

# HARMLESS DATA SCIENCE PROGRAMMING IN R FOR BEGINNERS

January 2021

---

<b>Instructor:</b>	Dr. Howard Liu	<b>Time:</b>	14:00 – 18:00 (GMT)
<b>Email:</b>	<a href="mailto:howard.liu@essex.ac.uk">howard.liu@essex.ac.uk</a>	<b>Place:</b>	Zoom

---

**Objectives:** This five-day workshop is designed to prepare students for future courses at the University of Essex when they have no previous programming experience in R. The instructor will teach you how to run R in RStudio (a R programming environment), how to work with basic objects and run operations, and also how to create/merge/subset a data frame. We will also cover some basic data visualization techniques. The workshop has two components daily. It will start with a programming lecture session and then a practice session where students' questions can be answered interactively and in real-time.

**Slack:** Using Slack, a online chat platform, to teach programming has been proven to be very successful both in my own experiences and the others. After having your email addresses, I will enroll you onto our Slack workplace where you can ask questions by pasting your scripts as well as error messages. The instructor and a teaching assistant would be able to help out in real-time.

**Office Hours:** Right after the lecture. The instructor will be available on Zoom as well as on Slack.

## Course content

### Day 1 (Jan. 10th): Introduction and R Basics

- Get familiar with RStudio Interface
- Basic mathematical operations
- R data types (e.g., character, numeric, logical, factor)
- Object and variable assignments
- Testing for missing data

### Day 2 (Jan. 11th): Common Functions in R (Essential for Daily Programming)

- if-else and else-if statements
- ifelse() function
- for Loop
- while Loop
- Break & next statement
- Creating functions
- the apply functions

### Day 3 (Jan. 12th): Basic Data Structures (Vector, Matrix, and List)

- Creating a Vector (a collection of “things”)
- Vector operations
- Creating a Matrix (a collection of “things” arranged in the format of rows and columns)
- Matrix operations
- Creating a List (a collection of “things” arranged in the format of an ordered list)
- Subsetting or slicing List

### Day 4 (Jan. 13th): Data Frame

- Import and export data frames
- Working with data frames
- Subsetting
- Merging
- Doing data manipulation with the easy-to-use dplyr function
- Pipeline operator

**Day 5 (Jan. 14th): Data Visualization and Beyond**

- Drawing nice figures with the friendly ggplot function
- Scatter plots
- Histograms
- Visualizing time-series data
- Visualizing cross-sectional data