

Capstone Project in Data Science

Tsung-Chih Lai

Fall 2025

Time: Monday 10:15 – 13:00

Office Hours: Wednesday 14:00 – 16:00

Classroom: College of Management 322

Office: College of Management 326

Website: <https://ecourse2.ccu.edu.tw/>

Email: tclai@ccu.edu.tw

Course Description

This course offers an introduction to the fundamentals of data science using the R programming language, with a focus on the tidyverse and tidymodels suites of packages. Through hands-on exercises and project-based work, students will develop core skills in data manipulation, visualization, modeling, and interpretation.

Learning Objectives

- Understand the core concepts and applications of data science.
- Use tidyverse tools for effective data wrangling and visualization.
- Build, tune, and evaluate predictive models using tidymodels.
- Interpret and communicate findings from data analyses and models.

Prerequisites

None. Prior exposure to basic statistics and programming is helpful but not required.

Textbooks

- Wickham, H., Çetinkaya-Rundel, M., and Grolemund, G. (2023). *R for Data Science*, 2nd Edition. <https://r4ds.hadley.nz/>
- Kuhn, M. and Silge, J. (2023). *Tidy Modeling with R*. <https://www.tmwr.org/>

Assessment

- In-class exercises (20%)
- Homework assignments (20%)
- Midterm proposal (20%)
- Final presentation (30%)
- Group meetings and participation (10%)

Requirements

- This is a computer-based course; students must bring a laptop to class.
- Course materials will be posted on the class website. Please check regularly.
- Absences from presentations or group meetings without valid justification will result in a zero grade for the missed component.

Course Outline

- Introduction to Data Science and R
- Data Visualization
- Data Transformation
- Data Tidying
- Exploratory Data Analysis
- Midterm Proposal Presentation (11/3)
- Modeling Basics
- Tools for Creating Effective Models
- Final Project Presentation (12/22)