CIS 325 Data Analysis and Visualization
Spring 2022

Textbook: Not Required

Instructor: Dr. Supraja Gurajala
Phone: 315-267-2091
Office: Dunn Hall 307
Email: gurajas@potsdam.edu

Office Hours: MWF - 9:00 am to 10:00 am
Tu Th - 9:30 am to 12:00 noon

Class Time/Place: Tu Th – 12:30pm to 1:45pm, Dunn Hall 206

Prerequisite: MATH 125 or CIS 125 or STAT 100 and upper division status

Final Exam: Wednesday, May 18, 8:00 – 10:00 a.m.

Course Description:

CIS 325 prepares students to manipulate, apply, and interpret probability and statistics in the context of complex, real-world data. Students will practice generating, visualizing, interpreting, and reporting analyses of large, complex data sets using an appropriate computational toolkit. Students' use of a toolkit assumes no prior programming experience but requires a level of understanding of probability and statistics. Examples will emphasize artificial intelligence, data science, and machine learning using large data sets.

Objectives:

All graduating students with a major in Computer Science shall demonstrate 10 specific learning outcomes. Learning outcomes that are applicable to CIS 325 are shown in bold, below. Students in the Computer Science program at SUNY Potsdam are expected to graduate with a:

1. **Knowledge of discrete and continuous mathematics --- including elementary probability and statistics – and the ability to apply logic and mathematical proof techniques to computing problems.**
   Students will learn the mathematical foundations of probability and statistics and apply this knowledge in practice on big data.
2. Knowledge of basic theory of computability and complexity of computation.
   Not emphasized in CIS 325.
3. **Knowledge of and the ability to apply programming fundamentals in at least two programming languages.**
   Not emphasized in CIS 325
4. **Knowledge of fundamental data structures and algorithms --- including analysis of their correctness and complexity --- related to various fields of computer science, and**

the ability to apply this knowledge to problems through the use of appropriate programming languages.
Students will apply knowledge of data structures in programming assignments and the project. Efficient handling of large data sets is emphasized.
5. Knowledge of computer architecture and organization, computer operating systems, and computer networks, and the ability to apply this knowledge to problems through the use of appropriate programming languages.
Not emphasized in CIS 325..
6. **Competence and effectiveness in technical oral, written, and visual communication, particularly as they apply to the dissemination of technical information on subjects dealing with computing technology and applications.**
Students will present the results of data analysis using various visualization techniques.
7. Knowledge of and skill in applying good practices in software engineering.
Not emphasized in CIS 325.
8. **The ability to function effectively in teams to accomplish a common goal.**
Students will work in teams to design and implement the course project.
9. An understanding of professional, ethical, legal, security, and social responsibilities and issues, including an awareness of impact of computing on individuals, organizations, and society.
Not emphasized in CIS 325.
10. **A commitment to continuing professional development.**
Students will extend their knowledge of data analysis and visualization, methods, and tools in the course project by using methods/tools that are not covered in other assignments.

Caring Community:

I recognize that this is an incredibly stressful time for you, your peers, and our community. Please know that there are resources available to you, both on and off campus, to support you during these very uncertain times. Our excellent Counseling Center staff are available to meet with you; more information can be found on their FAQ page accessed at: https://www.potsdam.edu/studentlife/wellness/counseling-center/coping-covid-19-pandemic/counseling-center-faqs. In addition, information on a variety of on- and off-campus resources can be found our Bear Care site: https://www.potsdam.edu/studentlife/wellness/bear-care. You are an incredibly important member of our Potsdam community; please take care of yourself, and each other.

Grading for the Course:

1. ***Weekly Quizzes:*** 10 %

   A ten-minute weekly quiz will be given once a week. It can be on any class day. It will be based on lectures and Homework problems assigned for you. There is no make-up quiz.

2. ***Homeworks and programming Assignments:*** 30%

   Several homeworks and small programming assignments will be given based on the concepts discussed in lectures. These homeworks will be the essential part of the course. HWs will be posted on moodle page along with the due date. Late work is penalized at

20% per calendar day that they are late. No late work is accepted beyond the cutoff date. Your final submitted HW should represent your individual work; it is, however, acceptable to discuss the solution approach with other students. You will be responsible for keeping track of due dates posted on moodle.

