

《数据分析与R语言》课程思政教学的探索和实践

李薇宇*, 程 瑶, 刘 增

苏州科技大学数学科学学院, 江苏 苏州

收稿日期: 2022年3月7日; 录用日期: 2022年4月6日; 发布日期: 2022年4月13日

摘 要

《数据分析与R语言》课程是统计专业课程,旨在从海量的、有缺失的或者是有噪声的实际数据中,借助R软件挖掘具有实用价值的信息,从而帮助学生了解复杂社会现象背后的规律,为国家培养社会急需的应用型人才。该课程针对实际案例进行数据分析,可以很好地结合社会热点进行思政教育。本文从教学目标、教学内容、课程考核、课程评价等方面出发,就如何挖掘思政元素进行探讨,并为面向理科生的思政教育提供了新的思路。

关键词

课程思政, 专业课程, 理科专业, 数据分析

Exploration and Practice of Ideological and Political Education in the Course of Data Analysis and R

Weiyu Li*, Yao Cheng, Zeng Liu

School of Mathematical Sciences, Suzhou University of Science and Technology, Suzhou Jiangsu

Received: Mar. 7th, 2022; accepted: Apr. 6th, 2022; published: Apr. 13th, 2022

Abstract

The Data Analysis and R is a professional course in the Department of Statistics. It aims to mine useful information from massive, missing or noisy data with the help of R software, so as to help

*通讯作者。

students understand the laws behind complex social phenomena, and cultivate applied talents urgently needed by the society for the country. The course conducts data analysis on real cases and can be well combined with social hotspots for ideological and political education. This paper provides a new idea to explore the ideological and political elements and to conduct ideological and political education for science students from several aspects, such as the teaching objectives, teaching content, course assessment, and course evaluation.

Keywords

Ideological and Political Education, Professional Course, Science Major, Data Analysis

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

当前,在国家政策引导和社会需求的双驱动下,以5G、大数据中心、人工智能、工业互联网等为代表的新型基础设施建设(简称:新基建)迎来发展热潮,急需一大批高水平创新型、复合型、应用型人才。高校教育与经济社会的共生关系决定了高校人才培养需要呼应社会需求,培养地方经济社会发展急需的创新复合型工程技术人才成为各地方高校面临的重要课题。

新基建对人才要求很高,除了需要具备扎实的理论基础、专业系统的知识体系、敏锐的科学直觉、解决问题的突出能力外,还需要有自由求真的探索精神以及为祖国科学献身的远大抱负和理想,这意味着高校在人才培养中要坚持知识性和价值性相统一[1],既要培养学生的专业知识和技能,又要塑造学生的社会主义核心价值观。因此,高校教育要将思政教育渗透到教学工作中,通过“立德树人”,实现知识传授、能力培养和价值引领的统一[2]。

“数据分析与R语言”是一门理科专业必修课程,旨在从海量的、有缺失的或者是有噪声的实际数据中,借助统计应用软件R挖掘具有实用价值的信息以及知识[3]。其中,挖掘到的信息和知识包括模型、统计规律、应用规则、使用模式等[4]。相对人文社科专业,课程思政在理科专业中的探索相对较少,可参考的成熟案例并不多[5]。因此,本文针对“数据分析与R语言”课程,探索课程思政融入该课程教学的建设思路与实施途径。

大数据时代已经到来,在商业、经济及其他领域中基于数据和分析去发现问题并做出科学、客观的决策越来越重要。本文将针对社会对应用型人才的培养需求,构建模块化教学,结合“国货之光”瑞幸咖啡造假事件[6]、“天河二号”受芯片制约[7]、以及抗击新冠肺炎疫情的中国行动等社会热点问题,就如何在“数据分析和R语言”课程中开展思政教育进行探索。本文旨在针对理科专业大学生“讲道理”的思维特点[5],构建沉浸式思政课堂,使思政课与社会现实紧密结合[8],用数据说话,避免灌输式的思政教育模式,尝试为高校推进理学课程思政改革提供参考借鉴。

