



華東師範大學  
East China Normal University

# 基于多源数据的学生核心素养发展评价方 法设计与学生画像研究

晏玉琴 51255903109

朱奕帆 51265903116

2023 年 12 月 15 日

## 摘 要

为了提升 21 世纪国家人才核心竞争力，2016 年教育部发布中国学生发展核心素养研究成果，将核心素养概括为三个方面的六大素养，并具体细化为十八个基本要点。然而，核心素养教育在实践中却面临种种困难与挑战，被广泛认为最具有挑战的问题就是如何针对学生的核心素养建立科学的评价体系。随着近年来智慧校园的不断发展和高校核心素养教育的层层推进，高校数据库中产生了海量拥有丰富价值的学生在校行为数据和素养教育数据。为了解决上述的挑战，我们提出了一套基于多源数据的学生核心素养发展评价方法并基于 DataEase 平台设计可视化网页为学生画像，帮助高校更直观地掌握学生核心素养的发展情况，从而辅助学校素养教育工作和培养方案制定工作的进行。案例分析结果显示，我们提出的方法对学生各个维度核心素养的评价较为准确且可解释性高，而且既可以对比不同学生间的核心素养发展情况，也可以对比同一学生不同时期的核心素养发展情况，评分结果对于高校素养教育决策具有一定参考价值和实践意义。

# 目录

<b>1 项目介绍</b>	<b>4</b>
1.1 项目背景	4
1.2 项目主题	6
<b>2 数据介绍</b>	<b>7</b>
2.1 校园卡消费数据	7
2.2 图书馆借阅数据	7
2.3 社团注册数据	7
2.4 第二课堂数据	8
2.4.1 第二课堂-学生活动详情	8
2.4.2 第二课堂-学生参与活动详细信息	8
2.4.3 第二课堂-学生培养方案进度	8
2.4.4 第二课堂-学生培养方案进度详情	8
2.4.5 第二课堂-学生已完成活动参与记录	8
<b>3 数据分析</b>	<b>9</b>
3.1 核心素养的数据化表达	9
3.2 核心素养发展的综合评分模型	12
3.2.1 数据预处理	12
3.2.2 多视角特征提取	12
3.2.3 基于因子分析的学生核心素养发展综合评分	13
3.3 案例分析	15
<b>4 数据可视化展示</b>	<b>17</b>
4.1 个人画像	17
4.2 多人画像和院系画像	20
4.3 全校总览	21
<b>5 总结与展望</b>	<b>23</b>
5.1 总结	23
5.2 展望	23

# 1 项目介绍

## 1.1 项目背景

近年来，为了落实立德树人根本任务，培养全面发展的人，提升 21 世纪国家人才核心竞争力，教育部和各大高校开展了大量相关研究和工作。

2014 年，教育部研制印发《关于全面深化课程改革落实立德树人根本任务的意见》，提出“教育部将组织研究提出各学段学生发展核心素养体系，明确学生应具备的适应终身发展和社会发展需要的必备品格和关键能力”。2016 年 9 月 13 日上午，中国学生发展核心素养研究成果发布会在北京举行。中国学生发展核心素养以培养“全面发展的人”为核心，充分反映新时期经济社会发展对人才培养的新要求，高度重视中华优秀传统文化的传承与发展，系统落实社会主义核心价值观。如图 1 所示，核心素养分为文化基础、自主发展、社会参与 3 个方面，综合表现为人文底蕴、科学精神、学会学习、健康生活、责任担当、实践创新等六大素养，并具体细化为国家认同等 18 个基本要点。各素养之间相互联系、互相补充、相互促进，在不同情境中整体发挥作用 [1]。

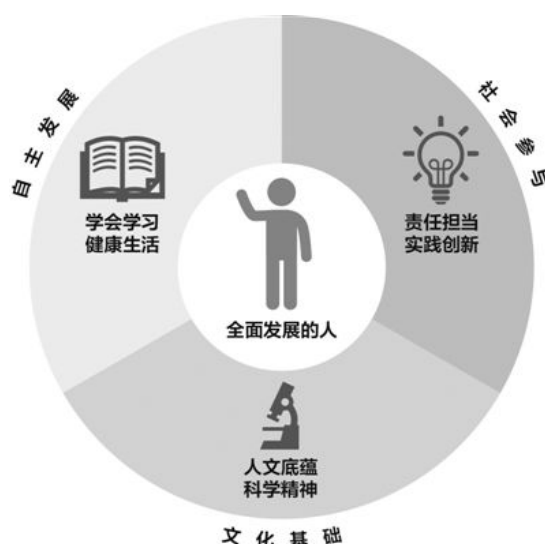


图 1: 中国学生发展核心素养

然而，核心素养在实践中具体落实时却步履维艰，这是因为素养教育可以反映在整个学校方方面面的教育教学之中，反映在学校的显性和隐形文化中，导致其具体落实时变得格外不易。推动课程改革、规范教师培训、建立科学评价缺一不可。而如何针对学生的核心素养建立科学的评价体系，是专家学者们一致认为最具有挑战的问题，也是本项目重点关注解决的问题。

2016 年 9 月 26 日，由中国教育学会主办，以“评价与学生核心素养”为主题的专题研讨会中指出未来探索的方向：要强化对于软技能和软环境的评价，善用互联网和大数据，对每个学生个体的整个受教育过程的数据和信息进行过程性保存和追踪，并进行深入分析，从而为核心素养的落实和综合素质评价提供技术支撑。



随着智慧校园的不断发展，学生在校学习生活的各种行为产生了海量的行为数据，这些学生的行为数据可以在一定程度上反映学生的行为模式，挖掘和分析这些数据可以支撑学校开展贫困生识别 [2]、社交关系挖掘 [3]、学生成绩预测 [4] 等各种研究，因此其中蕴含着丰富的数据价值。另外，在调研学生发展核心素养的评价体系建立相关研究后我们发现，大多数研究侧重于理论框架和技术路线的分析设计 [5, 6, 7]，而缺少真实数据驱动的方式设计学生发展核心素养的评价体系，这可能是因为核心素养的概念提出不久，且各高校关于核心素养的课程仍在不断改革，所以大部分高校在现阶段仍需要若干年的时间来积累足够的数据基础。

2021 年 1 月 9 日，华东师范大学发布《关于制订全育人理念下专业培养方案的指导意见》，提出学生在毕业时应具备明德乐群、基础扎实、反思探究、身心健康、持续发展和国际视野 6 个共同素养。并要求各专业制定培养方案时，应围绕人才培养目标，以学生的能力达成和素质发展为核心，加强第一课堂和第二课堂的深度融合。如图2展示的是华东师范大学第二课堂培养流程，首先院系在系统中为新生核查和培养方案，然后新生入学后，可以参与系统中或系统外的活动来认定第二课堂学分和培养方案进度，最后在学生毕业前，院系对第二课堂计划完成情况进行最终认定。

