

INST 627 – Data Analytics for Information Professionals

Section 0201

Location: Hornbake, Room 2119

Time: Tuesdays, 6-8:45PM

Instructor: Dr. Jessica Vitak
E-mail: jvitak@umd.edu
Phone: (301) 405-8617
Office: 2117G Hornbake, South Building
Office Hours: Thursdays 11AM-12PM (in my office), Mondays 1-2PM (virtual),
or by appointment

COURSE DESCRIPTION

Advances in hardware and software technologies have led to a rapid increase in the amount of data collected, with no end in sight. Decision making in the coming decades will depend, to an ever greater extent, on extracting meaning and knowledge from all that data. In this class we focus on one branch of statistics, inferential statistics, to help us reason about data. By gathering datasets, formulating proper statistical analyses and executing these analyses, information professionals play a significant role in bridging the gap between raw data and decision making.

This course will introduce basic concepts in data analytics including study design, measure construction, data exploration, hypothesis testing, and statistical analysis. The course also provides an overview of commonly used data manipulation and analytic tools. Through homework assignments, projects, and in-class activities, you will practice working with these techniques and develop statistical reasoning skills.

LEARNING OBJECTIVES

After completing this course you will be able to:

- Select and evaluate various types of data to use in decision making;
- Use prescriptive and descriptive analyses to reach defensible, data-driven conclusions;
- Select and apply appropriate statistical methods;
- Use MS Excel and SPSS for basic data manipulation and analysis; and
- Critically evaluate data analyses and develop strategies for making better decisions.

COURSE MATERIALS

Software

The following software is required for you to successfully complete the homework, exams, and project for this course.

1. **Microsoft Excel** 2007, 2010 or 2013 (or Excel 2011 for Macintosh) - If you do not have access to a computer that has Excel 2007, 2010 or 2013 (or Excel 2011 for Macintosh) installed, consider downloading Office 2013 Professional Plus (for Windows) or Office 2011 for Macintosh through the university's TERPware website (<https://terpware.umd.edu>).
2. **SPSS** Statistical Analysis software (also called PAWS): UMD students can download a free version from TERPware (<https://terpware.umd.edu>). The licensing fee is \$35 for one year.

Note: SPSS is a good statistical analysis program for people with no stats experience and with little to no programming experience. If you would like to learn [R](#) (a free, open-source programming language and software), I encourage you to do so, especially if you plan on taking INST-737, Digging Into Data. That said, I will only be using SPSS in class demonstrations.

Hardware

We will be using this software in many classes, so you should be bringing your laptop or tablet with you to every class. If this is a problem, please email me and we will find an alternative solution.

Readings and Online Resources

There are many good texts and online sources for information on decision-making, statistical techniques and data tools. The main textbook used for the class will be **The Online Stats Book** (<http://onlinestatbook.com/2/index.html>) developed primarily by Rice University. This book is thorough, easy to understand, and is available for free in several formats. In addition, we will be collectively keeping an updated list of tutorials and references on Canvas that prove to be useful during the class, and you may be asked to watch or read tutorials in addition to sections from the Rice textbook.

COURSE ACTIVITIES

Participation

While you will not be graded on your participation in the class, the nature of the material being taught requires that you attend lecture and are an active participant. I encourage students to ask questions whenever they do not understand a topic in class; however, if you do not attend lectures, do not expect me to be sympathetic when you perform poorly on a test. Your participation will largely determine your final grade in this course.

Homework

Most weeks you will have an assignment that is designed to assess your mastery of the topics and techniques covered the previous week and provide you feedback to improve your understanding of the material. There will be a total of 12 homework assignments. **Your 10 best scores on the homework assignments will be used to calculate your final grade**, while the lowest two scores will be dropped.

You may work with your colleagues to figure out the underlying concepts and problem-solving processes, but are expected to work *individually* to answer the specific problems that are assigned. Completed assignments will be submitted via ELMS and are due prior to the start of class for the week they are due. If an assignment due date is a religious holiday for you, please let me know at least one week in advance, so an alternate due date can be set.

Non-graded Quizzes

From time to time, we will hold quizzes at the start of the class to help me gauge the level of comprehension of a topic. These quizzes will not be graded but will enable me to spend more time on topics that students are struggling with and will help students recognize areas they will need more work in to ensure success in the class.

Group Project

In groups of 3, you will prepare a data-related analytic project. Over the course of the semester, you will identify an existing dataset (or generate your own), develop research questions that can be tested through statistical analyses, formulate an appropriate statistical analysis, carry out the analysis, and report on the results.

The group project is broken into five components with due dates throughout the semester.

