Prerequisite/Co-requisite: None

Audience: Any graduate student

Description:

The course introduces students to applied examples of data collection, processing, transformation, management, and analysis to provide students with a hands-on introduction to the data science experience. Students will explore key concepts related to data science, including applied statistics, information visualization, text mining, and machine learning. “R,” the open-source statistical analysis and visualization system, will be used throughout the course. R is reckoned by many to be the most popular choice among data analysts worldwide; having knowledge and skill with using it is considered a valuable and marketable job skill for most data scientists.

Credits: 3

Learning Objectives:

After taking this course, the students will be able to understand:

- Essential concepts and characteristics of data
- Scripting/code development for data management using R and R-Studio
- Principles and practices in data screening, cleaning, and linking
- Communication of results to decision makers

At the end of the course, students are expected to be able to:

- Identify a problem and the data needed for addressing the problem
- Perform basic computational scripting using R and other optional tools
- Transform data through processing, linking, aggregation, summarization, and searching
- Organize and manage data at various stages of a project life cycle
- Determine appropriate techniques for analyzing data

Bibliography/Texts/Supplies—Required:

Saltz, J. S., & Stanton, J. M. (2021). *Data science for business with R*. Sage Publications.

The book is available from Amazon in paperback and electronic version at: https://www.amazon.com/Data-Science-Business-Jeffrey-Saltz/dp/1544370458.

Bibliography/Texts/Supplies—Additional:

I will provide additional and supplemental readings in the digital learning system as electronic documents for downloading and printing. Students are expected to read the assigned materials for discussions and coursework.

Requirements:

The work for this class will involve the following:

- Homework (30%) is designed for you to practice the necessary skills in carrying out data processing, analysis, and management tasks.
- Participation (12%) includes attendance and participation in class.
- Homework Explainer Videos (5%) help you practice explaining data science coding concepts to others.
- The exam (25%) is designed to evaluate your mastery of concepts, methods, and tools in data analysis and management.
- Final project (28%): For the final project you work on a dataset provided, transform the data as needed, and provide a written analysis with visualizations (group of three to four students). Students will be assigned into a group. The grade is comprised of 25% final submission and 1% for each of the three project updates submitted.

Grading:

- Grade levels follow the scales below:

Highest	Lowest	Letter	Expectation
100.00%	93.00%	A	Your work is outstanding.
92.99%	90.00%	A−	
89.99%	87.00%	B+	Your work is about what would be expected of a serious student.
86.99%	83.00%	B	
82.99%	80.00%	B−	
79.99%	77.00%	C+	Your work falls below what is expected but is adequate.
76.99%	73.00%	C	
72.99%	70.00%	C−	
69.99%	0.00%	F	Your work is out of the picture.

- Each assigned work will be graded on the scale as specified for the component, which will be summed at the end of the semester.
- It is unethical to allow some students additional opportunities, such as extra credit assignments, without allowing the same options to all students.
- Students who wish to dispute a grade may resubmit the assignment for regrading with a one-page statement of explanation of why the paper should be regraded. If the student resubmits, the assignment will be regraded, which means the grade may go up, down, or stay the same. Except for extraordinary circumstances, no appeal for an individual

assignment or project will be considered later than 2 weeks after the assignment was graded.

General Requirement for Assignment Submissions
Assignments must be professionally prepared with computer applications. Unless otherwise stated, R source files need to be submitted to the digital learning system for each homework assignment (not Word documents or PDFs). Your project will be a presentation and a written report. If you have additional supporting materials, you must embed them into your MS Word or PDF document.

For each assignment, be sure to include the following information within the R source code (as a comment):

- Course number (i.e., IST 687)
- Name(s) of the author(s) or the team members
- Name of the assignment (e.g., "Homework 2")
- Date the assignment is due and date in which it is submitted

Course-Specific Policies on attendance, late work, makeup work, examinations if outside normal class time, etc.:

Homework assignments are due prior to the start of the next synchronous class. Homework assignments must be submitted by 11:59 pm ET (i.e., before midight Syracuse time) on the third night after the synchronous class. Late homework assignments will receive a 0 (no credit). If your final project is late, **I will deduct 10% of the original grade for the first day of lateness plus 15% for every subsequent day.** You will not receive full credit for topics/assignments presented in an unprofessional manner. Professionalism includes the proper use of grammar, punctuation, and limiting spelling mistakes. Professionalism also means adhering to given instructions. Failure to adhere to the assignment instructions will result in a reduction of the grade. If English is not your first language, set up an appointment with the writing program so they can help you improve your writing.

Academic Integrity Policy

Syracuse University's academic integrity policy reflects the high value that we, as a university community, place on honesty in academic work. The pilot policy in effect at the School of Information Studies defines our expectations for academic honesty and holds students accountable for the integrity of all work they submit. Students should understand that it is their responsibility to learn about course-specific expectations, as well as about university-wide academic integrity expectations. The pilot policy governs appropriate citation and use of sources, the integrity of work submitted in exams and assignments, and the veracity of signatures on attendance sheets and other verification of participation in class activities. The pilot policy also prohibits students from submitting the same work in more than one class without receiving written authorization in advance from both instructors. Under the pilot policy, students found in violation are subject to grade sanctions determined by the course instructor and non-grade sanctions determined by the School or College where the course is offered. SU students are required to read an online summary of the university's academic integrity expectations and provide an electronic signature agreeing to abide by them twice a year during pre-term check-in on MySlice. For more information and the pilot policy, see http://academicintegrity.syr.edu.

