南京审计大学

硕士研究生课程考核成绩评定表

**（适用于非笔试考核）**

**2023 —2024 学年第 1 学期**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 课程代码 | JS50210T | 课程名称 | Python高级程序设计 | | |
| 年级与专业 | 电子信息专硕 | 学号 | MP2309018 | 姓名 | 罗伟力 |
| 考核形式 | （√）课程论文 （ ）研究报告 （ ）项目方案设计  （ ）案例分析报告 （ ）口试 （ ）其他 | | | | |
| 任课教师 | 张凡龙 | | | | |
| 教师评语：  签名： 年 月 日 | | | | | |
| 成绩评定： 课程组长（任课教师）签名： 年 月 日 | | | | | |

数据之路：从数学建模赛道到计算机审计前沿

南京审计大学电子信息专硕（计算机学院）罗伟力

摘要：随着信息时代的到来，数据的产生、传输和存储已成为现代社会的重要组成部分。本文以个人在第二十届华为杯数学建模比赛中获得国家三等奖的经历为出发点，探讨了数学建模与计算机审计这两个看似独立领域之间的深刻关联。突显了数学建模比赛经历对于推动计算机审计前沿研究的启示。通过这一学术探讨，我们期望能够为未来的研究方向提供新的思路和启示。

关键词：数学建模；数据分析；前沿研究；计算机审计

**Abstract:**With the advent of the information age, the generation, transmission, and storage of data have become an important component of modern society. Starting from my personal experience of winning the third prize in the 20th Huawei Cup Mathematical Modeling Competition, this article explores the profound relationship between mathematical modeling and computer auditing, two seemingly independent fields. Highlighting the inspiration of mathematical modeling competition experience for advancing cutting-edge research in computer auditing. Through this academic discussion, we hope to provide new ideas and insights for future research directions.

**KeyWords:**Mathematics Modeling; Data Analysis; Frontier Research; Computer Audit

一、引言

在当今数字化浪潮的推动下，计算机审计和数学建模作为信息科技领域中的两个关键分支，日益凸显其在实际问题解决和决策支持中的重要性。计算机审计旨在确保信息系统的安全和有效运行，而数学建模通过数学和计算方法解决实际问题，两者似乎各自在自己的领域内发光发热。然而，是否存在一个无形的桥梁，将这两者紧密联系在一起？本文将探讨这一问题，着眼于个人在华为杯数学建模比赛中的经历，以及这一经验对计算机审计研究的启示。

在华为杯数学建模比赛中，我团队荣获国家三等奖的殊荣，这一经历在我个人学术和职业发展中留下了深刻的印记。这个比赛锻炼了我们的创意思维、团队协作和问题解决能力，为进一步深入计算机审计研究奠定了坚实的基础。我们开始思考，数学建模的创新精神是否能够为计算机审计带来新的视角和方法？本文旨在回答这一问题，探索数学建模与计算机审计之间的交叉点，并挖掘数学建模在计算机审计前沿的潜在贡献。

通过本研究，我们希望能够为数字化时代信息安全和数据治理提供新的思路和启示，进一步推动计算机审计领域的发展。在这个数据之路上，数学建模将成为我们的探照灯，引领我们走向计算机审计的前沿。

# 二、数学建模的精华

2.1数据建模比赛概述

数学建模比赛是一个以实际问题为基础的竞技活动，参与者通常在有限的时间内，通过运用数学建模方法解决给定的实际问题。这类比赛的目标不仅在于找到问题的解决方案，更注重解决方案的创新性和可行性。

2.2 创意思维的培养

数学建模比赛通过引导参与者从多角度思考问题，鼓励他们敢于挑战传统思维，培养了创意思维。在解决实际问题的过程中，参与者需要通过数学建模方法将问题抽象化，找到解决方案的新颖途径。

2.3 团队协作的重要性

数学建模比赛通常是一个团队协作的过程。团队成员需要共同探讨问题，分享不同的见解，并最终达成一致的解决方案。这种团队协作培养了沟通和合作的能力，使参与者能够更好地应对复杂问题。

2.4 问题解决的方法

数学建模中，解决问题的方法多种多样，涉及到数学建模的各个方面。常见的方法包括数学模型的建立、方程的求解、优化算法的应用等。这些方法为参与者提供了解决实际问题的有力工具。

2.5 数学建模实例

以我们参与的第二十届华为杯中国研究生数学建模比赛C题为例，我们面临了一个关于大规模创新类评审研究的实际问题。通过团队的协作和创意思维，我们构建了一个新颖的数学模型，利用优化算法求解，最终得到了令人满意的结果。这个实例将为后文的讨论提供一个具体的案例分析。

# 三、计算机审计数据采集、清理和分析的综述

3.1 引言

在数字化时代，计算机审计成为确保信息系统安全和合规性的关键手段。而数据的采集、清理和分析是计算机审计过程中不可或缺的环节。本节将综述计算机审计中数据处理的关键技术和方法。以下是一种常见的计算机审计步骤：

