

第一章 实验总则

一、课程基本信息

课程名称：经济数据数量化分析课程类别：集中性实践环节必修课先修课程：线性代数、统计学实验环境：计算机、Python 3.x 编程环境，包含 pandas、numpy、matplotlib、requests、pymysql、tushare 等第三方库。

二、课程目标

掌握金融分析相关基本理论，并能应用于具体经济领域。

能够运用 Python 及第三方数据包分析金融领域实际问题。

具备自主学习与解决金融投资分析、风险管理等实际问题的能力。

三、实验学时分配

数据采集及清洗：10 学时

金融数据库搭建及使用：6 学时

证券相关数据分析与挖掘：8 学时

Python 与常用统计计量方法的应用：8 学时

四、实验纪律与要求

严格遵守实验室管理规定，爱护设备，保持环境整洁。

实验前预习内容，明确目的与步骤。

独立完成实验，规范编写代码，严禁抄袭。

数据采集合法合规，禁止恶意高频访问服务器。

按时完成实验并提交实验报告。

第二章 实验一 数据采集及清洗

一、实验基本信息

实验项目：数据采集及清洗

实验类型：综合性实验

实验学时：10 学时

实验环境：Python 编程环境 (requests、json、pandas、numpy)

二、实验目的

了解接口请求方式 (get、post、请求头、请求体)，能用 Python 调用接口完成数据采集。

了解文本数据处理方法与应用场景，能进行金融文本采集与处理。

熟悉结构化数据清洗流程，能完成去空、去重、规范化、异常值处理。

三、实验原理

网络数据采集：通过 API 接口获取金融数据，遵循网络安全与数据合规要求。

数据清洗：对缺失值、重复值、异常值、格式错误进行处理，提升数据质量。

文本处理：对金融新闻、公告等文本进行提取、去噪、结构化转换。

四、实验内容与步骤

(一) 金融数据接口采集

安装并配置 Python 环境与相关库。

申请金融数据接口，编写 get 请求获取股票数据。

解析 JSON 数据并转换为 DataFrame 格式。

(二) 金融文本数据处理

爬取金融网页文本信息。

去除特殊符号、空格、无效字符等噪声。

文本关键信息提取与简单结构化。

(三) 金融数据清洗

读取数据，查看数据结构、缺失值、重复值。

缺失值处理：删除或填充。

重复值删除。

异常值识别与处理。

数据格式统一与字段规范化。

五、实验结果提交

可运行 Python 代码文件。

清洗前后数据对比表。

清洗完成的规范数据集。

六、注意事项

控制接口请求频率，禁止恶意爬取。

数据来源合法，尊重版权。

代码注释清晰，便于复现检查。

第三章 实验二 金融数据库搭建及使用

一、实验基本信息

实验项目：金融数据库搭建及使用

实验类型：综合性实验

实验学时：6 学时

实验环境：Python、MySQL、pymysql、SQLAlchemy

二、实验目的

了解常见数据库及应用场景，能根据业务选择数据库。

掌握 MySQL 中 SQL 建表、增删改查基础操作。

能用 Python 实现数据库连接、数据读写与查询分析。

三、实验原理

关系型数据库适合结构化金融数据存储与管理。

SQL 实现数据表定义与数据操作。

Python 通过数据库驱动实现与 MySQL 交互。

四、实验内容与步骤

（一）数据库环境搭建

安装 MySQL 及可视化工具。

创建金融专用数据库。

（二）SQL 基础操作

创建股票信息表、行情数据表。

完成数据插入、修改、删除。

条件查询、排序、聚合统计。

（三）Python 连接数据库

安装数据库连接库。

编写连接代码。

将清洗后数据写入数据库。

从数据库读取并做简单分析。

五、实验结果提交

数据库结构截图。

Python 连接与操作代码。

数据读写与查询结果。

六、注意事项

妥善保管数据库账号密码。

重要操作前备份数据。

合理设置字段类型与长度。

第四章 实验三 证券相关数据分析与挖掘

一、实验基本信息

实验项目：证券相关数据分析与挖掘

实验类型：综合性实验

实验学时：8 学时

实验环境：Python、pandas、numpy、matplotlib

二、实验目的

能通过接口获取常用证券金融数据。

掌握收益率计算、描述性统计、时间序列分析方法。

能完成投资组合相关性分析与可视化。

三、实验原理

收益率：简单收益率与对数收益率计算。

描述统计：集中趋势、离散程度、分布特征。

时间序列：趋势、波动、平稳性观察。

相关性分析：资产间联动关系。

四、实验内容与步骤

（一）数据获取与预处理

获取多只股票历史价格数据。

日期排序、缺失值处理、数据对齐。

（二）收益率计算与描述统计

计算日收益率。

计算均值、方差、最值、分位数等。

绘制直方图、时序图。

（三）时间序列分析

观察价格与收益率波动特征。

绘制移动平均线，分析趋势。

（四）相关性分析

计算多只股票收益率相关系数矩阵。

绘制热力图，分析相关性强弱。

五、实验结果提交

完整 Python 分析代码。

描述统计结果表、可视化图表。

相关性分析结果与解读。

六、注意事项

时间序列数据必须按日期有序排列。

图表标题、坐标轴、图例完整清晰。

结合金融理论解释统计结果。

第五章 实验四 Python 与常用统计计量方法的应用

一、实验基本信息

实验项目：Python 与常用统计计量方法的应用

实验类型：综合性实验

实验学时：8 学时

实验环境：Python、pandas、numpy、scipy、statsmodels

二、实验目的

掌握描述性统计、正态性检验、区间估计方法。

理解线性回归原理，完成参数估计与假设检验。

实现等权投资组合构建与分析。

三、实验原理

描述统计：均值、方差、偏度、峰度。

正态性检验：判断数据分布特征。

区间估计：均值置信区间计算。

线性回归：最小二乘估计、t 检验、F 检验。

四、实验内容与步骤

（一）单只股票描述统计与检验

选取股票收益率，计算核心指标。

正态性检验。

均值置信区间计算与解读。

（二）经典线性回归分析

设定自变量与因变量。

构建回归模型，输出系数、 R^2 、t 值、F 值。

系数显著性与方程整体显著性检验。

（三）等权投资组合实践

选取多只股票构建等权重组合。

计算组合收益率与风险指标。

对比个股与组合表现。

五、实验结果提交

描述统计与检验结果报告。

回归模型输出与检验结果。

等权组合收益风险分析。

六、注意事项

变量选择符合经济逻辑。

正确解读显著性水平与 P 值。

结合金融理论解释结果。

第六章 实验报告撰写规范

每份实验独立撰写，结构完整。

报告包含：实验名称、目的、原理、步骤、代码、结果、分析、总结。

代码可直接运行，关键步骤加注释。

图表清晰规范，标注完整。

结果分析结合金融与统计知识，逻辑严谨。

严禁抄袭，雷同报告按零分处理。

第七章 考核与成绩评定

一、考核方式

平时考查 (40%) + 期末实验报告 (60%) 平时考查包括：课堂互动 (30%)、实训操作 (40%)、研讨 (30%)。

二、评分标准

代码正确性与完整性：40%

实验结果与图表质量：30%

结果分析深度与规范性：20%

报告格式与态度：10%