

# Data Science

Lecture 1-1: Course Introduction and Organization



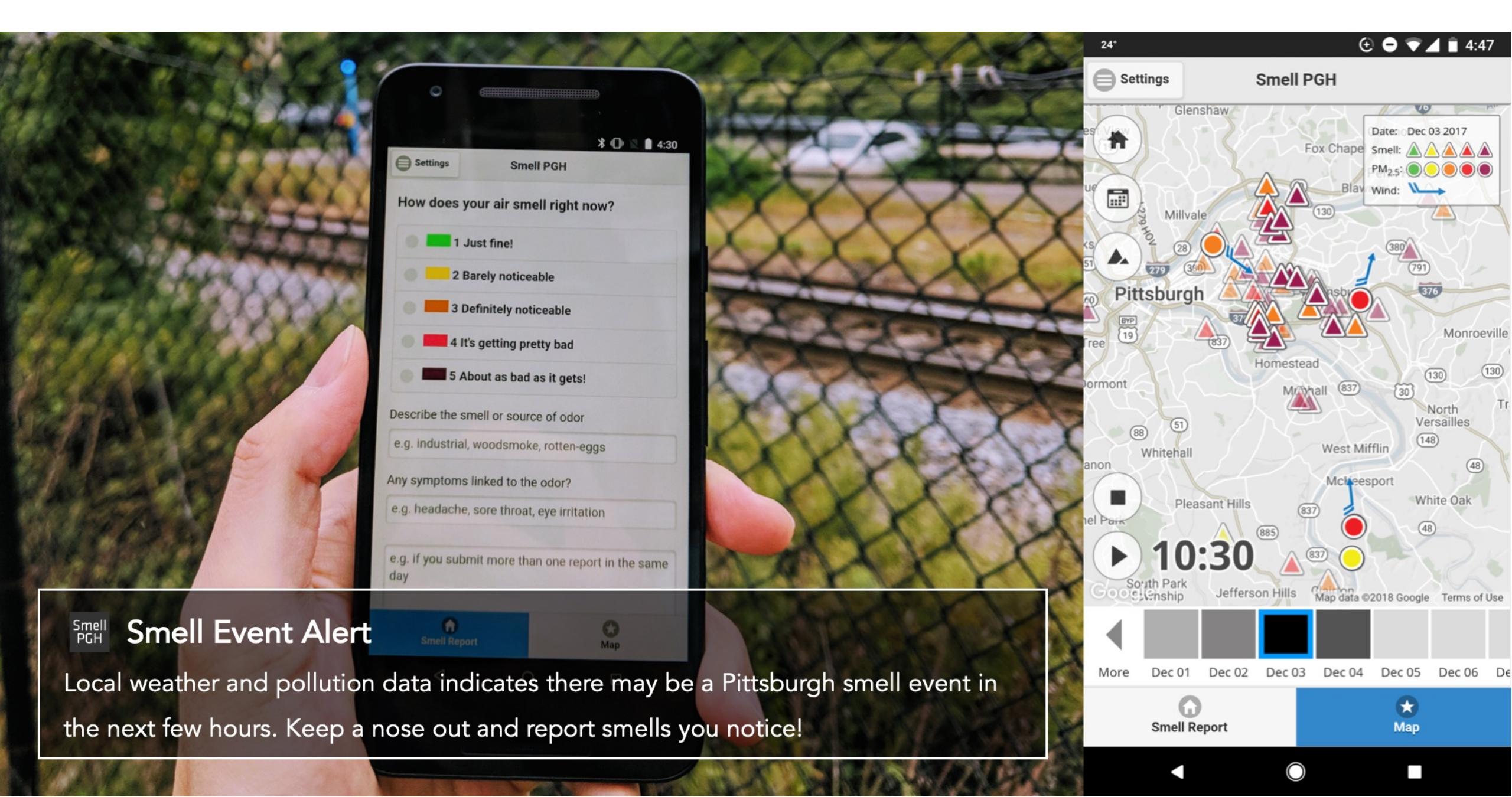
UNIVERSITY  
OF AMSTERDAM

Lecturer: Yen-Chia Hsu

Date: Feb 2023

Data science is about turning rich data into  
actionable insight and making data impactful!

