

哈尔滨石油学院首届教学设计大赛

一、课程基本信息

课程名称	Python 数据分析	授课教师	邓春伟
课程性质	专业课	课程学时	48
授课对象	数据科学与大数据技术	先修课程	面向对象程序设计 (Python)

二、课程简介

本课程是计算机科学与技术、大数据、人工智能等相关专业的一门核心课程。Python 语言是学生进行数据分析所需要掌握基础性语言和分析工具，是未来学生掌握数据分析技术的学习基础。本课程在教学内容方面着重以 Python 语言数据分析基础及 Pandas 数据预处理，Pandas 数据分析，Matplotlib 数据可视化为主。通过一系列的数据分析训练项目，培养学生具有一定的数据分析理解和应用实践能力。让学生掌握 Python 语言数据分析，Pandas 数据分析及 Matplotlib 数据可视化，为今后职业数据分析工作奠定基础。

三、学情分析

本课程主要教学对象是本科二年级数据科学与大数据技术专业学生，是本专业的专业核心课，课程总学时 48（理论 24 学时，实验 24 学时）学时，并配设 1 周课程设计。本课程需要在完成数据科学导论、Python 程序设计先修课程，此课程又为数据可视化、数据挖掘奠定了一定的理论及实践基础。

通过项目的分解、案例的融汇、知识的聚合、项目的复现等教学方法及手段、组织学生参加比赛、参与项目提升学生的自信心、内驱力，重思维培养，筑基知识能力，形成知识图谱，根据学生特点构成知识——应用层梯，理论筑基应用，应用展现理论，理论与应用融合提升能力。从而实现培养学生知识融合能力、用数据思维解决问题的能力、知识服务社会的能力、创新能力的目标。

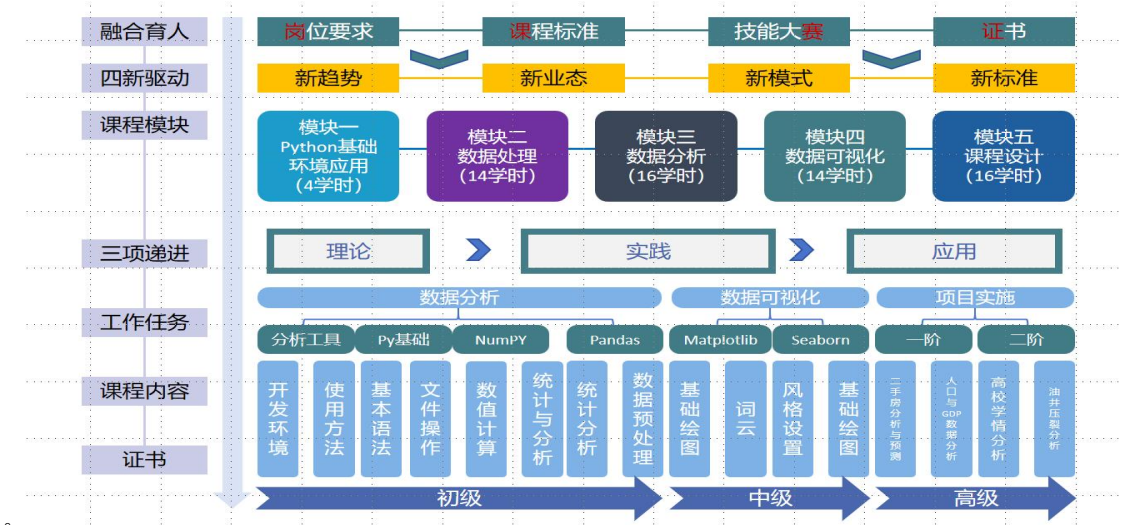
影响学习的最重要的因素是学生已经知道了什么,还有目前的从知识和技能基础、认知和实践能力、学习特点、心理特点四个维度分析，所以特从以下几方面进行分析。

维度	情况分析
知识和技能基础 (课程的“起点锚点”)	1. 基础差异大 (编程/数学背景分层) 2. 工具碎片化 (Python/SQL/ 统计工具应用零散)
认知和实践能力 (能力培养的“核心靶点”)	1. 抽象思维弱 (算法/模型理解困难) 2. 项目落地不足 (全流程实操能力欠缺)
学习特点 (教学方法的“适配依据”)	1. 理论耐受低 (纯理论学习兴趣低迷) 2. 学用脱节 (知识与应用场景割裂)
性格特点 (教学方法的“适配依据”)	1. 思想活跃, 逻辑分析能力有所欠缺 2. 畏难情绪 (复杂技术/ 代码调试恐惧) 3. 自信不足 (项目输出 / 表达能力薄弱)

	素质目标 (1) 通过对数据分析的了解，增强探索意识。 (2) 树立大数据思维和时代意识，自觉遵守职业道德和法律法规。 (3) 提高选择合适方法解决不同问题的能力。 (4) 养成分析问题、事前做好准备的良好习惯。 (5) 提高分析问题、针对不同问题选择合适方法的能力。 (6) 强化数据安全意识，提高信息技术应用能力。 (7) 理解国家大数据战略，构建大数据思维和时代意识。 (8) 锻炼具体问题具体分析的思维方法，培养一丝不苟的工作态度，增强积极主动寻求解决方法意识。
--	---

