, 2020, 16(01): 56-58+75. DOI: 10.14004/j.cnki.ckt.2020.0024.