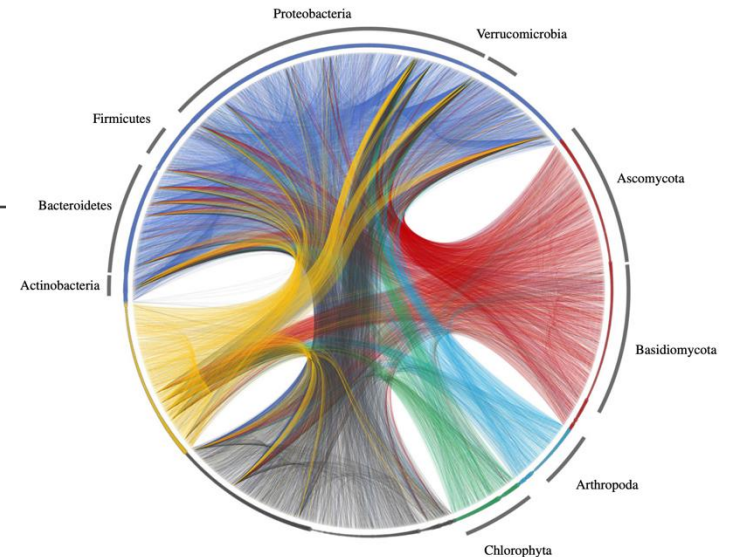
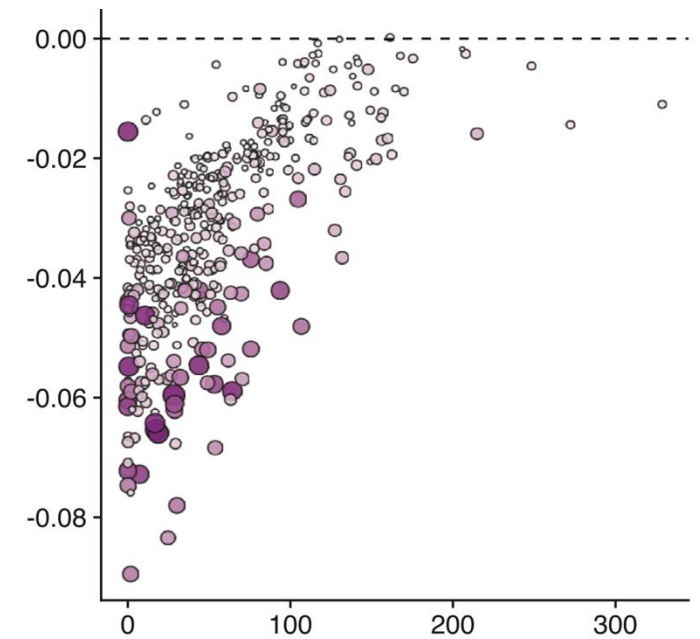


DASC 3240

Data visualization in Science

2024/25 SPRING SEMESTER

MASAYUKI USHIO, ASSISTANT PROFESSOR AT OCES
USHIO@UST.HK (RM CYT-2013)



Instructor

- Masayuki USHIO (ushio@ust.hk, Room CYT-2013)
- Assistant Professor at Department of Ocean Science (OCES)
- Ph.D (Science) at Kyoto University, Japan
- Joined OCES in September 2022



@HKUST



@Borneo





<https://www.youtube.com/watch?v=TQdTV1rAIWY>

Chem7, CC-BY2.0 (up)
Huangdan2060, CC-BY3.0 (bottom)

Today's outline

1. What is “Data Visualization”?
2. Course information
3. Questionnaire
4. (+ Basics in figure and data presentation)

What is “Data Visualization”?

Data Visualization

- “*Data visualization is the art and science of telling captivating stories with data.*” (Belorkar et al. 2022)
- Scientists have to communicate insights effectively using data visualization (Belorkar et al. 2022).
- Data visualization is the final step of the process of **collecting, analyzing, and summarizing data**. Its purpose is to extract and summarize essential information about the data and effectively convey a key message to the audience.
- A good visualization is necessary for good science. A bad visualization fails to convey a key message, or may even lead to a **misunderstanding** of the data. Conversely, understanding data visualization is also necessary to detect misuse and misinterpretation of data.

Intended Learning Outcomes (ILOs)

- Elaborate the basic concept of visualization tools.
- Do simple data processing given a variety of data formats with R/Python.
- Formulate a visualization solution to some real-data problems and interpret the results by graphs.
- Implement some visualization techniques with R or Python.
- Apply the conceptual and practical skills to interpret data in physics, chemistry, life science and ocean science as well as other disciplines with data.

