DSA2101

Essential Data Analytics Tools: Data Visualization

Yuting Huang

Course Policies

The teaching team

Lecturer:

- Dr. Huang Yuting (yhuang@nus.edu.sg)
- Office: S16 04-01
- Office hour: by appointment

Teaching assistants (TAs): Consultation hours of the TAs will be announced around Week 3.

	NUS email	Consultation hours
Shim Jaejun	e0878501@u.nus.edu	
Pan Yuting	pan.yuting@u.nus.edu	
Zhang Simian	zhang.simian@u.nus.edu	

Topics covered

Here is a list of topics we plan to cover this semester:

1. `R` programming
2. Importing data into `R`
3. Data manipulation with `R`
4. Principles of data visualization
5. The grammar of graphics
6. Exploring data through visualization

Tentative teaching plan (Subject to change)

- Weeks 1–2: Basics in `R` programming
- Weeks 3–4: Importing data into `R`
- Week 5: Data manipulation
- Week 6: Tidy data
- Week 7: Relational data
- **Week 8: In-class midterm test**
- Week 9: Principles of data visualization
- Weeks 10–11: The grammar of graphics
- Weeks 12-13: Exploring data through visualization

Lectures

- Tuesdays and Fridays **from 8 to 9:30am** at LT32.
- Although lectures will be recorded and web-cast, I encourage you to attend the live lectures in person.

Tutorials will begin in Week 3

- We have six tutorial slots in each week.
- **Tutorial attendance is compulsory**. You will also need to bring your own laptop.
- Tutorial worksheets will be released one week in advance. You are expected to **work on the worksheet beforehand**.
- During the actual tutorial, you may be randomly selected to share your code/plots/analyses.

Evaluation components

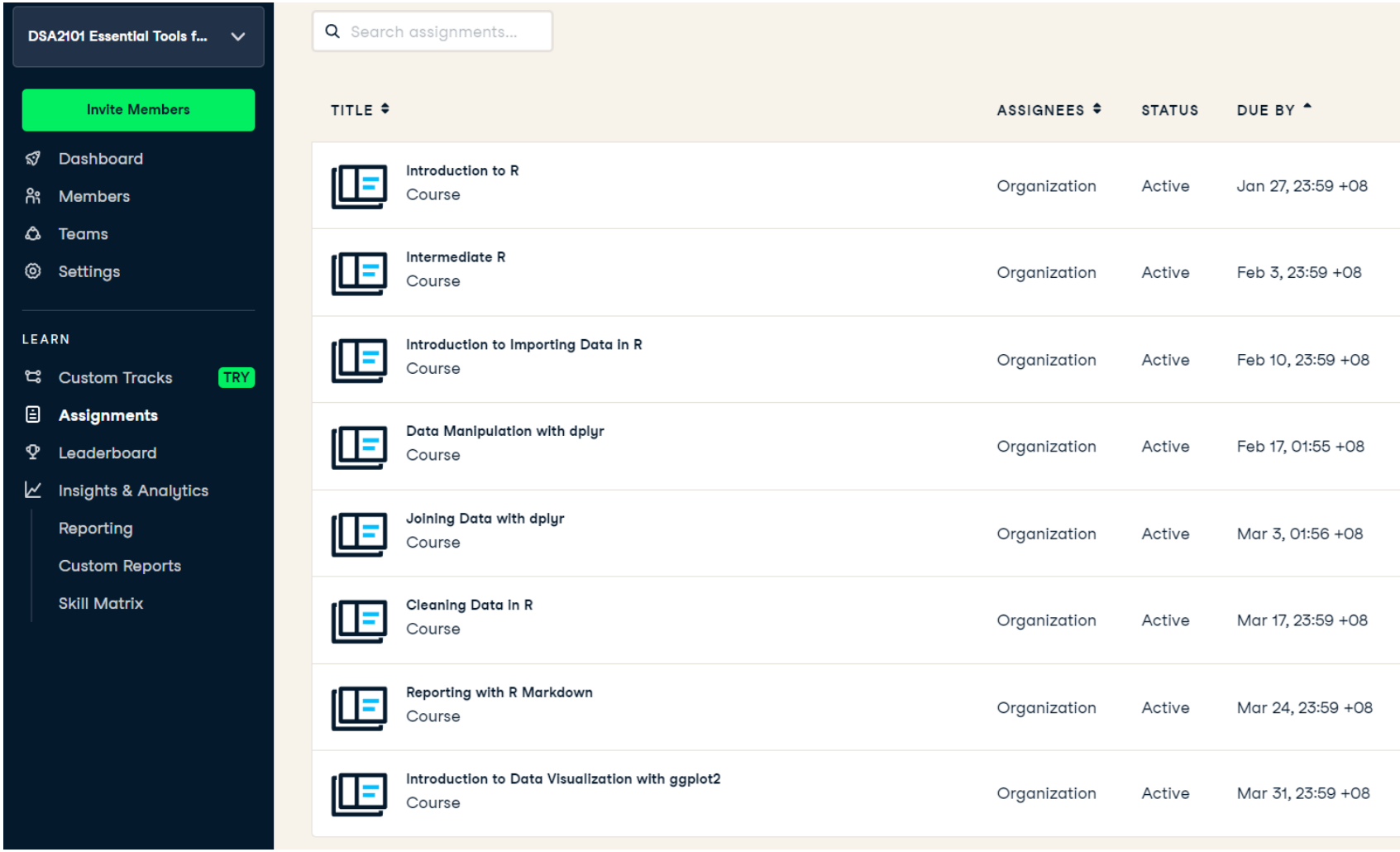
- Tutorial (attendance + presentation) 5%
- DataCamp assignments 10%
- Group project 15%
- Midterm test 30%
- Final exam 40%

Activate DataCamp Classroom account

As part of the class, you will be required to complete DataCamp assignments.

