

探讨 Deepseek 在《R 语言数据分析》课程中的创新应用

陈湘宁, 张鸿儒

北京农学院, 北京 102206

DOI: 10.61369/RTED.2025150019

摘 要 : 人工智能盛行时代下, 一系列 AI 软件的出现对高等教育产生巨大影响, 推动教育行业前进。以高校《R 语言数据分析》为例, 探讨人工智能时代下的课程教学改革措施, 挖掘人工智能对数据分析教学的影响。将 R 语言用于数据分析, 存在一定的困难, 需要掌握编程语言。对于非计算机专业学生而已, 难度大, 学习周期长, 严重影响软件的使用。因此, 本文重点介绍了 Deepseek 在 R 语言数据分析中的辅助教学应用。为人工智能时代教学智能转型, 培养学生的数据分析能力, 适应并运用 AI 技术需求提供参考。

关 键 词 : Deepseek; R 语言; 数据分析

Discussion on the innovative Application of Deepseek in the Course "R Language Data Analysis"

Chen Xiangning, Zhang Hongru
Beijing University of Agriculture, Beijing 102206

Abstract : In this era of widespread artificial intelligence, the emergence of a series of AI software has profoundly impacted higher education, propelling the advancement of the education sector. Taking the university course "Data Analysis with R" as an example, this paper explores teaching reform measures for the curriculum in the AI era and examines the influence of artificial intelligence on data analysis instruction. Employing R for data analysis presents certain challenges, requiring proficiency in the programming language. For non-computer science students, this presents considerable difficulty and a lengthy learning curve, severely hindering software adoption. Consequently, this paper focuses on the auxiliary teaching application of Deepseek within R language data analysis. It aims to provide insights for the intelligent transformation of teaching in the AI era, fostering students' data analysis capabilities to meet and apply AI technology requirements.

Keywords : Deepseek; R language; data analysis

引言

R 语言作为一门以统计分析与数据可视化为核心的编程语言, 在食品科学领域的研究与实践具有重要意义。R 语言拥有完整的数据准备、处理、分析、建模和可视化的系统, 对统计教学和科研实践起到很好的辅助作用^[1]。

对于食品专业学生而言, R 语言不仅能够支持常规的数据清洗、方差分析、回归建模和显著性检验, 还广泛应用于感官评价分析、营养成分数据处理、食品质量控制、多组学数据挖掘等多个研究环节。通过掌握 R 语言, 研究生可以更系统、更高效地处理实验数据, 提升科研的规范性和结果的可视化表达水平。同时, R 语言拥有大量食品与生物统计相关的扩展包, 如 ggplot2 (用于高质量图形绘制) 和 dplyr (用于数据清洗整理), 为食品领域的复杂数据处理提供了有力工具^[2,3]。在当前“数据驱动科研”趋势下, R 语言已成为食品专业学生开展科研、撰写论文和参与国际学术交流的重要技能支撑。基于 R 语言代码的数据分析及可视化是理解和交流实验发现的关键工具, 同时也确保了可扩展性和可重复性。然而, 复杂的编程语言对学生造成了重大障碍。R 语言虽然功能强大, 但其语法、函数调用和数据结构对初学者尤其是非计算机专业的食品、生物、管理等专业学生存在较高的学习门槛。将人工智能技术引入教学过程中, 能有效降低学生在代码编写与理解方面的障碍。

DeepSeek 能够作为一种 AI 辅助编程工具，能够高效地帮助学生和科研人员完成 R 语言环境下的数据清洗、统计分析、图形绘制等，降低 R 语言编程门槛，对初学者友好，减少语法错误和调试成本，提高科研效率，促进教学应用与课堂互动^[4]。

一、Deepseek 辅助 R 语言数据分析

（一）Deepseek 辅助 R 语言 dplyr 包

dplyr 是语言中常用的数据处理包，其典型的优势为具有语法简洁、操作直观、支持链式管道操作、具备强大分组汇总和数据清洗能力等特点，广泛应用于科研与教学中。然而，学生在使用过程中常因编程基础薄弱而面临多种困难，如函数命名混淆、管道语法不熟练、变量引用混乱、分组逻辑不清及报错难以调试等问题，特别是在缺乏真实数据背景与可视化联动的教学情境下，更容易产生畏难情绪与挫败感，影响数据分析能力的提升。DeepSeek 辅助 R 语言中 dplyr 包的使用具有显著优势，能够智能识别用户意图，快速生成符合语法的数据处理代码，降低管道操作和函数嵌套的学习门槛；同时具备错误提示与即时修改功能，帮助学生理解变量调用、分组汇总等核心逻辑，有效提升代码调试效率与学习成就感，特别适用于编程基础薄弱的专业学生在真实科研数据处理中的应用训练^[5]。

（二）Deepseek 辅助 R 语言 tidyr 包

tidyr 包是 R 语言中专用于数据整理与结构转换的核心工具，以“整洁数据”（tidydata）为理念，提供如 pivot_longer()、pivot_wider()、separate() 和 unite() 等函数，能高效完成宽表与长表转换、列拆分与合并等操作，极大提升数据分析前的预处理效率。其语法清晰、与 dplyr、ggplot2 等 tidyverse 套件高度兼容，适用于各类科研与教学场景。然而，学生在实际使用中常因数据结构理解不足、函数概念混淆及参数设置不熟悉等问题而产生困惑，尤其在处理多列批量转换或字符拆分时，容易出现逻辑错误或运行报错，影响数据分析流程的连贯性。

