



《Python程序设计基础教程（微课版）》

<http://dbl-lab.xmu.edu.cn/post/python>

第11章 基于数据库的持久化



林子雨 博士/副教授

厦门大学计算机科学与技术系

E-mail: ziyulin@xmu.edu.cn ▶▶

主页: <http://dbl-lab.xmu.edu.cn/linziyu>





主讲教师



2017年度厦门大学奖教金获得者

2020年度厦门大学奖教金获得者

主讲教师：厦门大学 林子雨 博士/副教授

中国高校首个“数字教师”提出者和建设者

2009年7月从事教师职业以来

累计**免费**网络发布超过**1500万**字高价值教学和科研资料

网络浏览量超过**1500万**次



提纲

- 11.1 关系数据库
- 11.2 关系数据库标准语言SQL
- 11.3 MySQL的安装和使用
- 11.4使用Python操作MySQL数据库

本PPT是如下教材的配套讲义：
《Python程序设计基础教程（微课版）》

厦门大学 林子雨,赵江声,陶继平 编著，人民邮电出版社
《Python程序设计基础教程（微课版）》教材官方网站：
<http://dblalab.xmu.edu.cn/post/python>

高等院校程序设计新形态精品系列

PYTHON

Python Programming Language

Python 程序设计基础教程

| 微课版 |

林子雨 赵江声 陶继平 编著



名师精品

多年计算机教学实践的厚积薄发



深入浅出

清晰呈现 Python 语言学习路径



实例丰富

有效提升编程语言的学习趣味



资源全面

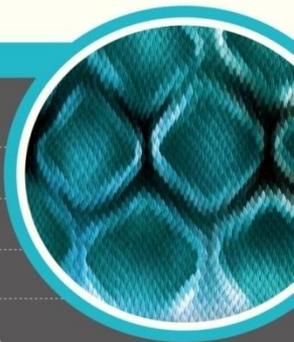
构建全方位一站式在线服务体系



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS





11.1 关系数据库

- 数据库是一种主流的数据存储和管理技术。数据库指的是以一定方式储存在一起、能为多个用户共享、具有尽可能小的冗余度、与应用程序彼此独立的数据集合。
- 对数据库进行统一管理的软件被称为“数据库管理系统”（**DBMS: Database Management System**），在不引起歧义的情况下，经常会混用“数据库”和“数据库管理系统”这两个概念。
- 在数据库的发展历史上，先后出现过网状数据库、层次数据库、关系数据库等不同类型的数据库，这些数据库分别采用了不同的数据模型（数据组织方式），目前比较主流的数据库是关系数据库，它采用了关系数据模型来组织和管理数据。



11.1 关系数据库

- 一个关系数据库可以看成是许多关系表的集合，每个关系表可以看成一张二维表格，如表11-1所示的学生信息表。

表11-1 学生信息表

学号	姓名	性别	年龄	考试成绩
95001	张三	男	21	88
95002	李四	男	22	95
95003	王梅	女	22	73
95004	林莉	女	21	96



11.1 关系数据库

- 目前市场上常见的关系数据库产品包括Oracle、SQL Server、MySQL、DB2等。因为关系数据库的数据通常具有规范的结构，因此，通常把保存在关系数据库中的数据称为“结构化数据”。
- 与此相对应，类似图片、视频、声音文件所包含的数据，没有规范的结构，被称为“非结构化数据”，而类似网页文件（HTML文件）这种具有一定结构但又不是完全规范化的数据，被称为“半结构化数据”。



11.1 关系数据库

总体而言，关系数据库具有如下特点：

（1）存储方式。关系数据库采用表格的储存方式，数据以行和列的方式进行存储，要读取和查询都十分方便。

（2）存储结构。关系数据库按照结构化的方法存储数据，每个数据表的结构都必须事先定义好（比如表的名称、字段名称、字段类型、约束等），然后再根据表的结构存入数据，这样做的好处就是，由于数据的形式和内容在存入数据之前就已经定义好了，所以，整个数据表的可靠性和稳定性都比较高，但是带来的问题就是，数据模型不够灵活，一旦存入数据后，如果需要修改数据表的结构就会十分困难。



11.1 关系数据库

(3) 存储规范。关系数据库为了规范化数据、减少重复数据以及充分利用好存储空间，把数据按照最小关系表的形式进行存储，这样数据管理就可以变得很清晰、一目了然。当存在多个表时，表和表之间通过主外键关系发生关联，并通过连接查询获得相关结果。

(4) 扩展方式。由于关系数据库将数据存储和数据表中，数据操作的瓶颈出现在多张数据表的操作中，而且数据表越多这个问题越严重。如果要缓解这个问题，只能提高处理能力，也就是选择速度更快、性能更高的计算机，这样的方法虽然具有一定的拓展空间，但是这样的拓展空间是非常有限的，也就是一般的关系型数据库只具备有限的纵向扩展能力。



11.1 关系数据库

(5) 查询方式。关系数据库采用结构化查询语言（即**SQL: Structured Query Language**）来对数据库进行查询。结构化查询语言是高级的非过程化编程语言，允许用户在高层数据结构上工作。它不要求用户指定对数据的存放方法，也不需要用户了解具体的数据存放方式，所以，各种具有完全不同底层结构的数据库系统，可以使用相同的结构化查询语言作为数据输入与管理的接口。结构化查询语言语句可以嵌套，这使它具有极大的灵活性和强大的功能。