3. ***Exams:*** 40%

    a. Midterm 1 – 13 %
    b. Midterm 2 – 13 %
    c. Final Exam – 14 %

    Exams will be closed book and closed notes unless specified otherwise. Any request for re-grading must be received in writing and within 3 days of receiving your graded exam back. Prior notice must be given to your instructor. No make-ups will be granted unless satisfactory documentation is produced to show an extenuating circumstance.

4. ***Project:*** 20%

    Team or individual project (depending on the number of students in the class) that gives students an opportunity to explore data analysis and visualization tools in applications to derive intelligence from a particular real-world dataset.

At the end of the semester I will calculate what fraction of the possible points you have earned, and your grade may be based on this distribution:

90% >=	A
80% - 90	B
70% - 80	C
60% - 70	D
< 60%	F

Note that final grades are determined using a class curve of the course-grade averages.

Due Dates

All due dates for the course will be strictly enforced. Prior approval will be required from the instructor for any late submission.

Tentative Schedule:

Week	Topics
1	Course Introduction Raw and Processed data Downloading and Reading Local Files / Excel / XML / JSON
2	Reading from MySQL / HDF5 / The Web / APIs Reading From Other Sources
3	Subsetting and Sorting

	Summarizing Data Reshaping Data Managing Data Frames with Pandas (pythons data analysis and manipulation tool) - intro Merging Data
4	Editing Text Variables Regular Expressions Regular Expressions II Working with Dates Data Resources
5	Principles of Analytic Graphics Exploratory Graphs Plotting Systems in Python Graphics Devices in Python/R
6	Lattice and ggplot2 Plotting System Midterm 1
7	The Human Memory Reasoning The Human Retina Perceiving Two Dimensions Perceiving Perspective
8	Dimension Reduction Working with Color in Python
9	2-D Graphics 2-D Drawing 3-D Graphics
10	Mapping Glyphs Parallel Coordinates Stacked Graphs
11	Projects proposal Midterm 2
12	Tufte's Design Rules Using Color Graphs and Networks
13	Embedding Planar Graphs Graph Visualization Tree Maps
14	Principal Component Analysis Multidimensional Scaling
15	Project Presentations

***Hardware**:* You will need a computer to do work on labs and programming assignments. You can use your personal computer or any machine in Dunn hall lab 302.

***Software**:* Here is a summary of the various software you may need for the course, in addition to the basics of a computer, browser, and typical software.

- **R**
- **Python**
- **Zoom** – In case we move to online. You can get a free Zoom account here https://zoom.us/.

Impact of extracurricular activities on class work

You make the choices about how you will spend your time, including investing your time in non-academic activities. As a student, you need to give priority to your academic work, and prevent extracurricular commitments from negatively impacting your work for classes. You are, of course, free to participate in activities that are meaningful to you; however, do not expect me to give special consideration because of time management issues that arise from those activities. You should not be missing class because of extracurricular activities, nor should you allow yourself to fall behind on assignments. **NOTE: I will not give extensions that relate to participation in extracurricular activities, even if the activity is related to Computer Science.**

Expectations for the Course

- You will be expected to come prepared to class and be an active participant in class discussions. You should plan on spending a significant time outside class in reviewing course material covered in class. It is critical that you keep up with the course material on a timely basis.
- Academic dishonesty: Students are expected follow the "SUNY Potsdam Academic Honor Code" (SUNY Potsdam 2014-2016 Undergraduate Catalog, p. 42) by doing their own work on quizzes, exams and programming assignments unless specifically directed otherwise by the instructor. Copying is strictly forbidden. Students caught cheating will receive a grade of 0 for that evaluation. Repeated offenses will result in dismissal from the course and possible disciplinary sanctions by the university. Academic Misconduct definitions, procedures, due process, and student rights are described on page 43 of the SUNY Potsdam 2014-2016 Undergraduate Catalog.
- Disability Assistance: Anyone who has special needs that must be accommodated to fulfill the course requirements should notify the instructor and the Director of Accommodative Services, 111 Sisson Hall, 267-3267. The college has resources available to assist qualified students with their academic studies.
- Food and Drink in Class and Lab**:** Beverages are allowed in the classroom as long you clean up after yourself and do not disturb others. In the Unix lab, food and drink are restricted to the coffee table. UNDER -NO- CIRCUMSTANCES ARE FOOD AND BEVERAGES (EVEN GUM) ALLOWED NEAR THE COMPUTERS.
- No devices are allowed during class. Notes must be hand-written
- Accommodation of Religious Observances: We will make reasonable accommodation for a student's religious beliefs. Please notify us within the first week of classes about any scheduled class date that conflicts with a religious observance.