1. **Group Formation** (due Week 3, no points): One person from each group should email me the names of all group members.
2. **Proposal** (due Week 4, 25 points): You and your group members should submit a 1-2 page single-spaced proposal in which you discuss your proposed project. This includes (a) a 1-2 paragraph overview of your topic, (b) at least **three** potential research questions to address the topic with a description of why you think they are valuable/interesting questions to ask; (b) the data set you want to use to analyze your research questions (note: if you're planning on collecting data, you should work closely with me early in the semester to facilitate data collection).
3. **Progress Report** (due Week 9, 25 points): This is a checkpoint for groups to make sure everyone is on the right path with the project. Groups should submit a 1-2 page, single spaced update on the project, including an updated description of the project and its significance, as well as re-sharing research questions or updating them, plus any challenges they've experienced or questions.
4. **Peer Review Feedback** (due Week 10, 25 points): Each student will be assigned to review another group's proposal and progress report and provide detailed feedback (250 words minimum). I will create discussion boards for this feedback to be shared.

5. **Poster** (due Week 15, 75 points): During the last meeting of the semester before the exam, each group will create a poster that highlights the project and key findings. Groups will be asked to give brief (3-5 minute) overviews of their project. In class peer evaluations will factor into grading.
6. **Final paper** (due Week 15, 150 points): Each group should submit a final report of their project, including an overview of the topic and research questions, as well as detailed analyses and interpretation of findings. More details will be provided in a separate document.

Exams

There will be a midterm and final exam, each worth 20% of your final grade. These exams provide an opportunity for you to test your understanding of the concepts, techniques, and problems associated with statistical reasoning. In order to learn and understand the material fully it is important to review and revisit it multiple times.

Grading

Final grades will be assigned based on the total number of points earned, using the following rubric. Please come and talk to me early if you think that there might be a problem. No extra credit will be given and grades will not be rounded up.

A	1000-900pts (A- 940-900pts)
B	899-800pts (B+ 899-870pts ; B- 830-800pts)
C	799-700pts (C+ 799-770pts ; C- 730-700pts)
D	699-600pts (D+ 699-670pts ; D- 630-600pts)
F	599-0pts

Category	Points
Homework (30%)	300
12 assignments (top 10 counted)	30 each
Group Project (30%)	300
Proposal (Due: 9/22)	25
Progress Report (Due: 10/10)	25
Progress Report Peer Feedback (Due: 10/17)	25
Poster (Due 12/8)	75
Paper	150
Exams (40%)	400
Midterm (October 20 class)	200
Final (December 18, 4pm)	200
TOTAL POINTS POSSIBLE	1000

Extensions

Timeliness is an essential component of graduate work, and extensions will only be available during personal emergencies. Students who need to request an extension should discuss the matter in advance with the professor. If an extension is granted, the work must be submitted within the extension period to avoid grade penalties. Unexcused delays in submission of the paper will result in a deduction of a letter grade for each day the paper is late, while unexcused delays in presentations will result in a deduction of a letter grade for each class meeting the presentation is late.

Late work

Unless approved by the professor in advance of the due date, late work will automatically be graded down by one step (i.e., 5%) for each day it is late (unless otherwise noted in the syllabus). For example, an assignment that would normally receive an A- if submitted on time would receive a B if it was submitted two days late. Homework assignments will be penalized two points per day late. Any assignment submitted more than five days late will not be accepted.

CONTACTING ME | MY EMAIL POLICY

If you have any questions about the course generally or specific questions about content, please check the syllabus and ELMS first. If these resources do not answer your question, email me directly at jvitak@umd.edu and include **INST627** in the subject line. If you would like to schedule an in-person meeting, please provide at least two days and times you are available.

I have a 48-hour email reply policy, which means I will reply to your email within two days of you sending it. The turnaround is usually much faster (especially if you include the course number in the subject line), but do not expect me to reply to late Monday night/early Tuesday morning questions about homework before class. It is your responsibility to check any assignments early to see if you have questions.

I also encourage you to visit me during office hours on Mondays (virtually) and Thursdays. This is an opportunity to ask questions about the material covered in the reading materials or in lecture. On Mondays, I will use the “Chat” feature in ELMS and, if needed, will provide a link to video (e.g., if I want to show how to do something in SPSS, I can share my screen).

If you are having trouble in the course, please talk to me as soon as possible. If you do poorly or lower than you expected on the first exam, it is imperative that you come to office ho that we can figure out the problem early.

ACADEMIC DISHONESTY

Cheating in any form (copying, falsifying signatures, plagiarism, etc.) will not be tolerated. It will result in a referral to the Office of Student Conduct irrespective of scope and circumstances, as required by university rules and regulations. There are severe consequences of academic misconduct, some of which are permanent and reflected on the student's transcript. If you have any questions regarding the University's policies on scholastic dishonesty, please see <http://osc.umd.edu/OSC/Default.aspx>.

It is very important that you complete your own assignments, and do not share files (excluding raw data), partial work or final work. Note that I do allow (and encourage) you to work with friends and classmates to learn the course concepts, but any assignments that appear to be copied straight from another will be treated as plagiarized.