11.1 关系数据库

(6) 事务性。关系数据库可以支持事务的ACID特性（原子性（Atomicity）、一致性（Consistency）、隔离性（Isolation）、持久性（Durability））。当事务被提交给了DBMS（数据库管理系统），则DBMS需要确保该事务中的所有操作都成功完成且其结果被永久保存在数据库中，如果事务中有的操作没有成功完成，则事务中的所有操作都需要被回滚，回到事务执行前的状态，从而确保数据库状态的一致性。

(7) 连接方式。不同的关系数据库产品都遵守一个统一的数据库连接接口标准，即ODBC（Open Database Connectivity）。ODBC的一个显著优点是，用它生成的程序是与具体的数据库产品无关的，这样可以为数据库用户和开发人员屏蔽不同数据库异构环境的复杂性。ODBC提供了数据库访问的统一接口，为应用程序实现与平台的无关性和可移植性提供了基础，因而获得了广泛的支持和应用。



11.2 关系数据库标准语言SQL

11.2.1 SQL简介

11.2.2 常用的SQL语句

本PPT是如下教材的配套讲义：

《Python程序设计基础教程（微课版）》

厦门大学 林子雨,赵江声,陶继平 编著，人民邮电出版社

《Python程序设计基础教程（微课版）》教材官方网站：

<http://dbl原因.xmu.edu.cn/post/python>

高等院校程序设计新形态精品系列

PYTHON

Python Programming Language

Python

程序设计基础教程

| 微课版 |

林子雨 赵江声 陶继平 编著



名师精品

多年计算机教学实践的厚积薄发



深入浅出

清晰呈现 Python 语言学习路径



实例丰富

有效提升编程语言的学习趣味



资源全面

构建全方位一站式在线服务体系



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS



11.2.1 SQL简介

自从SQL成为国际标准语言以后，各个数据库厂家纷纷推出各自的SQL软件或与SQL的接口软件。这就使大多数数据库均用SQL作为共同的数据存取语言和标准接口，使不同数据库系统之间的相互操作有了共同的基础。SQL已经成为数据库领域中的主流语言，其意义十分重大。SQL的主要特点如下：

- (1) 综合统一。
- (2) 高度非过程化。
- (3) 面向集合的操作方式。
- (4) 以同一种语法结构提供多种使用方式。
- (5) 语言简洁，易学易用。



11.2.2 常用的SQL语句

1.创建数据库

在使用数据库之前，需要创建数据库，具体语法如下：

CREATE DATABASE 数据库名称;

每条SQL语句的末尾用英文分号结束。

可以使用如下语句查看已经创建的所有数据库：

SHOW DATABASES;

创建好数据库以后，可以使用如下语句打开数据库：

USE 数据库名称;



11.2.2 常用的SQL语句

2.创建表

在一个数据库中，会包含多个表。创建一个表的语法如下：

```
CREATE TABLE 表名称  
(  
列名称1 数据类型,  
列名称2 数据类型,  
列名称3 数据类型,  
....  
);
```

可以使用如下SQL语句查看所有已经创建的表：

```
SHOW TABLES;
```



11.2.2 常用的SQL语句

3.插入数据

可以使用INSERT INTO语句向表中插入新的记录，其语法形式如下：

```
INSERT INTO 表名称 VALUES (值1, 值2,...);
```

也可以指定所要插入数据的列：

```
INSERT INTO表名称(列1, 列2,...) VALUES (值1, 值2,...);
```

4.查询数据

可以使用SELECT语句从数据库中查询数据，其语法形式如下：

```
SELECT 列名称 FROM 表名称;
```



11.2.2 常用的SQL语句

5.修改数据

可以使用UPDATE语句修改表中的数据，其语法形式如下：

```
UPDATE 表名称 SET 列名称 = 新值 WHERE 列名称 = 某值;
```

6.删除数据

可以使用DELETE语句从表中删除记录，其语法形式如下：

```
DELETE FROM 表名称 WHERE 列名称 = 某值;
```



11.2.2 常用的SQL语句

7.删除表

可以使用DROP TABLE语句从数据库中删除一个表，其语法形式如下：

DROP TABLE 表名称;

8.删除数据库

可以使用DROP DATABASE语句删除一个数据库，其语法形式如下：

DROP DATABASE 数据库名称;



11.3 MySQL的安装和使用

11.3.1 安装MySQL

11.3.2 MySQL的使用方法

本PPT是如下教材的配套讲义：

《Python程序设计基础教程（微课版）》

厦门大学 林子雨,赵江声,陶继平 编著，人民邮电出版社

《Python程序设计基础教程（微课版）》教材官方网站：

<http://dbl原因.xmu.edu.cn/post/python>

高等院校程序设计新形态精品系列

PYTHON

Python Programming Language

Python

程序设计基础教程

| 微课版 |

林子雨 赵江声 陶继平 编著



名师精品

多年计算机教学实践的厚积薄发



深入浅出

清晰呈现 Python 语言学习路径



实例丰富

有效提升编程语言的学习趣味



资源全面

构建全方位一站式在线服务体系



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS



11.3.1 安装MySQL

访问如下MySQL官网地址下载安装包：

<https://dev.mysql.com/downloads/windows/installer/8.0.html>

在MySQL下载页面中（如图11-1所示），选择“mysql-installer-community-8.0.23.0.msi”下载。



图11-1 MySQL下载网页



11.3.1 安装MySQL

使用安装包mysql-installer-community-8.0.23.0.msi开始安装，如果在安装过程中提示需要安装“.NET Framework 4.5.2”，则需要到如下网址下载.NET Framework 4.5.2的安装文件NDP452-KB2901907-x86-x64-AllOS-ENU.exe并安装：