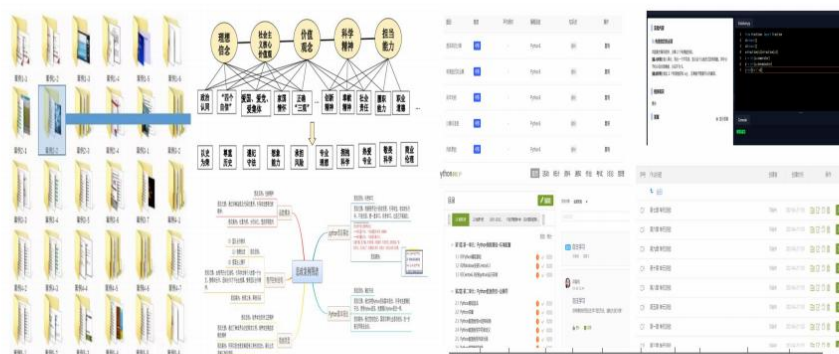
五、教学设计总体思路

《Python 数据分析》课程遵循**国家教学标准、人才培养方案和课程标准**，围绕“数据分析师”这一**专业核心岗位**，以真实的项目为**工作任务**，引入最新的**行业标准和岗位操作标准**，培养学生的核心技能，**课标融合**；将无数据分析师**职业资格证书标准**有机融入课程内容，**课证融合**。教学内容中，以“一数(多源数据处理)一析(建模分析应用)”为**知识引领**，以“一洗(数据清洗预处理)一用(分析建模落地)”为技能主线，以“一谨(心有严谨)一专(手有专攻)”为思政要素，形成**三项推进、四新驱动、课政融合、融合育人**的教学模式；在教学实施中，借助自主开发的数据分析案例库和故障调试题库、Python 可视化仿真平台、AR 数据可视化演示系统、Jupyter 在线实训环境、虚拟数据分析师工作场景仿真等**信息化手段**，以学生为中心，开展数据分析师、业务需求方、数据运维工程师等角色扮演，深化学生的职业认同感，采用“线上线下混合式”教学模式，强化学生的数据分析知识水平和实战应用技能水平。