University of Maryland Code of Academic Integrity

The University of Maryland, College Park has a nationally recognized Code of Academic Integrity, administered by the Student Honor Council. This Code sets standards for academic integrity at Maryland for all undergraduate and graduate students. As a student you are responsible for upholding these standards for this course. It is very important for you to be aware of the consequences of cheating, fabrication, facilitation, and plagiarism. For more information on the Code of Academic Integrity or the Student Honor Council, please visit <http://shc.umd.edu/SHC/Default.aspx>.

ACCOMODATIONS

Please come and see me as soon as possible if you think you might need any special accommodations for disabilities. In addition, please contact the Disability Support Services (301-314-7682 or <http://www.counseling.umd.edu/DSS/>). Disability Support Services will work with us to help create appropriate academic accommodations for any qualified students with disabilities. If you experience psychological distress during the course of the semester you can get professional help at the Counseling Center (301-314-7651 or <http://www.counseling.umd.edu/>).

EMERGENCY PREPAREDNESS

Please see the University's Emergency Preparedness Website (<http://www.umd.edu/emergencypreparedness>) for information about the current status of the campus. If a class session needs to be rescheduled, I will e-mail you as soon as possible.

Inclement Weather: In the event of inclement weather, students should check the UMD homepage (umd.edu) or call 301-405-SNOW (7669) to determine if there are delays or closures. Closures and delays will also be sent over the e2 Campus notification system. Follow the link to sign up for alerts: www.alert.umd.edu. Also make sure you either check your UMD email regularly or forward UMD emails to an account you do check regularly, in case I email out a class cancellation.

PRELIMINARY COURSE SCHEDULE

*Unless otherwise noted, all chapter readings refer to the Rice textbook ([available online](#)).

Week/ Date	Topics	Required Readings	Due
W1 9/1	Introduction & Course Overview	Chapter 1 , Sections on Descriptive and Inferential Statistics	<ul style="list-style-type: none"> Excel and SPSS Installed
W2 9/8	Measurement & Design	Chapter 1 , Sections on Variables and Levels of Measurement ; Chapter 6 , Section on Causation	<ul style="list-style-type: none"> Homework 1: Research Questions
W3 9/15	Descriptive Statistics Overview	Chapter 1 , Section on Distributions ; Chapter 2 , Section on Histograms ; Chapter 3 , Sections on Central Tendency (read both sections), Variability , Shapes of Distribution	<ul style="list-style-type: none"> Homework 2: Designing a Data Analytics Project
W4 9/22	Probability and Sampling	Chapter 7 , Section 3 (<i>History</i>); Chapter 9 , Section 2,6 (<i>Introduction & Sampling Distribution of the Mean</i>); Chapter 11 2 and 3 (<i>Introduction & Significance Testing</i>)	<ul style="list-style-type: none"> Homework 3: Measures of Central Tendency Project Proposal
W5 9/29	Hypothesis Testing (one-sample t-tests)	Chapter 11 , Sections 4-8 (<i>Type I & II Errors thru Steps in Hypothesis Testing</i>); Chapter 12 , Section 2 (<i>Single Mean</i>)	<ul style="list-style-type: none"> Homework 4: Probability
W6 10/6	Hypothesis Testing (two-sample t-tests)	Chapter 10 , Section 7,8,9,11 Chapter 12 , Section 4	<ul style="list-style-type: none"> Homework 5: Hypothesis Testing and Power
W7	Chi-squared and Midterm	Chapter 17 , Section 2,3	<ul style="list-style-type: none"> Homework 6: T-tests

W8 10/20	Midterm Exam	n/a	<ul style="list-style-type: none"> • Homework 7: Chi-square
W9 10/27	Qualitative Data Overview	<i>Qualitative Research for the Information Professional</i> (Chapter 3—PDF on ELMS)	<i>No homework</i>
W10 11/3	Analysis of Variance (ANOVA)	Chapter 15 , Section 2,3,4	<ul style="list-style-type: none"> • Homework 8: Qualitative Data • Progress Report
W11 11/10	Factorial Designs	Chapter 15 , Section 6,8	<ul style="list-style-type: none"> • Homework 9: One-way ANOVA • Progress report Peer Feedback
W12 11/17	Correlations & Linear Regression	Chapter 14 , Section 2,4,5,6	<ul style="list-style-type: none"> • Homework 10: Factorial Designs
W13 11/24	Multiple Regression	Chapter 14 , Section 9	<ul style="list-style-type: none"> • Homework 11: Regression
W14 12/1	Logistic Regression	Statistics review: Logistic regression Logistic Regression Versus Multiple Regression (PDFs on ELMS)	<ul style="list-style-type: none"> • Homework 12: Multiple Regression
W15 12/8	Group Presentations, Final Exam Review, and Course Evaluations	n/a	<ul style="list-style-type: none"> • Final Project Poster • Final Project Due
12/18	Final Exam: FRIDAY, DECEMBER 18, 4-6PM		

**Note: This schedule is for planning purposes and may change at any time.
See ELMS/Canvas for current information and deadlines.**