<https://www.microsoft.com/zh-CN/download/confirmation.aspx?id=42642>



11.3.1 安装MySQL

在安装MySQL过程中，当出现“Choosing a Setup Type”界面时，需要选择“Server only”（如图11-2所示）。

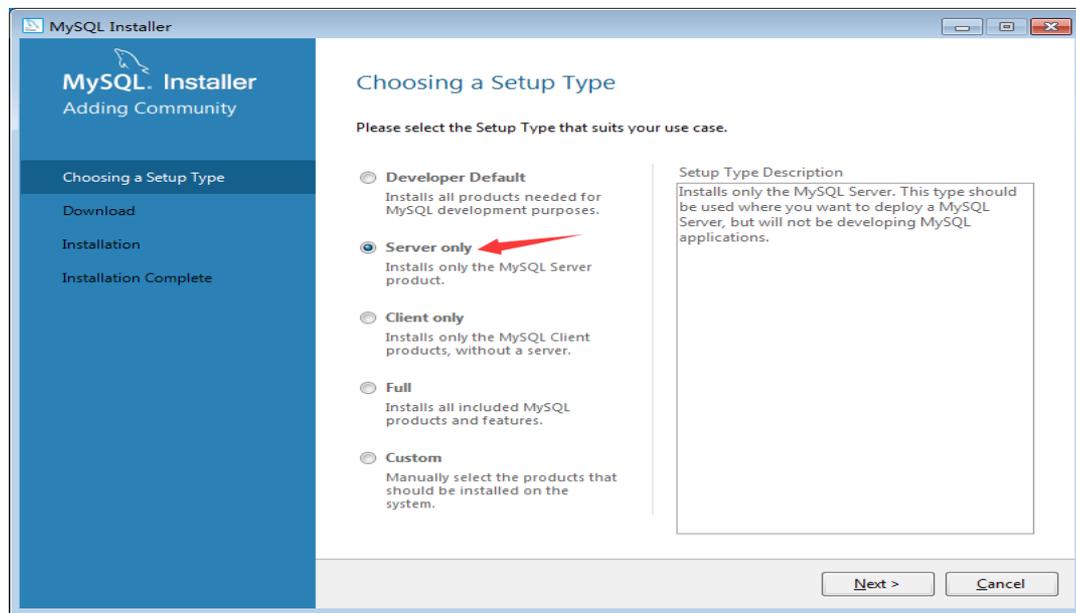


图11-2 选择安装类型界面



11.3.1 安装MySQL

在安装MySQL过程中，如果提示需要安装“Microsoft Visual C++ 2015-2019 Redistributable (x64) – 14.28.29325”时，选择同意安装即可（如图11-3所示）。

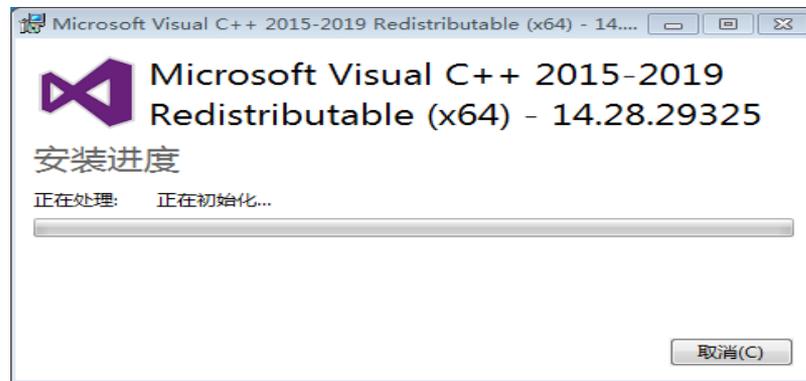


图11-3 安装过程中显示的界面



11.3.1 安装MySQL

安装完成以后，MySQL数据库的后台服务进程已经被自动启动，这时就需要使用一个客户端工具来操作MySQL数据库，我们可以使用MySQL安装时自带的命令行界面作为客户端工具来操作数据库。具体方法是，在Windows7的开始菜单中点击“MySQL 8.0 Command Line Client”图标，然后输入数据库密码（这个密码是在安装MySQL的过程中用户自己设置的），就会出现如图11-4所示界面。可以在命令提示符“mysql>”后面输入SQL语句来执行数据库的各种操作。

```
MySQL 8.0 Command Line Client
Enter password: *****
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 12
Server version: 8.0.23 MySQL Community Server - GPL

Copyright (c) 2000, 2021, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input
statement.

mysql> _
```

图11-4 MySQL的命令行界面



11.3.1 安装MySQL

需要说明的是，MySQL数据库服务进程启动以后，会占用一定的系统资源。实际上，我们平时在计算机上很少使用MySQL数据库，因此，为了减少对系统资源的占用，没有必要每次开机都自动启动MySQL数据库后台服务进程，可以设置为“手动”启动服务进程，这样，只有当需要用到MySQL数据库时，再去手动启动即可。这里以Windows7操作系统为例介绍如何把MySQL数据库服务设置为“手动”启动。