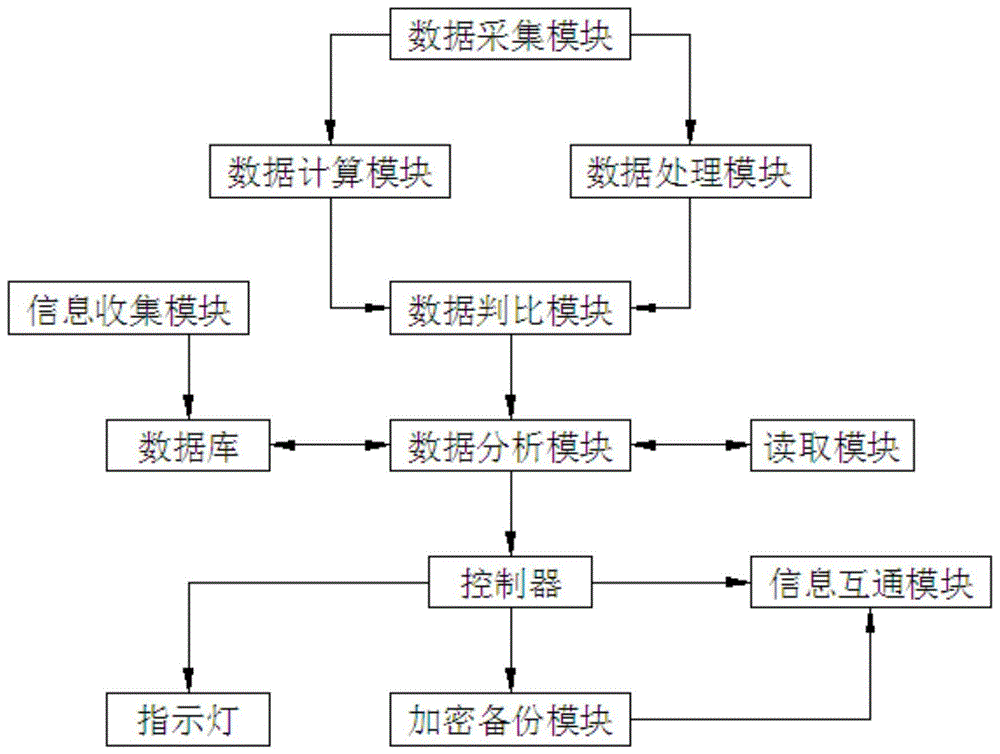


图1 计算机审计流程图

3.2 数据采集技术综述

数据采集是计算机审计的第一步，影响着后续审计分析的质量。常见的数据采集技术包括实地审计、日志文件采集和网络流量分析。实地审计通过实地访查和观察获得信息，而日志文件采集则通过记录系统活动来获取审计数据。网络流量分析则关注数据在网络中的传输情况，为审计提供了更全面的视角。

3.3 数据清理技术综述

数据清理是确保审计数据质量的关键步骤。数据清理技术包括数据标准化、异常值检测、缺失值处理等。数据标准化通过统一数据格式，提高数据一致性。异常值检测用于识别异常数据点，而缺失值处理则关注如何处理缺失的数据，以保证审计分析的准确性。

3.4 数据分析技术综述

数据分析是计算机审计的核心，涉及统计分析、数据挖掘和机器学习等技术。统计分析用于总结和描述数据特征，数据挖掘关注从大量数据中挖掘潜在的规律，而机器学习技术通过模型训练和预测来发现异常行为和风险。

3.5 技术整合与应用案例

在实际计算机审计中，这些数据处理技术相互交织，共同发挥作用。通过整合数据采集、清理和分析技术，审计人员能够更全面地了解系统运行状态，及时发现潜在问题。以某实际案例为例，通过采用综合技术，审计团队成功发现并解决了系统中的潜在安全风险。

3.6 未来发展趋势

未来，计算机审计数据处理技术将面临更多挑战和机遇。人工智能技术的不断发展将为审计提供更智能的数据处理工具，例如自动化的异常检测和智能决策支持系统。同时，大数据和云计算等新兴技术将使得审计处理更大规模和复杂性的数据成为可能。未来的趋势可能包括更强大的实时监控和预测分析，以更好地应对动态的信息系统环境。

3.7 结论

综合考虑数据采集、清理和分析的综述，不难看出它们在计算机审计中扮演着至关重要的角色。高效的数据处理技术为审计提供了更全面、精准的信息，使得审计人员能够更好地了解和应对潜在风险。未来的发展趋势表明，数据处理技术将继续发展，为计算机审计提供更强大的工具和方法。