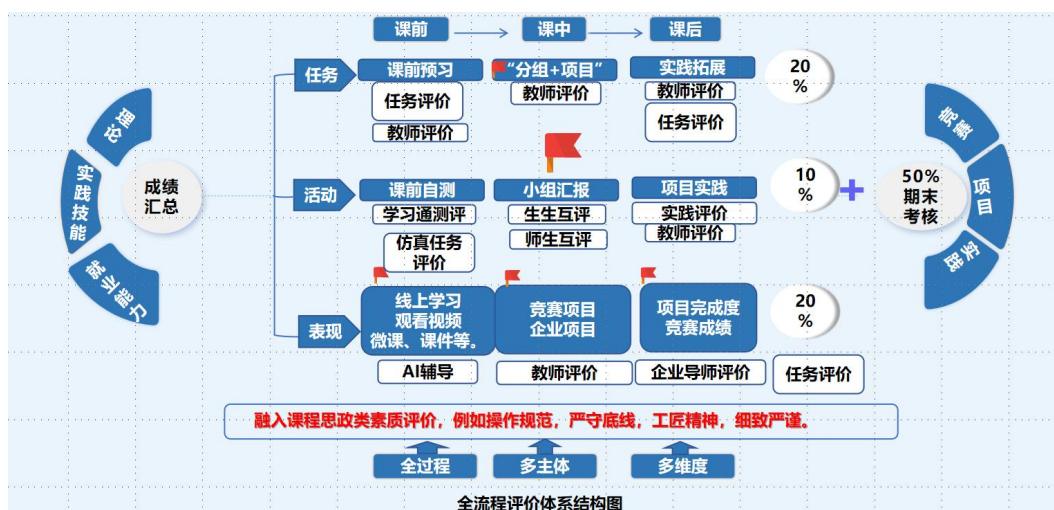
六、课程资源

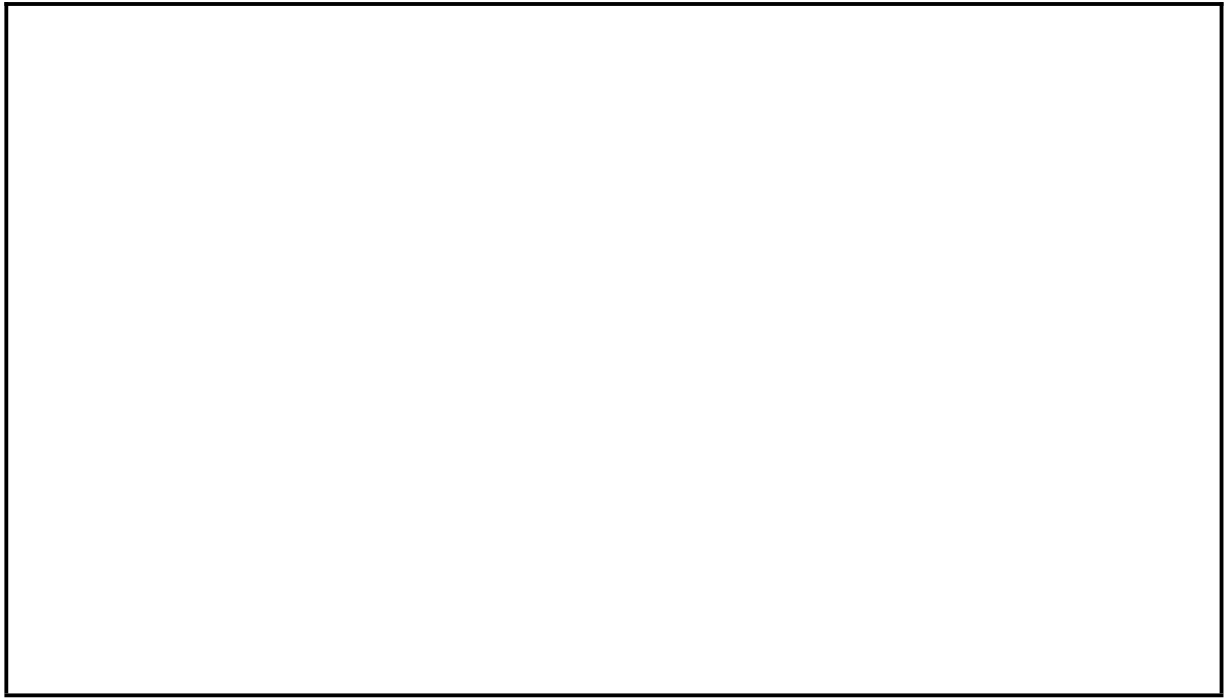
学习通 <https://mooc1.chaoxing.com/course/239814714.html>



七、课程评价

以学生为主体，以目标为导向，设置如下评价体系：





课题	第 20 课 二手房数据分析与房价预测需求分析
课时	2 课时 (90 min)
教学目标	知识目标： (1) 能够对案例进行需求分析 (2) 能根据需求对数据进行不同的预处理，选择合适的可视化图形进行分析和展示，并选择合适的相关变量进行预测 (3) 练习使用 Pandas 处理异常值、重复值、缺失值，离散化连续数据和编码字符型数据等 能力目标： (1) 能够参照软件标准规范，从开发角度对案例进行需求分析。 (2) 能够参照标准撰写开发文档。 素质目标： (1) 养成分析问题、事前规划的良好习惯 (2) 培养工匠精神 (3) 学会在生活、学习和工作中使用二八法则，抓主要矛盾，提高做事效率。
教学重难点	教学重点：对案例进行需求分析 突破手段：条理清晰，概括地写，可以从环节设置、信息化手段、学习资源等方面写 教学难点：使用 Pandas 处理异常值、重复值、缺失值，离散化连续数据和编码字符型数据 突破手段：条理清晰，概括地写，可以从环节设置、信息化手段、学习资源等方面写
教学方法	案例分析法、问答法、讨论法、讲授法
教学用具	计算机、投影仪、多媒体课件、教材、学习通
岗位对接	围绕 数据分析师 “数据处理 - 分析建模 - 可视化输出”的核心工作流程，将岗位要求拆解融入课程模块，实现知识学习与职业需求的同频对齐。
专业竞赛	以数据分析比赛的 Python 分析赛项为实践载体，将赛题拆解融入课程模块（如把赛事数据清洗环节对应 Python Pandas 教学、建模环节对应机器学习算法教学）；同时借鉴赛事的评分维度（如模型准确率、特征工程创新性、分析报告完整性），重构课程作业、项目实训的评价指标体系。
证书考核	以 CDA Level II 数据科学家认证对数据清洗、特征工程、建模的考察方式（如全流程实操案例拆解、能力分层评估）和标准要求为导向，将认证考核要点拆解融入课程各模块教学，让课堂实训、项目作业完全对标认证的能力评判标准。
思政教育	本课以“ 技术赋能民生，思政铸魂育人 ”为核心逻辑，将思政目标与专业实践节点精准对接：在数据处理环节渗透“诚信”素养，在模型构建与结果解读中强化“担当”意识，在区域差异分析中深化“公正”认知，最终让学生在解决房价这一核心民生问题的过程中，理解“数据服务社会”的专业价值，树立“以技立身、以善立心”的职业追求。
教学设计	本课程主要采用一标、二线、三段、四问、五动、六学的多维互动的教学法：