11.3.1 安装MySQL

在Windows系统桌面上的“计算机”图标上单击鼠标右键，在弹出的菜单中点击“管理”，在出现的计算机管理界面（如图11-5所示）中，在左侧栏目中点击“服务”，在右侧栏目中会出现很多服务进程，其中，就可以找到名称为“MySQL80”的服务进程，可以看到，该服务进程的状态为“已启动”，启动类型为“自动”。

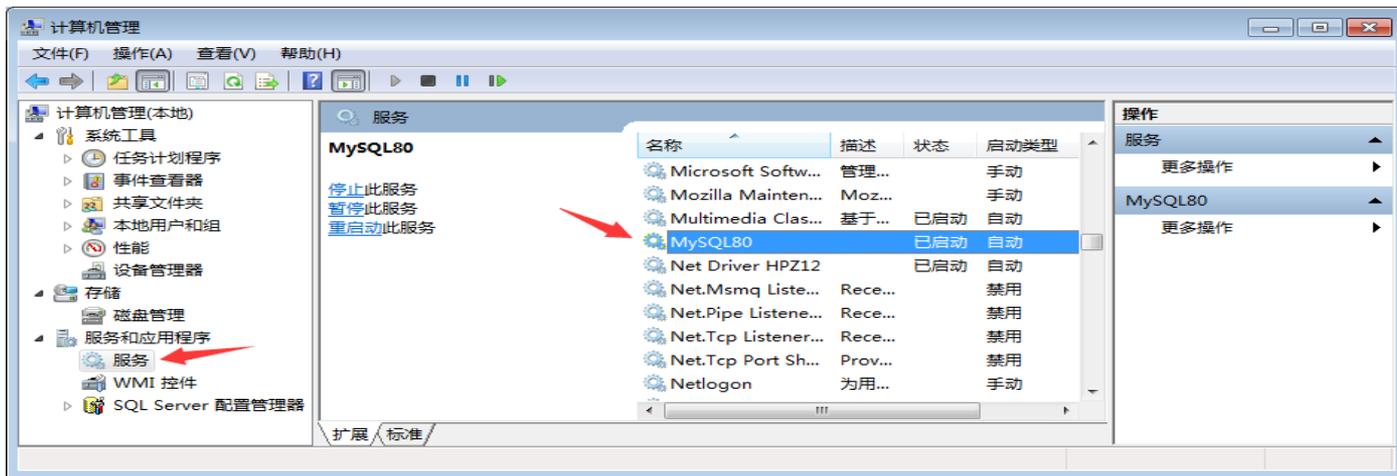


图11-5 计算机管理界面



11.3.1 安装MySQL

在“MySQL80”这一行上单击鼠标右键，在弹出的菜单中点击“属性”，会弹出如图11-6所示的界面，在这个界面中，在“服务状态”下面有三个按钮，即“启动”、“停止”和“暂停”，分别用来启动、停止和暂停MySQL服务进程。为了修改启动类型，可以在“启动类型”右侧的下拉列表中选择“手动”，最后单击“确定”按钮即可。

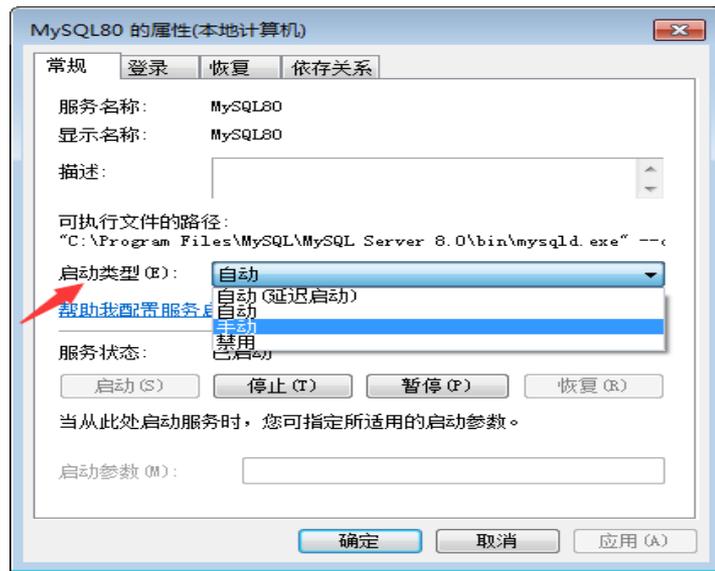


图11-6 MySQL启动类型设置界面



11.3.2 MySQL的使用方法

这里给出一个综合实例来演示MySQL数据库的用法，具体要求是，创建一个管理学生信息的数据库，把表11-3中的数据填充到数据库中，并完成相关的数据库操作。

表11-3 学生表

学号	姓名	性别	年龄
95001	王小明	男	21
95002	张梅梅	女	20



11.3.2 MySQL的使用方法

打开MySQL数据库的命令行界面，输入如下SQL语句创建数据库school:

```
mysql> CREATE DATABASE school;
```

需要注意的是，SQL语句中可以不用区分字母大小写。

可以使用如下SQL语句查看已经创建的所有数据库:

```
mysql> SHOW DATABASES;
```

创建好数据库school以后，可以使用如下SQL语句打开数据库:

```
mysql> USE school;
```



11.3.2 MySQL的使用方法

使用如下SQL语句在数据库school中创建一个表student:

```
mysql>CREATE TABLE student(  
    -> sno char(5),  
    -> sname char(10),  
    -> ssex char(2),  
    -> sage int);
```

