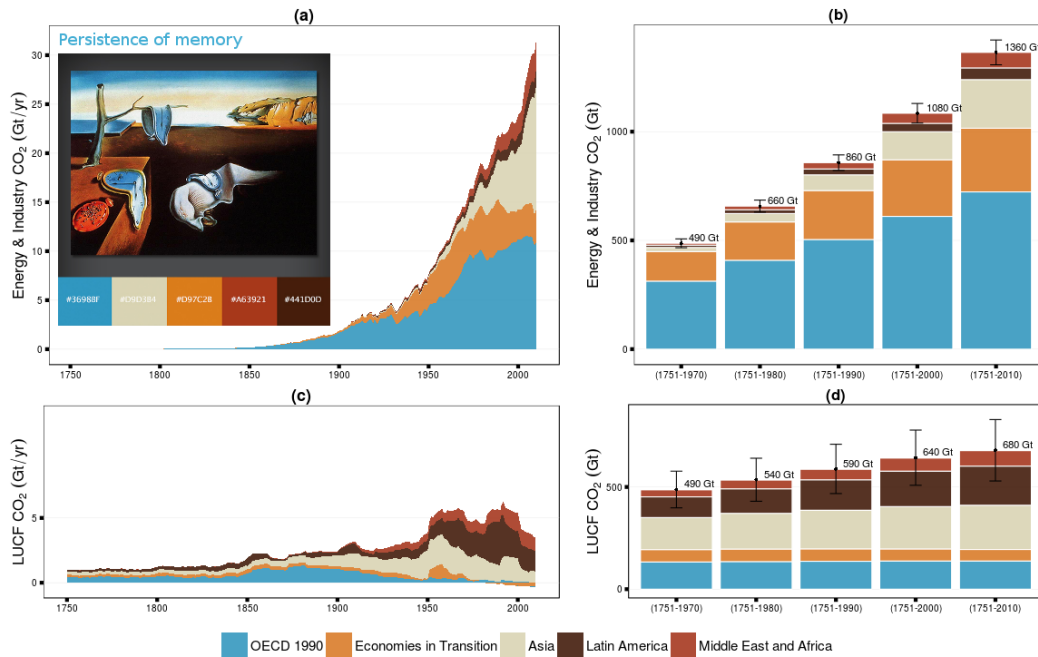# Syllabus Geography GEOG788Z: Topics in Data Visualization with Processing and Applications in R

Instructor: Giovanni Baiocchi (ʤoˈvanni baiɔkki)

- **Office**: 1133 LeFrak
- **Office Hours**: Thu 16:00–17:00
- ☎ +1 (301) 405 2342   ⊡ FAX +1 (301) 314 9299.   ✉ Baiocchi@umd.edu

1 Course aims and objectives

The aim of this visualization course is to effectively communicate information through graphical means. This course aims to provide an understanding of the theory underpinning the field of data visualisation and to provide an introduction to the practicalities of creating effective graphical representations of data.

Upon completion of this course students should be able to:

- Process different kind data for effective visualization using modern data science workflow tools supported in R

- Produce accurate, clear, and reproducible graphics using the R statistical software

- Critique visualizations and inform its design following Tufte's principles

- Design effective visualizations that tap into human's high bandwith channels to the cognitive centers

- Deconstruct visualizations following the Grammar of Graphics principles

- Design multi-plot displays (panels and posters) following graphic design principles

- Construct interactive data visualizations according to interaction design principles

This course combines a mix of lectures, practical work, and informal discussions around important topics in data visualisation. Students will be required to critique existing visualisations, find interesting data sets, and create their own visualisations.

2 Software

This course will use the popular open source statistical computer language R and packages (such as, tidyverse, knitr, and shiny) for practical workshops and assignments. Additional software that will be useful includes: GIMP, Inkscape, ImageJ, and geogebra.

2.1 R

The required software for this class is the **R statistical language**. R is the open source, freeware version of S (The commercial implementation is Splus). R is one of the most powerful and versatile statistical environments, and is available for free download for use on PC, Mac, UNIX and Linux systems. The software is available in the Lefrak/Geography Open Labs. If you have a laptop or home computer, get this software and download it immediately. An online tutorial of basic functions and syntax: http://www.ats.ucla.edu/stat/r/

There are many variants of the R computing environment (e.g. GUI-based). Please feel free to use whatever version you are most comfortable with on your personal computer. We recommend, however that you use **R Studio**.

3   Prerequisites

- Basic R

- Work with own computer

4   Course Material

There is no required text. Selected readings peer-reviewed journals and other material will be assigned.

4.1   Recommended textbooks

Not required, but mostly available free via UMD library website:

- Wickham, Hadley. *Advanced R*. Boca Raton, Florida: CRC Press. Available free on line: http://adv-r.had.co.nz/.

- Wickham, Hadley, und Garrett Grolemund. 2016. *R for Data Science: Visualize, Model, Transform, Tidy, and Import Data*. O'Reilly Media. Available free on line: http://r4ds.had.co.nz/.

- E. Tufte, *The Visual Display of Quantitative Information*, Graphics Press, 2001.

- W. S. Cleveland, *The Elements of Graphing Data*, Hobart Press, 1994.

- Wickham, Hadley. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2026. http://ggplot2.org.

- L. Wilkinson, *The Grammar of Graphics*, Springer-Verlag, 1999.

- C. Ware, *Information Visualization: perception for design*, 3rd Edition, Morgan Kaufmann, 2012. Available as ebook from the UMD library: https://umaryland.on.worldcat.org/oclc/784147337

5   Class Organization

The class will meet for 2+ hours every week (Tue 13:00 pm – 15:30 pm, room LEF 1171) . There will be a first part based more on lectures and in class discussions (first hour or so mode lecture, second hour practical). Later in the course emphasis will be on student presentations.

