

CHM 6165
INTRODUCTION TO CHEMOMETRICS
Fall Semester 2020, 3 Credits

Instructor: John Bowden, Basic Science Building 324, 352-294-4063, john.bowden@ufl.edu

Lectures: M W F, 3:00PM-3:50PM (Period 8), Zoom Synchronous

Office hours: After Zoom lecture period or by appointment (via Zoom).

Course objectives: To provide students with an introductory foundation in statistical and data management methods/tools often used by analytical chemists, using the programming language R as the teaching platform. Students will learn how to select the appropriate method, handle data, and effectively disseminate results.

Textbooks: Txt1 – “R in Action,” 2nd Edition, Robert I. Kabacoff; \$59.99 (list), \$52.52 (Amazon, new), \$41.56 (Amazon, used) AND Txt2 – “Practical Statistics for the Analytical Scientist”, 2nd Edition, Stephen Ellison, Vicki Barwick, Trevor Farrant, \$69.00, \$38.84 (Amazon new) or \$34.84 (Amazon used). Both books are not required, but strongly suggested and may serve as useful references.

Grading: Grades will be based out of 300 points: 95 points total for class presentations and accompanying reports, 30 points for completion of homework assignments, 75 points for mid-term exam and 100 points for final exam.

Grade scale: Course grades will be assigned based on the following grading scheme: A (100 – 94), A- (93.99-90), B+ (89.99 – 87), B (86.99 – 84), B- (83.99 – 80), C+ (79.99 – 77), C (76.99 – 74), C- (73.99 – 70), D+ (69.99 – 76), D (66.99 – 64), D- (63.99 – 61), E (60.99 – 0).

Problem sets: Homework assignments, to supplement the Zoom lectures, will be made available throughout the semester.

Class Presentations: In an effort to provide an interactive element, there will be four class presentations throughout the semester. **Selected topics for the presentations should be made in consultation with Dr. Bowden (deadlines will be set throughout the semester).**

Presentation #1 (9/18/20, 10 points) – students will select a case study on data ethics and will give a 5 min presentation to the class. The exercise will be graded based on the ability to communicate and present the material, with special attention to framing key concepts in data ethics.

Presentation #2 (10/9/20, 10 points) – student will select a publication using a chemometric method/approach and will present the article to the class (5 min). The exercise will be graded based on the ability to communicate and critique the material, while highlighting the advancement and potential applications.

Presentation #3 (11/9/20, 25 points) – students will select an R function/code (not covered in class) and will give a 10 presentation to the class. Accompanying the presentation, the student will also need to submit a 1-page instructional report on how to use the code that will be distributed to the class for handy reference.

Presentation #4 (12/7-12/9/20, 50 points) – students will select an open-source software program (a list of possible programs will be provided, for example, MixOmics) and will give a 10-minute presentation to the class, along with a 1-page cheat sheet for how to use the program.

Mid-term: The mid-term exam will be take-home and will be handed out **Friday, October 9th and will be due Wednesday, October 14th at the beginning of class.**

Final exam: The cumulative final exam will be take-home and will be handed out **Friday, December 7th and will be due Friday, December 11th at 1 pm.**

Information on current UF grading policies is online: (<https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx>).

Course policies:

Attendance will not be recorded, but participation in lectures and demonstration periods is important for assimilating the course material. Any request for make-up exams should be made to Dr. Bowden as far in advance as possible.

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, www.dso.ufl.edu/drc/) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

UF students are bound by The Honor Pledge which states: "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (<http://www.dso.ufl.edu/sccr/process/student-conduct-honorcode/>) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class."

Students are expected to provide feedback on the quality of instruction in this course by completing online evaluations at <https://evaluations.ufl.edu>. Evaluations are typically open during the last two or three weeks of the semester, but students will be given specific times when they are open. Summary results of these assessments are available to students at <https://evaluations.ufl.edu/results/>.

For counseling the following resources are available to students: **U Matter, We Care:** If you or a friend is in distress, please contact umatter@ufl.edu or 352-392- 1575. **Counseling and Wellness Center:** <http://www.counseling.ufl.edu/cwc/Default.aspx>, 352-392-1575; the University Police Department: 352-392-1111 or 911 for emergencies. **Sexual Assault Recovery Services (SARS):** Student Health Care Center, 352-392-1161.

COVID-19 Changes: CHM 6165 is typically an in-person course. Due to the COVID-19 pandemic, we will instead meet online via Zoom. Lectures will occur synchronously to encourage student participation (the small class size of a graduate course is amenable to in-class discussion). However, our class sessions may also be audio-visually recorded for students in the class to refer back and for enrolled students who are unable to attend live. These recordings will be uploaded to Canvas for asynchronous learning. If you experience technical issues with connecting, you should contact the UF Help Desk (<https://helpdesk.ufl.edu>, 352-392-4357).