	教师 学生 课前 教师：上传资料【助学】、发布任务【助学】、上传视频【助学】、布置预习任务、收集发现问题、问题分析整理、制定活动方案 学生：查阅资料【自学】、学习课程【自学】、观看视频【自学】、完成预习任务、整理提出问题、完成预习方案、反馈授课教师 教学平台 线上 课中 教师：提出问题、提示引导【导学】、释疑解惑、教学活动、讲学【评学】、归纳总结 学生：深度思考【想问、敢问、会问、善问】、分组讨论、【动脑、动嘴、动眼、动手、动脑】、提出方案 解决问题、得出结论 教学平台 线上 线下 课后 教师：布置任务作业、作业批改评定、成绩统计分析、学情分析整理、课后辅导答疑、因材施教【督学】 学生：提交任务作业、自我梳理知识、自我反思总结、自主探究学习、自主拓展学习、专项训练【乐学】 教学平台 线上 一标：紧紧围绕教学目标，确定教学重点和难点，提炼教学内容，采取恰当的教学方式，选择最优的教学方案。 二线：教学空间实现线上+线下相结合。 三段：教学时间划分课前、课中、课后三个时间段，使学生能够利用一切机会、时间，以各种形式参与到教学之中。 四问：教学过程融入“四问”，激发学习兴趣，使学生“想问”；突破思维限制，使学生“敢问”；构建问题情境，使学生“会问”；引导求知欲望，使学生“善问”。 五动：教学互动使学生能主动参与、主动探索、主动思考、主动实践、主动感悟；使学生做到动脑、动眼、动嘴、动手、动耳，在“五动”中，养成良好的学习习惯，掌握科学的学习方法，各方面的素质和諧全面地发展。 六学：教学活动与导学、助学、督学、评学、自学、乐学相融合，课堂内外相互依托，通过师生多种感官的全方位参与，促进课堂教学多向交流，充分体现了一种师生之间、学生之间及人机之间的多向互动反馈过程，充分展示出教学过程的情境性、创造性、和谐性和交互性。 第 1 节课：课前考核（2 min）→问题导入（5 min）→传授新知（28 min）→课堂实践（10 min） 第 2 节课：问题导入（8 min）→传授新知（20 min）→课堂实践（12 min）→课堂小结（3 min）→作业布置（2 min）	
学情分析	学生已经学习过 Excel 数据分析等相关软件，掌握了数据处理的常用函数和方法、生成图表的过程。为什么还要学习本门课程，招聘岗位要求，熟练掌握各种数据分析工具。 学生有为难情绪，学的过程已经完成，但是不知道什么时候用，该怎么用，但是缺少对实际问题分析处理能力。	
教学过程	主要教学内容及步骤	设计意图
第一节课		
课前考核 (2 min)	【教师】 使用学习通进行签到 【学生】 按照老师要求课前考核	培养学生的组织纪律性,掌握学生的出勤情况

问题导入 (5 min)	【教师】提出问题 我们已经学习了 python 数据处理的相关知识，那么关于二手房数据分析与房价预测如何进行需求分析呢？ 【学生】思考、举手回答	通过提问的方法，了解学生对案例的了解程度，进而引出新知
传授新知 (28 min)	4min 【教师】总结学生的回答，并引入新知，讲解案例需求分析的相关知识 7.1 需求分析 7.1.1 目标分析 ➔【教师】提出问题 默认排序 最新发布 房屋总价 房屋单价 房屋面积 共找到 87703 套北京二手房 太湖银河湾 3室1厅 南北 太湖银河湾 - 通州其他 3室1厅 79.74平米 南北 精装修 中楼层(共18层) 2018年建 板楼 0人关注 / 刚刚发布 550万 68,975元/平 安华里社区 1室1厅 西北 安华里社区 - 安贞 1室1厅 59.06平米 西北 其他 顶层(共18层) 1990年建 板楼 0人关注 / 刚刚发布 462万 78,226元/平 本案例的目标是什么？ ➔【学生】聆听、思考、主动回答问题 ➔【教师】总结学生的回答，并讲解新知 本章的目标是根据某房产网站中最新发布的北京二手房数据（共 100 页 3000 条数据，见图 7-1），包括区域、户型、面积、装修、房龄、总价、结构、朝向和单价等，分析各区二手房数量和均价、二手房面积和总价区间占比、二手房房龄和均价的相关性、是否靠近地铁的不同装修二手房均价等，然后预测不同条件下二手房的房价。 7.1.2 数据源 ➔【教师】多媒体出示“最新发布的北京二手房数据.xlsx”文件的内容图片，并详细讲解各部分表示的内容 数据源为通过网络爬虫爬取的某房产网站中最新发布的北京二手房数据，保存在“最新发布的北京二手房数据.xlsx”文件中。	分析数据源，辨别特征值，提出显示特征，确定将产生的隐形特征值