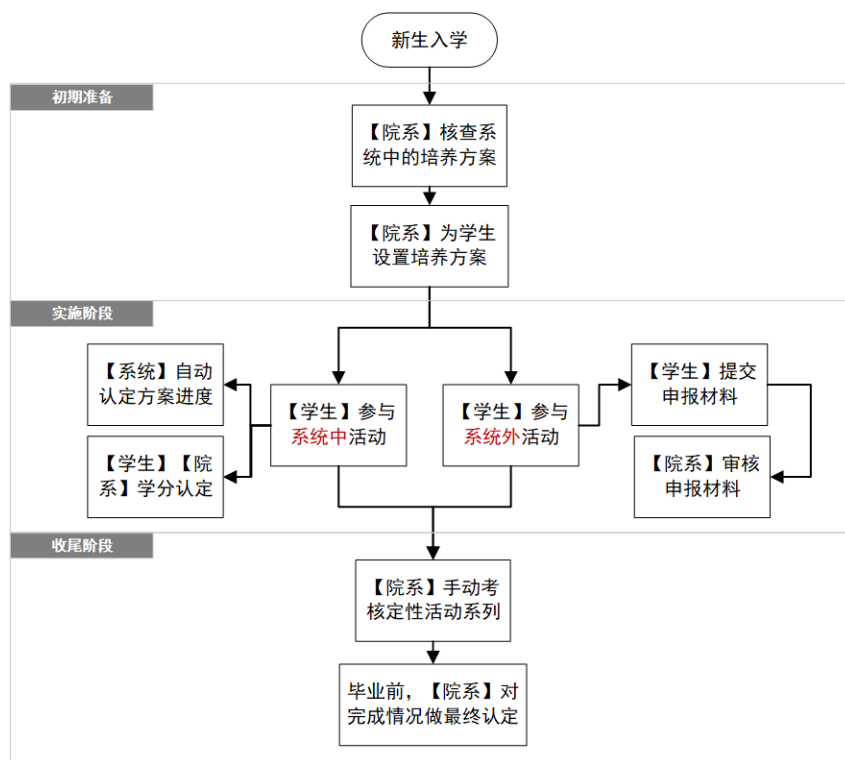


图 2: 第二课堂培养流程

图3是某院系 2023-2024 学年开展第二课堂活动的菜单。在学校关于第二课堂活动建设的过程中产生了丰富的数据，这些数据包含各院系开展不同类型第二课堂活动的详细次数、时间、参与人次、对应模块等，也包含每个学生参与各类第二课堂活动的详细次数、时间、评分评语、计划完成进度等。对这些数据进行挖掘和研究，可以帮助分析

学生的核心素养培养情况，帮助判断学生哪些维度的核心素养培养存在不足，因此这些数据可以为设计学生发展核心素养的评价体系提供丰富且真实的数据基础。

大夏书院第二课堂活动菜单						
2023-2024学年 ///						
活动模块	活动系列	活动时间	活动频率	内容简介	参与人数	类别
思想素质	校园公益系列	秋季 / 春季学期	2 期 / 学期	1. 校内公益服务 2. 书院主题活动志愿者	1. 个人、团队报名参加	定量
	场馆志愿者系列	秋季 / 春季学期	1 期 / 学期	1. 上海各类场馆志愿服务	1. 个人、团队报名参加	定量
志愿服务	向社区报到系列	寒假 / 暑假	1 期 / 学期	1. 在乡村、街道和社区参与基层治理、志愿服务和社会调研	1. 个人、团队报名参加	定性
	幸福研习系列	秋季 / 春季学期	3 期 / 学期	1. 幸福日记 2. 记录“小确幸” 3. 正念冥想体验	1. 个人、团队报名参加	定量
社会实践	新生团培营	秋季学期	1 期	1. 破冰与团队建设	2. 院系、班级组织参加	定量
	小班团建营	秋季学期	1 期	3. 面向院系、班级定制	3. 面向院系、班级定制	定量
	寝室 MINI 马拉松	秋季学期	1 期	1. 校内 5 公里马拉松	1. 个人、团队报名参加	定量
	书院杯篮球赛	秋季学期	1 期	1. 篮球比赛	2. 院系、班级组织参加	定量
	师生网球交流赛	秋季 / 春季学期	1 期	1. 网球比赛 2. 网球体验	1. 个人、团队报名参加	定量
	沪上高校校园体育交流赛	秋季 / 春季学期	视具体情况而定	1. 足球校际交流	1. 个人、团队报名参加	定量
	迎国庆 10.1 公里徒步	秋季学期	1 期	1. 浦江公园徒步	1. 个人、团队报名参加	定量
	社区体育竞赛	秋季 / 春季学期	12 次 / 学期	1. 网式足球 2. 太极拳	1. 个人、团队报名参加	定量
	国画、书法、水彩、陶艺、茶艺	秋季学期 / 春季学期	12 次 / 学期	1. 国画 2. 书法（软笔、硬笔） 3. 水彩 4. 陶艺 5. 茶艺	1. 个人、团队报名参加	定量
	拉丁舞、中国舞、现代舞	秋季学期 / 春季学期	12 次 / 学期	1. 拉丁舞 2. 中国舞 3. 现代舞	1. 个人、团队报名参加	定量
美育实践	全球胜任力	春季学期	1 期	1. 国际组织实习与任职 2. 外事礼仪 3. 跨文化沟通	1. 个人、团队报名参加	定量
	影子系列	秋季 / 春季学期	4 期 / 学期	1. 影子秘书长 2. 影子选调生	1. 个人、团队报名参加 3. 面向院系、班级定制	定量
生涯发展	企业参访系列	秋季 / 春季学期	4 期 / 学期	1. 在沪企业 2. 新闻媒体	1. 个人、团队报名参加 3. 面向院系、班级定制	定量
	生涯下午茶系列	春季学期	4 期 / 学期	1. 出国升学深造 2. 选调生选拔 3. 企业求职与创业政策 4. 各类教师选聘与成长	1. 个人、团队报名参加	定量
	赴国境外留学个性化咨询系列	秋季 / 春季学期	4 期 / 学期	1. 美国 2. 英国、新加坡、澳大利亚 3. 香港、澳门	1. 个人、团队报名参加	定量
	学术朋辈谈系列	秋季 / 春季学期	12 次 / 学期	1. 不是天生优秀，而是吃苦耐劳：我是这样过大学 2. 在集体中成长：我这样锻炼领导力 3. 学会适应大学的那些变化：我这样处理自己的情绪	1. 个人、团队报名参加	定量
人文素养	翰墨大夏系列	秋季学期	1 期 / 学期	1. 从读一本书到撰写读书笔记 2. 院长联合荐书与读书笔记征集	1. 个人、团队报名参加	定量
	书香大夏跨学科导读系列	秋季 / 春季学期	14 期 / 学期	1. 50 本书导读（每学期 25 本）	1. 个人、团队报名参加 3. 面向院系、班级定制	定量
	学术下午茶系列	秋季学期	4 期 / 学期	1. 中学 - 大学学习方法转换与学术入门 2. 从知识获得到提出问题	1. 个人、团队报名参加	定量
	中文演讲表达训练营	秋季 / 春季学期	1 期	1. 演讲的基本技能 2. 演讲表达实操	1. 个人、团队报名参加 3. 面向院系、班级定制	定量
	英文演讲表达训练营	秋季 / 春季学期	1 期	1. 英语口语 2. TED 演讲	1. 个人、团队报名参加	定量
创新创业	传统节日系列	秋季 / 春季学期	4 期 / 学期	1. 中秋 2. 冬至 3. 元宵 4. 端午	1. 个人、团队报名参加	定量
	创新创业项目分享	秋季 / 春季学期	4 期 / 学期	1. 项目孵化 2. 项目计划书撰写 3. 参赛经历	1. 个人、团队报名参加 3. 面向院系、班级定制	定量

