|  |  |  |  |
| --- | --- | --- | --- |
| 课题名称 | 第3章 数据分析工具Pandas | 计划学时 | 24 课时 |
| 内容分析 | Pandas是一个基于NumPy的Python库，专门为了解决数据分析任务而创建的，它不仅纳入了大量的库和一些标准的数据模型，而且提供了高效操作大型数据集所需的工具，被广泛地应用到很多领域中，包括经济、统计、分析等学术和商业领域。 |
| 教学目标及基本要求 | 1. Pandas的数据结构分析
2. Pandas索引操作及高级索引
3. 算术运算与数据对齐
4. 数据排序
5. 统计计算与描述
6. 层次化索引
7. 读写操作
8. 案例—北京高考分数线统计分析
 |
| 教学重点 | 1. Pandas的数据结构分析
2. Pandas索引操作及高级索引
3. 算术运算与数据对齐
4. 数据排序
5. 统计计算与描述
6. 层次化索引
7. 读写操作
 |
| 教学难点 | 1. Pandas的数据结构分析
2. Pandas索引操作及高级索引
3. 层次化索引
4. 读写操作
 |
| 教学方式 | 教学采用教师课堂讲授为主，使用教学PPT讲解 |
| 教学过程 | **第一课时****（Series，DataFrame，索引对象）****一、创设情境，导入Pandas**1. 教师介绍新的数据分析工具Pandas，并引出Pandas索引操作。

Pandas是一个基于NumPy的Python库，专门为了解决数据分析任务而创建的，它不仅纳入了大量的库和一些标准的数据模型，而且提供了高效操作大型数据集所需的工具，被广泛的应用到很多领域中，包括经济、统计、分析等学术和商业领域。1. 教师根据课件，介绍Pandas中常用的两种数据结构。

在Pandas中常用的两种数据结构分别是Series和DataFrame，Series是一种一维的数据结构，而DataFrame是一种二维的数据结构。1. 教师根据课件，介绍Pandas索引对象。

Pandas中的索引都是Index类对象，又称为索引对象，该对象是不可以进行修改的，以保障数据的安全。1. 明确学习目标
* 要求学生掌握Series
* 要求学生掌握DataFrame
* 要求学生熟悉索引对象

**二、进行重点知识的讲解**1. 教师根据课件，介绍Series

Series是一个一维数组对象，它能够保存任何类型的数据，比如整数、字符串、浮点数等，主要由一组数据与之相关的索引两部分构成。1. 教师根据课件，介绍DataFrame

DataFrame是一个二维数组对象，它每列的数据可以是不同的数据类型，与Series的结构相似，DataFrame的结构也是由索引和数据组成的，不同的是，DataFrame的索引不仅有行索引，还有列索引。1. 教师根据课件，介绍索引对象

Pandas中的索引对象都是Index类对象，又称为索引对象，该对象是不可以进行修改的，以保障数据的安全。**三、归纳总结，布置作业/随堂练习**1. 回顾上课前的学习目标，并对本节课的内容进行总结。

 教师总结本节课需要掌握的知识点，包括Series、DataFrame、索引对象。1. 使用日照职业技术学院在线教学平台下发课后作业。

**第二课时****（重置索引、索引操作、算术运算与数据对齐）****一、回顾上节课的内容，继续讲解本课时的索引操作和算术运算与数据对齐**1. 教师对学生们的疑问进行统一答疑。
2. 通过示例代码，回顾总结上节课内容，继续讲解本节内容。
3. 明确学习目标
* 要求学生掌握重置索引
* 要求学生掌握索引操作
* 要求学生熟悉算术运算与数据对齐

**二、进行重点知识的讲解**1. 教师根据课件，讲述重置索引操作。

Pandas中提供了一个重要的方法是reindex()，该方法的作用是原索引和新索引进行匹配，也就是说，新索引含有原索引的数据，而原索引数据按照新索引排序。如果新索引中没有原索引数据，那么程序不仅不会报错，而且会添加新的索引，并将值填充为NaN或者使用fill\_value()填充其他值。1. 教师根据课件，讲述Series和DataFrame其他的索引操作。

通过索引位置获取数据、通过索引名称获取数据、使用切片获取数据、获取不连续数据。1. 教师根据课件，讲述算术运算与数据对齐。

Pandas执行算术运算时，会先按照索引进行对齐，对齐以后进行相应的运算，没有对齐的位置会用NaN进行补齐。其中，Series是按行索引对齐的，DataFrmae是按行索引、列索引对齐的。**三、归纳总结，布置作业**1. 回顾学习目标，对本节课的内容进行总结。

教师带领学生总结本节课需要掌握的内容，包括重置索引和索引操作。1. 布置随堂练习，检查学生掌握情况。

 根据在线教学平台和随堂练习资源，给学生布置随堂练习，检测学生的掌握程度，并对学生出现的问题进行解决。 1. 使用日照职业技术学院在线教学平台下发课后作业。

**第三课时****（按索引排序、按值排序、常用的统计计算、统计描述）****一、回顾上节课内容，继续讲解本节课的内容**1. 教师对学生们的疑问进行统一答疑。
2. 回顾总结上节课内容，讲解数据排序和统计计算与描述知识。

在上节课中，介绍了Pandas的索引操作和算术运算与数据对齐，接下来，将介绍数据排序和统计计算与描述。1. 明确学习目标
* 要求学生掌握按索引排序
* 要求学生掌握按值排序
* 要求学生熟悉常用的统计计算
* 要求学生熟悉统计描述