2. 课程介绍和授课对象

《数据分析与R语言》课程是统计学专业第6学期开设的专业教育必修课,属于数据分析与建模能力模块。本课程共2学分32学时,其中课程讲授24学时,上级实验8学时。选用教材为张良均等[3]编著的《R语言与数据挖掘》。该教材具有丰富的案例和程序,可操作性强,比较适合没有R语言编程基

础和数据挖掘基础的学生,学生可以通过学习本书中的理论知识和上机实践,逐步掌握使用 R 语言进行数据挖掘的方法。

本课程的授课对象为理科专业学生。理科生的思维特点是“讲道理”,对社会问题相对缺乏敏感或关注,导致他们对于思想政治教育的认知水平和理解能力往往低于人文社会科学专业学生[5]。教师通过灌输式、说教式进行思想政治教育可能会引起学生抵触情绪,引发学生对思政教育持一种片面的、消极的态度。为了达到理想的教学效果,本课程在切入思政元素时应根据理科学生的思维特点,注重逻辑性、科学性、真理性。在此基础上,结合课程特点,寓价值观引导于实际数据的案例分析之中,用数据说话,落实立德树人根本任务。同时,在思政元素取材上要确保新鲜性。思政元素陈旧、缺少新鲜感是目前大学思政课程遇到的困境之一,因此,我们在挖掘本课程所蕴含的思政育人元素时,会适当地结合一些当前热点问题进行分析,这里强调适当是因为如果一堂课呈现的思政元素数量太多,就会冲淡专业课程内容,这不仅影响专业知识的传授,而且可能招致学生的反感,更谈不上取得思政育人的效果。下面我们将重点介绍如何开展《数据分析与 R 语言》课程的思想政治教育工作。

3. 思政教学的建设思路与举措

3.1. 在教学目标方面体现思政目标

本课程将通过课堂教学、讨论以及上机的方式,使学生掌握大数据分析的基本概念、基本原理和基本方法,掌握有关数据发掘、处理、建模和解释的基本原理和方法,了解和熟悉数据分析在社会科学研究、商业分析和公共管理等领域的实际案例。《数据分析与 R 语言》课程的教学目标主要分为以下三个方面:

1) 知识目标。使学生掌握用 R 语言进行数据挖掘的常用算法函数,了解算法背后的统计学原理及其他数学知识点。掌握数据分析的基本流程,包括从数据处理、数据分析到分析结果整理的过程。掌握数据分析的基本原理,包括描述统计分析,差异性和相关性分析,以及基本数据建模方法。掌握数据分析的基本工具,能够正确使用 R 软件和 R 包整理、分析数据。

2) 能力目标。不仅要向学生讲授各种函数的使用格式,还要介绍函数的应用场景和适用范围。通过将函数的应用实例和运行结果解析贯穿于整个教学过程,使学生逐步掌握各种函数的优劣及限制条件。通过上机实践,明晰函数应用于实际数据的编程实现过程,从而培养学生利用 R 语言解决大数据分析和决策的相关实际问题的能力。

3) 价值目标。在知识的传授中,培养学生质疑、批判、严谨、求真的科学文化素养以及热爱祖国、献身科学的家国情怀,帮助学生养成良好的社会责任感和职业道德,助力学生建立正确的科学观、世界观。

其中,价值目标的实现需要借助思政教育的力量,具体从以下三个方面体现:

1) 引导学生从图、表和数字等不同方式展示的数据中挖掘复杂社会现象背后的规律,从而培养学生的人文素养和实事求是的专业习惯、深化其严谨的科学思维素质、激发不畏艰难、尽心尽力的敬业精神。

2) 通过丰富的案例分析(例如,线性回归在碳排放方面的应用),引导学生关注时事和社会问题,从而培养学生的社会责任感和参与意识。

3) 依托数据总结国家发展取得的一些成就(例如,中国抗疫的成就),激发学生爱国主义情怀,内化社会主义制度自信。

3.2. 在模块化教学内容中切入思政元素

应用型、技术技能型人才的培养,不仅直接关乎经济社会发展,更是关乎国家安全命脉的重大问题。对此,本课程以“能力导向”为课程建设目标、以知识“应用”为课程评价标准,为国家培养具备初步研发能力的应用型人才。应用型人才培养的关键是要将知识转化为能力。然而,数据分析教学中使用的