最新发布的北京二手房数据.xlsx - Excel														文件	开始	插入	页面布局	公式	数据	审阅	视图	开发工具						
	A	B	C	D	E	F	G	H	I	J	K	L	M															
1	小区名	所在市道或镇	所在区	户型	面积	朝向	装修	楼层	年份	结构	总价	单价	房源标签															
2	台湾银河湾	通州其他	通州	3室1厅	79.74平米	南北	精装	中楼层(共18层)	2018年建	板楼	550万	68,975元/平																
3	安华里社区	安贞	朝阳	1室1厅	59.06平米	西北	其他	顶层(共18层)	1990年建	塔楼	462万	78,226元/平																
4	卢卡新天地	天通苑	昌平	2室1厅	90.03平米	东北	其他	18层	2014年建	板楼结合	492万	54,049元/平	近地铁															
5	鲁能·新院水岸公园	马坡	顺义	3室2厅	142.64平米	南北	精装	低楼层(共15层)	2014年建	板楼	520万	36,450元/平																
6	康嘉锦苑一期	丰台其他	丰台	2室2厅	99.49平米	南北	精装	中楼层(共6层)	2003年建	板楼	283万	28,443元/平																
7	东亚印象台湖	通州其他	通州	2室1厅	65.12平米	南	其他	中楼层(共12层)	2015年建	板楼	225万	34,552元/平																
8	培院小区	牡丹园	海淀	1室1厅	45.5平米	南	其他	低楼层(共6层)	1989年建	板楼	560万	123,077元/平	近地铁															
9	中德新城两区	亦庄开发区其他	亦庄开发区	2室1厅	72.59平米	南	其他	低楼层(共21层)	2013年建	板楼结合	320万	44,084元/平																
10	锦城	卢沟桥	丰台	3室1厅	123.02平米	西南	精装	中楼层(共26层)	2008年建	板楼结合	489万	39,750元/平																
11	华芳园	成寿寺	丰台	3室1厅	96.57平米	西南	简装	顶层(共24层)	2000年建	塔楼	430万	44,528元/平	近地铁															
12	马连道中街	马连道	西城	2室1厅	48.29平米	南	其他	中楼层(共5层)	1976年建	板楼	670万	138,746元/平																
13	北分厂	海淀北部新区	海淀	2室1厅	53.8平米	南北	精装	中楼层(共4层)	1980年建	板楼	368万	68,402元/平	近地铁															
14	红莲锦园	马连道	西城	2室1厅	77平米	南	简装	低楼层(共22层)	2000年建	塔楼	820万	106,494元/平																
15	锦雅家园北区	宋家庄	丰台	2室1厅	87.4平米	南北	精装	11层	2007年建	板楼	568万	64,989元/平																
16	甜水园东里	甜水园	朝阳	3室1厅	118.89平米	西北	简装	顶层(共18层)	1996年建	塔楼	899万	75,617元/平	近地铁															
17	龙翔路小区	牡丹园	海淀	2室1厅	61.4平米	东南	其他	低楼层(共16层)	1994年建	塔楼	634万	103,258元/平																
18	京兴北区	康义城	顺义	3室3厅	137.25平米	南北	精装	顶层(共6层)	2005年建	板楼	460万	33,516元/平																
19	万柳园小区	玉泉营	丰台	2室1厅	54.45平米	南北	其他	中楼层(共6层)	1992年建	板楼	325万	59,688元/平	近地铁															
20	广安园	草桥	丰台	1室1厅	51.23平米	南	其他	中楼层(共17层)	1998年建	塔楼	330万	64,416元/平																
21	龙湖嘉宁one	五道口	海淀	1室1厅	51.6平米	南	其他	10层	2011年建	板楼结合	350万	67,830元/平	近地铁															
22	049-新世代	万庄	丰台	2室1厅	57.79平米	北	精装	低楼层(共23层)	2005年建	塔楼	400万	70,255元/平	近地铁															
23	宏盛南园	小西天	海淀	2室1厅	58平米	南北	精装	高楼层(共6层)	1984年建	板楼	546万	94,138元/平	近地铁															
24	玉桥北里	乔庄	通州	1室1厅	43.2平米	南	其他	顶层(共6层)	1990年建	板楼	188万	43,519元/平																
25	润枫水尚西区	朝青	朝阳	2室1厅	96.04平米	东	其他	9层	2007年建	板楼	860万	89,547元/平	近地铁															

