# Python程序设计和数据分析

# 第一章 大数据环境下的Python语言

本章主要阐述数据的重要性、大数据的概念以及基于大数据的技术体系，以此引出Python语言在大数据环境下的重要地位。

## 1.1 大数据的概念

## 1.2 基于大数据的技术体系

### 1.2.1 大数据的架构和平台

Hadoop、Spark等平台的简单介绍。

### 1.2.2 大数据的存储

关系型数据库和NoSQL数据库。

### 1.2.3 大数据的编程语言

Java、Python、R等语言在大数据和数据分析环境下的优劣。

## 1.3 Python语言简介

### 1.3.1 Python语言的起源

### 1.3.2 Python语言的优势

### 1.3.3 大数据与Python语言

# 第二章 Python语言环境的搭建

主要介绍Python语言的版本、安装以及IDE的选择和环境配置等。

## 2.1 Python的版本

介绍Python的2.x和3.x的区别以及本教材选择的Python版本和理由。

## 2.2 Python语言环境的搭建

采用安装Anaconda的一站式解决方案。介绍Anaconda的安装以及选择国内镜像网站的配置。

### 2.2.1 安装Anaconda

### 2.2.2 配置Anaconda

### 2.2.3 安装PyCharm

### 2.3.4 使用Python Shell

### 2.3.5 Python第三方库的安装和使用

## 2.3 Jupyter Notebook的使用

重点介绍目前流行的Python交互环境Jupyter Notebook的使用，本书实例基本采用Jupyter Notebook环境。

### 2.3.1 什么是Jupyter Notebook

### 2.3.2 Markdown初步

### 2.3.3 Jupyter Notebook中程序的编写和运行

### 2.3.4 一个简单的实例

# 第三章 Python语言的数据类型

本章介绍Python语言的基本数据类型（数值、字符串、布尔类型）以及Python所特有的列表、元组、字典等数据类型

## 3.1 Python的基本类型

### 3.1.1 数值类型

### 3.1.2 布尔类型

### 3.1.3 变量

### 3.1.4 字符串

### 3.1.5 类型转换

## 3.2 列表

## 3.3 元组

## 3.4 字典

### 3.4.1 字典类型基础

### 3.4.2 字典类型操作

### 3.4.3 字典类型实例：英文文本单词词频的统计

## 3.5 关于数据类型的综合实例

# 第四章 Python语言的流程控制

本章主要介绍Python语言的选择结构和循环结构。

## 4.1 Python语言的选择结构

if else、if elif else结构以及一个综合实例

### 4.1.1 if else结构

### 4.1.2 if elif else结构

## 4.2 Python语言的循环结构

for、while循环结构以及一个综合实例

### 4.2.1 for循环结构

### 4.2.2 while循环结构

### 4.2.3 循环综合实例

## 4.3 Python语言的异常处理

try except else和try finally结构。

# 第五章 函数、输入输出和文件

主要介绍Python内建函数、自定义函数的创建和调用、文件的读写等。

## 5.1 Python的内建函数

## 5.2 自定义函数

### 5.2.1 函数的定义

### 5.2.2 返回值和参数

### 5.2.3 递归函数

## 5.3 函数创建和调用综合实例

## 5.4 Python语言的输入输出

## 5.5 文件读写

# 第六章 Python面向对象编程

简单介绍面向对象的基本概念、Python语言类的创建、继承和多态等概念。

## 6.1 面向对象编程基本概念

## 6.2 Python的类

## 6.3 类的继承

## 6.4 面向对象编程综合实例

# 第七章 Python数据分析

主要介绍Python数据分析相关的numpy、matplotlib、pandas等第三方库的应用。

## 7.1 基于Python的数据分析简介

## 7.2 Python数据分析环境

### 7.2.1 numpy初步

### 7.2.2 matplotlib初步

### 7.2.3 pandas初步

## 7.3 Python统计分析基础

### 7.3.1 描述性数据分析

### 7.3.2 探索性数据分析

## 7.4 基于Python的数据分析实例

# 第八章 Python爬虫与数据分析

主要介绍基于Python的爬虫程序设计，涉及beautifulsoup、pyQuery、scrapy等第三方库的使用。

## 8.1 简单的Python爬虫

## 8.2 基于beautifulsoup的Python爬虫

豆瓣电影Top250网站数据获取和分析。

## 8.3 基于pyquery的Python爬虫

淘宝商品搜索数据的获取和分析。

## 8.4 基于scrapy的Python爬虫

## 8.5 综合实例：豆瓣读书数据的爬取和分析

# 第九章 Python的图形用户界面

介绍基于内置的tkinter库和PyQt库的用户界面。

## 9.1 基于tkinter的用户界面

## 9.2 基于PyQt的用户界面

# 第十章 Python游戏编程

简单介绍基于pygame库的游戏编程。

# 第十一章 Python机器学习介绍

简介机器学习的现状和Python中机器学习相关的库的应用。

## 11.1 机器学习的概念

## 11.1 Python的机器学习库scikit-learn

# 第十二章 Python文本分析

基于Python的文本分析、词云的制作。

## 12.1 文本分析基础

## 12.2 Python文本分析

## 12.3 综合实例：宋词分析