使用如下SQL语句查看已经创建的表:

```
mysql> SHOW TABLES;
```

使用如下SQL语句向student表中插入两条记录:

```
mysql> INSERT INTO student VALUES('95001','王小明','男',21);  
mysql> INSERT INTO student VALUES('95002','张梅梅','女',20);
```



11.3.2 MySQL的使用方法

使用如下SQL语句查询student表中的记录:

```
mysql> SELECT * FROM student;
```

使用如下SQL语句修改表中的数据:

```
mysql> UPDATE student SET age =21 WHERE sno='95001';
```

使用如下SQL语句删除student表:

```
mysql> DROP TABLE student;
```



11.3.2 MySQL的使用方法

使用如下SQL语句查询数据库中还存在哪些表：

```
mysql> SHOW TABLES;
```

使用如下SQL语句删除数据库school：

```
mysql> DROP DATABASE school;
```

使用如下SQL语句查询系统中还存在哪些数据库：

```
mysql> SHOW DATABASES;
```



11.4使用Python操作MySQL数据库

使用Python操作MySQL数据库之前，需要安装PyMySQL，它是Python中操作MySQL的模块。在Windows操作系统的cmd中运行如下命令安装PyMySQL：

```
> pip install PyMySQL
```

11.4.1 连接数据库

11.4.2 创建表

11.4.3 插入数据

11.4.4 修改数据

11.4.5 查询数据

11.4.6 删除数据



11.4.1 连接数据库

首先打开MySQL数据库的命令行界面，在MySQL数据库中创建一个名称为school的数据库（如果已经存在该数据库，则需要先删除再创建），然后，编写如下代码发起对数据库的连接：



11.4.1 连接数据库

```
01 # mysql1.py
02 import pymysql.cursors
03 # 连接数据库
04 connect = pymysql.Connect(
05     host='localhost', # 主机名
06     port=3306, # 端口号
07     user='root', # 数据库用户名
08     passwd='123456', # 密码
09     db='school', # 数据库名称
10     charset='utf8' #编码格式
11 )
12 # 获取游标
13 cursor = connect.cursor()
14 # 执行SQL查询
15 cursor.execute("SELECT VERSION()")
16 # 获取单条数据
17 version = cursor.fetchone()
18 # 打印输出
19 print("MySQL数据库版本是: %s" % version)
20 # 关闭数据库连接
21 connect.close()
```

上面代码的执行结果如下：
MySQL数据库版本是： 8.0.23



11.4.2 创建表

在school数据库中创建一个表student，具体代码如下：

```
01      # mysql2.py
02      import pymysql.cursors
03      # 连接数据库
04      connect = pymysql.Connect(
05          host='localhost',
06          port=3306,
07          user='root',
08          passwd='123456'
09          db='school',
10          charset='utf8'
11      )
12      # 获取游标
13      cursor = connect.cursor()
14      # 如果表存在，则先删除
15      cursor.execute("DROP TABLE IF EXISTS student")
```



11.4.2 创建表

```
16      # 设定SQL语句
17      sql = ""
18      CREATE TABLE student(
19          sno char(5),
20          sname char(10),
21          ssex char(2),
22          sage int);
23      ""
24      # 执行SQL语句
25      cursor.execute(sql)
26      # 关闭数据库连接
27      connect.close()
```



11.4.3 插入数据

把表11-3中的两条数据插入到student表中，具体代码如下：

```
01      # mysql3.py
02      import pymysql.cursors
03      # 连接数据库
04      connect = pymysql.Connect(
05          host='localhost',
06          port=3306,
07          user='root',
08          passwd='123456',
09          db='school',
10          charset='utf8'
11      )
12      # 获取游标
13      cursor = connect.cursor()
14      # 插入数据
15      sql = "INSERT INTO student(sno,sname,ssex,sage) VALUES ('%s', '%s', '%s', %d)"
```



11.4.3 插入数据

```
16     data1 = ('95001','王小明','男',21)
17     data2 = ('95002','张梅梅','女',20)
18     cursor.execute(sql % data1)
19     cursor.execute(sql % data2)
20     connect.commit()
21     print('成功插入数据')
22     # 关闭数据库连接
23     connect.close()
```



11.4.4 修改数据

把学号为“95002”的学生的年龄修改为21岁，具体代码如下：

```
01      # mysql4.py
02      import pymysql.cursors
03      # 连接数据库
04      connect = pymysql.Connect(
05          host='localhost',
06          port=3306,
07          user='root',
08          passwd='123456',
09          db='school',
10          charset='utf8'
11      )
12      # 获取游标
13      cursor = connect.cursor()
```



11.4.4 修改数据

```
14      # 修改数据
15      sql = "UPDATE student SET sage = %d WHERE sno = '%s' "
16      data = (21, '95002')
17      cursor.execute(sql % data)
18      connect.commit()
19      print('成功修改数据')
20      # 关闭数据库连接
21      connect.close()
```