6   Assessment and Grading

Grades will be determined by a series of homework projects and student led discussions. You will complete up to 7 assignments: 5 focused homeworks and 2 self-directed projects.

- Lab assignments using R (60%)

- Independent project on a relevant visualization issue and presentation (40%)

Assignment plan (tentative)	
TBA	HW1: Basic R
TBA	HW2: Tidy data in R
TBA	HW3: Design Critique
TBA	HW4: Visual perception
TBA	HW5: Grammar of Graphics
TBA	Project 1: Research and communication
TBA	Project 2: Interactive graphics

7   Assessment Guiding Principle

This visualization course will employ a competency-based grading approach system that should provide a better control over the final grade and an improved transparency over the process. I will provide feedback on selected submission, and you will have the opportunity to resubmit a revised version.

8   Examination Dates

There are no traditional examinations (see Assessment above).

There may be unannounced in-class quizzes. The frequency of these is determined by our ability to stay on top of the required reading and understanding of concepts. Depending on how many are given they could count from 0% - 15% of the course. The weighting of the other elements will be proportionately reduced to achieve a total percentage of 100%. Thus, while class attendance is not mandatory, failure to attend may impact your grade because of missed quizzes.

9   Intended Course Coverage

- Data processing with R

- Data Visualization principles

- Grammar of graphics topics

- Special issues in visualizing, missing data, big data, time series, spatial data, social data.

10   Timetable (Tentative)

Basic R and Data Processing				
	Key Date		**Lecture Topic**	**Readings and assignments**
Week 1	31 Aug	5:00pm–7:00pm	Course introduction; setting up R, reproducible research in R, Rstudio, R markdown, and knitr	NA
Week 2	7 Sept	5:00pm–7:00pm	Basic R.	Set up R system; Read Foundations chapter sections on Data structures and Subsetting from *advanced R book*; read chapter 1 of *R for Data Science*
Week 3	14 Sept	5:00pm–7:00pm	Tidy data; Data Processing with dplyr and tidyr	HW1: TBA
Visualization Principles				
Week 4	21 Sept	5:00pm–7:00pm	Critique visualizations and inform its design following Tufte's principles	
Week 5	28 Sept	5:00pm–7:00pm	Design effective visualizations that tap into human's high bandwith channels to the cognitive centers	
Week 6	5 Oct	5:00pm–7:00pm	Grammar of Graphics principles Implementation in R: ggplot	
Design Principles				
Week 7	12 Oct	5:00pm–7:00pm	Color theory; Slippery slopes; Bubble trouble	
Week 8	19 Oct	5:00pm–7:00pm	Interactive graphics	
Week 9	26 Oct	No class	Fall Break	
Research and Communication				
Week 10	2 Nov	5:00pm–7:00pm	Big Data	
Week 11	9 Nov	5:00pm–7:00pm	Missing Data, Publication quality graphics	
Week 12	16 Nov	5:00pm–7:00pm	Maps; Projections	
Week 13	23 Nov	No class	Holiday - campus closed	
Special Topics and Presentations				
Week 14	30 Nov	5:00pm–7:00pm	TBA	
Week 15	7 Dec	5:00pm–7:00pm	TBA	

11 Official Notices

- Attendance and participation : You are strongly advised to attend and participate in all lectures since this will provide a basic understanding of the subject matter of the course particularly as no text book is assigned. The course grade is dependent upon successful completion of written papers. Active participation in discussions is an important part of this class. The campus policy on excused absences is stated on page 37 of the University of Maryland Undergraduate:

  (http://www.umd.edu/catalog/0405/chapter4.pdf)

- Disabilities : If you have a documented disability and wish to discuss academic accommodations, please contact the Instructor as early as possible. I will make every effort to accommodate students who are registered with the Disability Support Services (DSS) Office and who provide me with a University of Maryland DSS Accommodation form which has been updated for the 2012 Spring semester. This form must be presented to me no later than February 1, 2012. I am not able to accommodate students who are not registered with DSS or who do not provide me with documentation which has been reviewed by DSS.

- Academic Integrity Expectations : 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://www.studenthonorcouncil.umd.edu/whatis.html

  By registering for this course you indicate your acceptance of these provisions for academic integrity.

- Students are expected to treat each other with respect. Disruptive behavior of any kind will not be tolerated. Students who are unable to show civility with one another, or myself will be subject to being referred to the Office of Student Conduct or to Campus Police. You are expected to adhere to the Code of Student Conduct.

- Campus Senate policy requires students who are absent due to illness/injury to furnish documentary support to the instructor. I require students to contact me by email or by phone, where possible, prior to class time in which you indicate that you have an illness or an injury. You must provide written documentation verifying your illness/injury immediately upon your return to class. You will not be allowed to turn in missed assignments or make up quizzes, tests, papers, etc. if you have not provided this documentation. Documentation not presented to me in a timely manner will not be accepted. In addition, if it is found that you have falsified the documentation provided, I will refer you to the University's Student Conduct Office.

- **CourseEvalUM**. Your participation in the evaluation of courses through CourseEvalUM is a responsibility you hold as a student member of our academic community. Your feedback is confidential and important to the improvement of teaching and learning at the University as well as to the tenure and promotion process. The date from which CourseEvalUM is open for you to complete your evaluations will be announced by the University. Please go directly to the website (www.courseevalum.umd.edu) to complete your evaluations by the requested date. By completing all of your evaluations each semester, you will have the privilege of accessing online, at Testudo, the evaluation reports for the thousands of courses for which 70% or more students submitted their evaluations.

- 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://www.shc.umd.edu.