

# 《Python 与数据分析》大作业

## 任务说明：

- 1) 每个团队 2-3 人，其中 1 人担任组长
- 2) A 题和 B 题任选其一即可
- 3) 题目和示例数据已公布，参见“日新学堂-资料”，4 月 15 日公布全部数据（提示：可先基于示例数据进行代码和报告的撰写）
- 4) 每个团队需提交**代码**（Python 文件，如收集并使用了其他相关数据，请一并提交）、**视频**（5 分钟时长的 MP4 文件，所有团队成员均需出境，介绍项目核心内容、团队分工等，可以以腾讯会议录屏的形式）、**报告**（PDF 文件，参考格式如下）
- 5) 提交截止时间为 **2023 年 5 月 31 日 22:00（第 15 教学周周三）**，所有文件打包为 zip 压缩包，命名为“选题（A or B）-团队 ID-所有团队成员名”，发送至 miqing@bjut.edu.cn，**超过截止时间将没有分数**
- 6) 禁止任何形式的抄袭或者作假

报告的**参考格式**，可根据题目进行调整，必须包含目录，其他部分非强制

（优秀报告参见：<https://www.tipdm.org/yxzpl/index.jhtml>）

### 摘要

#### 一、绪论

- 1) 背景
- 2) 问题重述
- 3) 研究意义

#### 二、数据描述和预处理

- 1) 数据来源和特点（数据量、字段含义）
- 2) 数据预处理（合并、清洗、标准化、变换）

#### 三、需要解决的问题 1

利用 NumPy、matplotlib、pandas、scikit-learn (or MindSpore) 进行数据的分析、可视化展示与建模，提取有用信息并形成结论，可适当展示核心

代码

四、需要解决的问题 2

同上

五、需要解决的问题 3

同上

六、总结

结论、不足、未来工作

七、团队管理

介绍团队分工、贡献占比（以百分比的形式给出）、项目推进流程等

参考文献