# 四、数学建模与计算机审计的交叉点

4.1 Python及相关库在数据处理与分析中的作用

4.1.1 选择Python的原因

在数据处理与分析的领域，为何选择 Python 作为主要的编程语言有着多方面的原因：

1.广泛的库支持：Python拥有丰富的第三方库，包括NumPy、Pandas、Matplotlib、Scikit-learn等，这些库提供了强大的数据处理、分析和机器学习工具。这使得团队在数据科学项目中能够更高效地利用这些工具。

2易学易用：Python相对于其他编程语言来说，语法简洁清晰，易于学习和使用。这降低了团队成员的学习曲线，使得团队能够更快速地掌握必要的编程技能，迅速进入数据处理工作。

3.开发速度快：Python的语法设计和丰富的库使得开发速度相对较快。这对于数据处理和分析任务来说尤为重要，因为项目通常需要快速地得到结果和反馈。

4.社区支持：Python拥有庞大的开发者社区，这个社区提供了丰富的文档、教程和支持。团队成员可以在社区中获取帮助、分享经验，从而更好地解决在数据处理过程中遇到的问题。

5.跨平台性：Python是一种跨平台的编程语言，可以在多种操作系统上运行，包括Windows、Linux和macOS。这种跨平台性有助于确保团队成员在不同环境下都能够协同工作。

6.开源生态系统：Python是开源的，这意味着团队可以充分利用其丰富的开源工具和库。这也意味着团队可以根据需要定制和修改工具，以满足项目的具体要求。

7.适应性强：Python不仅仅局限于数据科学领域，还广泛应用于Web开发、人工智能、自动化脚本等领域。这种广泛的适应性意味着团队成员在其他项目中也可以充分发挥所学的Python技能。下图是使用Python中字典和列表结构实现数据收集：

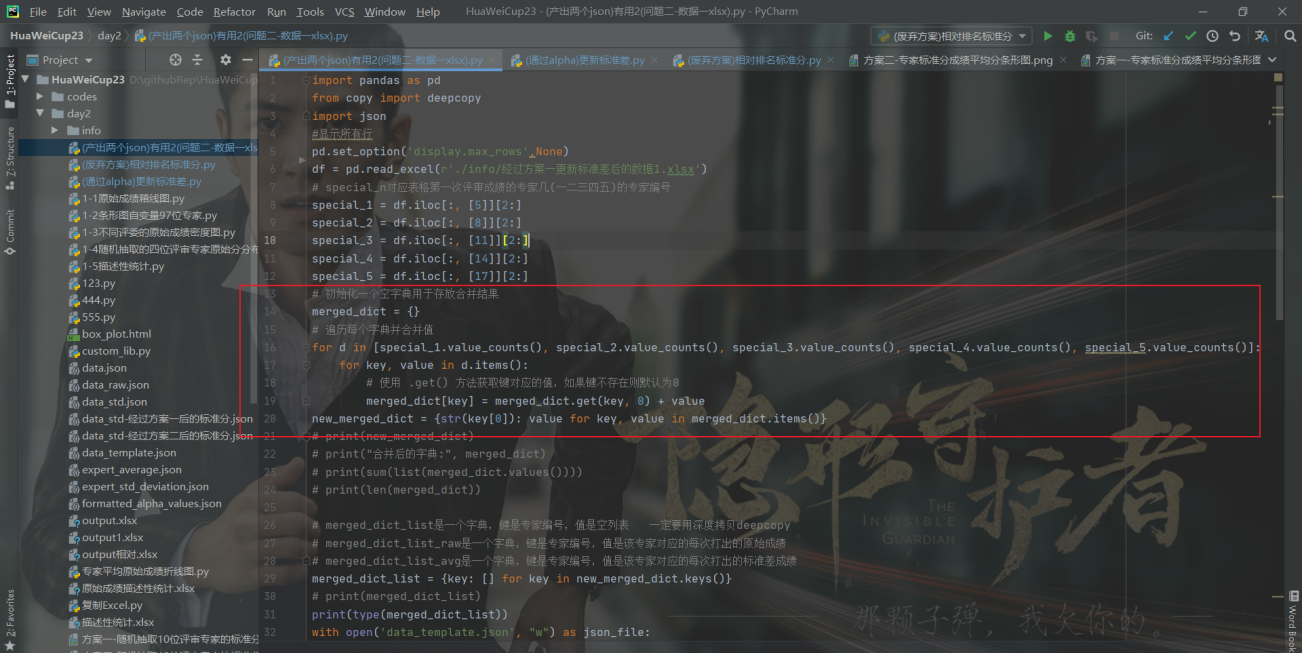


图2 原生语法

选择Python作为数据处理与分析的主要编程语言是出于其丰富的库支持、易学易用、开发速度快、社区支持等多方面的考虑。这些优势使得Python成为数据科学领域的首选工具之一