Course information

Course assessment

- Weekly Assignments (7–8): 40% (5% each, but may be adjusted)
- Final projects: 60%

Course structure

- 1.5h lecture × 2 in each week (**Mon & Wed 10:30–11:50 @ CYT G009B**):
 - 1st lecture in each week: Lecture
 - 2nd lecture in each week: Mostly hands-on + Assignment
 - **Students will be asked to submit the assignment each week starting on 26 Feb**
- 1h tutorial in each week (Thu 18:00–18:50 @ Room 2304 [Lift 17/18]):
 - Q&A, troubleshooting, face-to-face discussion, etc.
 - Students who need technical help and/or want to learn more about data visualization may attend the tutorial sessions

Lecture schedule (Mon/Wed 10:30–11:50, CYT G009B)

Wk	L#	Date	Topic	Format	Assignments
1	1	3 Feb (Mon)	Course introduction Basics in figure and data presentation I	Lecture	
	2	5 Feb (Wed)	Basics in figure and data presentation II	Lecture	
2	3	10 Feb (Mon)	Basics in figure and data presentation – Exercise	Hands-on	
	4	12 Feb (Wed)	R and RStudio – I. Overview and installation	Lecture + hands-on	
3	5	17 Feb (Mon)	R and RStudio – II. R basics	Lecture + hands-on	
	6	19 Feb (Wed)	Data manipulation and ggplot2	Lecture + hands-on	
4	7	24 Feb (Mon)	Plotting data – Basics and more	Lecture + hands-on	
	8	26 Feb (Wed)		Assignments	Yes
5	9	3 Mar (Mon)	Scatterplot and regression	Lecture + hands-on	
	10	5 Mar (Wed)		Assignments	Yes
6	11	10 Mar (Mon)	Visualization of high dimensional data: PCA and others	Lecture + hands-on	
	12	12 Mar (Wed)		Assignments	Yes
7	13	17 Mar (Mon)	Visualization of spatial data: Maps and images	Lecture + hands-on	
	14	19 Mar (Wed)		Assignments	Yes

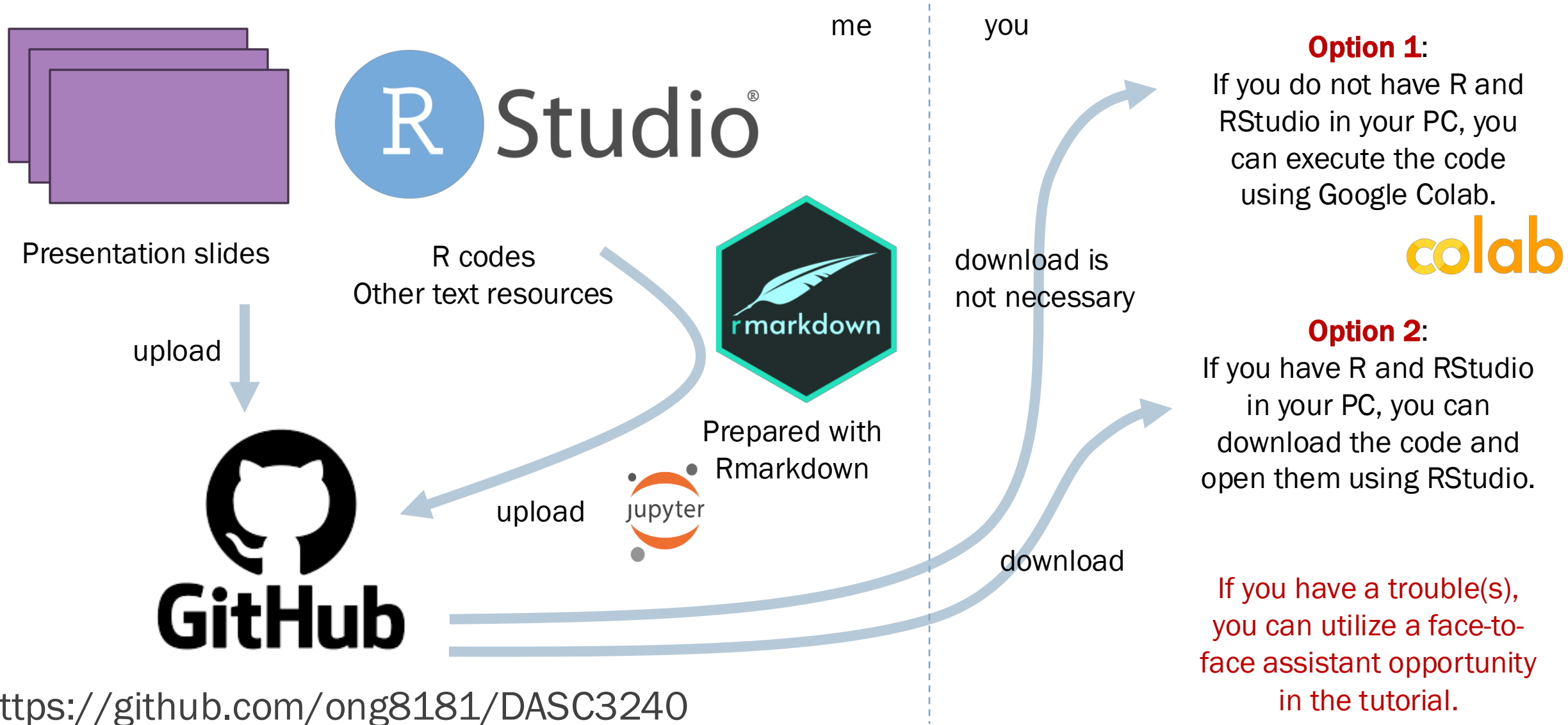
Lecture schedule (Mon/Wed 10:30–11:50, CYT G009B)

