

Python 数据分析

Python 中对于数据分析经常使用三件套，即 `numpy`、`pandas` 和 `matplotlib`。本课程的系列实验就是通过对三大库进行学习，并通过代码练习来掌握三大库的一些常规用法。

每个“Python 数据分析”实验的编写都包含了五组内容，即实验目的、实验学时、实验内容、实验环境、实验步骤。其中实验目的介绍了本实验的教学目标，学生该达到的能力目标；实验学时则根据情况列出课程实验的课时；实验环境说明了实验实施所需要的软件硬件平台；实验内容主要是对相关的 `jupyter notebook` 文件内容的概述；实验步骤则简要描述课程知识点与 `jupyter notebook` 文件操作的每一步输入及输出情况。

实验环境：汇萃人工智能教学实验平台，基于系统自带的 Python 环境或者 `jupyter notebook`

实验一 课程介绍

在本实验中将对 Python 数据分析做简单的介绍，并介绍怎么安装 `numpy`、`pandas` 和 `matplotlib` 三大库，在实际的实验环境中不需要学生手动安装，已经默认安装了相关的库、机器学习 和深度学习的库。

实验内容：掌握 Python 数据分析的基本方法，了解数据分析常用的三大库，并了解库的安装方法。

实验二 Anaconda 介绍

通过原生的 `python` 程序，再手工安装各种库来进行数据分析完全可以解决问题，但是效率不高，特别是在使用的过程中因为某些库暂时没有安装，还需要联网安装，比较麻烦；为了提升效率，整合各种资源，方便小白用户的使用，国外出现了多个包含数据分析库在内的发行版本，其中 `Anaconda` 就是非常好用的一个发行版本。本实验将对 `Anaconda` 进行介绍，并了解其安装及包的下载和管理。

实验内容：掌握 `Anaconda` 的安装及包管理，了解 `Anaconda` 的强大功能。

实验三 Jupyter 介绍

`Jupyter notebook` 在前面的实验中已经做介绍，本实验将对 `notebook` 做更为全面的介绍，帮助学生了解 `notebook` 的安装及使用界面等，在本教学环境中对 `notebook` 做了二次开发，有些菜单已经被屏蔽，学生可以在自己的 PC 上进行 `notebook` 的安装及使用。

实验内容：了解 `notebook` 的基本情况，学会安装 `notebook`，熟悉 `notebook` 的基本操作界面和代码执行方法。

实验四 Numpy 基本操作

`Numpy` 是一个强大的第三方库，主要用来支持向量与矩阵的运算，同时 `Numpy` 还对数组运算提供大量的数学函数。`Numpy` 是 `pandas`、`scikit-learn` 等的基本组成部分，用途非

常广泛。本实验将介绍 numpy 的基本情况 & ndarray，并进行一些简单的数组运算。

实验内容：了解 numpy 的基本情况，掌握其核心的 ndarray，并通过对数组的操作运算加深对 numpy 的理解

实验五 Numpy 索引与切片

索引与切片在 python 中使用非常广泛，也是其相比较其它语言来讲更具效率的地方，许多其它的语言需要通过循环和数组下标的方式来达到类似的功能，在 Python 中只需要一行代码即可实现。本实验将对索引与切片做介绍，同时因为索引与切片非常灵活，所以在日常的使用需要多练习。

实验内容：了解 Python 和 numpy 中的索引与切片功能，并对矩阵的一些变换做了解。

实验六 Numpy 数据处理

在了解了 numpy 的基本操作之后，本实验将对 numpy 的数据处理做讲解。主要有 numpy 中的基本操作、数组运算和 numpy 中的一些通用函数及文件操作。

实验内容：了解 numpy 数据处理的基本方法，并掌握 numpy 数据处理的一些常用函数。

实验七 Pandas 对象

相比较 numpy 而言 pandas 才是数据处理的利器，其实 pandas 是基于 numpy 的，提供了 Series 和 DataFrame 对象，用来对大数据进行数据处理，本实验将讲解 pandas 里面的基本的一些内容，在后续的实验中将继续对 pandas 的数据处理做介绍。

实验内容：了解 pandas 中的一些常用数据处理对象，掌握 Series、DataFrame、索引等对象，并掌握实际的操作方法。

实验八 Pandas 数据预处理

在了解了 pandas 的基本操作之后，本实验将对数据预处理做一些讲解，数据预处理在人工智能等应用中是非常重要的一个环节。

实验内容：通过 pandas 对数据做预处理，掌握数据预处理的基本概念及实际的操作流程。

实验九 Pandas 大数据处理

在大数据的分析处理中也经常用到 pandas，本实验将对该内容做简单的讲解，了解其中的一些聚合、过滤、转换及应用等。

实验内容：掌握 pandas 在大数据处理中的应用，学会通过 python 代码来实际处理数据。

实验十 Matplotlib 基础运用

matplotlib 是一个强大的绘图库，通过 matplotlib 可以编写几行代码就可以实现多种优

美的图形，本实验将对 matplotlib 做简单的介绍，后续的实验还将继续加深对 matplotlib 的使用。

实验内容：了解 matplotlib 的强大功能，编写实际的图形绘制代码。

实验十一 Matplotlib 绘图方法

前面对 matplotlib 做了简单的介绍，并绘制了一些图形，本实验还将继续深入学习 matplotlib 的绘图。

实验内容：使用 matplotlib 绘制折线图及散点图。

实验十二 Matplotlib 自定义绘图

前面都是采用了 matplotlib 自带的一些方式来绘图，本实验将通过自定义的方式来实现图形的绘制。

实验内容：学会 matplotlib 的自定义绘图。

尊敬的老师，如果您对我们的课程实验感兴趣，需要完整的课程实验资料，请填写“实验资源申请表”，并发邮件联系我们，我们将会尽快将您所选资源发送给您。谢谢。

联 系 人：张延迪

E-mail: zhangyd@hc-vision.cn

联系电话：17681806123

杭州汇萃智能科技有限公司教育部

附件：

课程资源申请表

尊敬的老师，感谢您对我们公司的支持！您可根据教学需要选择我司的课程资源，并认真填写下表，我们会尽快将您所选资源发送给您。

姓名	
学校	
院系	
职称/职务	
教龄	
E-mail	
手机号	
所授课程	

您对我司课程资源有什么建议和意见？