图 3: 某院系第二课堂活动菜单

## 1.2 项目主题

时至今日，第二课堂已经是师大卓越养成教育的一部分。从 2021 级本科生入学至今，学生参与第二课堂活动，以及完成个人养成教育方案的过程中，已经产生了丰富的记录数据，如何利用这些数据，并结合学生在校消费数据、图书馆借阅数据和社团注册记录数据等，为学生核心素养发展情况进行画像，帮助学校更直观地了解学生不同维度的核心素养的发展情况，从而辅助学校常规教学工作、第二课堂活动开展工作和培养方案制定工作的进行，是我们本次研究的主题。

## 2 数据介绍

本项目一共使用了 8 个数据集，分别是包含学生消费行为的校园卡消费数据、包含学生图书馆借书行为的图书馆借阅数据、包含学生参加社团行为的社团注册数据、包含各学院举办第二课堂活动信息的学生活动详情、包含学生参与（未必完成）第二课堂活动行为的学生参与活动详细信息、包含学生第二课堂方案完成情况的学生培养方案进度数据、包含学生各模块各系列第二课堂参与次数和时长的学生培养方案进度数据、包含学生成功完成第二课堂活动行为的学生已完成活动参与记录。下面详细介绍各数据集的字段和数据样例。

### 2.1 校园卡消费数据

包括约 3 亿条学生校园卡消费记录，每条记录包括用户 ID（学生学号）、消费时间、卡余额、交易金额、科目描述（消费类别）等字段，数据样例如表1所示：

表 1: 校园卡消费数据样例

用户 ID	消费时间	卡余额	交易金额	科目描述
614A	2020-09-13 18:15:41	37	12.5	餐费支出
614A	2020-09-13 21:30:10	34	3.0	淋浴支出

### 2.2 图书馆借阅数据

包括约 76 万条学生图书馆借阅记录，每条记录包括用户 ID（学生学号）、书名、借阅时间等字段，数据样例如表2所示：

表 2: 图书馆借阅数据样例

用户 ID	书名	借阅时间
25B3	法律的故事	2021-06-10 00:00:00
CE24	大学基础物理	2021-06-24 00:00:00

### 2.3 社团注册数据

包括约 2.4 万条学生社团注册记录，每条记录包括学生学号、所在院系、学期、学年、社团名称、社团指导单位、社团类型、社团校区等字段，数据样例如表3所示：

表 3: 社团注册数据样例

学号	所在院系	学期	学年	社团名称	社团指导单位	社团类型	社团校区
005B	数据科学与工程学院	第一学期	2023-2024	七草动漫社（中北）	教育学部	文化体育类	中山北路校区
902D	软件工程学院	第一学期	2023-2024	博远信息技术社	软件工程学院	学术科技类	中山北路校区

## 2.4 第二课堂数据

### 2.4.1 第二课堂-学生活动详情

包括 2783 条学生活动详情记录，每条记录包括活动 ID、创建时间、活动名称、活动主办方、开始时间、结束时间、适用模块等字段，数据样例如表4所示：

表 4: 学生活动详情数据样例

活动 ID	创建时间	活动名称	活动主办方	开始时间	结束时间	适用模块
EF76	2023-10-09 14:05:24	走进巴洛克音乐	音乐学院	2023-10-12 00:00:00	2023-10-12 23:59:59	美育实践

### 2.4.2 第二课堂-学生参与活动详细信息

包括 133774 条学生参与活动详细信息记录，每条记录包括学号、活动 ID、是否完成、是否报名、是否签到、学生评分、学生评语等字段，数据样例如表5所示：

表 5: 学生参与活动详细信息数据样例

学号	活动 ID	是否完成	是否报名	是否签到	学生评分	学生评语
CE24	885K	1	1	1	5	讲解由浅入深，获益匪浅

### 2.4.3 第二课堂-学生培养方案进度

包括 12648 条学生培养方案进度记录，每条记录包括学号、年级、院系名、专业名、参与活动次数、第二课堂方案完成进度等字段，数据样例如表6所示：

表 6: 学生培养方案进度数据样例

学号	年级	院系名	专业名	参与活动次数	第二课堂方案完成进度
C4A8	2021	计算机科学与技术学院	计算机科学与技术	22	23

### 2.4.4 第二课堂-学生培养方案进度详情

包括 44841 条学生培养方案进度详情记录，每条记录包括学号、模块名、系列名、次数、小时数等字段，数据样例如表7所示：

### 2.4.5 第二课堂-学生已完成活动参与记录

包括 104670 条学生已完成活动参与记录，每条记录包括学号、活动名称、活动开始时间、活动结束时间、发起机构、活动模块、小时数等字段，数据样例如表8所示：

表 7: 学生培养方案进度详情数据样例

学号	模块名	系列名	次数	小时数
FA6C	人文素养	“师大法学”系列学术报告、沙龙、论坛等	6	6

表 8: 学生已完成活动参与记录数据样例

学号	活动名称	活动开始时间	活动结束时间	发起机构	活动模块	小时数
733D	走进巴洛克音乐	2023-10-12 00:00:00	2023-10-12 23:59:59	音乐学院	美育实践	2

### 3 数据分析

在数据分析阶段，我们首先进行了核心素养的数据化表达，对每一个基本要点进行含义解读并分析其可提取的行为特征和对应的数据来源；然后我们提出了一个学生核心素养发展综合评分模型，该模型包含数据预处理、多视角特征提取和基于因子分析的学生核心素养发展综合评分三个模块；最后我们基于该评分模型对全校同学的六大核心素养进行评分并进行案例分析，结果显示该评分可以良好的反映同学们各个维度核心素养的发展情况。