Wk	L#	Date	Topic	Format	Assignments
8	15	24 Mar (Mon)	Understanding variations and uncertainties in plots	Lecture + hands-on	
	16	26 Mar (Wed)	Understanding and interpreting plots	Lecture + hands-on	
9	17	31 Mar (Mon)		Assignments	Yes
		2 Apr (Wed)	Mid-term break	No class	
10	18	7 Apr (Mon)	Random topics in data visualization and programming	Lecture + hands-on	
	19	9 Apr (Wed)	plotly – Interactive plots	Lecture + hands-on	
11	20	14 Apr (Mon)	gganimation – Creating animation	Lecture + hands-on	
	21	16 Apr (Wed)		Assignments	Yes
12		21 Apr (Mon)	Easter Monday	No class	
	22	23 Apr (Wed)	Integrating R and Python	Lecture + hands-on	
13	23	28 Apr (Mon)	Preparing publication-ready plots	Lecture + hands-on	
	24	30 Apr (Wed)		Assignments	Yes
14		5 May (Wed)	Buddha's Birthday	No Class	
	25	7 May (Wed)	Course Review and Q&A session		

Tutorial (Thu 18:00–18:50, Rm 2304)

Wk	T#	Date	Topic	
1	1	6 Feb (Thu)	Setting up R/RStudio environment	Technical support for the setup
2	2	13 Feb (Thu)	Setting up R/RStudio environment	Technical support for the setup
3	3	20 Feb (Thu)	R and basic statistics	
4	4	27 Feb (Thu)	R and basic statistics	
5	5	6 Mar (Thu)	Scatterplot and regression	Follow-up section for scatterplot
6	6	13 Mar (Thu)	Visualization of high dimensional data: PCA and others	Follow-up section for dimension reduction
7	7	20 Mar (Thu)	Maps and images	
8	8	27 Mar (Thu)	Understanding uncertainty	
9		3 Apr (Thu)	Mid-term break	No Class
10	9	10 Apr (Thu)	Setting up environment for animation	Technical support for the setup
11	10	17 Apr (Thu)	Setting up environment for Python	Technical support for the setup
12	11	24 Apr (Thu)	Setting up environment for Python	Technical support for the setup
13		1 May (Thu)	Labor Day	No Class
14	12	8 May (Thu)	Course Review and Q&A session	

Course materials



⌘ Course AI Policy

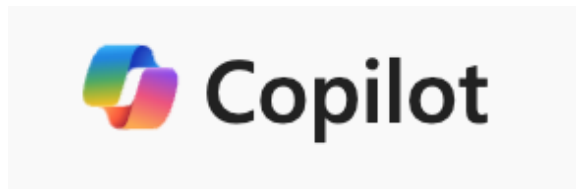
- In this course, students are allowed to use generative AI to assist them in various ways. However, appropriate credit must be given for any use of generative AI.
- Additionally, students must review, analyze, and revise the output from AI to ensure it is better suited for the assessments or projects.
- Students must not copy, paste, and submit the output as if it is entirely their own work.

⌘ ChatGPT and other AI tools

- Learning detailed algorithm is beyond this course. At this point, you have to know only one thing:
 - **They are very useful; sometimes they are correct, but sometimes NOT. They are statistical or machine learning “models,” so their answers are basically PREDICTIONS based on existing knowledge.**



ChatGPT



<https://chat.openai.com/>

<https://copilot.microsoft.com/>

<https://gemini.google.com/>

<https://www.deepseek.com/>

❌ Copilot with VS code, RStudio or other editors.

- Copilot (with VS code and RStudio)
 - AI programming assistant
 - GitHub Copilot is powered by a generative AI model developed by **GitHub**, **OpenAI**, and **Microsoft**.
 - Can be used with VS code, RStudio, or other editors.
 - Programming-oriented chatGPT?
- **You may use it** and **please declare** that you used copilot in your assignments.

