

LIS 622 / ICT 622

Data Analysis and Visualization

Term: Spring 2023

Credit hours: 3

Meeting days/time/location: online, asynchronous

Instructor Information

| | |
|------------------------------|---|
| Name: | Monica Rogers, Ph.D. |
| Email: | mlro271@uky.edu |
| Office hours: | Monday and Tuesday, 10am to Noon ET |
| Preferred method of contact: | Email/Canvas. I usually respond to questions within 24 hours (expect a delay during weekends and holidays.) |

Course Description

This course examines three major categories of topics in relation to data analysis and visualization. First, this course will cover the basic ways that data can be obtained from various sources, such as raw text files, web APIs, and data repositories. It will also cover the techniques of data cleaning and how to organize data for analysis. Second, the course will cover the essential techniques for analyzing quantitative data. It will teach prediction and clustering methods that are useful to solve various real data analysis tasks. In addition, students will learn major theories and recent methods in text analysis. Third, this course teaches how to create visualizations that effectively communicate the meanings behind data and information. The course will cover key practical skills in information visualization, such as plotting, mapping, and network visualization. This course will not be mathematically intensive. Instead, the course will use existing computational tools and programming libraries to solve various problems. You will use the R language and environment for data analysis and visualization.

Required Materials

No textbook. Lecture notes will be provided. All course reading materials will be available on Canvas. Students will need to install the R software toolkit on their own computer. Instructions will be made available.

Associated Expenses

N/A

Activities Outside of Regular Class Meetings

N/A

Skill and Technology Requirements

Distance learning students are expected to have at least a minimum level of computer literacy skills and the availability of technological resources. Students must have regular access to a computer with: reliable internet, audio capabilities, Firefox, Chrome, or Safari browsers. Please be certain that you can

install software on your computer, specifically R, and that you can view Adobe Reader documents. Microsoft Office 365 is provided free to students: <https://download.uky.edu/>

I am your first point of contact for technology issues, but if you need immediate assistance, contact UKIT.

For technical assistance, contact ITS Customer Services 24/7 at 859-218-HELP (4357) for urgent needs. For non-urgent matters, choose the option that works best for you at <https://techhelpcenter.uky.edu/gethelp>

Student Learning Outcomes

After completing this course, the student will be able to:

1. Explain basic concepts and major methods in data analysis.
2. Identify and obtain different types of data from various sources, including data repositories, social media and web and publicly open data (e.g., governmental websites).
3. Think critically about data and identify appropriate methods to solve given problems.
4. Apply appropriate data analysis procedures and visualization techniques to draw conclusions from analyses.
5. Utilize graphical and numerical summaries to effectively represent analysis results.
6. Identify appropriate visualization methods and apply them to different types of data.
7. Implement R packages to analyze and visualize data.
8. Interpret and author R programming.

Course Details

Tentative Course Schedule (subject to change)

| Week | Date | Topics | Assignment |
|-----------------------|-------------|--|--------------------------|
| 1 | 1/9 – 1/13 | Course Overview | |
| 2 | 1/16 – 1/20 | Introduction to R | |
| 3 | 1/23 – 1/27 | Visualization with R | |
| 4 | 1/30 – 2/3 | Basic Statistics with R | Assignment #1 (due 2/3) |
| 5 | 2/6 – 2/10 | Clustering and visualization | Assignment #2 (due 2/10) |
| 6 | 2/13 – 2/17 | Classification and visualization | |
| 7 | 2/20 – 2/24 | Text data processing and visualization | |
| 8 | 2/27 – 3/3 | LDA topic modeling | Assignment #3 (due 3/3) |
| 9 | 3/6 – 3/10 | Sentiment analysis | Assignment #4 (due 3/10) |
| 10 | 3/13 – 3/17 | Spring Break – no class or office hours | |
| 11 | 3/20 – 3/24 | Network analysis and visualization | |
| 12 | 3/27 – 3/31 | Bibliographic analysis and visualization | Assignment #5 (due 3/31) |
| 13 | 4/3 – 4/7 | Linear and non-linear regression models | |
| 14 | 4/10 – 4/14 | Time series analysis | |
| 15 | 4/17 – 4/21 | Geospatial data visualization | |
| 16 | 4/24 – 4/28 | Final project | |
| Final report due: 5/3 | | | |

Course Activities and Exams

Your final grade is determined by your performance on the following items: five assignments (50%), a final project (30%), and class participation (20%).

