GEOG 4195: Geospatial Data Analytics and Visualization
MW 2:00-3:20 pm

Dr. Lisa Nagaoka
ENV 210C or GAB 210K
MW 1:00-2:00 or by appt
Lisa.Nagaoka@unt.edu

Most data analysis courses fall into two categories – traditional ones that focus on statistical or modeling techniques and newer data science courses. The more traditional courses often assume that researchers have more experience working with data than they often do. Since datasets are often provided in these courses, training does not include how to collect and manage data. Data science courses, in contrast, tend to cover these topics. But since data science has emerged to handle large datasets generated with rise of digital technologies, the analytics often require extensive programming and modeling knowledge, which may be beyond the average researcher. This course fills the gap by focusing on the parts of data science that will allow researchers to better utilize traditional analytical techniques and collaborate with data scientists. In addition, we incorporate geospatial data, which often has specific requirements and structures that require different methods.

The content of the course focuses on the front and back end of the research process. On the front end, we will cover the steps of collecting data, building and organizing databases, and conducting exploratory data analysis that are often omitted from traditional data analysis courses. On the back end, we will explore the graphic design of and analytic content in visualizations (graphs, tables, and maps) that are used to communicate findings in print reports or online content. We will also be introducing students to a variety of software and applications that are being used in the workplace. Many of these programs are free or open source, and thus provide alternative options when commercial software packages are unavailable. In addition, because we will be focusing on geospatial datasets, the methods and tools used in these processes will often be different from those used with non-spatial datasets.

COURSE STRUCTURE
Pre-Class Assignments
The "lecture" material will be presented online in the form of readings and videos. The assignments will be either reading questions or Perusall. These resources will be either informational and instructional. They will provide content and information about a topic similar to a traditional lecture. The instructional videos will demonstrate how to use the various programs and tools we will be using in class. We highly recommend that you take notes on the informational videos and go through the steps in the instructional videos with the provided datasets.

In-Class Instruction
Most class meetings will consist of using the tools covered in the instructional videos with a new question and dataset. We will also discuss important concepts and answer specific questions about the tools and methods.

Attendance
Attendance is essential for success in this class. The course emphasizes skill development; class time will be used to introduce, review, and troubleshoot various analytics programs. Additionally, students will collaborate on course projects. Class sessions will be key opportunities for students to meet and discuss those projects.

Absences should be handled like absences at work or canceling a doctor's appointment. Contact the instructor as soon as possible. Excused absences are those situations that are beyond your control that prevent you from attending class. Illnesses do not need to be medically verified with a doctor's note.

AI Usage
In this course, we will be using generative AI to illustrate how AI can assist in analysis. Each assignment will clarify how AI should be used. Initial assignments on a topic will require you to use a particular program without AI assistance. Working through the process through traditional methods is an important step in learning analytics. Without it, you will not be able to evaluate the quality of the AI output, its strength and shortcomings.

GRADING
All assignments will be available on Canvas.

Assignments (50%)
You will have two different assignments. The pre-class assignments will include pre-class videos, questions about the content of the videos and exercises that will help you learn the basic skills that will be covered in class. In-class assignments are similar to the one demonstrated in the instructional videos but will be used to demonstrate proficiency. You will be required to submit those assignments to demonstrate proficiency with the skills.

Project (50%)
This semester's project will be on utilities data from UNT Facilities, who are our client. They are interested in understanding variability in utilities usage across buildings and time, with an eye to improving efficiencies. The project will be structured as both a group and individual project. Each student will be responsible for working with the data from one building. Groups will be formed around building type (e.g., dorms, academic, service). Thus, the analysis will dive deep into the utilities usage patterns of individual buildings, and then you will come together in a group to analyze similarities and differences within that building type. The group will also evaluate when and why utilities usage is high and what environmental or building factors may improve efficiency. The final product will be an individual story map for each building and a group story map for each building type.

COURSE TOPICS (subject to change)

Introduction
   Wk 1 Course Introduction; What is Data Science/Analytics?
   Wk 2 Generating Data

Section 1: Analytics Process Using Excel
   Wk 3 Data Acquisition and Preparation
   Wk 4 Data Organization and Management
   Wk 5 Exploratory Data Analysis
   Wk 6 Visualization
   Wk 7 Visualization Output

Section 2: Alternative Methods
Data Preparation
   Wk 8 Alteryx

Data Management
   Wk 9 Databases and SQL

Data Analytics
   Wk 10 EDA and ESDA in Power BI
   Wk 11 EDA with Tableau
   Wk 12 EDA and ESDA with GeoDa

Data Visualization

Wk 13 Tableau Dashboards
Wk 14
Wk 15 Final Project