Disability-Related Accommodations

Syracuse University values diversity and inclusion; we are committed to a climate of mutual respect and full participation. If you believe that you need accommodations for a disability, please contact the Office of Disability Services (ODS), disabilityservices.syr.edu, located at 804 University Avenue, room 309, or call 315.443.4498 for an appointment to discuss your needs and the process for requesting accommodations. ODS is responsible for coordinating disability-related accommodations and will issue "Accommodation Authorization Letters" to students as appropriate. Since accommodations may require early planning and generally are not provided retroactively, please contact ODS as soon as possible. Our goal at the iSchool is to create learning environments that are useable, equitable, inclusive and welcoming. If there are aspects of the instruction or design of this course that result in barriers to your inclusion or accurate assessment or achievement, please meet with me to discuss additional strategies beyond official accommodations that may be helpful to your success.

Religious Observances Notification and Policy

SU's religious observances policy, found at https://policies.syr.edu/policies/university-governance-ethics-integrity-and-legal-compliance/religious-observances-policy, recognizes the diversity of faiths represented in the campus community and protects the rights of students, faculty, and staff to observe religious holy days according to their tradition. Under the policy, students should have an opportunity to make up any examination, study, or work requirements that may be missed due to a religious observance provided they notify their instructors no later than the end of the second week of classes through an online notification form in MySlice listed under **Student Services/Enrollment/My Religious Observances/Add a Notification**.

Student Academic Work Policy

Student work prepared for University courses in any media may be used for educational purposes, if the course syllabus makes clear that such use may occur. You grant permission to have your work used in this manner by registering for, and by continuing to be enrolled in, courses where such use of student work is announced in the course syllabus.

I intend to use academic work that you complete this semester for educational purposes in this course during this semester. Your registration and continued enrollment constitute your permission.

I intend to use academic work that you complete this semester in subsequent semesters for educational purposes. Before using your work for that purpose, I will either get your written permission or render the work anonymous by removing all your personal identification.

Course Evaluations

There will be an end of course evaluation for you to complete this term, described below. This evaluation will be conducted online and is entirely anonymous. You will receive a notification from the Syracuse University Office of Institutional Research & Assessment (OIRA) department in your email account with the evaluation website link and your passcode.

- End of semester evaluation will be available for completion in Week 10 prior to your final exams week. This evaluation is slightly longer and it is used to gauge the instructor performance and make adjustments to the course to ensure it meets our student needs.

We faculty work hard to do the best possible job when preparing and delivering courses for our students. Please understand that not only does the school use the course evaluations to make decisions about the curriculum in order to improve where necessary, but they also use them to

make decisions about faculty members. Please take the time and fill out this evaluation as your feedback and support of this assessment effort is very much appreciated.

School Library Media Program Assessment

The School Library Media Program is accredited by the Council for the Accreditation of Educator Preparation (CAEP) through the Syracuse University School of Education. As a part of that accreditation, the School Library Media Program must assess student performance on the competencies that correlate to program outcomes. The competencies which are assessed are identical to the items on your Competency Checklist. As a part of School Library Media Program planning, course-embedded assessments have been aligned with student competencies. For CAEP reporting, each faculty member with competency-based, course-embedded assessments is asked to rate (1=Ineffective, 2=Developing, 3=Effective or 4=Highly Effective) candidates 'performance on the respective competencies. This is the same rating scale students use when completing the competency checklist. Scoring is conducted for key assignments and not all assignments for a course.

What this means for you:
Your individual score is NOT a grade and it is part of an aggregate report. If a student is performing at an Ineffective or Developing level, a comment is submitted with the score, which is also aggregated. Individual scores and comments are not associated with specific student names.

Schedule

This schedule is subject to change.

Date	Readings**	Topics+	Topic During Lab	Homework (HW)*
Week 0				Install R
Week 1	Introduction: Data Science—Many Skills□Ch 1: Begin at the Beginning With R	What Is Data Science and R Overview	Basic R Coding□(Vectors , Conditionals)	HW 1: Working With Vectors
Week 2	Ch 2: Rows and Columns Ch 3: Data Munging	Using R to Manipulate Data	Data Frames and Sorting	HW 2: Manipulating Data Frames
Week 3	Ch 4: What's My Function? Ch 5: Beer, Farms, and Peas Ch 6: Sample in a Jar	Descriptive Statistics, Inferential Statistics, and Functions	Descriptive Statistics and Functions	HW 3: Distributions and Writing Functions
Week 4	Ch 7: Storage Wars	Connecting with External Data Sources	External Data: JSON	HW 4: Getting Data
Week 5	Ch 8: Pictures vs. Numbers	Introduction to Visualization	Data Viz	HW 5: Visualizations

Week 6	Ch 9: Map Mashup	Working With Map Data	Maps	Project Update 1*** HW 6: Working With Maps
Week 7	Ch 10: Lining Up Our Models	Linear Modeling	Exam***	HW 7: Linear Modeling
Week 8	Ch 11: What's Your Vector, Victor? Ch 14: Shiny Web Apps	Support Vector Machines, Decision Trees, and Interactive Applications	Supervised Learning via SVM and Trees	Project Update 2*** HW 8: Support Vector Machines
Week 9	Ch 12: Hi Ho, Hi Ho—Data Mining We Go Ch 13: Text Mining	Association Rule Mining and Text Mining	Using Text Mining	HW 9: Text Mining
Week 10	Ch 15: Time for a Deep Dive	Deep Learning	Deep Mining	Project Update 3*** HW 10: Shiny
Week 11			Final Project	Final Project due***

+Topics discussed in video lectures.
*Assignments must be submitted by midnight on the third night after the synchronous class.
**Readings should be completed prior to video lecture.
***Details to be provided in class