CMU ONLINE
Carnegie Mellon University

Syllabus Overview Sample: Applications of Real-World Data Science: A Capstone Experience

Program: Applied Data Science

Faculty: Dr. Weijing Tang, Assistant Professor, Department of Statistics & Data Science

Course Number: 36-644 **Units:** 6

Zoom Class: Weekday Evening, ET

Course Description

This is your final course in the Foundations of Data Science Online Graduate Certificate program. In this course, you will apply the skills and knowledge you've acquired throughout the program to real-world data when carrying out a seven-week capstone project.

This course will train you how to create, and to communicate, impactful data science results, thereby preparing you to be an effective data scientist in the context of real-world applications.

While working on your project, you will attend weekly lectures to get training in data science communication, as well as statistical guidance. A unique opportunity of the capstone course is that you will gain experience working on real data science problems; however, a unique challenge is that there will not necessarily be a single "correct" means to solve these problems… so you will need to determine those data science methods that would be the most appropriate for your project and goals. Furthermore, after implementing data science methods, you will need to communicate your results in a way that broad audiences can understand. The weekly lectures will involve regular workshopping to help you implement methods and present results, such that you will be able to make real-world impacts with your data science work after completing the certificate program.

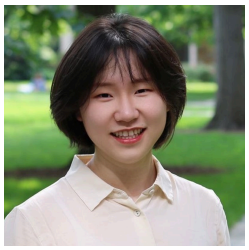
Course Learning Objectives

By the end of this course, students will be able to:

- Formulate research questions that can be addressed with data-driven inquiry, and identify how information in a given dataset is related to those research questions.
- Translate research questions into statistical tasks, including (but not limited to) data preprocessing, feature engineering, visualization, and statistical modeling and inference.

- Pinpoint notable characteristics about the data via exploratory data analysis, including appropriate and well-designed visualizations.
- Implement statistical methods to produce accurate and relevant results that address a project's research questions.
- Communicate and document methods, results, and conclusions in a way that is accurate, compelling, and useful for statistical and non-statistical audiences.
- Compose and iteratively update a project's goals and timeline, in order to do the above while abiding by strict deadlines.

Faculty Introduction



Dr. Weijing Tang is an Assistant Professor in the Department of Statistics and Data Science at Carnegie Mellon University. She earned her bachelor's degree in mathematics from Tsinghua University and her Ph.D. in statistics from the University of Michigan. Before joining Carnegie Mellon, she was a postdoctoral research fellow in biostatistics at Harvard University. Her research focuses on developing statistical methods and theory for data valuation in data-centric AI, multi-source statistical learning, network analysis, and survival analysis, with applications in the health and social sciences. In this course, she aims to help students integrate and apply the skills they have developed throughout the program to a real-world capstone project and communicate their methods, findings, and conclusions clearly to diverse audiences.

Course Topics

Topic	Learning Objectives
Workflow of Capstone Projects, and Potential Projects	• Describe the main stages of statistical consulting and solving a data science problem, from start to finish. • Determine a project to potentially work on for the capstone class.
Initial Project Details and EDA	• Describe the four steps to problem-solving in data science. • Explain characteristics/behaviors that differentiate an expert from a novice. • Present a data science project overview at a high level
Data Science Review and Further Workshopping	• Give an effective overview of what EDA they've done so far.

	● Give effective and constructive feedback to others' presentations.
Advanced Workshopping	● Workshop specific tasks for your project. ● Give effective and constructive feedback to others' presentations.
Practice Presentations and Draft Executive Summary	● Prepare a 10 minute practice presentation for your project. ● Write a draft one-page executive summary for your project.
Practice Presentations and Draft Executive Summary Cont.	● Prepare a 10 minute practice presentation for your project. ● Write a draft one-page executive summary for your project.
Final Presentations and Executive Summary	● Prepare a 10 minute presentation for your project. ● Write a final one-page executive summary for your project.

NOTE: This is a sample designed to provide an overview of topics covered in the course based on previous course syllabi. All students will be provided an updated and final syllabus at the start of the course. Faculty, content, and schedule are all subject to change.