4.1.2 Numpy的作用

NumPy（Numerical Python）是Python中用于科学计算的基础库，主要用于处理多维数组和矩阵。在数据处理与分析中，Numpy提供了多维数组对象（numpy.ndarray），这是其最核心的数据结构。这种数组结构支持高效的数学和逻辑操作，使得数据的处理速度相比传统的Python列表更快。Numpy引入了广播（broadcasting）的概念，使得对不同形状的数组进行通用函数操作变得更加容易。这意味着可以对形状不同的数组进行操作，而不需要显式地扩展数组。Numpy内置了大量的数学函数，包括基本的数学运算、三角函数、指数和对数函数等。这些函数能够方便地用于数据处理和科学计算任务。Numpy提供了灵活的内存管理功能，能够高效地存储和操作大规模的数据集。Numpy数组的内存布局使得对数组的操作更加高效，尤其是在处理大型数据时。Numpy包含了线性代数运算所需的函数，如矩阵乘法、特征值分解等。此外，Numpy还提供了伪随机数生成器，用于模拟和随机实验。Numpy与其他科学计算库（如SciPy）紧密整合，共同构成了Python科学计算生态系统的基础。SciPy扩展了Numpy的功能，提供了更多高级的科学计算工具。Numpy支持多种文件格式的输入和输出，包括文本文件、二进制文件以及常见的科学数据格式（如NetCDF）。这方便了从不同来源读取数据并将处理结果保存到文件。

Numpy作为Python科学计算的基石，为数据处理提供了高效的数据结构和强大的数学功能。其广泛的应用范围使得数据科学家和分析师能够在处理和分析数据时更加便捷、高效。

4.1.3 Pandas的作用

Pandas是Python中用于数据分析和处理的强大库，主要建立在NumPy的基础上。以下是Pandas在数据处理与分析中的关键作用：

1.数据结构：Pandas引入了两个主要的数据结构：Series和DataFrame。Series是一维标签数组，而DataFrame是二维表格，类似于关系型数据库中的表格。这些数据结构使得数据的表示更加直观和灵活。

2.数据清理和准备：Pandas提供了丰富的数据清理工具，包括处理缺失值、重复值、异常值等。通过内置的方法，数据分析师能够轻松地进行数据清理和预处理，提高数据质量。

3.数据索引和选择：Pandas的强大之处在于其灵活的索引系统。通过标签或位置进行选择，以及通过布尔索引进行条件选择，使得数据的筛选和提取变得非常方便。

4.数据聚合和变换：Pandas支持数据的分组和聚合操作，例如groupby功能。这使得可以对数据进行按组的统计分析，同时也提供了灵活的数据变换方法，如apply函数。

5.时间序列处理：Pandas对时间序列数据提供了专门的支持。可以轻松处理时间戳、时间间隔、日期范围等，以及执行时间序列上的各种操作。

6.数据合并和连接：Pandas提供了多种合并和连接数据的方法，包括merge和concat。这使得数据可以方便地从不同的源合并到一起，形成更完整的数据集。

7.可视化支持：Pandas结合Matplotlib和Seaborn等可视化库，提供了数据可视化的支持。通过直接调用plot函数，用户能够轻松生成各种类型的图表。

8.灵活的数据输入和输出：Pandas支持多种文件格式的数据输入和输出，包括CSV、Excel、SQL数据库、JSON等。这方便了从不同来源读取数据和将处理结果保存到文件。

Pandas作为数据处理和分析的利器，提供了高层次、易用且功能丰富的工具，使得数据科学家和分析师能够更轻松地进行数据处理、分析和探索。以下是使用Pandas读取数学建模提供的Excel表格以及使用DataFrame去操作数据：