DeepSeek 在辅助学生使用 R 语言中 tidyr 包时，能够根据数据结构智能推荐适用的整理函数（如 pivot_longer()、pivot_wider()、separate() 等），自动生成标准语法代码，降低数据“整洁化”处理的理解与操作难度；同时具备错误定位与交互式解释功能，帮助学生快速掌握数据重构逻辑，有效提升数据预处理效率与可视化准备质量，特别适用于教学与科研中的复杂数据清洗任务。

二、Deepseek 辅助 R 语言数据可视化

（一）Deepseek 辅助 R 语言 ggplot

R 语言 ggplot2 包具有图层叠加理念（LayeredGrammar）图像由多个图层（layers）组成，例如：基础图层、几何图层（如点、线、柱状）、坐标系、标签、主题等，允许用户“像搭积木一样”逐层构建图形，便于个性化控制与美学表达。该包支持坐标轴标签、图例、颜色、形状、字体、主题等各类细节的精细化

调节，可以绘制科学论文级别的高质量图形。

学生在学习 ggplot2 可视化模块时普遍面临较大困难，尤其是编程基础薄弱的学生。在 ggplot2 中，图形的构建依赖于“图层叠加”与“美学映射”逻辑，需要学生具备较强的逻辑思维与代码表达能力。然而，由于大多数食品、生物等专业学生并非计算机背景，缺乏编程训练，常在基本语法（如函数嵌套、参数传递、引号使用）上出现频繁错误。例如，aes() 函数内外变量混淆、图层顺序错误等问题，极易导致图形无法正确呈现，挫败感强。加之 ggplot2 的语法高度抽象、参数多样，学生在理解和复用他人代码方面也存在困难，缺乏调试与自我修复能力。对于这类编程基础薄弱的学生群体，传统的黑板教学或 PPT 展示往往难以满足差异化学习需求，迫切需要借助人工智能辅助工具、可视化教学平台与渐进式代码训练方法，降低学习门槛，提高图形构建与数据表达的能力。

DeepSeek 在辅助学生使用 R 语言中的 ggplot2 包进行数据可视化方面优势显著。它能够根据学生的自然语言输入智能识别绘图需求，自动生成符合 ggplot2 语法规范的代码，帮助学生快速掌握图形构建的基本结构和逻辑。同时，DeepSeek 可以根据数据特征推荐合适的图表类型（如散点图、条形图、箱线图）和所需的图层（如 geom_*、facet_wrap、labs 等），减少学生对复杂语法的依赖。此外，系统还能针对配色、坐标轴标签、图例样式等视觉元素提出优化建议，帮助学生生成美观规范的图形，提升图表表达质量。针对初学者常见的语法错误，DeepSeek 提供清晰的提示和纠错建议，提升学生的调试能力和自学效率。

（二）Deepseek 辅助 R 语言 plotly

R 语言中 plotly 包具有强大的交互式可视化能力，能与 ggplot2 无缝集成并生成动态图形，适用于科研演示和网页嵌入，但学生在使用过程中常因其语法结构复杂、输出环境特殊及交互参数设置较多而感到困惑，影响图形调试与实用性理解^[6]。

DeepSeek 在辅助使用 R 语言中 plotly 包时，能够根据学生绘图要求智能生成交互式图形代码，自动匹配美学参数与布局设置，显著降低 hover 信息定制、图例控制和多图联动等复杂操作的门槛；同时具备即时纠错与语法优化能力，帮助学生快速掌握交互式图形结构与输出格式，提升科研演示质量与学习效率，尤其适合网页展示与教学可视化应用场景。

三、结语与展望

在数据科学教育日益普及的背景下，DeepSeek 等大语言模型在 R 语言教学中的引入，为高校课程改革与教育创新注入了新的动力。DeepSeek 通过智能代码生成能力，有效弥合了理论知

识与实践操作之间的鸿沟，帮助学生更高效地掌握 R 语言的数据分析与可视化技能。其即时反馈、智能纠错、跨模块知识关联等特点，不仅提升了教学效率，也激发了学生的学习兴趣与探索主动性。

展望未来，DeepSeek 有望在教学过程中实现更加个性化的

学习路径推荐与智能辅导，逐步构建“AI 教师 + 人类教师”的混合式教学新模式。同时，随着模型对教学内容理解能力的增强，DeepSeek 还可助力课程内容更新、案例生成与实验设计，成为高校数据分析课程智能化转型的重要推动力。如何科学嵌入教学体系，规范其使用边界，也将成为教育者持续探索的重要课题。

参考文献

[1] 龚高, 刘玲玲, 陈秋明, 刘武军. R 语言在《生物统计附试验设计》实验教学中应用 [J]. 现代畜牧科技, 2025(07): 167 - 170.

[2] 袁佳. R 语言及 ggplot2 在环境空气监测数据可视化中的应用 [J]. 中国高新技术企业, 2015(16): 88 - 91.

[3] 刘红伟, 邓晓伟, 段同庆, 李长平, 马骏. data.table 和 dplyr 软件包在数据操作方面效率的评价 [J]. 中国卫生统计, 2020, 37(01): 141 - 144.

[4] 陈世浩. 基于 DeepSeek 的智能编程教学系统的设计与实现 [J]. 信息与电脑, 2025, 37(12): 102 - 105.

[5] Jayas D S. Food Dehydration[M]//Reference Module in Food Science. Elsevier, 2016.

[6] 应轩宇, 蔡强, 纪伟. 基于 R 语言的可视化技术在食品生产经营主体风险分级数据上的应用 [J]. 食品安全导刊, 2023(27): 162 - 164.