

# GTECH 38520 / 78520

## Data Analysis and Visualization with R

Spring 2021, Tuesday 2:10 - 5:00 PM online  
Instructors: Allan Frei, Peter Marcotullio, Shipeng Sun

### Instructor Information

#### Allan Frei

Office: 1005, Hunter North  
Email: afrei@hunter.cuny.edu

Office hours: See Blackboard  
Phone: 212-772-5322

#### Peter J. Marcotullio

Office: 1003e, Hunter North  
Email: pmarcotu@hunter.cuny.edu

Office hours: See Blackboard  
Phone: 212-772-5264

#### Shipeng Sun

Office: 1043, Hunter North  
Email: shipeng.sun@hunter.cuny.edu

Office hours: See Blackboard  
Phone: 212-396-6039

### Course Description and Objectives

This course introduces the basics of R and the practical knowledge of data cleaning, re-organization, modeling, statistics, and analysis for research and visualization, particularly in geospatial fields. The class meets once a week for 3 hours per session with hands-on practices. Three different members of the Department of Geography and Environmental Science, as identified above, lead the class. The goal of the course is to introduce students to the use of R programming for univariate and multivariate analysis and visualization, mapping and spatial analysis.

### Course Online Methodology

This course will be completely online and taught in a synchronous fashion. This means that lectures will be live. We encourage student participation by asking questions and responding to instructor questions. There will also be exercises in breakout groups during class which are considered important for student learning.

### Learning Outcomes

At the end of the semester, students will be able to

- 1) Use RStudio to perform basic data analysis functions including Input/Output, basic Exploratory Data Analysis (EDA), and graphical output.
- 2) Use RStudio to develop, test, and execute R script.
- 3) Use advanced R programming to import, clean, transform, and summarize data
- 4) Use *ggplot2* to visualize data in points, lines, area charts and smoothed curves
- 5) Import and map spatial data using R *sf* and *ggplot2* packages
- 6) Conduct basic cluster analysis and regression using spatial data

## Pre-requisite

Required: STAT 113. Suggested but not required: GTECH 201, GTECH 301, or GTECH 702  
A computer language background is not required.

## Schedule

The course has 6 learning sections, one midterm, and one final exam. Each professor teaches two of the following sections.

1. Introduction to R, getting started, Exploratory Data Analysis (EDA)
2. Exploration and Manipulation of Large Data Sets,  
<https://www.analyticsvidhya.com/blog/2015/12/faster-data-manipulation-7-packages/>
3. Efficient management and manipulation of data frames
4. Visualization of non-spatial data with *ggplot2* package
5. Spatial data organization, I/O, format conversion, and spatial reference systems
6. Visualization of spatial data, mapping, and basic spatial analysis

## Course Materials

Required Textbook: None

Recommended Books:

- *ggplot2, Elegant Graphics for Data Analysis* (2<sup>nd</sup> Edition), by Hadley Wickham, Springer, (2016);
- *R for Data Science, Import, Tidy, Transform, Visualize and Model Data*, (1<sup>st</sup> Edition) by Hadely Wickham and Garrett Grolemund, O'Reilly (2016).
- *Geocomputation with R* by Robin Lovelace, Jakob Nowosad, Jannes Muenchow (2019). Available at <https://geocompr.robinlovelace.net/>.
- *Spatial Data Science with R* by Robert J. Hijmans (2019). Available at <https://rspatial.org/raster/index.html>

## Software

- R 4.0 or above. Download free at [www.r-project.com](http://www.r-project.com). Must be installed prior to installing RStudio
- RStudio Desktop 1.3 or above (Open-Source License, not the Commercial License). Download Free at [www.rstudio.com](http://www.rstudio.com)

## Grading Scale

Academic performance evaluation criteria include assignments, exams and online activity participation with breakdown as follows.

Components	Percentage
Lab Assignments	60%
Exams	30%
Participation	10%

## Assignments

Every week students will be assigned assignments to practice the skills learned in the previous lesson. Assignments are due prior to the beginning of the next class. NO LATE ASSIGNMENTS WILL BE ACCEPTED UNLESS EXPLICITELY GRANTED BY THE INSTRUCTOR.

## Course Calendar & Content

This schedule is subject to change.