11.4.5 查询数据

找出学号为“95001”的学生的具体信息，具体代码如下：

```
01      # mysql5.py
02      import pymysql.cursors
03      # 连接数据库
04      connect = pymysql.Connect(
05          host='localhost',
06          port=3306,
07          user='root',
08          passwd='123456',
09          db='school',
10          charset='utf8'
11      )
12      # 获取游标
13      cursor = connect.cursor()
```



11.4.5 查询数据

```
14     # 查询数据
15     sql = "SELECT sno,sname,ssex,sage FROM student WHERE sno = '%s' "
16     data = ('95001',) #元组中只有一个元素的时候需要加一个逗号
17     cursor.execute(sql % data)
18     for row in cursor.fetchall():
19         print("学号:%s\t姓名:%s\t性别:%s\t年龄:%d" % row)
20     print('共查找出', cursor.rowcount, '条数据')
21     # 关闭数据库连接
22     connect.close()
```



11.4.6 删除数据

删除学号为“95002”的学生记录，具体代码如下：

```
01      # mysql6.py
02      import pymysql.cursors
03      # 连接数据库
04      connect = pymysql.Connect(
05          host='localhost',
06          port=3306,
07          user='root',
08          passwd='123456',
09          db='school',
10          charset='utf8'
11      )
12      # 获取游标
13      cursor = connect.cursor()
```



11.4.6 删除数据

```
14     # 删除数据
15     sql = "DELETE FROM student WHERE sno = '%s'"
16     data = ('95002',) #元组中只有一个元素的时候需要加一个逗号
17     cursor.execute(sql % data)
18     connect.commit()
19     print('成功删除', cursor.rowcount, '条数据')
20     # 关闭数据库连接
21     connect.close()
```



