

University of Wisconsin-Stevens Point
 College of Letters and Science
 Department of Computing and New Media Technologies
 Fall 2018 - Version 1.5

Course: Introduction to Data Analysis (DAC 101) #82432
 Books: Introduction to Data Science (2017) by Jeffrey S. Saltz & Jeffrey M. Stanton.
 (ISBN-13: 978-1506377537) Available through the bookstore
 Book Website <https://study.sagepub.com/saltz>
 Class Time: TR 12:30-1:45 p.m. (Room: Sci A224)
 Professor: Dr. Kurt A. Pflughoeft (Floog'heft)
 Office: Science B231
 Office hours: M 2-3 pm, W 12-1 pm, TR 11-12 pm and by appointment
 Contact: kpflugho@uwsp.edu

Course Description: Introduction to the concepts of how different areas make use of data analytics, and the stages of analyzing data; learn to interpret data, communicate insights, and leverage data. The course demonstrates applied examples of data collection, processing, transformation, management, and analysis to provide students with a hands-on introduction to the data analysis experience. Students will explore key concepts related to data analysis, including applied statistics, information visualization, text mining, and machine learning. “R”, the open source statistical analysis and visualization system, is used throughout the course. R is reckoned 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 data scientists.

<u>Week #</u>	<u>Dates</u>	<u>Topic*</u>	<u>Reading/Assignments</u>
1.	Sep 4, 6	Data Science Intro <i>Video: Ch 3: Intro to R coding</i>	Ch 1-3, Lab 1
2.	Sep 11, 13	Structured Data <i>Videos: Ch 5: Rows & Columns, Ch 6: Sorting Dataframes</i>	Ch 4-6, Lab 2, HW 1
3.	Sep 18, 20	Summarization and functions <i>Videos: Ch 8 Functions, Ch 9 Distributions</i>	Ch 7-9, Lab 3, Q1
4.	Sep 25, 27	Sampling and Inference <i>Videos: Ch10 Samples, Sampling</i>	Ch 10, Lab 4, HW 2
5.	Oct 2, 4	External Data <i>Videos: Ch11 SQL, JSON</i>	Ch 11, Lab 5, Q2
6.	Oct 9, 11	Visualization <i>Videos: Ch 12 Basic GGPlot, Bar/Scatter Plots</i>	Ch 12, Lab 6
7.	Oct 16, 18	Mapping <i>Videos: Ch 13 Points on a Map, Map Zooming</i>	Ch 13, Lab 7, Q3
8.	Oct 23, 25#	Review, #Midterm Exam	.
9.	Oct 30, Nov 1	Linear Relations <i>Video: Linear Modeling</i>	Ch 16, Lab 8, HW 4.

<u>Week #</u>	<u>Dates</u>	<u>Topic*</u>	<u>Reading/Assignments</u>
10.	Nov 6, 8	Rules <i>Video: Rule Mining</i>	Ch 17, Lab 9, Q4
11.	Nov 13, 15	Vectors <i>Video: Ch19 SVM</i>	Ch 18, Lab 10, HW 5.
12.	Nov 20, 22	Predictions, Thanksgiving	Q5
13.	Nov 27, 29	Text mining <i>Video: Ch 14 Word Cloud, Ch 15 Sentiment Analysis</i>	Ch 14-15, Lab 11, HW 6.
14.	Dec 4, 6	Interactive/Web	Ch 19, Lab 12, Q6
15.	Dec 11, 13	Big Data	Ch 20, Lab 13
16.	Dec 19	Final Exam 10:15-12:15 pm	Sci A224

Schedule Footnotes:

This schedule is a guide to the coverage of topics. The instructor reserves the right to alter the presentation schedule as necessary to benefit the class.

- No prior experience in programming or statistics is needed.
- Case studies may be presented with some topics.
- Chapter readings, videos, handouts, and lectures are SEPARATE sources for information.
- Videos are available on the book's website; videos are normally 15 minutes or less and are to be viewed outside of class time.

Learning Objectives

At the end of the course, students are expected to understand:

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

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

- Cite examples of data analysis successes and failures for different fields.
- 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, visualization, and searching
- Interpret basic statistical measures
- Organize and manage data at various stages of a project lifecycle
- Determine appropriate techniques for analyzing data

Distribution of Points

Midterm	:15%
Final	:20%
Homeworks	:25%
Lab Exercises	:20%
Quizzes	:15%
Attendance	:05% (Includes Lecture and Lab attendance)

POLICIES

Academic Standards - UW-Stevens Point values a safe, honest, respectful, and inviting learning environment. In order to ensure that each student has the opportunity to succeed, we have developed a set of expectations for all students and instructors. This set of expectations is known as the Community Rights and Responsibilities document, and it is intended to help establish a positive living and learning environment at UWSP. Click here for more information: <http://www.uwsp.edu/dos/Pages/AcademicMisconduct.aspx> Academic integrity is central to the mission of higher education in general and UWSP in particular. Academic dishonesty (cheating, plagiarism, etc.) is taken very seriously. Don't do it! The minimum penalty for a violation of academic integrity is a failure (zero) for the assignment. For more information, see the "Student Academic Standards and Disciplinary Procedures" section of the Community Rights and Responsibilities document, UWSP Chapter 14. This can be accessed at: <http://www.uwsp.edu/dos/Documents/CommunityRights.pdf> - page=11

ADA Statement - The Americans with Disabilities Act (ADA) is a federal law requiring educational institutions to provide reasonable accommodations for students with disabilities. For more information about UWSP's policies, check here: <http://www.uwsp.edu/disability/Pages/faculty/lawAndPolicy.aspx>. If you have a disability and require classroom and/or exam accommodations, please register with the Disability and Assistive Technology Center at the beginning of the course and then contact me. I am happy to help in any way that I can. For more information, please visit the Disability and Assistive Technology Center, located on the 6th floor of the Learning Resource Center (the Library). You can also find more information here: <http://www.uwsp.edu/disability/Pages/default.aspx>.

