

EDLD 652: Introduction to Course

Cengiz Zopluoglu

College of Education, University of Oregon

Jan 9, 2023

Eugene, OR

Today's Goals:

- **Know your classmates**
- **Know your instructor**
- **Course overview**
- **The RGraph Gallery**
- **RTutor**
- **Introduction to Processing Text Data**

Know Your Classmates

- Go to the course website
- Click **Discussions** tab.
- Find the discussion **Introduce Yourself to Your Classmates**
- Submit a text that describes yourself in 15 minutes.
- By the end of week, respond to at least two of your classmates.
- This is a graded activity. You will receive 1% extra credit as part of this activity.

15:00

Know Your Instructor

- How to pronounce my name?

<https://namedrop.io/cengizzopluoglu>

- Call me whatever you find most convenient (with one exception). I don't have a preference.
 - Cengiz
 - Dr. Cengiz
 - Dr. Z
 - Professor
 - ~~Hey Doc~~
- Pronouns: He/Him/His
 - My native language (Turkish) is a gender-neutral language so please excuse my misuse of pronouns in daily language. My apologies in advance.

A little bit about myself...

- was born and raised in Istanbul (Turkey),
- received my BA in Mathematics Teaching from Abant Izzet Baysal University (Bolu, Turkey) in 2005,
- taught math for a year at 6th grade,
- came to the United States in 2007 for graduate education with a loan from Turkish government,
- received my degree in 2013 from University of Minnesota,
- was a faculty member at the University of Miami (2013-2020),
- moved to Eugene in Summer 2020,
- am married with two kids (8 and 11).

Course Overview

This review doesn't replace reading the syllabus!

Course Modality

- This course is an in-person course: which means that, unlike asynchronous online/ASYNC WEB courses, we will meet during scheduled class meeting times in Lokey 115.
- Having said that, please do not attend in person if you are not feeling well, even if you think it is not COVID.
- There is no Zoom link to attend the classes synchronously.
- The lectures will be recorded using Panopto and the links will be posted on CANVAS.
- If you need accommodation related to a medical or other disability, you can set those up through AEC (<https://aec.uoregon.edu>).

Course Description

EDLD 652: Data Visualization for Educational Data Science

- is the second course in a sequence of courses on educational data science (EDS) specialization,
- is taught using free and open-source statistical programming languages,
- aims to provide students with the foundational principles and practice of data visualization, mainly applied to scientific and technical data
- The primary learning outcomes:
 - transform data in a variety of ways to create effective data visualizations
 - understand and be able to apply basic string operations and work with textual data
 - understand best practices in data visualization
 - customize ggplot2 graphics by reordering factors, creating themes, etc.
 - create an online data visualization portfolio using [distill](#), [flexdashboard](#), [pagedown](#), or [bookdown](#) to demonstrate key learning

I assume you

- understand the R package ecosystem (how to find, install, load, and learn about them)
- Can read "flat" (i.e., rectangular) datasets into R
 - should be using RStudio Projects & the [here](#) package
 - see [Jenny Bryan's blog post](#) for why.
- can perform basic data wrangling and transformations in R, using the tidyverse or base packages
 - leverage appropriate functions for introductory data science tasks (pipeline)
 - "clean up" the dataset using scripts and reproducible workflows
- use version control with R via git and GitHub
- use R Markdown to create reproducible dynamic reports
- [EDLD 651 by Dr. Nese](#)

Examples

Below are some links to final projects from students who have taken this class previously.

Dashboards

- [Alexis Adams-Clark](#)
- [Brendan Cullen](#)

Blog post

- [Teresa Chen](#)
- [Ouafaa Hmaddi](#)
- [Murat Kezer](#)

Office hours

- Online office hours:
 - Monday, 4pm – 5pm (in-person & Zoom, check the syllabus for the link)
 - If possible, notify me in advance to avoid waiting time.
- Please use Canvas course website for communication related to this class.

Textbook

All course readings are freely available online

- [Happy git with R, Jenny Bryan](#)
- [R for Data Science, Hadley Wickham Garrett Grolemund](#)
- [Data Visualization: A Practical Introduction, Kieran Healy](#)
- [Fundamentals of Data Visualization, Claus O. Wilke](#)

Course Structure (see the syllabus for details)

- In-person lectures
- Readings
- Four Lab Assignments (40%, 10% each)
- Data visualizations 'in the wild' presentation (5%)
- A final project (55%)

Lab assignments

- Labs must be submitted within one week of being assigned.
- Labs will include designated in-class time (if possible), but any work not completed during class must be completed on your own.
- Labs will be scored on a “best honest effort” basis, which generally implies zero or full credit.
- Assignments will be posted as .Rmd files. Each student must turn in the completed lab work electronically on Canvas as a Word or Pdf file (after knitting your .Rmd file)
- If you are hosting your work online as a GitHub repo, you can also post the link to the GitHub repo

Data visualizations 'in the wild' presentation

- On the first day of the class, you will be randomly assigned a date to share data visualizations you have found in publications or websites or anywhere else IRL (in real life).
- This is not a formal presentation. You share the links to the data visualizations with me, and I will project them.
- There will be a discussion post on Canvas dedicated to each week. When it is your turn, you must post your reflection and discussion here before the class.
- You will then present your discussion about what you like/dislike about the plot you chose.
- We will then have a brief discussion as a class.
- Note that this project is aimed at identifying a range of data visualizations. It would be ideal if both good and poor data visualizations were presented from various sources.

Presentation Order

Date	Presenters
Week 4 - Jan 30	Tony, Adriana, Deanna
Week 5 - Feb	Lucy, Dominik, Santiago
Week 6 - Feb 13	Mallory, Sam, Seulbi
Week 7 - Feb 20	Tram, Sabreen, Tian
Week 8 - Feb 27	Maria, Denicia, Zhen
Week 9 - Mar 6	Lauren, Haitham, Brittany

These are already assigned on Canvas.

Final Project

- The final project includes multiple components, culminating in a data visualization portfolio.
- The portfolio must include at least three data visualizations, communicate the history of each visualization, and be deployed to the web.
- The due dates for each component are as follows:
 - Proposal: Mon, Jan 30 (Week 4)
 - Draft: Mon, Feb 20 (Week 7)
 - Peer Review: Mon, Mar 6 (Week 9)
 - Presentation: Mon, Mar 13 (Week 10)
 - Product: Mon, Mar 20

The due date for the proposal can be (somewhat) flexible. However, unlike the labs and homework, the remaining aspects of the final project cannot be changed, and you will lose points if your work is submitted late. It is primarily because of concerns about peer review and complete the project by the end of the term.

Review the document for final project on Canvas

BREAK

15:00

The R Graph Gallery

RTutor

Introduction to textual data

Review of R script on the website

Web Scraping Practice In Case of Additional Time