附录A：主讲教师林子雨简介



主讲教师：林子雨

单位：厦门大学计算机科学与技术系

E-mail: ziyulin@xmu.edu.cn

个人网页: <http://dmlab.xmu.edu.cn/post/linziyu>

数据库实验室网站: <http://dmlab.xmu.edu.cn>

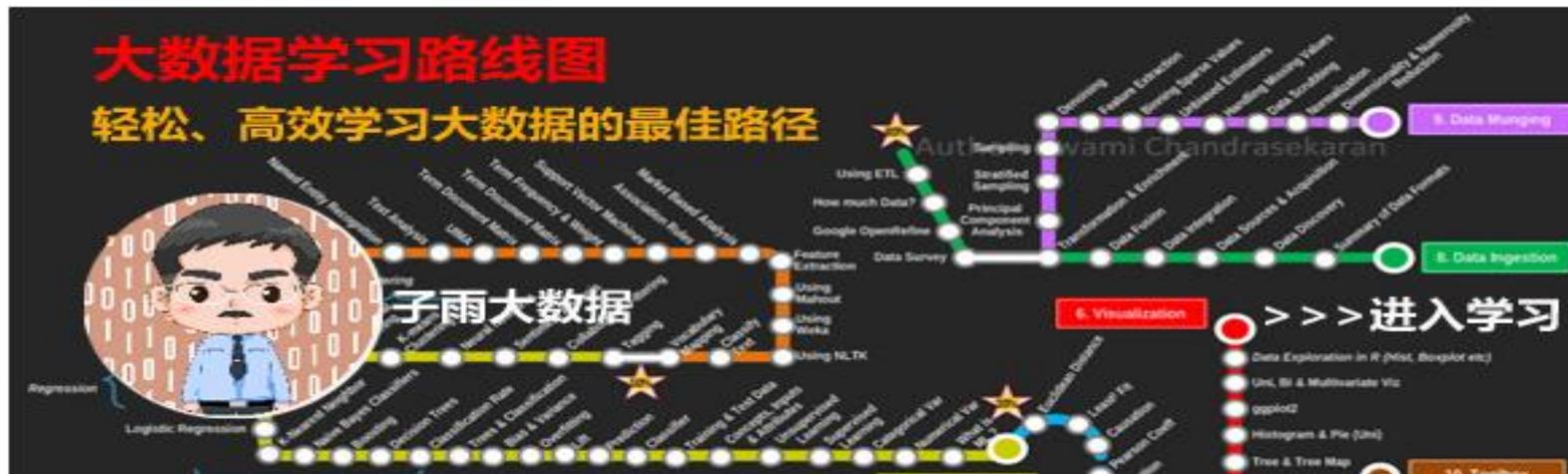


扫一扫访问个人主页

林子雨，男，1978年出生，博士（毕业于北京大学），全国高校知名大数据教师，现为厦门大学计算机科学系副教授，厦门大学信息学院实验教学中心主任，曾任厦门大学信息科学与技术学院院长助理、晋江市发展和改革局副局长。中国计算机学会数据库专业委员会委员，中国计算机学会信息系统专业委员会委员。国内高校首个“数字教师”提出者和建设者，厦门大学数据库实验室负责人，厦门大学云计算与大数据研究中心主要建设者和骨干成员，2013年度、2017年度和2020年度厦门大学教学类奖教金获得者，荣获2019年福建省精品在线开放课程、2018年厦门大学高等教育成果特等奖、2018年福建省高等教育教学成果二等奖、2018年国家精品在线开放课程。主要研究方向为数据库、数据仓库、数据挖掘、大数据、云计算和物联网，并以第一作者身份在《软件学报》《计算机学报》和《计算机研究与发展》等国家重点期刊以及国际学术会议上发表多篇学术论文。作为项目负责人主持的科研项目包括1项国家自然科学基金青年基金项目(No.61303004)、1项福建省自然科学基金青年基金项目(No.2013J05099)和1项中央高校基本科研业务费项目(No.2011121049)，主持的教改课题包括1项2016年福建省教改课题和1项2016年教育部产学协作育人项目，同时，作为课题负责人完成了国家发改委城市信息化重大课题、国家物联网重大应用示范工程区域试点泉州市工作方案、2015泉州市互联网经济调研等课题。中国高校首个“数字教师”提出者和建设者，2009年至今，“数字教师”大平台累计向网络免费发布超过1000万字高价值的研究和教学资料，累计网络访问量超过1000万次。打造了中国高校大数据教学知名品牌，编著出版了中国高校第一本系统介绍大数据知识的专业教材《大数据技术原理与应用》，并成为京东、当当网等网店畅销书籍；建设了国内高校首个大数据课程公共服务平台，为教师教学和学生学习大数据课程提供全方位、一站式服务，年访问量超过400万次，累计访问量超过1500万次。



附录B：大数据学习路线图



大数据学习路线图访问地址：<http://dblab.xmu.edu.cn/post/10164/>



附录C：林子雨大数据系列教材



林子雨大数据系列教材

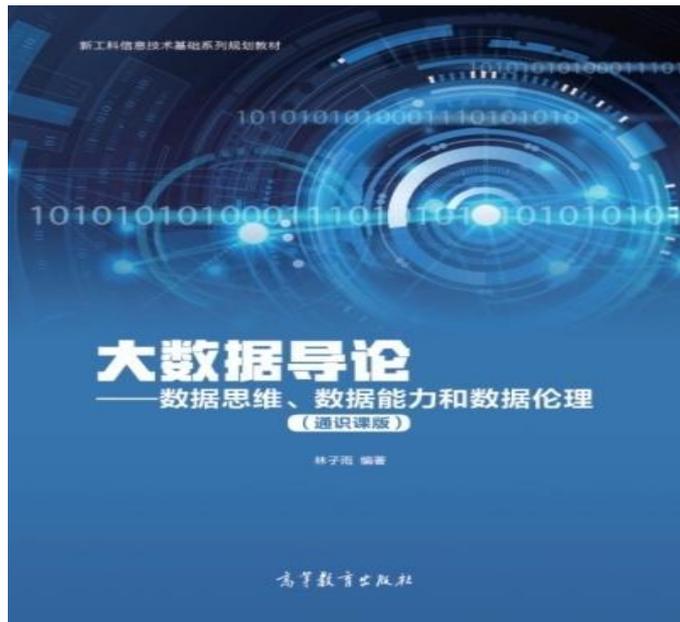
用于导论课、专业课、实训课、公共课

了解全部教材信息：<http://dbllab.xmu.edu.cn/post/bigdatabook/>



附录D：《大数据导论（通识课版）》教材

开设全校公共选修课的优质教材



本课程旨在实现以下几个培养目标：

- 引导学生步入大数据时代，积极投身大数据的变革浪潮之中
- 了解大数据概念，培养大数据思维，养成数据安全意识
- 认识大数据伦理，努力使自己的行为符合大数据伦理规范要求
- 熟悉大数据应用，探寻大数据与自己专业的应用结合点
- 激发学生基于大数据的创新创业热情

高等教育出版社 ISBN:978-7-04-053577-8 定价：32元 版次：2020年2月第1版
教材官网：<http://dbl原因.xmu.edu.cn/post/bigdataintroduction/>



附录E：《大数据导论》教材

- 林子雨 编著《大数据导论》
- 人民邮电出版社，2020年9月第1版
- ISBN:978-7-115-54446-9 定价：49.80元

教材官网：<http://dblab.xmu.edu.cn/post/bigdata-introduction/>



开设大数据专业导论课的优质教材



扫一扫访问教材官网



附录F：《大数据技术原理与应用（第3版）》教材

《大数据技术原理与应用——概念、存储、处理、分析与应用（第3版）》，由厦门大学计算机科学系林子雨博士编著，是国内高校第一本系统介绍大数据知识的专业教材。人民邮电出版社 ISBN:978-7-115-54405-6 定价：59.80元

全书共有17章，系统地论述了大数据的基本概念、大数据处理架构Hadoop、分布式文件系统HDFS、分布式数据库HBase、NoSQL数据库、云数据库、分布式并行编程模型MapReduce、Spark、流计算、Flink、图计算、数据可视化以及大数据在互联网、生物医学和物流等各个领域的应用。在Hadoop、HDFS、HBase、MapReduce、Spark和Flink等重要章节，安排了入门级的实践操作，让读者更好地学习和掌握大数据关键技术。

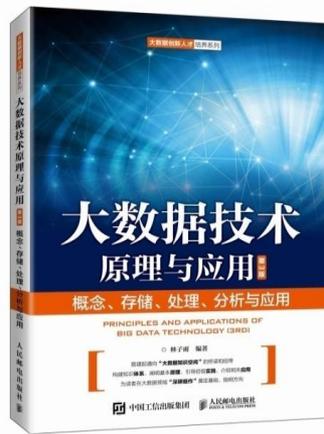
本书可以作为高等院校计算机专业、信息管理等相关专业的大数据课程教材，也可供相关技术人员参考、学习、培训之用。