Wee k	Session	Date	Topic	Instructor
1	1	02/02/21	Introduction to this course; Introduction to R and RStudio	Frei
2	2	02/09/21	Univariate Exploratory Data Analysis	Frei
3	3	02/16/21	Bivariate Exploratory Data Analysis	Frei
4	4	02/23/21	Graphical Analysis in R Base Package	Frei
5	5	03/02/21	Applying what we have learned so far to a research question	Frei
6	6	03/09/21	<i>tidyverse</i> , Part I: Management and manipulation of data frames	Marcotullio
7	7	03/16/21	<i>tidyverse</i> , Part II: Management and manipulation of data frames	Marcotullio
8	8	03/23/21	<b>Midterm</b>	
9		03/30/21	<b>Spring Recess</b>	
10	9	04/06/21	<i>ggplot2</i> , Part I: Visualization of non-spatial data	Marcotullio
11	10	04/13/21	<i>ggplot2</i> , Part II: Visualization of non-spatial data	Marcotullio
12	11	04/20/21	GIS Basics, <i>sf</i> package, spatial data types, spatial data I/O	Sun
13	12	04/27/21	Re-projecting and mapping spatial data with <i>ggplot2</i>	Sun
14	13	05/04/21	Spatial Join and Basic Spatial Analysis	Sun
15	14	05/11/21	Application cases of R-Spatial	Sun
16		05/18/21	<b>Reading Day</b>	
17	15	05/25/21	<b>Final Exam</b>	

## **Policies**

### Attendance

A significant part of your grade depends upon class participation. Class participation includes attendance at online live lectures, participation in organized class discussions, and timely submission of assignments and exams. We expect respectfully behavior towards instructors and the other students in the virtual world, through non-threatening language during discussions, and by allowing others to speak.

### Course Blackboard Site

*Web-enhancement* in the context of this course means that everything pertaining to this course will be communicated through BlackBoard. You are required to check the BlackBoard course site on a daily basis. All changes to the syllabus will be announced on the course home page. All lecture and lab materials are accessible through BlackBoard, and Blackboard is also the place where you upload your assignments. Your exams and lab assignments will be graded based on what you have uploaded to BlackBoard and that is where you will find your grades and may access course statistics that help you to assess your standing at any given time.

### Communication

Students should try to communicate via official Hunter or CUNY emails. Announcements will also be sent through the emails registered in the Blackboard system. Make sure your email there is updated. Many are still having the emails that they used for admission applications. Professionalism and "netiquette" are expected in the communication through emails (check out those links). If your emails are not replied to in a timely fashion, please consider rewriting your emails in a better way.

### Hunter College Policy on Academic Integrity

Hunter College regards acts of academic dishonesty (e.g., plagiarism, cheating on examinations, obtaining unfair advantage, and falsification of records and official documents) as serious offenses against the values of intellectual honesty. The College is committed to enforcing the CUNY Policy on Academic Integrity and will pursue cases of academic dishonesty according to the Hunter College Academic Integrity Procedures.

### ADA Policy

In compliance with the American Disability Act of 1990 (ADA) and with Section 504 of the Rehabilitation Act of 1973, Hunter College is committed to ensuring educational parity and accommodations for all students with documented disabilities and/or medical conditions. It is recommended that all students with documented disabilities (Emotional, Medical, Physical, and/or Learning) consult the Office of AccessABILITY, located in Room E1214B, to secure necessary academic accommodations. For further information and assistance, please call: (212) 772-4857 or (212) 650-3230.

## Hunter College Policy on Sexual Misconduct

In compliance with the CUNY Policy on Sexual Misconduct, Hunter College reaffirms the prohibition of any sexual misconduct, which includes sexual violence, sexual harassment, and gender-based harassment retaliation against students, employees, or visitors, as well as certain intimate relationships. Students who have experienced any form of sexual violence on or off campus (including CUNY-sponsored trips and events) are entitled to the rights outlined in the Bill of Rights for Hunter College.

a. Sexual Violence: Students are strongly encouraged to immediately report the incident by calling 911, contacting NYPD Special Victims Division Hotline (646-610-7272) or their local police precinct, or contacting the College's Public Safety Office (212-772-4444).

b. All Other Forms of Sexual Misconduct: Students are also encouraged to contact the College's Title IX Campus Coordinator, Dean John Rose (jtrose@hunter.cuny.edu or 212-650-3262) or Colleen Barry (colleen.barry@hunter.cuny.edu or 212-772-4534) and seek complimentary services through the Counseling and Wellness Services Office, Hunter East 1123.

CUNY Policy on Sexual Misconduct Link:

<http://www.cuny.edu/about!administrationloffices/Ja/Policy-on-SexualMisconduct-12-1-14-with-links.pdf>

## Syllabus Change Policy

Except for changes that substantially affect implementation of the evaluation (grading) statement, this syllabus is a guide for the course and is subject to change with advance notice. Changes will be announced in class and on Blackboard, which students are expected to check regularly during the semester.