Students who participate with their camera engaged or utilize a profile image are agreeing to have their video or image recorded. If you are unwilling to consent to have your profile or video image recorded, be sure to keep your camera off and do not use a profile image. Likewise, students who un-mute during class and participate verbally are agreeing to have their voices recorded.

If you are not willing to consent to have your voice recorded during class, you will need to keep your mute button activated and communicate exclusively using the "chat" feature, which allows students to type questions and comments live. The chat will not be recorded or shared. As in all courses, unauthorized recording and unauthorized sharing of recorded materials is prohibited.

In order to ensure student privacy, communication concerning grades will only be conducted through Canvas or online via a Zoom meeting.

| Date | Lecture | Topic | Txt 1 | Txt 2 |
|------------|--------------|-------|---|----------------|
| M | 8/31 | 1 | Overview and Introduction to Chemometrics | |
| W | 9/2 | 2 | Introduction to R/Functions | CH 1-5 |
| F | 9/4 | 3 | Tidyverse and Data Manipulation | CH 1-5 |
| M | 9/7 | | Labor Day (no class) | |
| W | 9/9 | 4 | ggplot2 for Basic Graphs | CH 3,6 CH 2 |
| F | 9/11 | 5 | ggplot2 for Basic Graphs | CH 6 CH 2 |
| M | 9/14 | 6 | ggplot2 for Intermediate Graphs | CH 11 |
| W | 9/16 | 7 | Basic Statistics | CH 7 CH 1, 4 |
| F | 9/18 | | Presentation #1: Research Ethics | |
| M | 9/21 | 8 | Distribution and Normality Tests | CH 3, 4 |
| W | 9/23 | 9 | Analysis of Variance | CH 9 CH 6 |
| F | 9/25 | 10 | Analysis of Variance | CH 9 CH 6 |
| M | 9/28 | 11 | Regression | CH 8 Ch 7 |
| W | 9/30 | 12 | Regression | CH 8 CH 7 |
| F | 10/2 | | Catch up day | |
| M | 10/5 | 13 | Experimental Design/Sampling Strategies/Power | CH 10 CH 8, 13 |
| W | 10/7 | 14 | Outliers, Filtering, Missing Data | CH 18 CH 5 |
| F | 10/9 | | Presentation #2: Publication from Journal | |
| M | 10/12 | 15 | Calibration | |
| W | 10/14 | 16 | Validation/Method Performance (Exam 1) | CH 9 |
| F | 10/16 | 17 | Non-Detects | |
| M | 10/19 | 18 | Signal Processing | |
| W | 10/21 | 19 | Multivariate Methods: Principal Component Analysis | CH 14 |
| F | 10/23 | 20 | Multivariate Methods: Principal Component Analysis | CH 14 |
| M | 10/26 | 21 | Multivariate Methods: PLS-DA and Other Methods | CH 14 |
| W | 10/28 | 22 | MetaboAnalyst | |
| F | 10/30 | 23 | Categorical Distribution and Models | |
| M | 11/2 | 24 | Big Data Handling (Guest, Jeremy Koelmel, Yale) | |
| W | 11/4 | 25 | Presentation #3: Introduce R script to Class | |
| F | 11/6 | 26 | Presentation #3: Introduce R script to Class | |
| M | 11/9 | | ggplot2 for Advanced Graphs | CH 19 |
| W | 11/11 | | Veteran's Day (no class) | |
| F | 11/13 | 27 | Machine Learning Intro (Guest, Alina Zare, UF) | |
| M | 11/16 | 28 | Machine Learning App (Brian Stucky, UF) | |
| W | 11/18 | 29 | Machine Learning App (Brian Stucky, UF) | |
| F | 11/20 | 30 | Time Series | CH 15 |
| M | 11/23 | 31 | Chemical Metrology (Katrice Lippa, NIST) | |
| W | 11/25 | | Thanksgiving (no class) | |
| F | 11/27 | | Thanksgiving (no class) | |
| M | 11/30 | 32 | Assessing Quality of Analytical Measurements | CH 11, 12 |
| W | 12/2 | 33 | QA/QC in the Analytical Laboratory | |
| F | 12/4 | 34 | Data Management Plans | |
| M | 12/7 | | Presentation #4: Introduce Software to Class | |
| W | 12/9 | | Presentation #4: Introduce Software to Class | |
| Final Exam | | | | |