教材普遍是分章节介绍的,这样看上去课程体系完整,但不利于培养学生运用所学知识解决实际问题的能力。在实际教学过程中,我们根据数据分析流程对课程知识进行整合,以能力为本位构建模块化教学体系,打通理论与实践教学环节。最终,把九个章节教学内容转化为四大模块:数据收集模块、数据管理模块、数据建模模块、数据分析报告模块。我们在四大模块教学过程中,将结合教学内容,通过各模块的案例分析,分别融入严谨的科学思维,实事求是的专业素养,学术道德与学术规范,不畏艰难、尽心尽力的敬业精神,社会热点选题中的爱国主义情怀、社会主义制度自信,实现民族复兴的理想和责任等思政教育元素。模块化教学内容中思政元素的切入详见表 1。

Table 1. Introduction of ideological and political elements in the modular teaching mode
表 1. 模块化教学模式下思政元素的切入

模块	教学内容	思政案例	思政元素
数据收集模块	1) 数据类型和数据结构。 2) 数据收集的方法与技能,包括数据埋点、爬虫、第三方数据统计平台、问卷调查等。	1) “国货之光”瑞幸咖啡造假事件造成恶劣影响,导致大量员工失业,重创了中概股在国际股票市场上的声誉[6]。“直播带货一姐”薇娅偷逃税被大数据分析发现。 2) 薇娅的走红也反映了数字经济时代下,数据成为关键的生产要素,成为重要的价值创造来源。2020 年我国数字经济规模达到 39.2 万亿元,占 GDP 比重为 38.6%。	1) 以数据造假问题告诫学生不好的数据不可怕,隐瞒篡改数据才可怕,提升学生的学术道德和诚信意识,培养学生严谨的科学思维与实事求是的专业素养。 2) 使学生意识到数字经济时代,应以大数据应用为发展目标,培养自己在数据采集、存储、处理、挖掘、应用、展示、衍生等方面的能力,从而帮助学生明晰未来目标。
数据管理模块	1) R 中常用的数据管理函数,涉及合并数据、清洗数据、标准化数据、转换数据等。 2) R 语言数据可视化函数,特别 ggplot2 绘图函数演示。	1) 长期以来,中国超级计算机已经超过美国,如:2010~2015 年,“天河二号”超级计算机连续六年在全球超级计算机 500 强榜单中以每秒 3.386×10^{16} 次计算称雄;“天河二号”已应用于多个领域,包括海量数据处理、生物医药、天气预报、智慧城市、云计算与大数据等,为经济社会转型升级提供重要支撑。 2) “天河二号”一期工程是以美国英特尔公司的处理芯片为主,自主研发的处理芯片为辅。在芯片技术出口受到限制之后,美国超级计算机重新排在第一位[7]。	1) 介绍“天河二号”计算机超强的计算和处理数据的能力,强调其在多个领域取得了显著的经济效益和社会效益,借此激发学生对中国的科技文化和其发展充满自信。 2) 从“天河二号”受芯片制约,延伸到中美贸易战中的“中兴事件”“华为事件”,使学生认识到中国芯片技术在整体上相对落后的现状,意识到只有掌握核心技术,才能够不受制于人,激励学生奋发图强,树立为国担当的信念。
数据建模模块	1) 介绍在数据挖掘中经常用到的几种类型的建模函数。 2) 对真实数据进行分析,确定变量间的相关关系,建立数学模型。 3) 对变量变化趋势进行反映,并对模型进行评价。	1) 2020 年中国武汉爆发新冠肺炎病毒疫情,中国政府实事求是,及时公布疫情,并且听取医学专家的意见,将武汉及时“封城”,有效地阻止了疫情的传播和肆虐。医学专家对已经掌握的数据进行了建模,并预测了人为干预即“封城”后,可以大幅减少感染人群对外的传播。以小组为单位让学生对美国疫情数据进行分析 and 预测,对比中国抗疫情况。 2) 课上,以近期宣传正能量的主题电影如“战狼”为例,获取影评信息,分析电影评论,以此获得电影好评和差评的原因。课下,要求学生以小组为单位完成几项综合任务:以政府工作报告为素材,让学生获取“红海行动”、“长津湖”等影评信息以及评论人所在地、年龄等信息,分析电影评论中给予的好评和差评是否和个人所在区域、年龄等相关。	1) 通过中美抗疫表现对比,体会新冠肺炎疫情下社会主义体制的优越性,激发学生科技报国的家国情怀和使命担当。 2) 在课堂中,通过小组合作,培养学生团队协作能力,让学生亲历建模过程,从实际数据中体会国家誓死捍卫人民生命的态度和决心。

