

# Python 数据分析与应用

## 【Python Data Analysis and Application】

### 一、基本信息（必填项）

课程代码：【按教学计划填写】

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 数据分析从小白到专家，田越 著，电子工业出版社，2021.04

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，黄红梅，张良均 著，人民邮电出版社，2018.4

课程网站网址：<https://mooc1.chaoxing.com/course-ans/courseportal/225141767.html>

先修课程：【面向对象程序设计 2050251 (3)】

### 二、课程简介（必填项）

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（必填项）

此处给出该课程的选课建议，例如适合专业、年级、学习基础要求等建议。

本课程适合全校各专业，尤其是计算机大类专业、大数据技术类相关专业学生，要求选学者掌握一门面向过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）

课程预期学习成果必须写到指标点，只写三级编码即可。“课程目标”这列要写清楚细化的预期学习成果，撰写时必须用适当的行为动词引导。

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	L0112	1. 能够以口头或书面形式，运用计算机专业的专业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
1	L0212	1. 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
		2. 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
		3. 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
3	L0512	1. 能够说出数据分析的前沿技术、工具等；	边讲边练；讨论	讨论、作业、在线评测、实验报告
		2. 能够解释数据分析的应用案例	边讲边练；讨论	讨论、作业、在线评测、实验报告

#### 五、课程内容（必填项）

此处分单元列出教学的知识点和能力要求。知识点用布鲁姆认知能力的 6 种层次：“知道”、“理解”、“运用”、“分析”、“综合”、“评价”来表达对学生学习要求上的差异。能力要求必须选用合适的行为动词来表达。用文字说明教学的难点所在。

##### 第一讲 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：1 学时，实验课时 1 学时。

## 第二讲 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：6 学时。

## 第三讲 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：3 学时，实验课时 3 学时。

## 第四讲 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：2 学时，实验课时 4 学时。

## 第 5 讲 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、饼图和箱线图。

重点：pyplot 的基础语法；

难点：绘制散点图、折线图、直方图、饼图和箱线图。

理论课时：2 学时，实验课时 4 学时。

## 第 6 讲 数据分析案例

通过本单元学习，学生能够加载数据文件、划分数据集；构建和评价回归模型；构建和评价分类模型；

重点：构建和评价回归模型；构建和评价分类模型；

难点：构建和评价回归模型；构建和评价分类模型；

理论课时：2 学时，实验课时 4 学时。

## 六、课内实验名称及基本要求（选填，适用于课内实验）

列出课程实验的名称、学时数、实验类型（演示型、验证型、设计型、综合型）及每个实验的内容简述。

序号	实验名称	主要内容	实验学时数	实验类型	备注
1	Python 数据分析环境搭建		1	验证型	
2	Numpy 数值计算基础		3	综合性	
3	Pandas 统计分析基础		4	综合性	
4	Matplotlib 数据可视化基础		4	综合性	
5	数据分析和预测案例		4	综合性	

## 七、评价方式与成绩（必填项）

总评构成（1+X）	评价方式	占比
1	在线考试（开卷）	40%
X1	平时作业和表现	30%
X2	实验报告	30%

“1”一般为总结性评价，“X”为过程性评价，“X”的次数一般不少于3次，无论是“1”、还是“X”，都可以是纸笔测试，也可以是表现性评价。与能力本位相适应的课程评价方式，较少采用纸笔测试，较多采用表现性评价。

常用的评价方式有：课堂展示、口头报告、论文、日志、反思、调查报告、个人项目报告、小组项目报告、实验报告、读书报告、作品（选集）、口试、课堂小测验、期终闭卷考、期终开卷考、工作现场评估、自我评估、同辈评估等等。**一般课外扩展阅读的检查评价应该成为“X”中的一部分。**

同一门课程由多个教师共同授课的，由课程组共同讨论决定X的内容、次数及比例。

撰写人：万永权

系主任审核签名：戴智明

审核时间：2022年4月