Attendance

Regular attendance is critical for your success in this course. You are responsible for updating yourself with announcements made in class concerning material covered, home works, and any changes in course syllabus, due dates, or other course-related issues.

SUNY Potsdam Department of Computer Science Code of Professional Conduct

1. ***Preamble***
   All members of the ACM, including the Computer Science faculty of SUNY Potsdam, are committed to ethical professional conduct as specified in the ACM Code of Ethics and Professional Conduct. Students, taking courses from the faculty, are bound by our commitment.

   All members of the Department are obliged to remind one another to behave professionally. Violations should be reported promptly; however, capricious or malicious reporting of violations is, itself, a violation. When reporting, bring all relevant aspects of the incident to the faculty's attention.
2. ***Moral Imperatives***
   As a Computer Science student I will...

   2.1. Respect all members of the Department.

      2.1.1. Be professional in face-to-face and electronic interactions.

      2.1.2. Be fair so everyone is free to work and learn.

      2.1.3. Be active in preventing discrimination in physical and electronic spaces frequented by Department members.

   2.2. Accept and provide appropriate feedback.

      2.2.1. Avoid starting or spreading rumors.

      2.2.2. Respect confidentiality.

   2.3. Be honest, trustworthy, and respect intellectual property.

      2.3.1. Only take credit for my own work.

      2.3.2. Respect the privacy of others.

      2.3.3. Access computing resources only when authorized and report any access risks discovered.

   2.4. Contribute to society and human well-being.

      2.4.1. Improve public understanding of computing and its consequences.

      2.4.2. Consider both the direct and indirect impacts of my actions.

Based on the ACM Code of Ethics and Professional Conduct, retrieved from https://www.acm.org/about-acm/acm-code-of-ethics-and-professional-conduct 11 August 2017

Student Support

Every student in this class is a valued individual. If you are struggling with issues outside of the classroom, please know that there are professionals both on and off campus who can assist you. If you need immediate assistance, please contact our campus Counseling Center (with free counseling) at (315) 267-2330 or visit their website. Links to other resources are provided below:

Rachel Bayliss- Title IX Support Staff & Title IX Core Team
Draime Extension S184, (315) 267-2350
VanHousen Extension, Rm. 392, (315) 267-2516
http://www.potsdam.edu/offices/hr/titleix

Bias Incident Reporting-
http://www.potsdam.edu/about/diversity/biasincident

Center for Diversity
223 Sisson Hall
(315) 267-2184
http://www.potsdam.edu/studentlife/diversity

University Police
Van Housen Extension
(315) 267-2222 (number for non-emergencies; for an emergency please dial 911)

Student Conduct and Community Standards
208 Barrington Student Union
http://www.potsdam.edu/studentlife/studentconduct/codeofconduct
Reachout (24-hour crisis hotline) ▪ (315) 265-2422

Renewal House (for victims of domestic violence)
SUNY Potsdam Campus Office: Van Housen Extension 390 (open Wednesdays, 9-5:00)
(315) 379-9845 (24-hour crisis hotline)
Renewalhouse_campus@Verizon.net

And please: if you see something, say something. If you see that someone that you care about is struggling, please encourage them to seek help. If they are unwilling to do so, Care Enough to Call has guidelines on whom to contact. Everyone has the responsibility of creating a college climate of compassion.