Attendance Policy - Attendance will be taken randomly in lecture/lab and will count towards your grade! I rarely lecture "STRAIGHT FROM" the book.

Average Time Investment/Workload Policy Statement

DAC 101 meets twice a week; each meeting is 75 minutes or 2.5 hours per week or 37.5 hours per semester. Additionally, you should expect to spend at least another 5 hours per week, on average, on outside class work including videos and chapter reading assignments.

Audio/Visual Recording Policy - Electronic recording of lectures (taping) is prohibited unless receiving prior written approval from the instructor. Approval will be granted only for self-study purposes. You are allowed to take pictures of whiteboards, blackboards or screens of lecture materials, if need be.

Classroom conduct – Please mute cell phones and any audible device during classes. Please do not hold private conversations or text while I am lecturing as it is a distraction.

D2L –

- Recorded grades as well as lecture materials (syllabus, PowerPoint class outlines, etc.) will be available on our course 205 D2L course site.
- It is your responsibility to check that your grades are posted correctly on D2L. Questions about any posted grade must be raised within TWO weeks of posting. Beyond this time frame, all grade postings are considered correct and final. The D2L site is not available after the final exam.
- It's best to download attachments rather than view directly in the browser.
- USE the OneDrive to temporarily save your files – if need be.
- **News feed** on D2L is the main communication tool (not email!) - consider the text msg option
- The UW system is in the process of moving to a different LMS called Canvas

Drop Policy - In accordance with the rules stated by the College of Letters and Science. I will **NOT** personally drop a student - you are responsible for filling out all the forms.

Exam Policy - Except for documented emergencies, no late or makeup in-class exercises, exams and quizzes will be given.

Grade Policy - The following scale can always be used to estimate your grade

Percentage breakdown for semester grades (weighted point totals)

A = 93-100%	B- = 80-82.99%	D+ = 67-69.99%
A- = 90-92.99%	C+ = 77-79.99%	D = 63-66.99%
B+ = 87-89.99%	C = 73-76.99%	D- = 60-62.99%
B = 83-86.99%	C- = 70-72.99%	F = < 60%

*Instructor reserves the right to implement a curve which is beneficial to the students.

Homework Policy – Homework assignments are listed on the syllabus as HW; they are usually due two weeks after the assigned date and are posted on D2L. Electronic copies of the completed homeworks must be posted on the D2L drop box by the required date/time. Format for the electronic file is listed in the lab bullet. Late assignments are not accepted. Note: You have OneDrive account to save files to if you need to access them later; alternatively you can save materials on a flash drive but remember to take it with you. Please note that the book's web site lists labs and homeworks but **ALWAYS** use what is posted on D2L. Late assignments are not accepted unless you have some sort of documented emergency; the professor determines if the situation constitutes an emergency.

Labs –are scheduled on Thursdays and usually have in-class exercises. Lab time may be redirected as lecture time at the discretion of the instructor. For lab assignments, you should turn in a Word document which lists your code, and one or more screen shots (or relevant copy/pastes) of the program's output to demonstrate the program works correctly. If you have extra lab time, you are encouraged to work on your homework assignments.

Lecture Notes – electronic version of the notes is available for some topics, however, I strongly encourage you to take good notes as that has been shown to reinforce memory recall.

News – Always check the news item on D2L to find the latest announcements concerning the class.

Software – Lab Virtual Desktop or install RStudio on your PC, we will use Excel for some assignments. I can help you with the R install on your laptop. 8 gig of RAM should be sufficient for an academic environment but data analysts in industry often require more as all data must be loaded in primary storage for R.

Plagiarism Policy - All assignments and tests should represent YOUR work otherwise you will not receive any credit for that portion of your grade. Disciplinary actions will be pursued for serious offenses.

Quiz Policy – quizzes are meant to test your understanding about topics that were currently presented. Quizzes will be take-home but you are NOT allowed to collaborate with others. You may use other resources such as google. For open-ended questions, be careful not to plagiarize.

University Emergency Preparedness –

In the event of a medical emergency call 9-1-1 or use Red Emergency Phones. Offer assistance if trained and willing to do so. Guide emergency responders to victims.

In the event of a tornado warning, proceed to the lowest level interior room without window exposure. See www.uwsp.edu/rmgt/Pages/em/procedures/other/floor-plans.aspx for floor plans showing severe weather shelters on campus. Avoid widespan structures (gyms, pools or large classrooms.)

In the event of a fire alarm, evacuate the building in a calm manner. Stay 200 yards away from the building. Notify instructor or emergency command personnel of any missing individuals.

Active Shooter – RUN/ESCAPE, hide, fight. If trapped hide, lock doors, turn off lights, spread out and remain quiet. Call 9-1-1 when it is safe to do so. Follow the instructions of emergency responders.

See UW-Stevens Point Emergency plan at <https://www.uwsp.edu/rmgt>