#### 3.1 核心素养的数据化表达

中国学生发展核心素养 [1] 的三大基本内涵、六大核心素养、18 个基本要点和对应的主要表现如表9所示。为了从数据中分析学生各个维度的核心素养的发展情况，我们对每个基本要点的表现描述进行总结并阐述为含义解读，并且根据这些描述分析了每个基本要点可提取的相关行为特征，以及对提取这些特征有帮助的相关数据来源，具体信息如表10所示。

表 9: 中国学生发展核心素养基本要点和主要表现

基本内涵	核心素养	基本要点	主要表现描述
文化基础	人文底蕴	人文积淀	具有古今中外人文领域基本知识和成果的积累；能理解和掌握人文思想中所蕴含的认识方法和实践方法等。
		人文情怀	具有以人为本的意识，尊重、维护人的尊严和价值；能关切人的生存、发展和幸福等。
		审美情趣	具有艺术知识、技能与方法的积累；能理解和尊重文化艺术的多样性，具有发现、感知、欣赏、评价美的意识和基本能力；具有健康的审美价值取向；具有艺术表达和创意表现的兴趣和意识，能在生活中拓展和升华美等。
	科学精神	理性思维	崇尚真知，能理解和掌握基本的科学原理和方法；尊重事实和证据，有实证意识和严谨的求知态度；逻辑清晰，能运用科学的思维方式认识事物、解决问题、指导行为等。
		批判质疑	具有问题意识；能独立思考、独立判断；思维缜密，能多角度、辩证地分析问题，作出选择和决定等。



		勇于探究	具有好奇心和想象力; 能不畏困难, 有坚持不懈的探索精神; 能大胆尝试, 积极寻求有效的问题解决方法等。
自主发展	学会学习	乐学善学	能正确认识和理解学习的价值, 具有积极的学习态度和浓厚的学习兴趣; 能养成良好的学习习惯, 掌握适合自身的学习方法; 能自主学习, 具有终身学习的意识和能力等。
		勤于反思	具有对自己的学习状态进行审视的意识和习惯, 善于总结经验; 能够根据不同情境和自身实际, 选择或调整学习策略和方法等。
		信息意识	能自觉、有效地获取、评估、鉴别、使用信息; 具有数字化生存能力, 主动适应“互联网 + ”等社会信息化发展趋势; 具有网络伦理道德与信息安全意识等。
	健康生活	珍爱生命	理解生命意义和人生价值; 具有安全意识与自我保护能力; 掌握适合自身的运动方法和技能, 养成健康文明的行为习惯和生活方式等。
		健全人格	具有积极的心理品质, 自信自爱, 坚韧乐观; 有自制力, 能调节和管理自己的情绪, 具有抗挫折能力等。
		自我管理	能正确认识与评估自我; 依据自身个性和潜质选择适合的发展方向; 合理分配和使用时间与精力; 具有达成目标的持续行动力等。
社会参与	责任担当	社会责任	自尊自律, 文明礼貌, 诚信友善, 宽和待人; 孝亲敬长, 有感恩之心; 热心公益和志愿服务, 敬业奉献, 具有团队意识和互助精神; 能主动作为, 履职尽责, 对自我和他人负责; 能明辨是非, 具有规则与法治意识, 积极履行公民义务, 理性行使公民权利; 崇尚自由平等, 能维护社会公平正义; 热爱并尊重自然, 具有绿色生活方式和可持续发展理念及行动等。
		国家认同	具有国家意识, 了解国情历史, 认同国民身份, 能自觉捍卫国家主权、尊严和利益; 具有文化自信, 尊重中华民族的优秀文明成果, 能传播弘扬中华优秀传统文化和社会主义先进文化; 了解中国共产党的历史和光荣传统, 具有热爱党、拥护党的意识和行动; 理解、接受并自觉践行社会主义核心价值观, 具有中国特色社会主义共同理想, 有为实现中华民族伟大复兴中国梦而不懈奋斗的信念和行动。
		国际理解	具有全球意识和开放的心态, 了解人类文明进程和世界发展动态; 能尊重世界多元文化的多样性和差异性, 积极参与跨文化交流; 关注人类面临的全球性挑战, 理解人类命运共同体的内涵与价值等。
	实践创新	劳动意识	尊重劳动, 具有积极的劳动态度和良好的劳动习惯; 具有动手操作能力, 掌握一定的劳动技能; 在主动参加的家务劳动、生产劳动、公益活动和社会实践中, 具有改进和创新劳动方式、提高劳动效率的意识; 具有通过诚实合法劳动创造成功生活的意识和行动等。
		问题解决	善于发现和提出问题, 有解决问题的兴趣和热情; 能依据特定情境和具体条件, 选择制订合理的解决方案; 具有在复杂环境中行动的能力等。
		技术应用	理解技术与人类文明的有机联系, 具有学习掌握技术的兴趣和意愿; 具有工程思维, 能将创意和方案转化为有形物品或对已有物品进行改进与优化等。

表 10: 核心素养含义解读及相关行为特征和数据来源

基本内涵	核心素养	基本要点	含义解读	可提取的行为特征	数据来源
文化基础	人文底蕴	人文积淀	表现在学生人文知识面广而深	参与人文素养/文化传承第二课堂活动、图书馆借阅人文领域书籍	培养方案进度详情、已完成活动参与记录、图书馆借阅数据
		人文情怀	表现在学生的人文关怀行为	参与志愿公益类社团、志愿服务/师德涵养第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录
		审美情趣	表现在学生的审美能力	参与文化体育类社团、美育实践/艺术素养第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录
	科学精神	理性思维	表现在学生思维科学且严谨	参与学术科技类社团、思想素质/专业学术/专业技能/专业理论/专业素养第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录
		批判质疑	表现在学生可以提出和分析问题	参与学术科技类社团、专题演讲会/学术研究院/专业拓展第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录
		勇于探究	表现在学生解决问题能力	参与学术科技/综合实践类社团、创新创业/专业拓展/职业训练营第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录
自主发展	学会学习	乐学善学	表现在学生主动学习行为	参与文化体育/学术科技社团、培养计划外/跨院系第二课堂活动、借阅课外书籍	社团注册数据、培养方案进度详情、已完成活动参与记录、图书馆借阅数据
		勤于反思	表现在学生反思调整行为	第二课堂活动参与频次分布、培养方案进度完成情况	培养方案进度、培养方案进度详情、已完成活动参与记录
		信息意识	表现在学生网络信息综合能力	参与创新创业类社团、创新创业/安全教育第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录
	健康生活	珍爱生命	表现在学生健康行为习惯	学生用餐特征(早餐频次及规律用餐)、淋浴特征、参与体育类社团、安全教育第二课堂活动	校园卡消费数据、社团注册数据、培养方案进度详情、已完成活动参与记录
		健全人格	表现在学生心理健康状况	社交行为特征、参与志愿公益类社团、心理健康/体育运动/身心健康第二课堂活动	校园卡消费数据、社团注册数据、培养方案进度详情、已完成活动参与记录
		自我管理	表现在学生自我管理能力	参与自律互助类社团、劳育实践第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录