**二、进行重点知识的讲解**1. 教师根据课件，讲述数据排序——按索引排序。

Pandas中按索引排序使用的sort\_index()方法，该方法可以用行索引或者列索引进行排序。1. 教师根据课件，讲述数据排序——按值排序。

Pandas中按值排序使用的sort\_values()方法。1. 教师根据课件，讲述常用的统计描述计算。

通过几个简单的示例演示常用的统计描述方法，如max、min、sum等。1. 教师根据课件，讲述统计描述。

如果希望一次性输出多个统计指标，比如平均值、最大值、最小值、求和等，则我们可以调用describe()方法实现，而不用再单独地逐个调用相应的统计方法。**三、归纳总结，布置作业**1. 回顾学习目标，对本节课的内容进行总结

教师带领学生总结数据排序的方法。1. 布置随堂练习，检查学生掌握情况。

 根据在线教学平台和随堂练习资源，给学生布置随堂练习，检测学生的掌握程度，并对学生出现的问题进行解决。1. 使用日照职业技术学院在线教学平台下发课后作业。

**第四课时****（认识层次化索引、层次化索引操作、读写文本文件）****一、回顾上节课内容，讲解本节内容**1. 教师讲解上节课作业，并对学生的疑问进行统一答疑。
2. 回顾上节课的内容，讲解本节课新的知识点。

本节课主要的新内容有认识层次化索引，层次化索引操作，以及读写文本文件。1. 明确学习目标
* 要求学生熟悉层次化索引
* 要求学生掌握层次化索引的操作
* 要求学生掌握读写文本文件操作

**二、进行重点知识的讲解**（1）教师通过图例引出什么层次化索引。当Pandas的Series或DataFrame对象只有一层索引时，可称其为单层索引，而层次化索引则可以理解为单层索引的延伸，即在一个轴方向上具有多层索引。根据示例，讲解创建层次化索引的几种方法（from\_tuples()、from\_arrays()、from\_product()）。（2）教师根据课件，讲解层次化索引的操作？常用的层次化索引操作有选取子集操作、交换分层顺序、排序分层。 （3）教师讲解完层次化索引的知识，讲解Pandas读写数据操作。在对数据分析时，通常不会将需要分析的数据写入到程序中，而是以文件或数据库的形式保存，所以接下来介绍Pandas中的读写操作。在本节中主要讲解Pandas的读写文本的操作to\_csv、read\_csv、read\_table**三、归纳总结，布置作业**（1）回顾学习目标，对本节课的内容进行总结。 教师带领学生总结本节课的内容，内容包括了解层次化索引、掌握层次化索引的操作、以及文本文件的读写操作。 （2）使用日照职业技术学院在线教学平台下发随堂练习，检测学生对课堂内容的掌握情况。**第五课时****（读写Excel文件、读取HTML表格数据、读写数据库）****一、回顾上节课内容，讲解本节内容**1. 教师讲解上节课作业，并对学生的疑问进行统一答疑。
2. 回顾上节课的内容，继续讲解读写数据的其他操作。

本节课主要的新内容有读写Excel文件、读取HTML表格数据、读写数据库、案例实现。1. 明确学习目标
* 要求学生掌握读写Excel文件
* 要求学生掌握读取HTML表格数据
* 要求学生掌握读写数据库

**二、进行重点知识的讲解**（1）教师通过介绍excel文件的格式，引出使用Pandas读写Excel文件。Pandas中提供了对Excel文件进行读写操作的方法，分别为to\_excel()和read\_excel()。（2）教师根据课件，讲解读取HTML表格数据。通过以网上某个表格数据进行实时读取引出使用read\_html()读取表格数据。 （3）教师根据课件，讲解读写数据库数据。对于存储在数据库中的数据，Pandas提供了read\_sql()函数用来读取，同时也提供了用于写入数据库的to\_sql()方法，教材中以mysql为例演示使用to\_sql和read\_sql 读取数据库数据。**三、归纳总结，布置作业**（1）回顾学习目标，对本节课的内容进行总结。 教师带领学生总结本节课的内容，内容包括掌握读写Excel、读取HTML表格数据、读写数据库。 （2）使用日照职业技术学院在线教学平台下发随堂练习，检测学生对课堂内容的掌握情况。**第六课时****（上机练习）**上机练习主要针对本章中需要重点掌握的知识点，以及在程序中容易出错的内容进行练习，通过上机练习可以考察同学对知识点的掌握情况，对代码的熟练程度。**上机一：（ 练习教材中的示例代码）****形式：单独完成****第七课时****（上机练习）**上机练习主要针对本章中需要重点掌握的知识点，以及在程序中容易出错的内容进行练习，通过上机练习可以考察同学对知识点的掌握情况，对代码的熟练程度。**上机一：（练习教材中的示例代码）****形式：单独完成****上机二：（完成案例—北京高考分数线统计分析）****形式：单独完成** |
| 思考题和习题 | 见教材第3章配套的习题 |
| 教学后记 |  |