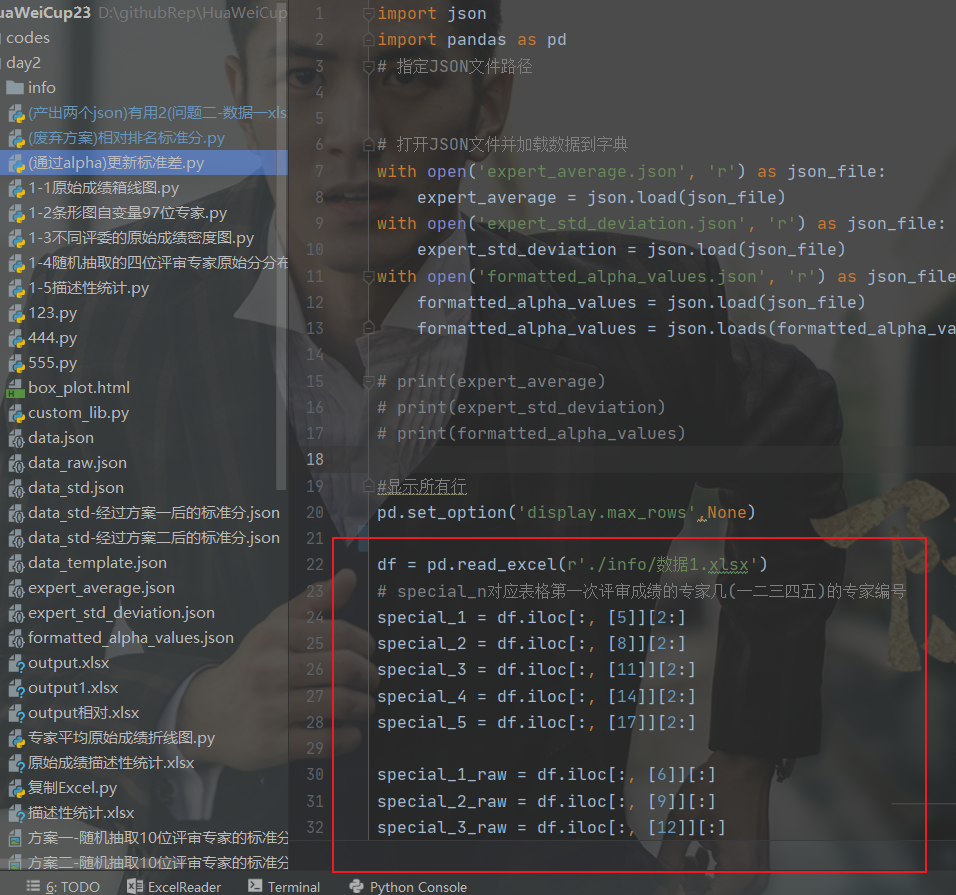


图3 Pandas实例

4.1.4 Matplotlib和Seaborn的作用

Matplotlib和Seaborn是Python中用于数据可视化的两个重要库，它们在数据处理和分析中的作用主要体现在以下几个方面：

1.绘制各种图表：

Matplotlib和Seaborn提供了丰富的图表绘制功能，包括折线图、散点图、柱状图、饼图、箱线图等。这些图表能够直观地展示数据的分布、趋势和关系，为数据分析提供直观的视觉表达。

2.数据分布可视化：

通过绘制直方图、核密度图等，Matplotlib和Seaborn能够有效地展示数据的分布情况。这对于了解数据的形状、集中趋势和离散度非常有帮助。

3.数据关系展示：

散点图、热力图等图表可以展示不同变量之间的关系。这有助于发现变量之间的相关性、趋势和异常值。

4.时间序列可视化：

Matplotlib和Seaborn支持时间序列数据的可视化，包括折线图、面积图等。这使得能够更清晰地观察时间序列的趋势和周期性。

5.高级可视化功能：

Seaborn是建立在Matplotlib基础上的高级可视化库，提供了更简洁的API和更美观的默认样式。Seaborn还具有更高级的可视化功能，如矩阵图、联合分布图等，使得数据科学家能够以更专业的方式展示分析结果。

6.可交互性：

Matplotlib和Seaborn可以与Jupyter Notebook等交互式计算环境集成，使得用户能够实时观察和调整图表，更好地与数据进行互动。

7.定制化和主题：

Matplotlib和Seaborn允许用户对图表进行高度的定制化，包括颜色、标签、图例等。此外，Seaborn提供了预定义的主题，使得图表具有更一致和专业的外观。

8.多图组合：

Matplotlib和Seaborn支持在一个图中组合多个子图，这对于同时展示多个数据视图或对比不同数据集非常有用。

Matplotlib和Seaborn作为数据可视化的利器，为数据分析提供了丰富的可视化工具和图表类型，使得数据科学家能够更好地理解和传达数据的意义。以下图表是在数据建模比赛中使用这两个库生成的图表：