社会参与	责任担当	社会责任	表现在学生的社会贡献	参与志愿公益类社团、志愿服务/公教项目/师范实训/社会实践第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录
		国家认同	表现在学生的思想觉悟	参与思想政治类社团、思想素质/国情文化/全球胜任力第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录
		国际理解	表现在学生的国际视野	参与全球胜任力/国情文化/跨文化交际第二课堂活动	培养方案进度详情、已完成活动参与记录
	实践创新	劳动意识	表现在学生参与劳动实践	参与综合实践/志愿公益类社团、劳育实践/社会实践/志愿服务第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录
		问题解决	表现在学生行动力	参与综合实践/创新创业类社团、劳育实践/生涯发展/职业技能第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录
		技术应用	表现在学生创新能力	参与创新创业/学术科技类社团、创新创业/职业技能/专业技能第二课堂活动	社团注册数据、培养方案进度详情、已完成活动参与记录

## 3.2 核心素养发展的综合评分模型

我们提出了一个基于多源数据的学生核心素养发展综合评分模型，模型框架如图4所示，该模型包括数据预处理、多视角特征提取和基于因子分析的学生核心素养发展综合评分三个模块。

### 3.2.1 数据预处理

因为从 2021 级本科生开始各个学院将第二课堂纳入学生毕业要求，所以首先将研究对象确定为 2021、2022、2023 这三个年级的本科生，且仅保留 2021 年 9 月 1 日之后开展的第二课堂活动数据。另外，对于校园卡消费数据进行数据清洗和数据修正，这包括异常值删除、合并一餐多刷记录、将异常打水数据转换为淋浴数据（许多学生在宿舍打热水淋浴而不是前往澡堂淋浴）。

### 3.2.2 多视角特征提取

基于表10的分析，核心素养的每一个基本要点都拥有来自若干数据源的相关行为特征值得提取。

在涉及学生参与某些模块第二课堂活动的行为时，考虑到学生参与频次越多、时长越长、分布越均匀，代表学生在该模块的素养培养状况越良好，因此提取学生参与对应



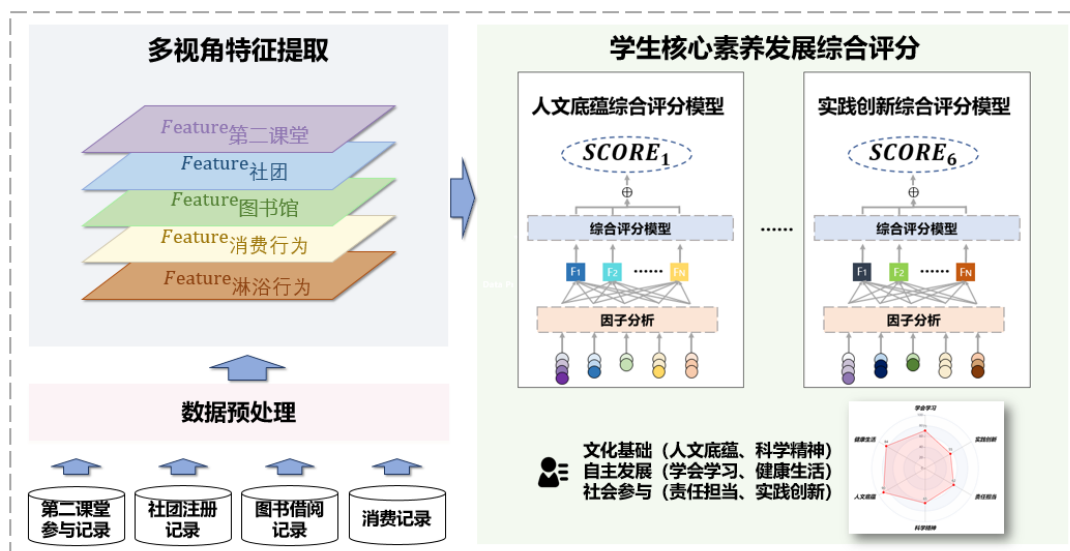


图 4: 学生核心素养发展综合评分模型框架图

第二课堂活动的频次、时长，并进一步计算以周为时间窗口的频次分布统计量（平均值、方差），用以描述学生参与第二课堂活动的行为，另外，考虑到学生参与培养计划外的、跨院系的第二课堂活动可能更多是出于自身兴趣而非计划要求，所以还提取了学生参与这些第二课堂活动频次和时长特征。

在涉及学生参与某些社团的行为时，提取学生参与对应类型社团的总次数，用以描述学生参与社团的行为。

在涉及学生在图书馆借阅书籍的行为时，考虑到学生借阅次数越多、借阅次数分布越均匀、借阅书籍种类越多，代表学生涉猎越广泛、主动学习能力越优秀，因此提取学生借阅书籍的频次、并计算以周为时间窗口的频次分布统计量（平均值、方差），用以描述学生的图书馆借阅行为。

在涉及学生用餐和购物消费行为时，考虑到学生用早餐次数多以及学生用餐规律可以代表学生作息习惯良好、学生校内用餐次数多代表学生校外用餐或点外卖次数少，因此提取学生用早餐的频次、正常时间用餐的频次、校内用餐的频次和金额、并进一步计算以周为时间窗口的频次分布统计量（平均值、方差），用以描述学生的校内消费行为。

在涉及学生的淋浴行为时，考虑到学生淋浴频次越多、淋浴金额越多，代表学生更可能关注个人卫生情况，因此以周为时间窗口提取学生周均淋浴频次和周均淋浴金额，用以描述学生的淋浴行为。

### 3.2.3 基于因子分析的学生核心素养发展综合评分

因子分析是主成分分析的一种推广与发展。因子分析法通过研究各变量间的相关系数矩阵，把这些变量间错综复杂的关系归结成少数几个公因子。由于归结出的公因子个数少于原始变量的个数，但它们又可以包含原始变量的绝大部分信息，所以因子分析的过程也是一种降维的过程，且相比于主成分分析，因子分析得到的公因子往往有更强的

可解释性。

因子分析的前提条件是观测变量间有较强的相关性，从表10中不难看出，由于语义上的相似性和可提取的行为特征的相似，同一个核心素养的三个基本要点之间或多或少存在一定的相关性，所以可以采用因子分析的方式对每个核心素养的若干行为特征进行降维。例如，学生参与学术科技类社团这一行为，对于科学精神核心素养下的理性思维、批判质疑和勇于探究三个基本要点都有一定的素养体现，此时就可以用降维的思想将这些基本要点重合的部分提取成公因子，而学生参与创新创业模块的第二课堂活动这一行为，则更多体现出勇于探究的素养，在理性思维和批判质疑上较少或没有体现。