- The activation link is now available on Canvas. Use it to activate your DataCamp Classroom account.
  - You may be asked to sign in with your NUS email (with domain @u.nus.edu).
- After activation, you will have access to most courses there for 6 months. Do make full use of it!

For each DataCamp assignment, you will receive a binary grade - either full or none. They are graded based on completion by the due date, not by XP points.



Group project

Project guideline will be released after the midterm exam.

- You will work as a team of **up to 5 persons**.
- More details will be announce in due course.

Late submission policy

- Late submissions will not be graded.
- Email submissions will not be graded.
- Only submissions through Canvas will be graded.

Software

For this class, we will be using the following software.

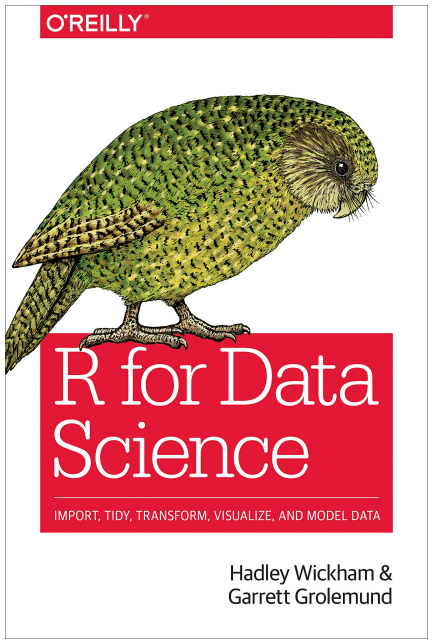
1. `R`
2. RStudio
3. Examplify

Make sure that you have a laptop (Windows or Mac), and the latest versions of software above.

Main references

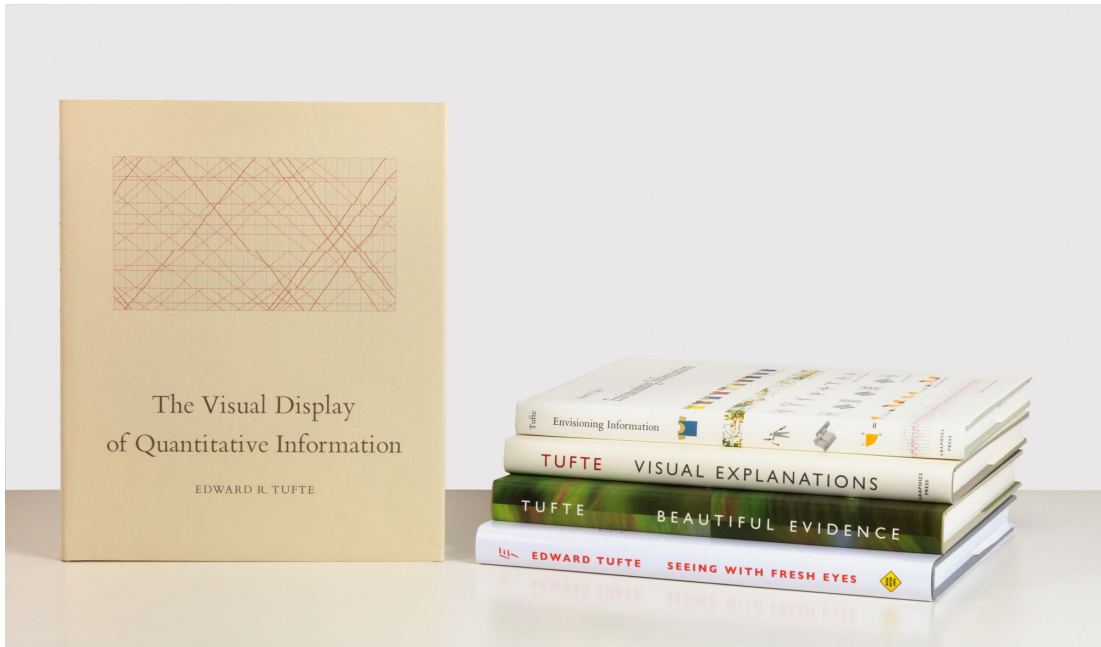
R for Data Science - by Hadley Wickham & Garrett Grolemund

- Online textbook: https://r4ds.had.co.nz/
- Hard copy is also available for purchase.
- Great resource for `R` using the `tidyverse` syntax.
- Lots of exercises and examples.



Due to time constraint, we won't be able to cover all the chapters. If you are serious about the field of data science, we encourage you to read through the book as carefully as you can.

The Visual Display of Quantitative Information - by Edward Tufte.



- Edward Tufte has written a series of books on presenting information visually.
- We will focus on some of the visualization principles he describes.

How to reach us?

You can talk to us before/after lectures and tutorials.

- Email us using your NUS email account.

How to ask a coding question?

- Copy and paste your code and the entire error message.
- A code that allows us to reproduce the error is ideal.
- **Do not send screenshots.**