IT223 Data Analysis, Spring 2022

Instructor: Noriko Tomuro
Class hours: Monday and Wednesday 11:50 AM – 1:20 PM, Lewis 1110
Office: CDM 648
Email: tomuro@cs.depaul.edu
Office Hours: Fridays. 5:00 PM - 7:00 PM, by Zoom. (Schedule Advising Appointments through BlueStar.)
Course website: https://d2l.depaul.edu/

1. Course Description

Application of statistical concepts and techniques to a variety of problems in IT areas and other disciplines, using a statistical package for simple data analysis. Course topics include descriptive statistics, elementary probability rules, sampling, distributions, confidence intervals, correlation, regression, and hypothesis testing.

Prerequisites: LSP 120 or MAT 120 or (MAT 130 or above or equivalent or Mathematics Diagnostic test placement into MAT 140).

Course objectives:

- To develop an understanding of the basic concepts of probability and statistics,
- To help students to be informed and critical readers of quantitative arguments,
- To provide sufficient skills to apply simple statistical techniques with the aid of a computer,
- To appreciate the role of statistics in empirical research and scientific study, and
- To gain flexible problem-solving skills applicable to unfamiliar statistical settings.

2. Textbook, Software and Other Resources

Textbook: [Required] "*Open Intro Statistics, 4th Edition*" (2019). David Diez, Mine Cetinkaya-Rundel, Christopher D Barr, ISBN-13: 978-1943450077, ISBN-10: 1943450072. (Freedownload available at https://www.openintro.org/stat/textbook.php)

Software: We will use a statistical software called R. R is a widely used and stable software that is free. RStudio is a user-friendly interface for R. You install R on your local machine and run it there. Alternatively, computers in all DePaul labs have R installed, so you can work there.

3. Coursework

- **Quizzes** will be given roughly weekly (a total of 8 quizzes). Quizzes are online, posted on D2L.
  Each quiz will have 6-8 multiple choice questions, and you have 2 hours to complete and submit. You can make only one submission attempt. If no submission attempt is made in two hours, a score of 0 will be given.
- **Assignments** (or 'homeworks') will be given roughly every other week (a total of 4 homeworks). Each homework will have both written and programming/lab questions. Written questions are intended to develop students critical thinking skills to analyze the results and articulate them in writing (in words), whereas programming questions are intended to increase students' familiarity with the statistical software.

- **Midterm** will be either an online exam or a project. Details will be announced later.
- **Final** will be a pencil-and-paper written and in-class exam. It will be held on <mark>June 6 (Mon), 11:30 am – 1:45 pm</mark>, in the regular classroom. Topics will be comprehensive from the beginning of the course. One page cheat-sheet and a calculator will be allowed to bring in the exam but nothing else.

4. Attendance Requirement

As this is an On-Campus section of the course, class attendance is **mandatory**. Attendance is measured on **weekly basis**. <mark>Students who miss class due to illness or other significant personal circumstance in which documentation has been obtained, must complete the Absence Notification Process through the Dean of Students Office</mark> (DOS; https://offices.depaul.edu/student-affairs/support-services/academic/Pages/absence-notification.aspx). After the instructor receives notification from the DOS, or in instances where there is no documentation, it will be up to the instructor's discretion to excuse the absence.

5. Grading

The course grade will be computed as follows:
- ✓ **Attendance** (5%)
- ✓ **Quizzes** (30%) – No late submissions are accepted. The **quiz with the lowest score will dropped** from the calculation of the final course score.
- ✓ **Assignments** (20%) – Late submissions are accepted up to three days late. However, a **10%-point penalty will be applied for each day** that is late.
- ✓ **Midterm** Exam (15%) – No late submissions are accepted.
- ✓ **Final** Exam (30%) – No late submissions are accepted.

Students receiving more than 90% of possible points are guaranteed at least an A-, more than 80% at least a B-, more than 70% at least a C-, and more than 60% at least a D.

6. Policy on Working Together

Collaboration on assignments IS permitted. Students are encouraged to discuss problems in the assignments and solution ideas with other students or tutors. However, each assignment handed in must be done by the person submitting it. All homework submissions will be run through TurnItIn which is plagiarism detection software. This software compares your assignment with all other submissions ever submitted for this course.

Collaboration on exams is NOT permitted. See the University and School policies on plagiarism and incompletes.

7. Policy on Submission Materials

ALL submission materials (except for online quizzes/exams) must have the **student's name**, the **course name** ("IT 223 Data Analysis"), and the **assignment number** (e.g. HW#1) written/typed at the top of the submission files. <mark>Submissions without the information will not be graded and receive a score of 0</mark>.