对每一个核心素养，采用因子分析对其包含的三个基本要点的所有行为特征进行降维后，对提取得到的公因子进行方差最大化的正交旋转，得到旋转后的因子载荷矩阵，可以看出原始变量对于每一个公因子的权重载荷，利用回归法可以得到各公因子的得分系数矩阵，再依据旋转后各公因子对方差的贡献比重，就可以计算出该核心素养的综合评分。



图 5: 以人文底蕴为例进行因子分析的旋转后因子载荷矩阵

以人文底蕴这一核心素养为例，对研究群体的每位学生提取 A0、A1、A2、A3、A4 分别是“人文积淀”这一基本要点对应模块的图书借阅次数、已完成活动参与记录和培养方案进度详情中频次和时长特征，B0、B1、B2、B3、B4 分别是“人文情怀”这一基本要点对应模块的社团注册次数、已完成活动参与记录和培养方案进度详情中频次和时长特征，C0、C1、C2、C3、C4 分别是“审美情趣”这一基本要点对应模块的社团注册次数、已完成活动参与记录和培养方案进度详情中频次和时长特征。原始变量共 15 个，采用因子分析并选择提取 3 个公因子，得到旋转后的因子载荷矩阵如图5所示，可以看出第一个公因子更多被“人文积淀”的相关特征表示，第二个公因子更多被“审美情趣”的相关特征表示，第三个公因子更多被“人文情怀”的相关特征表示。进一步，将

每个公因子对方差的贡献比重作为各公因子在综合评分时的权重，即可得到每个学生人文底蕴这一核心素养的综合评分。为了便于后续进行可视化展示，最后一步基于改进的 Sigmoid 函数将综合评分映射到 (0,100) 区间范围内。

### 3.3 案例分析

为了验证我们提出的学生核心素养发展综合评分模型的合理性，首先我们对计算机科学与技术学院的三名同学展开了案例分析，图6是这三位同学的核心素养评分雷达图，可以看出 67D7 同学的六个核心素养评分都非常高；1E19 同学在科学精神和实践创新上的评分高于 0654 同学，但是在人文底蕴、责任担当、健康生活和学会学习四个维度上的评分低于 0654 同学。

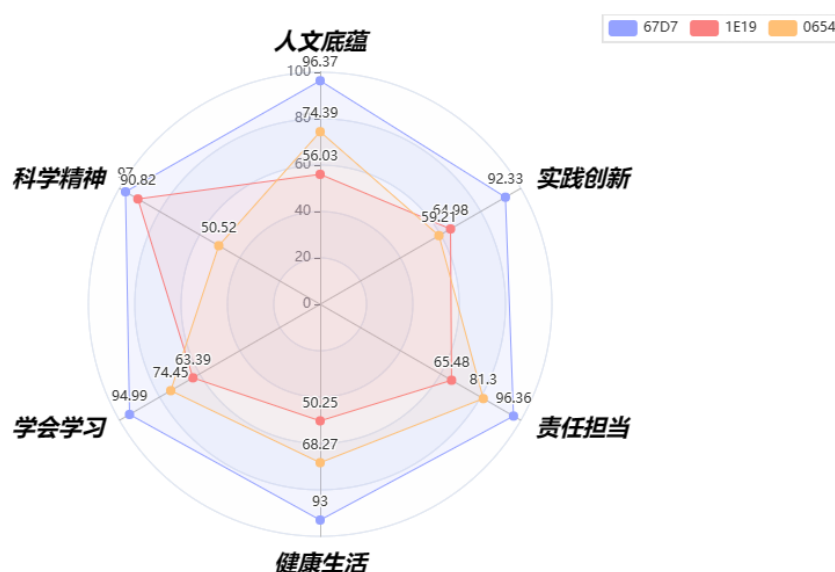


图 6: 三位同学的核心素养评分雷达图

分析这三位同学的各模块第二课堂活动参与情况，如图7和图8所示，可以看出 67D7 同学在各种模块的第二课堂活动参与次数和参与时长几乎都大于等于另外两位同学，而 1E19 同学参加思想素质创新创业两个模块的第二课堂活动次数相比 0654 同学略胜一筹，这可能是导致 1E19 同学科学精神和实践创新上评分更高的一个原因。另外，社团注册数据显示 67D7 同学参加过 4 次社团，1E19 同学参加过 2 次社团而 0654 同学从未参加过社团。

然后我们针对 67D7 同学进一步进行案例分析，该同学来自计算机科学与技术学院 2021 级。图9展示了这位同学在 2021-2022 学年、2022-2023 学年和 2023-2024 学年分别的核心素养发展评分，可以看出随着该同学在校学习生活的时间越来越久，其六个维度的核心素养评分均呈现出不同幅度的提升，尤其是科学精神的核心素养评分提升巨大。一个原因是该同学在校积极参与各模块的第二课堂活动，参与的次数和时长在不断扩大，图10展示了截止 2023 年 11 月 30 日，该同学参与各模块第二课堂活动的总时长分布情况。

