

高等职业教育财经类专业群 **智慧财经** 系列教材

高等职业教育财务会计类专业 **“岗课赛证”** 融通教材



高等职业教育在线开放课程新形态一体化教材



大数据技术 在财务中的应用 (Python版)

张卫平 徐大诏 主编

- | 大数据与会计
- | 大数据与财务管理
- | 大数据与审计
- | 会计信息管理
- | 业财数据应用与管理
- | 财税大数据应用



中国教育出版传媒集团
高等教育出版社

内容提要

本书是高等职业教育财经类专业群智慧财经系列教材之一，也是高等职业教育财务会计类专业“岗课赛证”融通教材。

本书以企业财务人员常用的工作场景为载体，基于项目引领、任务驱动、理实结合、教学做一体的教学模式，采用由浅入深、层次递进的“层次化”策略，将本书设为7个项目，包括数据分析认知、Python 程序设计基础、利用 Numpy 对财务信息进行数值计算、利用 Pandas 进行财务数据预处理、利用 Pandas 进行财务数据分析、利用 Matplotlib 和 Pyecharts 实现财务数据可视化、利用 Pandas 进行财务报表分析。

本书同步建设有数字课程“财务大数据分析”，围绕主要知识技能点配套建设了微课、视频讲解、教学课件、全书程序源代码、练习题参考答案等教学资源。可登录“智慧职教”平台，进行线上学习；还可通过书中的二维码实现移动化随扫随学。授课教师如需获取本书授课用教学课件、全书程序源代码、参考答案等配套资源，请登录“高等教育出版社产品信息检索系统”（xuanshu.hep.com.cn）免费下载。

本书可以作为高等职业教育专科、职业本科院校财务会计类专业及相关财经商贸类专业的教材，也可以作为财务工作从业人员和社会人士学习“大数据技术在财务中的应用”的参考用书。

图书在版编目（CIP）数据

大数据技术在财务中的应用：Python版 / 张卫平，
徐大诏主编. -- 北京：高等教育出版社，2025. 2.
ISBN 978-7-04-063189-0
I. F275-39
中国国家版本馆CIP数据核字第2024PR0214号

大数据技术在财务中的应用（Python版）
DASHUJU JISHU ZAI CAIWU ZHONG DE YINGYONG (Python BAN)

策划编辑 马 一	责任编辑 马 一	封面设计 李树龙	版式设计 马 云
责任绘图 邓 超	责任校对 吕红颖	责任印制 张益豪	

出版发行 高等教育出版社
社 址 北京市西城区德外大街4号
邮政编码 100120
印 刷 唐山嘉德印刷有限公司
开 本 787mm×1092mm 1/16
印 张 24.5
字 数 410千字
插 页 1
购书热线 010-58581118

咨询电话 400-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.hepmall.com.cn>
<http://www.hepmall.com>
<http://www.hepmall.cn>
版 次 2025年2月第1版
印 次 2025年7月第2次印刷
定 价 49.80元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换
版权所有 侵权必究
物 料 号 63189-00

目 录

项目一 数据分析认知 / 1

任务一 数据分析基本内容认知 / 3

任务二 Python 开发环境的搭建 / 7

项目练习 / 23

项目二 Python 程序设计基础 / 25

任务一 Python 的基础语法 / 27

任务二 Python 的数据类型 / 37

任务三 Python 的流程控制语句 / 57

任务四 Python 的函数 / 73

项目练习 / 85

项目三 利用 Numpy 对财务信息进行数值计算 / 91

任务一 Numpy 的基本操作 / 93

任务二 Numpy 的统计分析 / 109

项目练习 / 122

项目四 利用 Pandas 进行财务数据预处理 / 125

任务一 Pandas 的基本操作 / 127

任务二 利用 Pandas 进行财务数据筛选 / 153

任务三 利用 Pandas 高阶函数进行财务数据处理 / 165

任务四 利用 Pandas 进行财务数据合并 / 178

任务五 利用 Pandas 进行财务数据清洗 / 193

项目练习 / 211

项目五 利用 Pandas 进行财务数据分析 / 217

任务一 利用 Pandas 对财务数据进行排序与排名分析 / 219

任务二 利用 Pandas 对财务数据进行统计与描述分析 / 228

任务三 利用 Pandas 对财务数据进行分组与分段分析 / 236

任务四 利用 Pandas 对财务数据进行交叉与透视分析 / 248

项目练习 / 257

项目六 利用 Matplotlib 和 Pyecharts 实现财务数据可视化 / 261

任务一 利用 Matplotlib 实现财务数据可视化 / 263

任务二 利用 Pyecharts 实现财务数据可视化 / 295

项目练习 / 325

项目七 利用 Pandas 进行财务报表分析 / 329

任务一 利用 Pandas 对财务报表进行偿债能力与营运能力分析 / 331

任务二 利用 Pandas 对财务报表进行盈利能力与发展能力分析 / 361

项目练习 / 377

用价值的信息的过程，可为管理决策提供支持。

（二）数据分析的流程

数据分析应按照一定的流程进行，主要包括需求分析、数据采集、数据预处理、分析与建模、数据可视化和数据部署应用。

1. 需求分析

数据分析的本质目标是满足分析需求，如果没有确定分析需求，将会导致分析无法落地。因此，需要相应的背景知识帮助分析人员明晰分析目标，确定用户需求。

2. 数据采集

根据确定的分析目标，采集相关数据。在数字经济时代，数据的种类、复杂性与来源都不断增长，包括财务报表数据、销售业务数据、证券价格数据、环境监测数据、医疗保健数据、政府公开的统计数据、期刊数据，以及社会化媒体数据等。对来源丰富、种类多样的数据，不仅要从多个途径进行采集，而且还要对格式多样的数据进行处理与合并。

3. 数据预处理

初步采集来的数据往往格式多样、形式复杂，通常不能直接用于数据分析。因此，对数据进行预处理是必不可少的步骤，数据预处理过程包括数据集成、数据清洗、数据变换，以及数据规约等。

4. 分析与建模

结合分析目标对预处理后的数据进行建模。首先，对数据进行比较基础的统计分析，了解数据的基本规律与特征，常见的方法包括对比分析法、同比分析法、环比分析法、结构分析法、因素分析法等。在了解数据的基本规律与特征后，再通过包括相关分析、聚类分析、回归分析、分类分析、关联分析，以及各种机器学习算法，对数据建立模型。

5. 数据可视化

为了便于理解，要对数据分析的结果通过图形的方式进行呈现，这就是数据可视化。将数据可视化，能够提高人们阅读数据分析结果的能力。常见的数据可视化工具包括 Python、R 语言、Power BI 等。

6. 数据部署应用

数据部署应用是指将数据分析结果应用到实际业务中，是数据产生实际价值的

4 大数据技术在财务中的应用（Python 版）