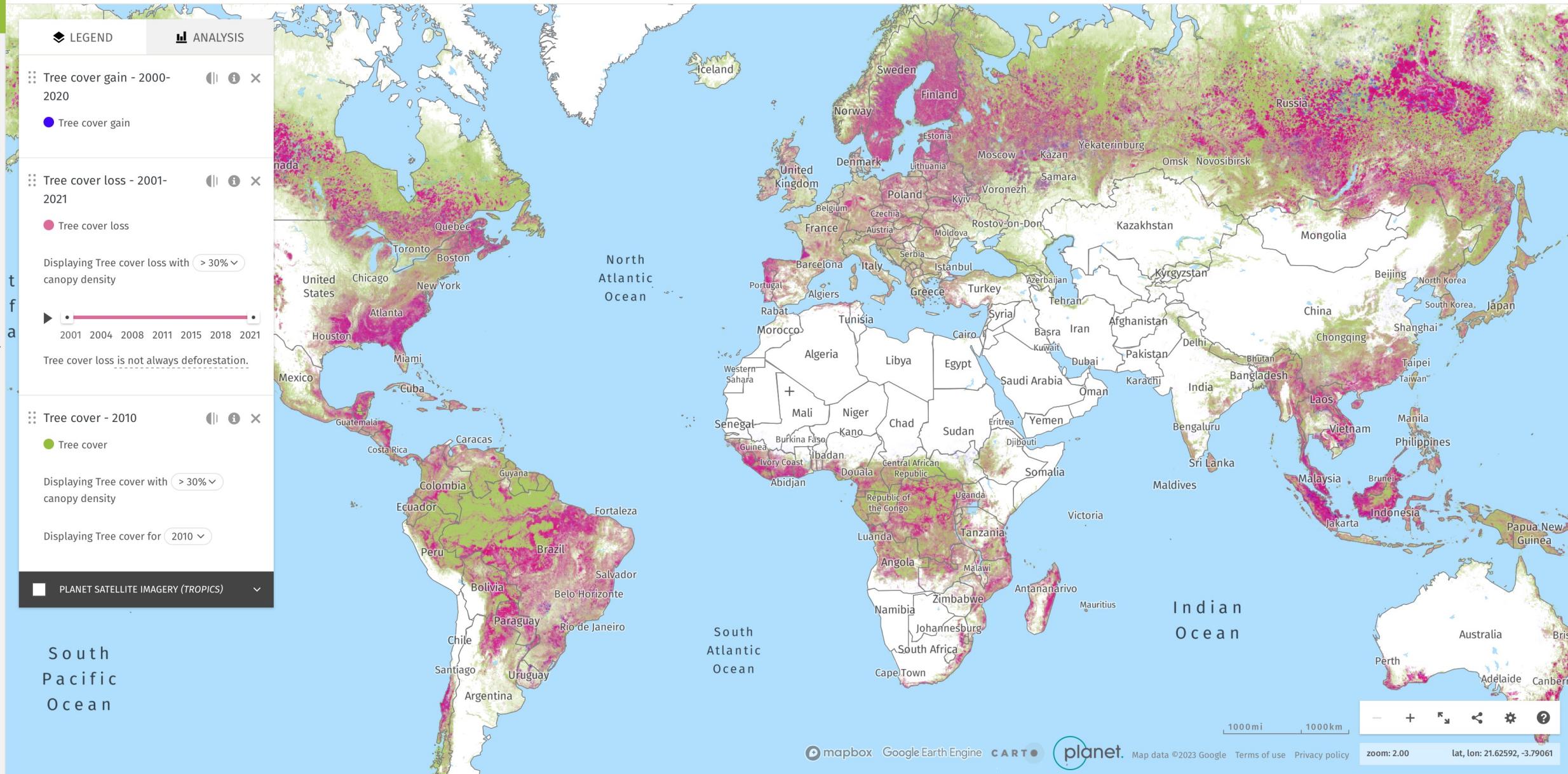
This course aims to familiarize you with  
various data science pipelines, including  
structured and unstructured data.