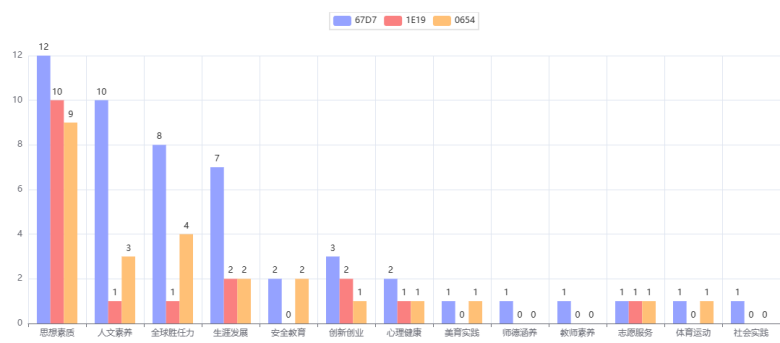


图 7: 三位同学的各模块第二课堂活动参与总次数

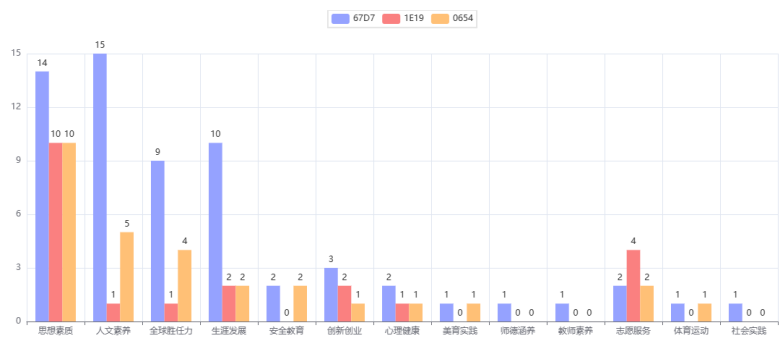


图 8: 三位同学的各模块第二课堂活动参与总时长

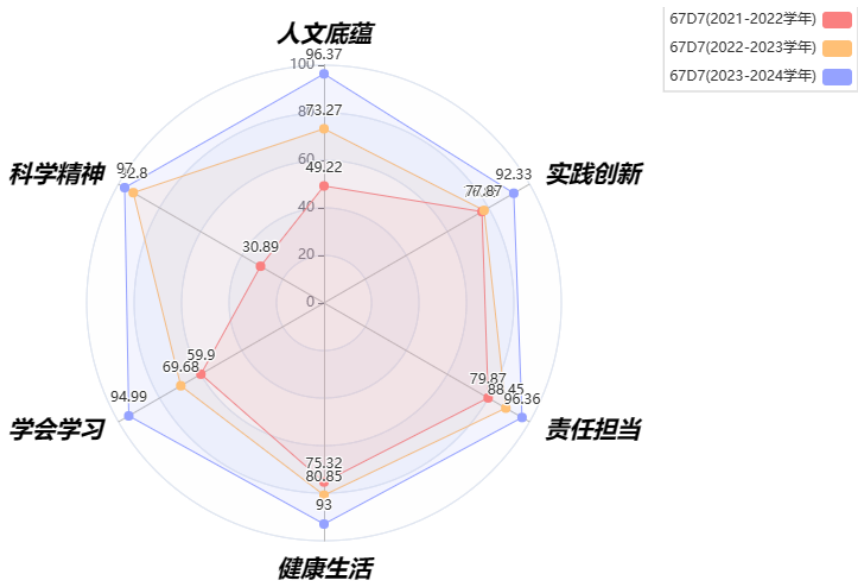


图 9: 某位同学每学年的核心素养评分雷达图

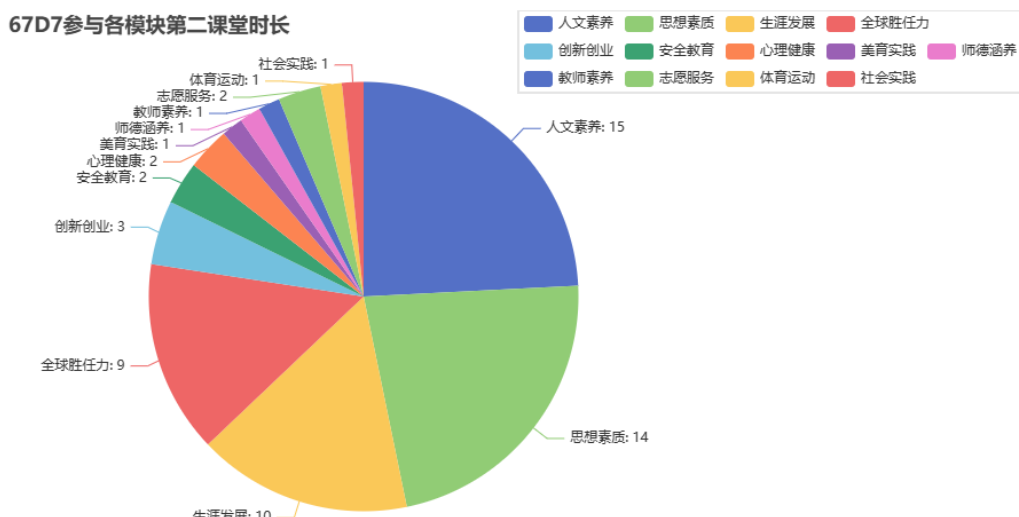


图 10: 某位同学参与各模块第二课堂总时长分布饼状图

综上，我们提出的学生核心素养发展综合评分模型对学生各个维度核心素养的评价较为准确且可解释性高，而且既可以对比不同学生间的核心素养发展情况，也可以对比同一学生不同时期的核心素养发展情况，且评分结果是具有一定参考价值和实践意义的。从每个学生的核心素养评分雷达图可以直观地看出学生各个维度的核心素养的发展情况，对于某个维度得分较低的学生，可以按照他们的培养方案为他们增加相关第二课堂的推送，对于长期各维度得分较低的学生，应当采取更多手段比如给予预警来提醒他们提高重视程度、辅导员谈话了解学生得分低的原因等，同时结合学生的反馈，适当调整学院培养方案的制定和第二课堂活动的制定等工作。

## 4 数据可视化展示

在这一部分中，我们基于 DataEase 平台设计并实现了一个可视化网页，并进行了发布，旨在为学生、老师和学校其他教育者提供帮助和支持。其中，可视化主界面如下图11所示，整体又划分为四个看板：个人画像看板、多人画像看板、院系画像看板以及全校总览看板。

### 4.1 个人画像

首先，个人画像看板对应的是单个学生详细情况的可视化，如图12、图13和图14所示，每个学生可以通过个人信息，如学号，查询到自己的核心素养得分详情、第二课堂完成详情以及消费行为模式的可视化，便于同学们根据自己的详细情况调整或者规划未来的学习和职业发展方向，从而提高学习效率和职业发展的准确性。在个人画像看板中，基于输入的学号，我们将报告被查询学生的六大素养得分详情，并展示其柱状图和六维雷达图，在雷达图中，我们还标注了全校的六大核心素养得分的平均分





图 11: 可视化主界面

(红色部分)，从而便于同学们针对自己薄弱的方面进行提高，也为后续同学们在选择社团或者是第二课堂活动时提供一定的参考价值。例如，图中展示的是一位学号为“67D7A97BE6DE80792ABDA70FE8FC8CF4245D7B5D”的学生，那么对应左上角会更新这位同学的年级以及所属院系，同时，可以看到这位同学的六大核心素养得分都接近于满分，总得分也达到了 0.95 的高分。同时，在第二课堂模块也能发现，这位同学高分的原因，在于他积极参与第二课堂的多模块活动，不仅有多项体育运动类也有志愿服务、全球胜任力等各个方面的多项活动，至此，我们可以得出这位同学在学校提供的第二课堂中具有更全面的发展。同时，我们还报告了这位同学的年度消费情况，如图14，可以看到在 2022 年第 8-11 以及 19-20 周他的消费会达到全年最高且比较平稳，也就是在春季秋季学期期中考试时间段，这可能是因为这位同学要考试因此会更多地在校内食堂消费，而不是出去用餐。



图 12: 个人画像看板 1



图 13: 个人画像看板 2



图 14: 个人画像看板 3

## 4.2 多人画像和院系画像

其次，基于单人的可视化详情，我们考虑为学院的老师们提供了解同学们的多人画像看板以及院系画像看板，以帮助学院老师及早发现和帮助严重“偏科”的同学。如图15所示，学院的老师可以通过选择需要关注的学生群体，例如 22 级所有学生或者某个寝室的学生，得到这个群体对应的核心素养平均情况。