Continued

数据分析 报告模块	1) 解读建模结果, 提炼出重要信息, 透过信息, 总结出管理决策。	1) 邀请大数据领域从业者对行业发展现状作报告, 通过特邀嘉宾对于亲身经历的讲述, 以及与学生现场互动的方式, 将目前大数据领域的技术岗位需求鲜活地展现在学生面前, 加强沉浸式“大思政课”的交互感。	1) 目前大数据领域的技术岗位主要集中在大数据平台开发、大数据应用开发、大数据分析和大数据运维四个岗位上, 本科生往往会从事大数据应用开发岗位。将行业现状以亲历者身份呈现给学生, 让学生提前做好职业规划, 加强对自己专业知识和技能的锻炼, 努力让自己的职业素养与行业领域对于大数据人才的需求相契合。
	2) 以数据为出发点解决大数据分析和决策的相关问题, 并配以图表展示。		

3.3. 在课程考核环节突出思政因素

《数据分析与 R 语言》以培养学生能力为导向, 以知识应用为课程评价标准, 对学生的学习效果评价主要源于学生在课堂讲授与课程实验两个环节的表现。根据这两个环节的教学层次和教学内容, 我们建立了一套适应思政育人的新型考核体系。

1) 设立既公平又体现差异化的考试评价体系。学生得到公平的对待是激发学生学习兴趣最重要的手段之一, 本课程在评分机制和评分细则两方面体现公平性。采取“平时成绩 + 期末考试”的模式对学生进行考核, 其中平时成绩占总评成绩的 40%, 主要由四项构成, 分别是课后作业(占 30%)、一个团队合作项目(占 50%)、和课堂参与度(占 20%)。课后作业考查学生的课堂知识消化能力; 团队合作项目通过挖掘思政案例、布置多样化开放性的题目考查学生的动手实践能力和团队合作能力; 课堂参与度由学生在出勤率、问题回答、小型汇报以及同学互评上的综合表现决定, 主要考查学生的学习态度以及口头表达能力。期末考试占比为 60%, 其中基本知识点占总体的 50%, 上机课中知识点的扩展占比 40%、开放性试题占比 10%。为了在保障公平性的同时体现差异性, 本课程采用了细化的五级制评分标准, 见表 2。

Table 2. Curriculum grading criteria
表 2. 课程评分标准

成绩(百分制)	等级	成绩(百分制)	等级
97~100	A+	77~79	C+
93~96	A	73~76	C
90~92	A-	70~72	C-
87~89	B+	67~69	D+
83~86	B	60~66	D
80~82	B-	0~59	F

2) 思政元素考核。在团队合作项目和期末考试开放性试题环节渗透对思政元素的考核。教师在学期初提供一些与思政相关的主题, 学生根据兴趣进行选题并按全国大学生数学建模竞赛和统计建模大赛规定的 3 人一组进行组队。学生针对每个模块所学内容, 对所选主题进行实战演练, 最终形成一份数据分析报告。学生须汇报研究背景、研究内容、主要成果、应用前景、小组分工以及体现的思政元素, 教师根据各小组研究成果、汇报表现以及关于思政元素的论述进行打分。此外, 学生在期末开放性试题解答中须结合当下热点和思政元素进行论述和关联。