7-2 数据源

7.2 数据预处理

7.2.1 数据解析

➔【教师】提出问题

数据解析的步骤有哪些？

➔【学生】聆听、思考、主动回答问题



➔【教师】总结学生的回答，并讲解新知



图 数据解析流程图

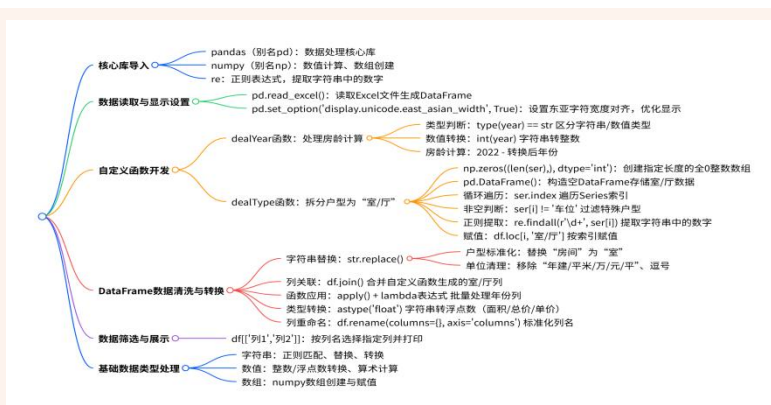


图 知识点思维导图

		➔【学生】观察、记录、理解 ➔【学生】教师演示完成后，学生按照教材提示和教师演示进行实际操作，并对比程序运行结果 ➔【教师】巡视课堂记录，对学生进行个别指导	小组讨论 ,分析结果																																																																																																																																																																
2min		7.2.2 异常值处理 ➔【教师】按照问题分析并按照步骤进行处理异常值程序演示： 包含车位的行： <table><thead><tr><th></th><th>小区名</th><th>所在街道或镇</th><th>所在区</th><th>户型</th><th>...</th><th>单价（元/平方米）</th><th>房源标签</th><th>室</th><th>厅</th></tr></thead><tbody><tr><td>523</td><td>嘉铭铜城A区</td><td>亚运村小营</td><td>朝阳</td><td>车位</td><td>...</td><td>54902.0</td><td>近地铁</td><td>0</td><td>0</td></tr><tr><td>748</td><td>奥林匹克花园三期</td><td>东坝</td><td>朝阳</td><td>车位</td><td>...</td><td>11394.0</td><td>NaN</td><td>0</td><td>0</td></tr><tr><td>1190</td><td>嘉铭铜城F区</td><td>亚运村小营</td><td>朝阳</td><td>车位</td><td>...</td><td>18634.0</td><td>近地铁</td><td>0</td><td>0</td></tr><tr><td>1258</td><td>三环新城7号院</td><td>玉泉营</td><td>丰台</td><td>车位</td><td>...</td><td>9172.0</td><td>近地铁</td><td>0</td><td>0</td></tr><tr><td>1467</td><td>鲁树湾二期</td><td>清河</td><td>海淀</td><td>车位</td><td>...</td><td>21331.0</td><td>NaN</td><td>0</td><td>0</td></tr><tr><td>1536</td><td>冠城大通瀛石</td><td>太阳宫</td><td>朝阳</td><td>车位</td><td>...</td><td>21484.0</td><td>近地铁</td><td>0</td><td>0</td></tr><tr><td>1569</td><td>中海寰宇天下熙山府</td><td>古城</td><td>石景山</td><td>车位</td><td>...</td><td>10922.0</td><td>近地铁</td><td>0</td><td>0</td></tr><tr><td>1642</td><td>三环新城6号院</td><td>玉泉营</td><td>丰台</td><td>车位</td><td>...</td><td>9223.0</td><td>NaN</td><td>0</td><td>0</td></tr><tr><td>1712</td><td>长阳半岛4号院</td><td>长阳</td><td>房山</td><td>车位</td><td>...</td><td>10000.0</td><td>近地铁</td><td>0</td><td>0</td></tr><tr><td>1946</td><td>北街家园六区</td><td>沙河</td><td>昌平</td><td>车位</td><td>...</td><td>4112.0</td><td>NaN</td><td>0</td><td>0</td></tr><tr><td>2250</td><td>首开国风美唐三期</td><td>回龙观</td><td>昌平</td><td>车位</td><td>...</td><td>6950.0</td><td>近地铁</td><td>0</td><td>0</td></tr><tr><td>2532</td><td>润枫领尚</td><td>通州其他</td><td>通州</td><td>车位</td><td>...</td><td>8987.0</td><td>近地铁</td><td>0</td><td>0</td></tr><tr><td>2578</td><td>清上园</td><td>安宁庄</td><td>海淀</td><td>车位</td><td>...</td><td>16598.0</td><td>NaN</td><td>0</td><td>0</td></tr><tr><td>2583</td><td>保利西山林语</td><td>海淀北部新区</td><td>海淀</td><td>车位</td><td>...</td><td>7750.0</td><td>NaN</td><td>0</td><td>0</td></tr><tr><td>2701</td><td>阳光新干线</td><td>亚运村</td><td>朝阳</td><td>车位</td><td>...</td><td>7000.0</td><td>近地铁</td><td>0</td><td>0</td></tr></tbody></table> [15 rows x 15 columns] 删除户型异常值前数据的行数： 3000 删除户型异常值后数据的行数： 2985		小区名	所在街道或镇	所在区	户型	...	单价（元/平方米）	房源标签	室	厅	523	嘉铭铜城A区	亚运村小营	朝阳	车位	...	54902.0	近地铁	0	0	748	奥林匹克花园三期	东坝	朝阳	车位	...	11394.0	NaN	0	0	1190	嘉铭铜城F区	亚运村小营	朝阳	车位	...	18634.0	近地铁	0	0	1258	三环新城7号院	玉泉营	丰台	车位	...	9172.0	近地铁	0	0	1467	鲁树湾二期	清河	海淀	车位	...	21331.0	NaN	0	0	1536	冠城大通瀛石	太阳宫	朝阳	车位	...	21484.0	近地铁	0	0	1569	中海寰宇天下熙山府	古城	石景山	车位	...	10922.0	近地铁	0	0	1642	三环新城6号院	玉泉营	丰台	车位	...	9223.0	NaN	0	0	1712	长阳半岛4号院	长阳	房山	车位	...	10000.0	近地铁	0	0	1946	北街家园六区	沙河	昌平	车位	...	4112.0	NaN	0	0	2250	首开国风美唐三期	回龙观	昌平	车位	...	6950.0	近地铁	0	0	2532	润枫领尚	通州其他	通州	车位	...	8987.0	近地铁	0	0	2578	清上园	安宁庄	海淀	车位	...	16598.0	NaN	0	0	2583	保利西山林语	海淀北部新区	海淀	车位	...	7750.0	NaN	0	0	2701	阳光新干线	亚运村	朝阳	车位	...	7000.0	近地铁	0	0	提出问题什么是异常值,异常值的处理方法
	小区名	所在街道或镇	所在区	户型	...	单价（元/平方米）	房源标签	室	厅																																																																																																																																																										
523	嘉铭铜城A区	亚运村小营	朝阳	车位	...	54902.0	近地铁	0	0																																																																																																																																																										
748	奥林匹克花园三期	东坝	朝阳	车位	...	11394.0	NaN	0	0																																																																																																																																																										
1190	嘉铭铜城F区	亚运村小营	朝阳	车位	...	18634.0	近地铁	0	0																																																																																																																																																										
1258	三环新城7号院	玉泉营	丰台	车位	...	9172.0	近地铁	0	0																																																																																																																																																										
1467	鲁树湾二期	清河	海淀	车位	...	21331.0	NaN	0	0																																																																																																																																																										
1536	冠城大通瀛石	太阳宫	朝阳	车位	...	21484.0	近地铁	0	0																																																																																																																																																										
1569	中海寰宇天下熙山府	古城	石景山	车位	...	10922.0	近地铁	0	0																																																																																																																																																										
1642	三环新城6号院	玉泉营	丰台	车位	...	9223.0	NaN	0	0																																																																																																																																																										
1712	长阳半岛4号院	长阳	房山	车位	...	10000.0	近地铁	0	0																																																																																																																																																										
1946	北街家园六区	沙河	昌平	车位	...	4112.0	NaN	0	0																																																																																																																																																										
2250	首开国风美唐三期	回龙观	昌平	车位	...	6950.0	近地铁	0	0																																																																																																																																																										
2532	润枫领尚	通州其他	通州	车位	...	8987.0	近地铁	0	0																																																																																																																																																										
2578	清上园	安宁庄	海淀	车位	...	16598.0	NaN	0	0																																																																																																																																																										
2583	保利西山林语	海淀北部新区	海淀	车位	...	7750.0	NaN	0	0																																																																																																																																																										
2701	阳光新干线	亚运村	朝阳	车位	...	7000.0	近地铁	0	0																																																																																																																																																										
3min		➔【学生】观察、记录、理解																																																																																																																																																																	
4min		➔【学生】教师演示完成后，学生按照教材提示和教师演示进行实际操作，并对比程序运行结果（详见教材） ➔【教师】巡视课堂记录，对学生进行个别指导																																																																																																																																																																	
		【学生】聆听、思考、理解、记忆																																																																																																																																																																	
课堂实践（10 min）	2min	【教师】提出问题 根据学过的知识及异常值的处理办法，请同学讨论尝试进行重复值处理																																																																																																																																																																	
	4min	【学生】聆听、思考、讨论、编写程序 7.2.3 重复值处理 房产网站中会重复发布相同的房源信息，此时，须将相同的房源数据删除。此处，检查所有列都相同的重复值并输出；然后，删除除包含重复值的第一行外其他包含重复值的行，并输出删除前、后数据的行数。 	通过课堂讨论，加深学生对处理数据重复值的理解																																																																																																																																																																
	2min	分析结果																																																																																																																																																																	
	2min	【教师】检查并总结学生程序编写																																																																																																																																																																	