图 15: 多人画像看板



图 16: 院系画像看板 1





图 17: 院系画像看板 2

### 4.3 全校总览

最后，我们聚合得到最后的全校总体情况看板，如图18、图19和图20，看板主要包括全校级别的信息概览、核心素养、第二课堂完成情况三个部分。该看板可以帮助学校的教育工作者，如学工老师，全面了解学生的发展方向以及每个方向的具体程度，从而更合理地调度学校资源以促进每一位学生全面而自由地发展。从看板中我们得知全校有10,924名在校本科生，31个院系，并且还能从柱状图中获取每个学院对应的总人数以及各个年级的人数分布。在核心素养模块，我们能看到全校核心素养人均得分排在前三的是院系分别是计算机科学与技术学院、孟宪承书院以及统计学院。并且，对应到第二课堂的完成情况，这三个学院在第二课堂人均完成进度、人均参与次数跨院系参与活动数量的排名也是在位于全校前列的，这也在侧面说明了我们设计的核心素养多维计算方法是符合实际情况的。另外，我们还能发现孟宪承书院由于本科生人数是全校最多的，高达2726名，远远超过其他学院的人数，因此其第二课堂活动的主办次数也是全校最多的，共计633次，远超了第2名大夏书院的活动主办次数。



图 18: 学校总览看板 1



图 19: 学校总览看板 2

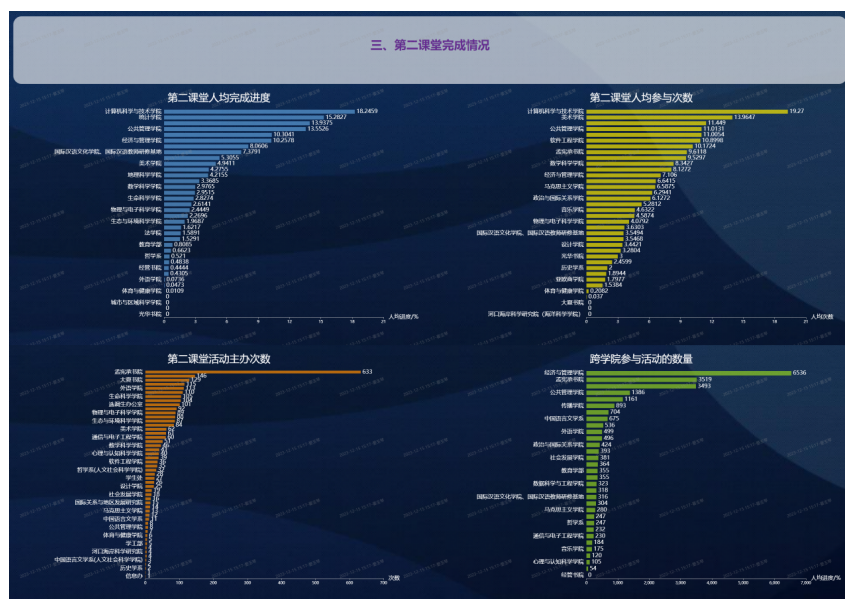


图 20: 学校总览看板 3

## 5 总结与展望

### 5.1 总结

本项目针对高校学生核心素养教育缺少科学的评价体系这一问题，分析了学生核心素养中每一个基本要点的含义和可提取的相关行为特征，提出了基于多源数据的学生核心素养发展评价方法，并构建可视化网页呈现学生画像来全面展示学生核心素养发展情况。本项目主要工作有：

(1) 总结了学生发展核心素养的每个基本要点并阐述为含义解读，并逐一构建每个基本要点和学生素养教育数据、行为数据的关联，为数据驱动的核心素养发展情况评价方法提供理论基础。

(2) 提出了一套基于多源数据的学生核心素养发展评价方法，具体分为数据预处理、多视角特征提取和基于因子分析的学生核心素养发展综合评分三个模块。

(3) 基于 DataEase 平台设计实现了一个呈现学生核心素养画像的可视化网页，包括个人画像看板、多人画像看板、院系画像看板以及全校总览看板。

本项目的创新点在于：

(1) 针对学生核心素养发展缺失科学评价体系的问题，创新性地提出基于多源数据的学生核心素养发展评价方法，并通过案例分析验证了该方法的可行性和可解释性。

(2) 设计实现了便于直观了解学生核心素养发展的可视化网页，可以帮助学生根据自己的详细情况调整或者规划未来的学习和职业发展方向，帮助老师及早发现和帮助严重“偏科”的同学。

### 5.2 展望

本项目提出的学生核心素养发展综合评分模型旨在对学生核心素养发展情况进行全面且完整的评价，但是仍存在可以进一步研究的问题，主要体现在：

(1) 数据来源受限。除了本项目利用的数据，学生在校还会产生许多其他行为数据，例如校园网使用数据、获得奖助学金/荣誉称号数据、选课数据和成绩数据等，这些数据中同样蕴含大量学生的行为信息可以帮助评价学生的核心素养发展情况。后续的研究可以利用更多源的数据构建评价模型对这一点进行优化。

(2) 评价的实时性不足。评价模型中涉及学生参与第二课堂的大量相关特征以学期为时间窗单位进行提取，因此对学生各个核心素养发展进行评分在学期末才能进行，而不能在学期中实时查看当前学期的评价结果。后续的研究可以通过改进特征提取的方式对这一点进行优化。

(3) 缺少不同专业学生的针对性评价。核心素养以培养“全面发展的人”为核心，但不同专业在培养方案和人才需求上存在客观的不同，后续的研究可以设计更具针对性的评价模型对这一点进行优化。例如，对文科专业的学生和理工科的学生在某些维度的核心素养可以采取个性化、差异化的评价方式。

## 参考文献

- [1] 核心素养研究课题组. 中国学生发展核心素养. 中国教育学刊, (1-3), 2016.
- [2] 夏杨. 基于一卡通数据的校园大数据分析平台的构建与应用. 华中师范大学学报 (自然科学版), (146-151), 2017.
- [3] Huaxiu Yao, Min Nie, Han Su, Hu Xia, and Defu Lian. Predicting academic performance via semi-supervised learning with constructed campus social network. In *DASFAA*, pages 597–609, 2017.
- [4] Ankit Kumar Saha, Abhishek Kumar Sharma, Soumya Sahoo, Shaikh Ejaz Hussain, and Nikhil Kumar Sahoo. Machine learning based prediction of student's performance based on psychological and behavioral data. In *International Conference on Mining Intelligence and Knowledge Exploration*, pages 396–408, 2023.
- [5] 张治. 基于大数据的多源多维综合素质评价模型的构建. 中国电化教育, (69-77+97), 2017.
- [6] 杨鸿. 大数据时代学生综合素质评价: 方法论、价值与实践导向. 中国电化教育, (27-34), 2018.
- [7] 钟苇笛. 数据驱动的核心素养评价: 本源、困境与破局. 中国电化教育, (35-43), 2022.