



数据建模课程资料

课程介绍

数据建模课程是计算机科学与工程学院开设的一门面向大一学生的基础课程,课程实验内容涵盖 Python编程入门、线性规划及统计学基础,结合Jupyter Notebook进行实验。课程实验指导开发环 境安装、代码管理、实验报告生成及工具使用,提供丰富的在线课程与书籍进行参考,助力学生掌 握数据分析与建模技能。

课程实验内容

- 1. Python基础
- 2. 线性规划
- 3. 统计学基础

开发环境的安装

- 1. 安装Python SDK 3.10+, 下载安装地址: www.python.org/downloads/
- 2. 下载和安装Python虚拟环境Anaconda: 下载链接
- 3. 运行"Anaconda Prompt"命令行, 切换到课程的根目录,运行命令: pip install -r requirements.txt
- 4. 安装集成开发环境VScode,下载安装地址: code.visualstudio.com/download
- 5. 在VSCode中安装课程需要的插件:
 - Python Extension Pack
 - Jupyter

- Code Runner
- o Data Wrangler
- Markdown All in One
- Markdown Math
- markdownlint
- o Print
- vscode-icons
- o code spell checker
- Github copilot
- 6. 安装和配置版本控制工具git, 下载和安装地址: https://git-scm.com/downloads

git配置方法如下:

```
git config --global user.name "[firstname lastname]"
git config --global user.email "[valid-email]"
```

- 7. 为了能够自动生成实验报告的pdf文件还需要安装两个辅助软件,下载并安装wkhtmltopdf: 下载地址,并将 wkthmltopdf 的默认安装目录 C:\Program Files\wkhtmltopdf\bin 添加到Windows的环境变量 Path 中。请查看添加Windows环境变量的方法
- 8. 下载并安装 pandoc:下载地址

课程材料的下载

- 1. 访问课程仓库主页: github.com/zhoujing204/data-modeling-course-student
- 2. 运行下面的命令克降课程仓库:

```
git clone https://github.com/zhoujing204/data-modeling-course-student.git
```

如果没有安装qit,也可以直接通过链接下载zip文件:下载链接

实验代码的管理

由于实验室机房的电脑安装了还原卡,电脑重启后会恢复到启动时的状态,请及时在实验室电脑外的地方妥善保管好你编写的代码,这里介绍如何利用Github和Git管理你的代码:

- 1. 注册Github账号,地址:<u>github.com/signup</u>,也可以码云来作为远程仓库:<u>gitee.com</u>
- 2. 在编写完成一段代码后,使用下面的命令(需要安装配置好git)保存到本地仓库(注意本地仓库的代码会被还原卡还原,不安全):

```
git add .
git commit -m "编写文字描述你完成的代码"
```

使用下面的命令本地仓库恢复你保存的代码:

git checkout filename

3. 在VSCode集成开发环境中登录你的Github账号,运行下面的命令将本地仓库同步到远程仓库:

git push

ſĊ

ſĠ

4. 使用下面的命令可以从远程仓库下载和更新代码到本地仓库:

git pull

Q

生成并提交实验报告

方法一:

- 1. 完成Jupyter Notebook文件 (.ipynb文件) 中的习题并保存。
- 2. 直接利用课程提供的脚本自动生成实验报告pdf文件。
- 3. 将实验报告pdf文件提交到对应的班级和实验的腾讯文档收集表:文档汇总地址

方法二:

- 1. 完成Jupyter Notebook文件 (.ipynb文件) 中的习题并保存。
- 2. 将ipynb文件上传到网站转换为pdf文件: www.vertopal.com/en/convert/ipynb-to-pdf
- 3. 将转换的pdf文件提交到对应的班级和实验的腾讯文档收集表:文档汇总地址

参考资料

1. 在线课程

○ Python课程: 课程主页

○ 统计学课程:课程主页

○ 数据科学课程:课程主页

。 Git课程: 课程主页

。 线性规划课程: <u>课程主页</u>

2. 参考书籍

- Python编程:从入门到实践 (第3版) -- [美] 埃里克·马瑟斯 (Eric Matthes)
- 流畅的 Python (第2版) -- [巴西] 卢西亚诺 拉马略
- o Practical Statistics for Data Scientists 50+ Essential Concepts Using R and Python
- Python for Data Analysis(第2版)