

# ChBE 4745/6745: Data Analytics for Chemical Engineers

## Official Course Syllabus – Fall 2026

**Instructor:** A.J. Medford

**Format:** Fully Online

---

### Course Description

The primary objective of this course is to introduce data science and machine learning techniques for analyzing, interpreting, and visualizing datasets commonly encountered in chemical and biomolecular engineering. Students will learn practical tools needed to understand which data analysis approaches are appropriate for different datasets, how to apply them, and correctly interpret their results. Topics covered will include: regression, classification, cross-validation, data storage and retrieval, dimensionality reduction, clustering, and time series analysis. The course will be taught in the Python programming language and will include a short introduction for students familiar with MATLAB or other languages. In addition to lectures, the course will feature homeworks that will step students through the application of the data analysis techniques introduced in the course to datasets common in chemical engineering including spectroscopy data, materials property predictions, and chemical process data.

### Course Objectives / Learning Outcomes

Once completed, the students should have the following capabilities:

- Utilize Python libraries to analyze, visualize, and organize data
- Construct analysis-ready data sets from spreadsheets and databases
- Evaluate the accuracy of analytics models using quantitative metrics
- Select an appropriate model and parameters using cross-validation
- Identify a supervised learning model with optimum complexity
- Apply unsupervised and supervised techniques to improve model performance through feature engineering
- Inspect high-dimensional data in low-dimensional space using visualization and quantitative metrics
- Apply data science techniques to real engineering data sets

## Required Course Materials

### Class Text:

- “Machine Learning Refined” – Watt, Borhani, & Katsaggelos (2016). ISBN: 978-1107123526. Partially available on the [Machine Learning Refined GitHub repository](#).

### Supplementary Texts:

- “Python for Computations in Science and Engineering” – Kitchin (2022+). Available on the [Point Breeze Publishing website](#).
- “Scientific Computing with Python” – Johansson (2015). Available on the [Scientific Python Lectures GitHub repository](#).
- “The Elements of Statistical Learning” – Hastie, Tibshirani, & Friedman (2008). ISBN: 978-0387848570. Available as a [free PDF from Stanford](#).
- “Numerical Linear Algebra” – Trefethen & Bau (1997). ISBN: 978-0898713619. Partially available on the [author’s website](#).

## Grading Policy

Component	6745	4745
Knowledge Checks	5%	10%
Group Evaluations	5%	10%
Homeworks	25%	20%
Self-written Problems	30%	30%
Final Project	35%	30%

Note: Homeworks, self-written problems, and project will have different questions/requirements for 4745/6745.

## Attendance Policy

- This is a fully online course.
- Login on a regular basis to complete your work, so that you do not have to spend a lot of time reviewing and refreshing yourself regarding the content.
- You will have a group that you are expected to interact with on a weekly basis. The core of the “Group Evaluation” portion of your grade will be related to your attendance of group meetings and participation.

## Additional Criteria for Successful Completion

Letter grade cutoffs will be assigned based on the final grade distribution. The following guidelines will be used to assign cutoffs:

- **A:** Technical mastery of most topics in the course and strong conceptual understanding of all topics. Full participation in course assignments and projects. Maximum cutoff: 90%.
- **B:** Conceptual understanding of most topics in the course, but significant issues with technical details of implementations. Full participation in course assignments and project. Maximum cutoff: 80%.
- **C:** Significant issues with conceptual understanding and technical implementation of many topics in the course. Partial participation in course assignments and project. Maximum cutoff: 70%.
- **D:** Minimal evidence of conceptual understanding or technical skills covered in the course. Partial participation in course assignments and project. Maximum cutoff: 60%.
- **F:** No evidence of conceptual understanding or technical skills covered in the course. Little or no participation in course assignments and project. Maximum cutoff: 50%.

Cutoffs may be reduced from the maximum, and adjustments may occur separately for each section, or a single cutoff may be utilized.

## Academic Honesty / Academic Integrity Statement

All students should abide by the Georgia Tech Student Honor Code.

- Review the [Georgia Tech Student Honor Code](#).
- Any student suspected of behavior in violation of the Georgia Tech Honor Code will be referred to Georgia Tech's Office of Student Integrity.

## Student-Faculty Expectations Agreement

Students and instructors are expected to abide by the Georgia Tech Student-Faculty Expectations Agreement. This agreement outlines the expectations for both students and instructors with regard to academic integrity, communication, and mutual respect. More information is available in the [Georgia Tech Catalog, Rule 22](#).

## Disability Services

Georgia Tech offers accommodations to students with disabilities. If you need a classroom accommodation, please make an appointment with the [ADAPTS office](#).

## Core IMPACTS Statement

This course is not currently designated as a Core IMPACTS course.