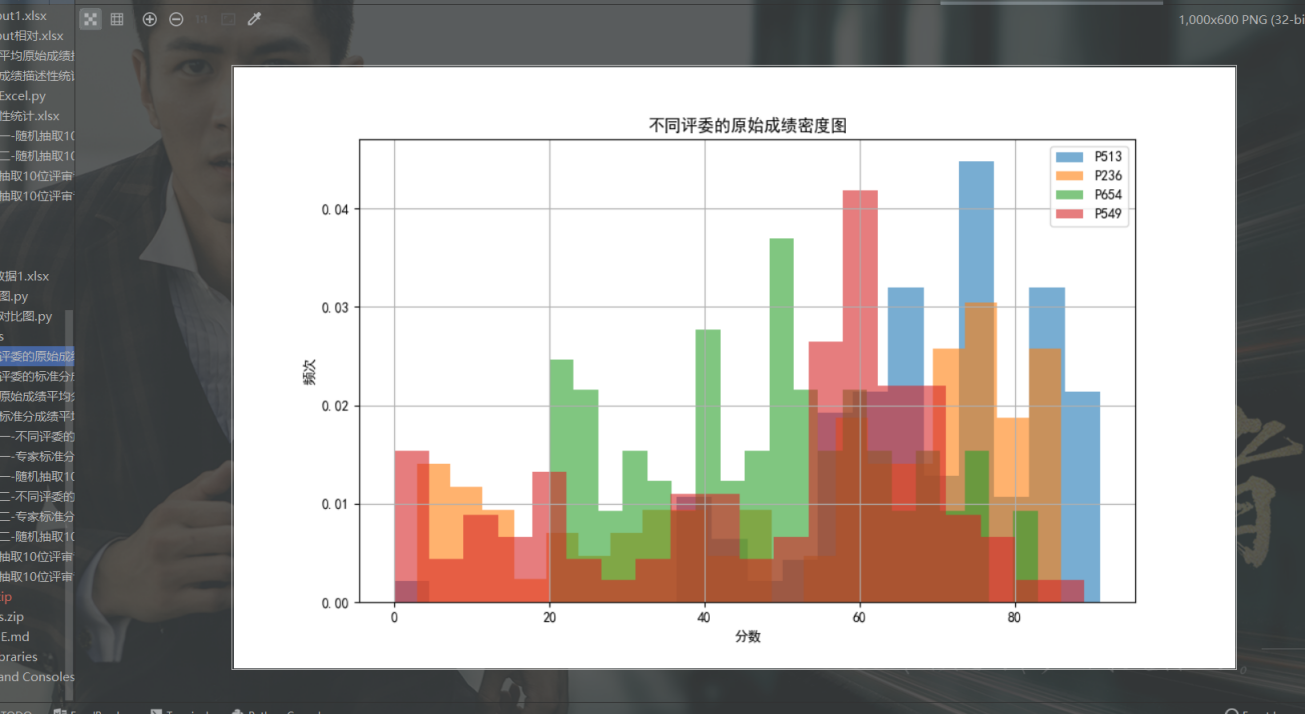


图4 不同评委的原始成绩密度图

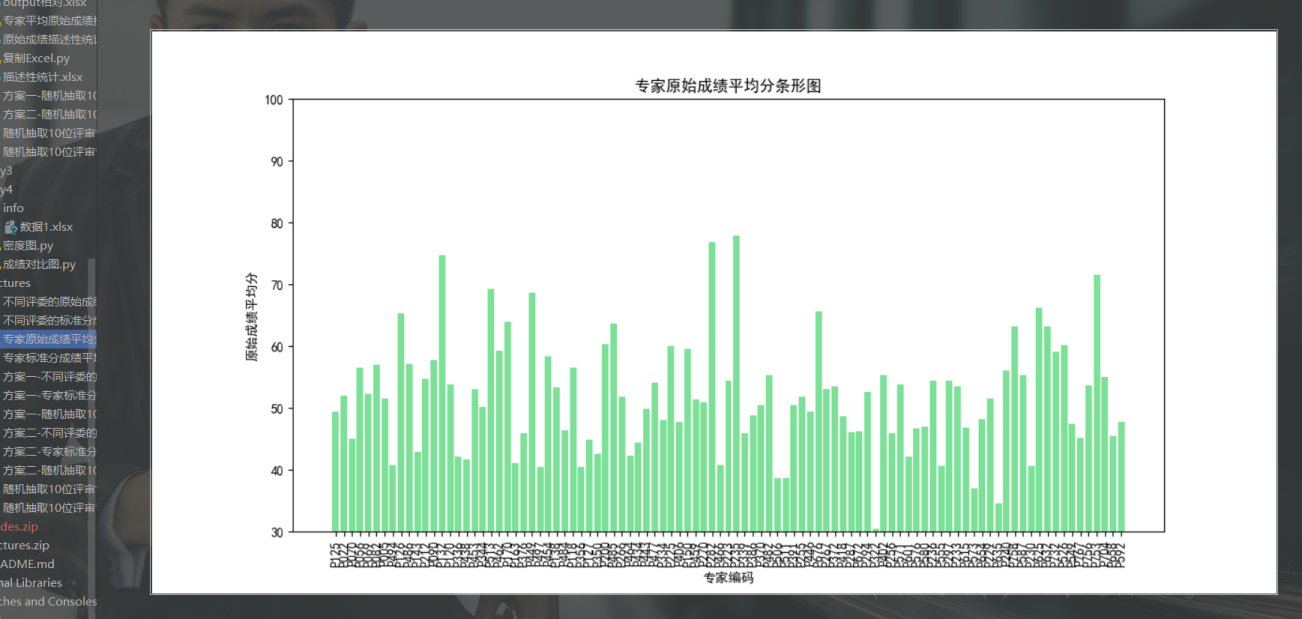


图5 专家原始成绩平均分条形图

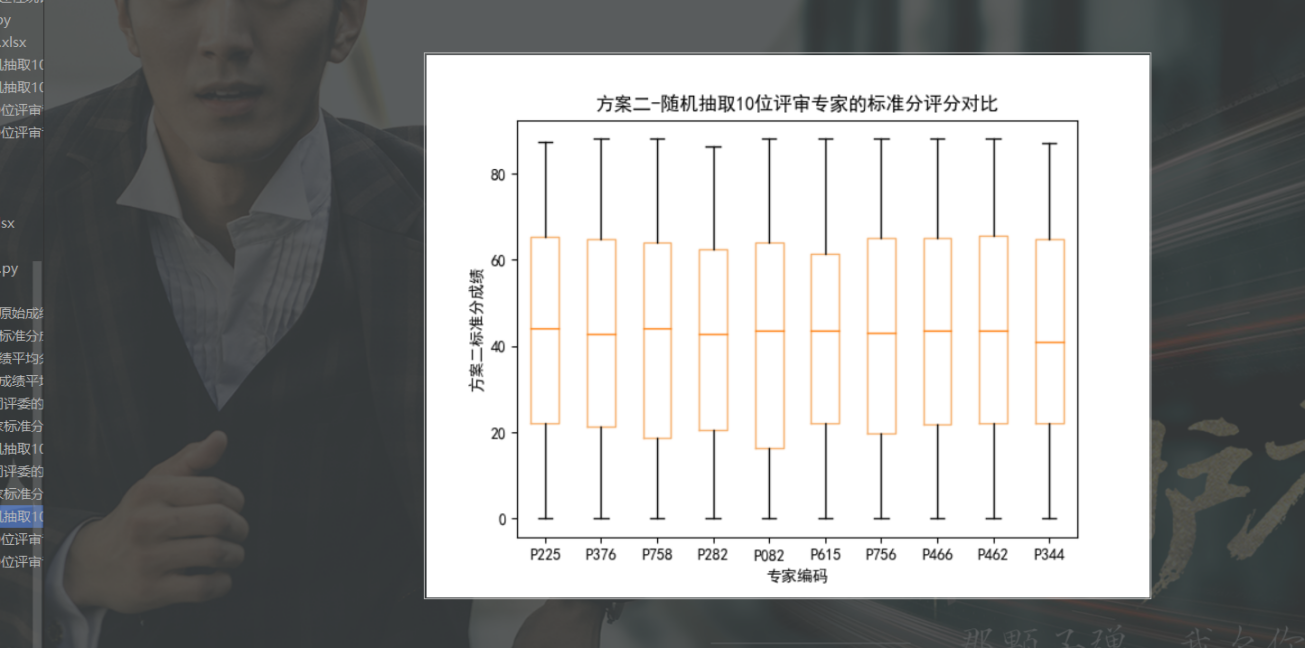


图6 使用方案二的评分对比图

4.2 AI工具在计算机审计中的应用

AI工具在计算机审计中扮演了重要的角色，其中包括chatgpt等自然语言处理工具。这些工具能够自动化审计报告的生成，使得审计人员能够更专注于关键问题的解决。此外，AI工具还可用于异常检测，通过智能算法识别潜在的安全风险和异常行为，为审计工作提供了新的方法。

4.2.1 自然语言处理

AI工具通过自然语言处理技术能够分析和理解大量的文本数据，包括审计报告、通信记录以及系统日志。自然语言处理使得审计人员能够更迅速地理解文本信息，提取关键信息，识别潜在问题。通过对文本的情感分析和实体识别，AI工具能够辅助审计人员更全面地理解系统运作中的风险和挑战。

4.2.2 智能辅助决策

AI工具在智能辅助决策方面通过机器学习算法和数据分析为审计人员提供决策支持。例如，对潜在风险的预测、评估和建议。智能辅助决策使得审计人员能够更准确地评估潜在风险，做出更明智的决策。通过对大量历史数据的学习，AI工具能够提供关于可能发生的安全事件的预测，帮助审计人员及时采取相应的防范措施。

4.2.3 自动化审计过程

自动化审计过程是AI工具的一项重要应用，可以涵盖日志分析、安全策略检测、异常行为识别等。通过自动执行重复性高的审计任务，提高审计效率。自动化审计过程使得审计人员能够将更多精力集中在高级别的分析和解决问题上，而不是花费大量时间在繁琐的任务上。通过机器学习和算法，AI工具能够发现和标记潜在的异常行为，提供更及时的审计反馈。



图7 智能审计示意图

这些AI工具在计算机审计中的应用方面都体现了智能化、高效化的特点，使得审计人员能够更好地应对复杂的信息系统环境，更迅速、准确地发现潜在风险和问题。同时，这些工具也在提高审计过程的效率和精度上发挥了关键作用。