Summary Description of Course Assignments

- Assignment #1 (10%) – Basic statistics with R
- Assignment #2 (10%) – Clustering analysis and visualization
- Assignment #3 (10%) – Textual analysis and visualization
- Assignment #4 (10%) – Sentiment analysis and visualization
- Assignment #5 (10%) – Network analysis and visualization
- Final Project (30%)
- Class participation (20%)

Grading Scale

Grades are based on the quality of work submitted.

90 – 100% = A

80 – 89% = B

70 – 79% = C

Below 70%= E

Attendance Policy/Acceptable Documentation

This is an online course, which requires asynchronous class discussion via Canvas to facilitate a sense of community. Canvas will also be used for making course announcements, distributing reading materials, submitting assignments and posting grades. Please visit <https://www.uky.edu/canvas> for course homepage.

Course related communications will occur mainly in Canvas (online discussions, questions and answers, etc.). In addition, all announcements will be posted in Canvas. Students thus are required to check the course shell in Canvas on a regular basis. Failure to receive such announcements cannot be used as an excuse for not being informed. I welcome emails, please feel free to email me at mlro271@uky.edu if you have any questions or concerns. In ordinary circumstances, it is expected that the instructor will respond within 24 hours during weekdays. Please expect a delay during weekends and holidays.

Excused Absences

Students need to notify the professor of absences prior to class when possible. *Senate Rules 5.2.4.2* defines the following as acceptable reasons for excused absences: (a) serious illness, (b) illness or death of family member, (c) University-related trips, (d) major religious holidays, and (e) other circumstances found to fit “reasonable cause for nonattendance” by the professor.

Students anticipating an absence for a major religious holiday are responsible for notifying the instructor in writing of anticipated absences due to their observance of such holidays no later than the last day in the semester to add a class. Two weeks prior to the absence is reasonable, but should not be given any later. Information regarding major religious holidays may be obtained through the Ombud (859-257-3737, http://www.uky.edu/Ombud/ForStudents_ExcusedAbsences.php)

Per *Senate Rule 5.2.4.2*, students missing any graded work due to an excused absence are responsible: for informing the Instructor of Record about their excused absence within one week following the period of the excused absence (except where prior notification is required); and for making up the missed work. The professor must give the student an opportunity to make up the work and/or the exams missed due to an excused absence, and shall do so, if feasible, during the semester in which the absence occurred.

Verification of Absences

Students may be asked to verify their absences in order for them to be considered excused. *Senate Rule 5.2.4.2* states that faculty have the right to request “appropriate verification” when students claim an excused absence because of illness, or death in the family. Appropriate notification of absences due to University-related trips is required prior to the absence when feasible and in no case more than one week after the absence.

Assignment Policies

Assignment Submissions

All assignments should be submitted online via Canvas. Unless due to an excused absence, an overdue assignment will get a penalty of 10% of total points for each day late. No assignment and project will be accepted after seven days.

Academic Policy Statements

Please review the Senate’s Academic Policy Statement: [Academic Policy Statements](#)

Academic Offenses (Cheating, Plagiarism, and Falsification or Misuse of Academic Records)

Please review the Senate-maintained webpage: [Rules Regarding Academic Offenses](#)

Resources

[Distance Learning Library Services](#)

[Tutoring and Coaching Resources](#)

Diversity, Equity, and Inclusion

Please review the Senate Council-approved: [Syllabus Statement on Diversity, Equity, and Inclusion \(DEI\)](#)

~~~~~