Example: making sense of air pollution data in Pittsburgh -- <https://smellpgh.org/>



Example: understanding parking patterns and law enforcement -- <https://algoritmeregister.amsterdam.nl/en/automated-parking-control/>



Example: analyzing the situation of global forest situation -- <https://www.globalforestwatch.org/map/>

amazon.nl Hello Select your address All nintendo switch

All Best Sellers Customer Service Today's Deals Gift Ideas Prime Your Amazon.nl Gift Cards Kindle Books Sell on Amazon

Video Games Best Sellers New Releases Trending Gaming Store Nintendo Switch PlayStation 5 PlayStation 4 Xbox Series X & S Xbox One PC

Electronics Deals Week Last day

Back to results



## Nintendo Switch Console, Grijs (Nintendo Switch)

Visit the Nintendo Store

Platform: Nintendo Switch

4.5 stars 15,331 ratings

€329<sup>99</sup>

All prices include VAT.

Platform For Display: Nintendo Switch

Edition: Grijs

Grijs Rood/Blauw

### About this item

- Slimme keuze voor dagelijkse behoeften
- Gemakkelijk mee te nemen, compact ontwerp
- Gemaakt met de nieuwste technologie
- De tool voor een reeks creatieve activiteiten voor iedereen
- Je favoriete content staat altijd op de voorgrond
- Een visuele ervaring van hoge kwaliteit

€329<sup>99</sup>

€9.11 delivery February 6 - 9.

Details

Select delivery location

In stock.

Quantity: 1

Add to Basket

Buy now

Secure transaction

Dispatches from TechLead NL  
Sold by TechLead NL

Add to List

Add other items:

### What other items do customers buy after viewing this item?



Nintendo Switch console (OLED model) with white Joy-Con docking station

Nintendo ★★★★★ 8,078

Nintendo Switch

€324.95

Get it as soon as Wednesday, Feb 1

FREE Delivery on orders dispatched by Amazon over €20



Nintendo Switch Mario Kart 8 Deluxe

Nintendo ★★★★★ 79

Nintendo Switch

€49.90

Get it as soon as Wednesday, Feb 1

FREE Delivery on orders dispatched by Amazon over €20



SanDisk MicroSDXC UHS-I Card for Nintendo Switch, true red

★★★★★ 255,010

€19.29



Orzly Carrying Case Compatible with Nintendo Switch and New Switch OLED...

★★★★★ 56,320

#1 Best Seller in Nintendo Switch Cases

41% off Deal

€13.53

Was: €23.09

Get it as soon as Thursday, Feb 2



New Super Mario Bros. U Deluxe (Nintendo Switch)

Nintendo ★★★★★ 12,488

Nintendo Switch

€48.95

Get it as soon as Monday, Feb 6

FREE Delivery on orders dispatched by Amazon over €20



Nintendo Switch console (OLED-model): nieuwe versie, intense kleuren, 7 inch scherm - met een...

Nintendo ★★★★★ 2,099

Nintendo Switch

€338.00



Nintendo Kart 8 : Deluxe (Nintendo Switch)