第二节课

问题导入
(8 min)

【教师】提出问题

我们已经学习了数据重复值，异常值预处理方式，那么缺失值要如何处理呢？

【学生】聆听、思考、回答问题

通过提问的方法,引导学生主动思考,激发学生的学习兴趣

2min

【教师】总结学生的回答，并引入新知，讲解数据预处理的相关知识

7.2.4 缺失值处理

→【教师】进行问题分析并按照步骤进行处理缺失值程序演示：

删除房龄缺失值前数据的行数： 2949

删除房龄缺失值后数据的行数： 2909

房源标签替换缺失值后的数据：

	朝向	装修	楼层	房龄	...	单价(元/平方米)	房源标签	室	厅	
0	南	北	精装	中楼层(共18层)	4.0	...	68975.0	不近地铁	3 1	
1	西	北	其他	顶层(共18层)	32.0	...	78226.0	不近地铁	1 1	
2	东	北	其他	18层	8.0	...	54649.0	近地铁	2 1	
3	南	北	精装	低楼层(共15层)	8.0	...	36456.0	不近地铁	3 2	
4	南	北	精装	中楼层(共6层)	19.0	...	28463.0	不近地铁	2 2	
...	...	...		...	...	...	...	...	...	
2995	西	精装	中楼层(共27层)	18.0	...	92780.0	近地铁	1 1		
2996	南	北	简装	15层	10.0	...	22262.0	不近地铁	3 1	
2997	南	北	其他	顶层(共27层)	15.0	...	93376.0	不近地铁	4 2	
2998	南	北	精装	6层	21.0	...	47690.0	不近地铁	2 2	
2999	南	西	北	其他	低楼层(共32层)	11.0	...	85848.0	近地铁	4 2