3.4. 在教学效果评价方面注重思政育人考核

教学质量评估是判断《数据分析与 R 语言》课程思想政治教学效果的重要环节。为了保障教学质量

评估结果的客观性, 我们将从同行评价和学生评价两个方面进行考查。

在同行评价中, 增加对课程思政教学理念、方法、手段及实施效果的评价, 具体的评价内容包括: 课程思政设计系统、主题凝练、与本专业其他课程的协同性、课程标准融入课程思政的合理性、课程团队建设及课程教学资源建设的效果。

在学生评价中, 借助思政元素教学评价表(见表 3), 增加学生对教师思政育人的有效评价, 旨在获取学生对融合了思想政治元素的“数据分析与 R 语言”课程的评价和反馈, 其中, “是”代表有促进作用, “否”代表没有影响, “其他”表示放弃。

Table 3. Teaching evaluation of ideological and political elements

表 3. 思政元素教学评价表

课程授课情况	是	否	其他
激发了奋发图强的昂扬斗志			
深入理解社会主义核心价值观			
增强民族自豪感			
增强了爱国主义情感			
增强了利用所学服务社会、服务人民、甘于奉献的信念			
对中国的科技文化和其发展充满自信			
体会到了重大灾难面前社会主义体制的优越性			
重塑工匠精神, 培养自己的专注力和潜心做好一件事的决心			
激发了自己科技报国的家国情怀和使命担当			
对自己的未来目标明晰, 充满信心			

通过学生的课程评价问卷调查结果得知, 大部分学生对于本课程评价良好, 学习热情有所升温, 数据分析的基本思维有所提高, 实际动手能力和团结协作能力有所加强。以学生典型评价为例: “这门课程比较实用, 有许多理论与实践结合的点, 同时还会就知识点引出当下热点案例进行分析讨论, 通过本课程的学习, 能够掌握数据分析的基本流程, 包括从数据处理, 数据分析到分析结果整理的过程”。

4. 结语

将课程思政渗透到教学之中是新时代高校教师的使命。“数据分析与 R 语言”课程是统计专业课程, 旨在培养学生用统计学和 R 语言解决实际问题的能力。该课程与实际联系紧密, 授课方式应生动, 在每个模块都有很多与学生生活密切相关的真实案例, 可以很好地结合社会热点进行思政教育。为了发挥思想政治理论课与本课程的协同效应, 教师需要不断学习, 提升自身的思政素养, 不断思考如何在该课程的教学过程中自然而又有趣地融入思政元素, 努力为国家培养大数据时代所需要的具备对海量数据进行专业化处理能力的创新型应用型人才。

基金项目

苏州科技大学数学科学学院课程思政示范课程建设项目(MKCSZ202101); 苏州科技大学教学改革与研究项目(2021JG-15)。

参考文献

- [1] 康玉梅, 王延邦, 陈猛. 基于类别思政元素的专业课课程思政改革与实践[J]. 辽宁教育行政学院学报, 2020,

37(5): 52-55.

- [2] 龚一鸣. 课程思政的知与行[J]. 中国大学教学, 2021, 369(5): 77-84.
- [3] 张良均, 谢佳标, 杨坦, 肖刚. R 语言与数据挖掘[M]. 北京: 机械工业出版社, 2016.
- [4] 郑晓川. 大数据智能分析及数据挖掘探讨[J]. 中国科技信息, 2021(21): 35-36.
- [5] 王宝军. 大学理科专业课程思政的特点和教学设计[J]. 中国大学教学, 2019(10): 37-40.
- [6] 范芯菱, 胡北忠. 上市公司财务舞弊动因及对策探析——以瑞幸咖啡造假事件为例[J]. 中国管理信息化, 2021, 24(23): 30-31.
- [7] 赵阳辉, 陈方舟, 温运城. 国之重器: “天河”高性能计算机发展历程[J]. 科学, 2016, 68(3): 50-53+64.
- [8] 许瑞芳, 张宜萱. 沉浸式“大思政课”的价值意蕴及建构理路[J]. 思想理论教育导刊, 2021, 275(11): 83-88.