

【Python 数据分析与应用】

【Python Data Analysis and Application】

一、基本信息

课程代码:【2058145】

课程学分:【2】

面向专业:【全校本科生】

课程性质:【通识教育选修课】

开课院系: 信息技术学院

使用教材:

教材: Python 数据分析与可视化, 魏伟一、李晓红、高志玲 主编, 清华大学出版社, 2021.07

参考书目:

1. Python 程序设计基础与应用 (第 2 版), 董付国 著, 机械工业出版社, 2022.1
2. Python 数据分析从入门到实践, 明日科技 著, 吉林大学出版社, 2022.2
3. Python 数据分析与应用 (第 2 版), 曾文权, 张良均 主编, 人民邮电出版社, 2021.11

课程网站网址: https://www.mosoteach.cn/web/index.php?c=interaction&m=index&clazz_course_id=4089FC14-07E2-11EE-8539-1C34DA7B3F7C

先修课程:【面向对象程序设计 2050251 (3)】

二、课程简介

数据分析是对数据进行检查、清理、转换和建模的过程, 有助于从数据中发现规律并制定科学的决策, 目前已广泛应用于自然科学、社会科学和管理科学的各个领域。Python 是当前数据分析工作中常用的编程语言, 有简洁的说明文档、丰富的第三方库。掌握 Python 编程的基础知识有助于各专业人员更好地开展数据分析相关的工作, 更好地解决实际工作中的问题。本课程面向各类编程学习者, 通过学习使用 Numpy, Pandas, Matplotlib 三个模块, 能够进行基本的数据统计分析; 通过基本数据分析操作技术结合数据预处理过程从学习, 能够构建分类、回归、归类模型。通过本课程的学习, 使学生学会使用 Python 进行科学计算, 数据处理、分析和建模, 可视化绘图, 完成对小型问题解决方案的分析与设计, 利用 Python 语言编程得到可视化结果; 正确分析和解释实验中的结果、现象和问题, 培养计算思维、数据思维及采用程序设计方法解决计算问题的实战能力。

三、选课建议

本课程适合全校各专业, 尤其是计算机大类相关专业的大二以上的学生, 要求选学者至少掌握一门面向过程编程或者面向对象编程语言。

四、课程目标/课程预期学习成果（必填项）（预期学习成果要可测量/能够证明）

序号	课程预期学习成果	课程目标 (细化的预期学习成果)	教与学方式	评价方式
1	L03 表达沟通	能够以口头或书面形式，运用计算机专业的知识和术语解释数据分析问题的需求；	边讲边练；讨论	讨论、作业、在线评测、实验报告
2	L02 专业能力	能够安装 Python 数据分析开发环境和第三方库	边讲边练；讨论	讨论、作业、在线评测、实验报告
		能够查看 Python 和第三方库的 API 说明文档，了解详细编程方法；	边讲边练；讨论	讨论、作业、在线评测、实验报告
		能够运用 Numpy、Pandas、Matplotlib 完成数据读取、分析和展示；	边讲边练；讨论	讨论、作业、在线评测、实验报告

五、课程内容

第 1 单元 Python 数据分析概述

通过本单元学习，学生能掌握数据分析的概念、数据分析的流程；了解数据分析的应用场景、数据分析的常用工具；能够解释 Python 数据分析的优势、说出 Python 数据分析的常用第三方库；能够安装和使用 Python 的开发工具 Pycharm、Anaconda；掌握 Jupyter Notebook 的基础功能；
 重点： Python 数据分析的概念、流程和应用场景；了解 Python 常用的数据分析库；
 难点：掌握 Pycharm、Anaconda 的安装与配置；掌握 Jupyter Notebook 的常用功能；
 理论课时：2，实践课时 2。

第 2 单元 Python 编程基础

通过本单元学习，学生能够掌握 Python 的基础语法，掌握基础数据类型、流程控制结构语句的用法；掌握 Python 各种序列（列表、字典、集合）的运用；创建 Python 类、对象、方法。
 重点：掌握 Python 基础语法，掌握 Python 各种序列（列表、字典、集合）的运用
 难点：掌握 Python 各种序列（列表、字典、集合）的运用；创建 Python 类、对象、方法。

 理论课时：4，
 实践课时：4。

第 3 单元 Numpy 数值计算基础

通过本单元学习，学生能够创建 Numpy 数组对象，生成随机数，通过索引访问数组，编号数组的形态，创建 Numpy 矩阵，使用数组进行简单的统计分析；

理论课时：2，
实践课时：2。

第4单元 pandas 统计分析基础

通过本单元学习，学生能够读取文本文件、excel 文件；查看 DataFrame 的常用属性；增删改查 DataFrame 数据；描述分析 DataFrame 数据；使用 Groupby 方法拆分数据；使用 agg 方法聚合数据；

重点：读取文本文件、excel 文件；增删改查 DataFrame 数据；

难点：增删改查 DataFrame 数据；描述分析 DataFrame 数据；使用 Groupby 方法拆分数据；使用 agg 方法聚合数据；

理论课时：4，
实践课时：4。

第5单元 Matplotlib 数据可视化基础

通过本单元学习，学生能掌握 pyplot 的基础语法，设置 pyplot 的动态 rc 参数；能够绘制散点图、折线图、直方图、饼图。

重点：pyplot 的基础语法；

难点：绘制散点图、折线图、直方图、饼图。

理论课时：4，
实践课时：4。

六、课内实验名称及基本要求

序号	实验名称	主要内容	实验时数	实验类型	备注
1	Numpy 数值计算基础	Numpy 数据处理	4	综合性	
2	Pandas 统计分析基础	Pandas 读取文本文件的数据，并进行数据过滤、处理	6	综合性	
3	Matplotlib 数据可视化基础	matplotlib 作图：散点图、折线图	6	综合性	

总评构成 (1+X)	评价方式	占比
------------	------	----

X1	大作业	40%
X2	章节测验	30%
X3	实验报告	15%
X4	作业和课堂表现	15%

七、评价方式与成绩

撰写人：万永权

系主任审核签名：戴智明

审核时间：2023年9月