8. Email Communication

Any email to the instructor should begin the **subject line starting with "IT 223"**, so that your message can

be easily identified. Failure to do so will delay the response time from the instructor.

9. Tentative Schedule of Topics

The following schedule is tentative. The reading assignments are from the course textbook.

Lecture (Class date)	Topic	Reading assignment
Week 1 (3/28, 30)	Introduction. Exploratory data analysis. Analyzing univariate distribution using graphs (histograms, bar charts, pie charts, and boxplots) and summary statistics for center and spread.	Chapter 1, Section 1.1, 1.2 Chapter 2, Section 2.1
Week 2 (4/4, 6)	Density functions and normal distribution. Using the normal distribution to approximate symmetric distributions, normal quantile plots to test normality assumptions.	Chapter 4, Section 4.1
Week 3 (4/11, 13)	Data Relationships: Scatter plots and correlation. Introduction to regression analysis. Model fitting and diagnostic, residual analysis.	Chapter 8, Section 8.1-8.4
Week 4 (4/18, 20)	Design of experiments, observational studies. Sample surveys.	Chapter 1, Section 1.3, 1.4
Week 5 (4/25, 27)	Randomness and the language of probability. Probability rules. Random variables, expected value and variability.	Chapter 3, Section 3.1, 3.3, 3.4
Week 6 (5/2, 4)	Binomial distribution, normal approximation of binomial distribution	Chapter 4, Section 4.3
Week 7 (5/9, 11)	Toward statistical inference. Sampling distribution forsample averages. Estimating averages using Confidence Intervals	Chapter 5, Section 5.1
Week 8 (5/16, 18)	Sampling distributions for proportions. Estimating proportions using confidence intervals.	Chapter 5, Section 5.2 Chapter 6, Section 6.1.1, 6.1.2, Section 6.2
Week 9 (5/23, 25)	Hypothesis Testing. Significance tests on averages and proportions	Chapter 5, Section 5.3 Chapter 7, Section 7.1
Week 10 (6/1)	**5/30 Memorial Day – NO CLASS** Final exam review	
Week 11 (6/6)	<mark>**Final Exam, 11:30 am – 1:45 pm**</mark>	

10. Important Dates

- o 3/28 (Mon) -- First day of class
- o 4/1 (Fri) -- 11:59 PM Deadline to add classes to SQ2022 schedule

- 4/8 (Fri) -- Last day to drop SQ2022 classes with no penalty (100% refund of tuition if applicable and no grade on transcript)
- 4/15 (Fri) -- Good Friday - University officially closed
- 4/16 (Sat) – 17 (Sun) -- Easter Holiday - University officially closed
- 5/13 (Fri) -- Last day to withdraw from SQ2022 classes
- 5/30 (Mon) -- Memorial Day - University officially closed
- **6/6 (Mon), 11:30 am – 1:45 pm – Final Exam**

11. School and University Policies

School policies (on Changes to Syllabus, Online Course Evaluations, Academic Integrity and Plagiarism, Academic Policies, Students with Disabilities) are found at
https://www.cdm.depaul.edu/academics/pages/classinfo.aspx?Term=20223&ClassNbr=37398&fid=339459

University policies and student expectations are below.

Attendance
Much learning happens while working with peers, discussion, and participation in a community. These are learning experiences that are almost impossible to make up individually. If you have concerns about your ability to attend class sessions, please reach out to me via email to discuss your concerns.

Class Discussion
Student participation in class discussions will be measured in two ways. First, students are highly encouraged to ask questions and offer comments relevant to the day's topic. Participation allows the instructor to "hear" the student's voice when grading papers. Secondly, students will be called upon by the instructor to offer comments about lecture topics. Students must pay the fullest attention to the lecture during the class to keep up with the materials and to participate in class discussion.

Attitude
A professional and academic attitude is expected throughout this course. Measurable examples of non-academic or unprofessional attitude include but are not limited to: talking or disrupting class when the instructor is speaking, mocking another's opinion, and cell phones ringing while being in the class. If any issues arise a student may be asked to leave the classroom. The professor will work with the Dean of Students Office to navigate such student issues.

Civil Discourse
DePaul University is a community that thrives on open discourse that challenges students, both intellectually and personally, to be Socially Responsible Leaders. It is the expectation that all dialogue in this course is civil and respectful of the dignity of each student. Any instances of disrespect or hostility can jeopardize a student's ability to be successful in the course. The professor will partner with the Dean of Students Office to assist in managing such issues.

Cell Phones/On Call
If you bring a cell phone to the class, it must be off or set to a silent mode. Should you need to answer a call during class, students must leave the classroom in an undisruptive manner.