[2909 rows x 10 columns]

→【学生】观察、记录、理解

→【学生】教师演示完成后，学生按照提示和教师演示进行实际操作，并对比程序运行结果



→【教师】巡视课堂记录，对学生进行个别指导

7.2.5 连续数据离散化

在使用模型对房价进行预测时，无法使用字符型数据，须对其进行编码。此处，将所在区、装修、结构、房源标签和朝向进行编码，然后将数据保存到“最新发布的北京二手房数据_预处理.xlsx”文件中。其中，编码前须重新设置连续的行索引；结构编码前须删除其中包含暂无数据的行；定义 get_dummies_dirt() 函数用于东、南、西、北、东北、东南、西南和西北 8 个方位的编码。

→【教师】按照教材进行问题分析并按照步骤进行处理连续数据离散化程序演示：

3min

通过教师讲解和课堂联系,使学生了解异常缺失值，时间信息处理,预处理数据保存等的相关知识

提出什么是数据离散化
讨论此项目离散化的目的

传授新知
(20 min)

	<pre> data = np.zeros((len(ser)), dtype='int')#创建数值为 0 的数组 df = pd.DataFrame({'东': data, '南': data, '西': data, '北': data, '东北': data, '东南': data, '西南': data, '西北': data}, index=ser.index) for i in ser.index: #分割字符串 rec = ser[i].strip().split(' ') #遍历每条数据分割后的方位 for dirt in rec: #如果方位信息包含在 dirt 中，则该方位列对应的行赋值为 1 if dirt in dirts: df[dirt][i] = 1 return df df = df.join(get_dummies_dirt(df['朝向'])) df.to_excel('最新发布的北京二手房数据_预处理.xlsx', index=False)。</pre>	
课堂小结 (3 min)	【教师】简要总结本节课的要点 本节课学习了需求分析与数据预处理的相关知识，包括数据解析，缺失值，重复值，异常值，连续数据离散化及字符串编码等相关内容，希望大家在课下多加练习，巩固所学知识。 【学生】总结回顾知识点	总结知识点 巩固学生对相关知识的印象
作业布置 (2 min)	【教师】布置课后作业 (1) 完成课后 “本章考核 7” 的习题。 (2) 预习第 8 章的内容。 【学生】完成课后任务	通过课后作业复习巩固学到的知识 提高理论接收能力
教学效果达成情况	本课次根据课程性质、学生实情、专业人才培养要求，针对以往教学中存在的痛点和难点，采取了独具特色的基于 “学情预判+全链式智慧教学即时反馈” 的混合式柔性定制式教学策略，进行了以下教学方法改革： 1.OBE 教学法:以学生为中心的开展教学组织与实施，课前发布预习，根据学生自学检测及主观题完成情况设计课堂教学知识图谱，设计综合错误改错题让学生共同改错加深概念的理解和认知，高光时刻分享优秀作业达到学生相互促进的目的，全员参与的填空题对关键语法进行重点强调，程序设计及互评实现了编程能力检测和生生互动，硬件实验体验启发学生的创新思维，基于自主开发的 VPL 实验平台的挑战性编程训练提高学生的编程实践能力，上机实践课后拓展则进一步开拓学生的视野，加深及扩展所学知识。	

2.BOPPPS 有效教学法:以雨课堂预习任务为导课-提出课程目标-不同知识点前测-参与式互动学习-反思后测,归纳总结,知识扩展,避免了学生游离于学习外的情况。

3.PBL 教学法:通过学生自学反馈设置问题—师生共同进行问题探索—讲述其解决过程及机制进行同类问题解答—强调概念进行问题归结—学生互动进行问题讨论—学生利用所学知识自主进行问题解决,自己得出结论,使学生成为学习的主体。

4.引导式教学法:通过学习通平台和虚拟仿真平台、发布课程知识点的运用,让学生用自己的智慧主动去思考如何进行该任务的设计和实现,引导学生自主学习。

5.探究式教学法:让学生顺着计算机处理问题的思路抽丝剥茧,理解计算思维的精髓,体会研究的乐趣,提高学生的科学研究精神。

综上所述,本课次的教学设计将“以学生为中心”的教学理念落到了实处,激发了学生的学习内驱力,“两性一度”特征鲜明,解决了传统课程教学中的痛点和难点,很好地达到了预期教学目标。

课件设计



教学反思

通过“诚信筑基、担当立心、公正领航”的思政融入,二手房数据分析与房价预测教学不再是单纯的技术训练,而是成为“理解民生、服务社会”的价值培育载体,让学生在掌握数据采集、建模、预测技能的同时,筑牢职业诚信、强化民生担当、深化公正认知,实现“专业能力”与“价值素养”的协同提升。

这节课整体节奏较好,但学生在课堂练习环节中的表现差距较大。每个学生对于程序的认知和感受不同,接受能力也不同,教师应因材施教,对于优秀学生应使其在达到基本要求后,向更高层次迈进,对于后进生,应打好其基础,努力使其达到基本要求,并通过学生间的互助,让优秀学生带动后进生。