

# 《专业技能实训 I》课程标准

学 分：3 学分

学 时：48 学时 （其中：讲课学时 0 上机学时：48）

先修课程：《程序设计基础》

后续课程：Python 数据挖掘

适用专业：人工智能技术应用及其计算机相关专业

开课院（系）：信息工程系

## 一、课程的性质与目标

本专业技能实训涉及科学计算库 NumPy、数据分析工具 Pandas、数据可视化、时间序列分析和文本数据分析。

通过对本课程的学习，学生能够熟悉数据分析的流程和思想，可以利用数据分析技术解决特定业务领域的问题。

## 二、教学条件要求

操作系统：Windows7 版本

Python 环境：Anaconda 4.5.9

开发工具：Jupyter Notebook

## 三、课程的主要内容及基本要求

### 第 1 章 数据分析概述

章名	数据分析概述	学时	2	
学习目标	1. 了解数据分析的背景及应用场景 2. 掌握什么是数据分析以及数据分析的流程 3. 会创建 Python 环境，使用 Anaconda 管理 Python 包 4. 会简单使用 Jupyter Notebook 5. 认识常见的数据分析工具			
知识点	了解	熟悉	掌握	重点
数据分析的背景	√			

什么是数据分析	√			
数据分析的应用场景	√			
数据分析的流程				√
Python 做数据分析的优势	√			
Anconda 的介绍		√		
Anconda 的安装及使用				√
启用 Jupyter Notebook				√
Jupyter Notebook 界面介绍及使用				√
常见的数据分析工具		√		

## 第 2 章 科学计算库 NumPy

章名	科学计算库 NumPy	学时	5	
学习目标	1. 认识 NumPy 数组对象，会创建 NumPy 数组 2. 熟悉 ndarray 对象的数据类型，并会转换数据类型 3. 掌握数组运算方式 4. 掌握数组的索引和切片 5. 会使用数组进行数据处理 6. 熟悉线性代数模块和随机数模块的使用			
知识点	了解	熟悉	掌握	重点
认识 NumPy 数组对象			√	
创建 NumPy 数组				√
ndarray 对象的数据类型				√
数组运算		√		
整数索引和切片的基本使用				√
花式（数组）索引的基本使用				√
布尔型索引的基本使用				√
数组的转置和轴对称			√	
NumPy 通用函数		√		
将条件逻辑转为数组运算			√	
数组统计运算		√		
数组排序		√		
检索数组元素			√	
线性代数模块	√			
随机数模块			√	
案例—酒鬼漫步		√		

## 第 3 章 数据分析工具 Pandas

章名	数据分析工具 Pandas	学时	7
----	---------------	----	---

学习目标	<ol style="list-style-type: none"> <li>1. 掌握 Pandas 的两种数据结构</li> <li>2. 掌握 Pandas 索引的相关操作</li> <li>3. 掌握 Pandas 的常见操作，包括算术运算、排序、统计计算</li> <li>4. 掌握 Pandas 读写数据的方式</li> </ol>			
知识点	了解	熟悉	掌握	重点
Series				√
DataFrame				√
索引对象		√		
重置索引			√	
索引操作			√	
算术运算与数据对齐		√		
数据排序			√	
统计计算与描述	√			
认识层次化索引		√		
层次化索引的操作				√
读写文本文件				√
读写 Excel 文件				√
读写 HTML 表格数据				√
读写数据库				√
案例—北京高考分数线统计分析		√		

## 第 4 章 数据预处理

章名	数据预处理	学时	6	
学习目标	<ol style="list-style-type: none"> <li>1. 掌握数据清洗常用操作，会检查和处理各类有问题的数据</li> <li>2. 掌握数据合并的常用方法，会使用不同的方式合并数据</li> <li>3. 掌握数据重塑的常见操作，会重塑 Pandas 对象的结构</li> <li>4. 掌握数据转换的常见操作，可以实现离散化和哑变量处理</li> </ol>			
知识点	了解	熟悉	掌握	重点
空值和缺失值的处理			√	
重复值的处理			√	
异常值的处理			√	
更改数据类型			√	
轴向堆叠数据				√
主键合并数据				√
根据行索引合并数据				√

合并重塑数据				√
重塑层次化索引				√
轴向旋转			√	
重命名轴索引			√	
离散化连续数据		√		
哑变量处理类别型数据	√			
案例—预处理部分地区信息		√		

## 四、学时分配

章目	上机	合计
第1章 数据分析概述	4 学时	4 学时
第2章 科学计算库 NumPy	12 学时	12 学时
第3章 数据分析工具 Pandas	16 学时	16 学时
第4章 数据预处理	14 学时	14 学时
第5章 考查—北京租房数据统计分析	2 学时	2 学时
合计	48 学时	48 学时

## 五、考核模式与成绩评定办法

本课程为考查课程，期末考试采用百分制的项目考查模式。学生的考试成绩由平时成绩（30%）和期末考查（70%）组成，其中，平时成绩包括出勤（5%）、作业（5%）、上机成绩（20%）。

## 六、选用教材和主要参考书

本大纲是根据教材《Python 快速编程入门》所设计的。

参考书籍：

传智播客.《Python 快速编程入门》人民邮电出版社. 201709

## 七、课程标准说明

本课程的授课模式为：课堂授课+上机，其中，课堂主要采用多媒体的方式进行授课，并且会通过测试题阶段测试学生的掌握程度；上机主要是编写程序，要求学生动手完成指定

的程序设计或验证。

撰写人：

审定人：

批准人：

执行时间：