

课程详述

COURSE SPECIFICATION

以下课程信息可能根据实际授课需要或在课程检讨之后产生变动。如对课程有任何疑问，请联系授课教师。

The course information as follows may be subject to change, either during the session because of unforeseen circumstances, or following review of the course at the end of the session. Queries about the course should be directed to the course instructor.

1.	课程名称 Course Title	数据科学实践 Data Science Projects and Presentation
2.	授课院系 Originating Department	统计与数据科学系 Department of Statistics and Data Science
3.	课程编号 Course Code	STA322
4.	课程学分 Credit Value	2
5.	课程类别 Course Type	专业核心课 Major Core Courses(2020&2021)\专业选修课 Major Elective Courses(2022)
6.	授课学期 Semester	春季 Spring
7.	授课语言 Teaching Language	中英双语 English & Chinese
8.	授课教师、所属学系、联系方式 (如属团队授课, 请列明其他授课教师) Instructor(s), Affiliation & Contact (For team teaching, please list all instructors)	胡延庆、统计与数据科学系、huyq@sustech.edu.cn
9.	实验员/助教、所属学系、联系方式 Tutor/TA(s), Contact	待公布 To be announced
10.	选课人数限额(可不填) Maximum Enrolment (Optional)	

11. 授课方式 Delivery Method	讲授 Lectures	习题/辅导/讨论 Tutorials	实验/实习 Lab/Practical	其它(请具体注明) Other (Please specify)	总学时 Total
学时数 Credit Hours	16		32		48
12. 先修课程、其它学习要求 Pre-requisites or Other Academic Requirements	MA204 数理统计 (Mathematical Statistics), STA321 分布式存储与并行计算 (Distributed Storage and Parallel Computing)				
13. 后续课程、其它学习规划 Courses for which this course is a pre-requisite					
14. 其它要求修读本课程的学系 Cross-listing Dept.					

教学大纲及教学日历 SYLLABUS

15. 教学目标 Course Objectives

本课程探讨应用数据科学的设计、制定和执行的概念和技术。涵盖的主题将包括实验设计，统计分析和预测建模，机器学习，数据可视化，科学写作和演示。组织学生分组团队合作讨论，以项目组的形式完成小组项目任务。通过本课程的学习，使学生提高解决实际问题的能力，并锻炼表达沟通能力。

This course explores concepts and techniques for design, formulation, and execution of practical, applied data science. Topics covered will include experimental design, statistical analysis and predictive modeling, machine learning, data visualization, scientific writing, and presentation. In addition, we will organize students to discuss and cooperate in groups and complete projects. Through the study of this course, students can and will improve their ability to solve practical problems and exercise their ability of expression and communication.

16. 预达学习成果 Learning Outcomes

通过本课程的学习，学生能够熟悉和掌握数据科学常用的方法和技巧，包括统计分析模型、预测模型、机器学习算法与实践、数据可视化方法以及科学写作、展示的能力。通过学生间的项目合作，培养数据分析、协作的能力。

Through the study of this course, students can be familiar with and master the methods and techniques commonly used in data science, including statistical analysis models, predictive models, machine learning algorithms and practices, data visualization methods, and the ability of scientific writing and presentation. Cultivate the ability of data analysis and collaboration through project cooperation between students.

17. 课程内容及教学日历 (如授课语言以英文为主，则课程内容介绍可以用英文；如团队教学或模块教学，教学日历须注明主讲人)

Course Contents (in Parts/Chapters/Sections/Weeks. Please notify name of instructor for course section(s), if this is a team teaching or module course.)

1. 搜索引擎、推荐系统及可视化，4 小时理论+ 6 小时实验。（第 1-4 周）
 2. 股票市场数据处理与分析，2 小时理论+ 4 小时实验。（第 5-6 周）
 3. 基于主成分分析的公共基金综合评价，2 小时理论+ 4 小时实验。（第 7-8 周）
 4. 自选题目期中汇报, 4 小时（第 8 周）
 5. 医学图像恢复和重建，4 小时理论+ 2 小时实验。（第 9-11 周）
 6. 地理空间数据分析和建模，2 小时理论+ 4 小时实验。（第 12-13 周）
 7. 数据清洗、补全方法及应用，2 小时理论+ 4 小时实验。（第 14-15 周）
 8. 自选题目期终汇报，4 小时（第 16 周）
-
1. Search engine, Recommendation system and Visualisation, 4 hours + 6 lab hours.
 2. Stock Market Data Process and Analysis: 2 lecture hours + 4 lab hours.
 3. Comprehensive Evaluation of Public Funds Based on Principal Component Analysis: 2 hours + 4 lab hours.
 4. Assessment 1, 4 hours
 5. Medical image restoration and reconstruction. 4 lecture hours + 2 lab hours.
 6. Geospatial data analysis and modelling. 2 lecture hours + 4 lab hours.
 7. Methods and applications of data cleaning and imputation: 2 hours + 4 lab hours.
 8. Assessment 2, 4 hours

18. 教材及其它参考资料 Textbook and Supplementary Readings

暂无教材，教师自制教材

课程评估 ASSESSMENT

19. 评估形式 Type of Assessment	评估时间 Time	占考试总成绩百分比 % of final score	违纪处罚 Penalty	备注 Notes
出勤 Attendance		20		
课堂表现 Class				

Performance				
小测验 Quiz				
课程项目 Projects		40		中期汇报 (Presentation)
平时作业 Assignments				
期中考试 Mid-Term Test				
期末考试 Final Exam				
期末报告 Final Presentation		40		
其它 (可根据需要 改写以上评估方 式) Others (The above may be modified as necessary)				

20. 记分方式 GRADING SYSTEM

- A. 十三级等级制 Letter Grading
 B. 二级记分制 (通过/不通过) Pass/Fail Grading

课程审批 REVIEW AND APPROVAL

21. 本课程设置已经过以下责任人/委员会审议通过
This Course has been approved by the following person or committee of authority

田国梁