欢迎访问《大数据技术原理与应用——概念、存储、处理、分析与应用》教材官方网站：

<http://dbllab.xmu.edu.cn/post/bigdata3>



扫一扫访问教材官网





附录G：《大数据基础编程、实验和案例教程（第2版）》

本书是与《大数据技术原理与应用（第3版）》教材配套的唯一指定实验指导书

大数据教材



1+1黄金组合
厦门大学林子雨编著

配套实验指导书



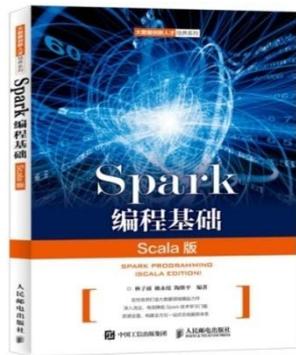
- 步步引导，循序渐进，详尽的安装指南为顺利搭建大数据实验环境铺平道路
- 深入浅出，去粗取精，丰富的代码实例帮助快速掌握大数据基础编程方法
- 精心设计，巧妙融合，八套大数据实验题目促进理论与编程知识的消化和吸收
- 结合理论，联系实际，大数据课程综合实验案例精彩呈现大数据分析全流程

林子雨编著《大数据基础编程、实验和案例教程（第2版）》

清华大学出版社 ISBN:978-7-302-55977-1 定价：69元 2020年10月第2版



附录H: 《Spark编程基础 (Scala版)》



《Spark编程基础 (Scala版)》

厦门大学 林子雨, 赖永炫, 陶继平 编著

披荆斩棘, 在大数据丛林中开辟学习捷径
填沟削坎, 为快速学习Spark技术铺平道路
深入浅出, 有效降低Spark技术学习门槛
资源全面, 构建全方位一站式在线服务体系

人民邮电出版社出版发行, ISBN:978-7-115-48816-9

教材官网: <http://dbllab.xmu.edu.cn/post/spark/>

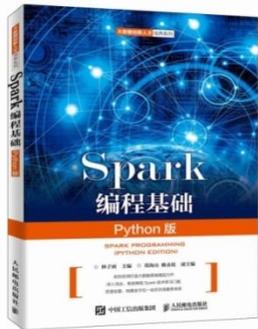


本书以Scala作为开发Spark应用程序的编程语言, 系统介绍了Spark编程的基础知识。全书共8章, 内容包括大数据技术概述、Scala语言基础、Spark的设计与运行原理、Spark环境搭建和使用方法、RDD编程、Spark SQL、Spark Streaming、Spark MLlib等。本书每个章节都安排了入门级的编程实践操作, 以便读者更好地学习和掌握Spark编程方法。本书官网免费提供了全套的在线教学资源, 包括讲义PPT、习题、源代码、软件、数据集、授课视频、上机实验指南等。



附录I: 《Spark编程基础 (Python版)》

《Spark编程基础 (Python版)》



厦门大学 林子雨, 郑海山, 赖永炫 编著

披荆斩棘, 在大数据丛林中开辟学习捷径
填沟削坎, 为快速学习Spark技术铺平道路
深入浅出, 有效降低Spark技术学习门槛
资源全面, 构建全方位一站式在线服务体系



人民邮电出版社出版发行, ISBN:978-7-115-52439-3

教材官网: <http://dbllab.xmu.edu.cn/post/spark-python/>

本书以Python作为开发Spark应用程序的编程语言, 系统介绍了Spark编程的基础知识。全书共8章, 内容包括大数据技术概述、Spark的设计与运行原理、Spark环境搭建和使用方法、RDD编程、Spark SQL、Spark Streaming、Structured Streaming、Spark MLlib等。本书每个章节都安排了入门级的编程实践操作, 以便读者更好地学习和掌握Spark编程方法。本书官网免费提供了全套的在线教学资源, 包括讲义PPT、习题、源代码、软件、数据集、上机实验指南等。



附录J：高校大数据课程公共服务平台



高校大数据课程

公 共 服 务 平 台

<http://dblab.xmu.edu.cn/post/bigdata-teaching-platform/>



扫一扫访问平台主页 扫一扫观看3分钟FLASH动画宣传片



附录K：高校大数据实训课程系列案例教材

为了更好地满足高校开设大数据实训课程的教材需求，厦门大学数据库实验室林子雨老师团队联合企业共同开发了《高校大数据实训课程系列案例》，目前已经完成开发的系列案例包括：

- 《电影推荐系统》（已经于2019年5月出版）
- 《电信用户行为分析》（已经于2019年5月出版）
- 《实时日志流处理分析》
- 《微博用户情感分析》
- 《互联网广告预测分析》
- 《网站日志处理分析》



系列案例教材将于2019年陆续出版发行，教材相关信息，敬请关注网页后续更新！

<http://dblab.xmu.edu.cn/post/shixunkecheng/>



扫一扫访问大数据实训课程系列案例教材主页

The background is a solid blue color with faint, light-blue silhouettes of people. At the top, there are two groups of people holding hands, suggesting a community or team. On the right side, there is a silhouette of a person standing with their hand on their head. In the bottom left, there are silhouettes of people sitting at a table, possibly in a meeting or classroom setting.

Thank You!

Department of Computer Science, Xiamen University, 2022