4.3 数学建模经验对计算机审计的启示

从数学建模比赛中获得的经验为计算机审计提供了新的视角。通过将AI工具和编程语言融入数学建模的实践中，我们发现这些工具不仅仅是解决数学问题的工具，同时也可以为计算机审计带来创新的方法。数学建模比赛培养了我们对数据的敏感性，使我们能够更好地应对审计中的数据挑战，提高审计的精确性和效率。数学建模经验对计算机审计的启示主要体现在问题抽象与建模、数据分析与模式识别、变量关系与因果推理、模型评估与优化、实验设计与验证、跨学科思维以及创新思维与问题解决等方面。

首先，在问题抽象与建模方面，数学建模强调将实际问题进行抽象，这对计算机审计很重要。审计人员需要将复杂的信息系统、流程和数据抽象成可分析的模型，以更系统地理解审计目标，定义明确的审计任务和目标。

其次，数据分析与模式识别方面，数学建模经验鼓励审计人员借助数据分析工具，通过深度分析和模式识别发现潜在的安全问题、异常行为或趋势。这强调了利用数据提取有用信息的重要性。

在变量关系与因果推理方面，数学建模强调建立变量之间的关系，有助于进行因果推理。在计算机审计中，理解系统中不同变量之间的关系能够帮助审计人员追踪潜在问题的根本原因。

模型评估与优化方面，数学建模要求对模型进行评估和优化，以确保模型对实际问题的拟合度和预测准确性。在计算机审计中，审计人员同样需要不断评估审计方法和工具的有效性，优化审计流程，以确保能够充分覆盖潜在风险和问题。

在实验设计与验证方面，数学建模涉及科学实验的设计和验证，这对计算机审计也很重要。审计人员可以采用模拟实验、漏洞测试等手段，验证系统的安全性，确保系统在不同场景下的稳健性。

跨学科思维方面，数学建模要求将数学方法应用于解决实际问题。在计算机审计中，审计人员同样需要具备跨学科的思维，将信息技术、数据分析和审计知识相结合，更全面地理解和解决复杂的审计问题。

最后，在创新思维与问题解决方面，数学建模要求创新思维，以提出新颖、有效的模型解决问题。在计算机审计中，审计人员同样需要具备创新思维，不断探索新的审计方法和工具，以适应不断变化的信息技术环境。这种创新思维有助于更灵活地运用不同领域的知识来解决审计问题。

# 五、总结

本论文综合了计算机审计、数学建模经验以及人工智能在审计领域的应用。从问题抽象与建模、数据分析与模式识别、变量关系与因果推理、模型评估与优化、实验设计与验证、跨学科思维以及创新思维与问题解决等方面，深入探讨了数学建模对计算机审计的启示。论文以“数据之路：从数学建模赛道到计算机审计前沿”为题，将华为杯数学建模比赛的经验与计算机审计领域相结合，突显了跨学科思维的重要性。论文阐述了数学建模经验如何引导审计人员进行问题抽象、建模，强调了数据分析和模式识别在发现潜在问题中的作用。

在数学建模的框架下，论文详细探讨了变量关系、因果推理以及模型的评估与优化。通过实验设计和验证的角度，展示了科学实验如何帮助审计人员验证系统的安全性。与此同时，创新思维在提出新型审计方法和工具方面发挥了关键作用。论文还深入分析了人工智能工具在计算机审计中的应用，特别关注了自然语言处理、智能辅助决策和自动化审计过程等方面。强调了这些技术如何加速审计流程，提高审计效率，并对潜在的风险提供更准确的识别和评估。

最后，通过对华为杯数学建模比赛获奖经历的整合，论文展现了作者在比赛中所用到的工具和技术，如Python、Numpy、Pandas等，以及如何将这些经验成功融入到计算机审计中。总体而言，这篇论文通过数学建模的视角深刻剖析了计算机审计的关键方面，并在人工智能的框架下提出了创新性的解决方案，为审计领域的未来发展提供了有益的参考。

参考文献

[1]翟舒婷.大数据环境下内部审计数字化转型问题研究[D].哈尔滨商业大学,2023.DOI:10.27787/d.cnki.ghrbs.2023.000602

[2]蒋文琴.大数据环境下农业上市公司财务舞弊审计模式优化研究[D].安徽财经大学,2022.DOI:10.26916/d.cnki.gahcc.2022.000244

[3]王珊珊.计算机通信及其他电子设备制造企业审计风险[D].北京交通大学,2021.DOI:10.26944/d.cnki.gbfju.2021.002285

[4]张加会.基于数据挖掘技术的线上招聘信息分析与应用[D].阜阳师范大学,2023.DOI:10.27846/d.cnki.gfysf.2023.000235.

[5]胡磊.大数据与会计专业Python课程思政的实践与思考[J].现代商贸工业,2024,45(01):146-148.DOI:10.19311/j.cnki.1672-3198.2024.01.046