Nintendo ★★★★★ 39,054

Nintendo Switch

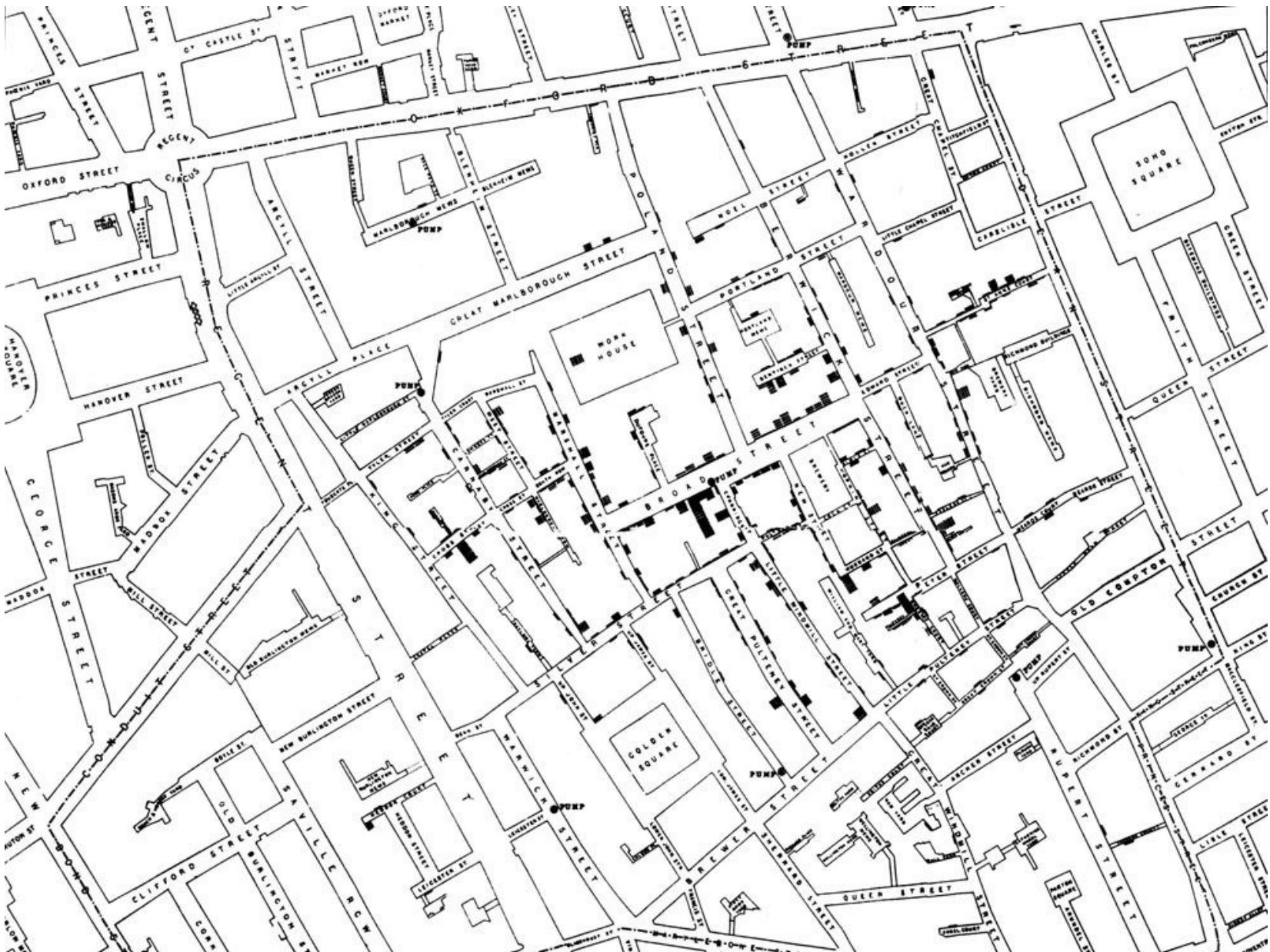
€49.95

Page 1 of 6

# Cholera Map

This map of London was created by John Snow in 1854. London was experiencing a deadly cholera epidemic, when Snow tracked the cases on this map. The cholera cases are highlighted in black. Using this map, Snow and other scientists were able to trace the cholera outbreak to a single infected water pump.

ILLUSTRATION BY JOHN SNOW, IMAGE COURTESY FINAVON, WIKIPEDIA

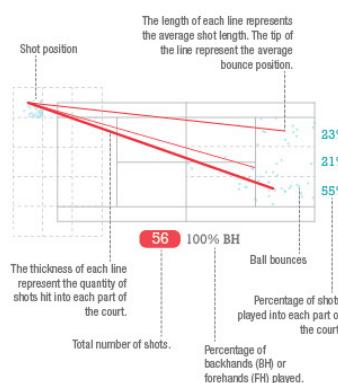




# Kei Nishikori Shot Charts

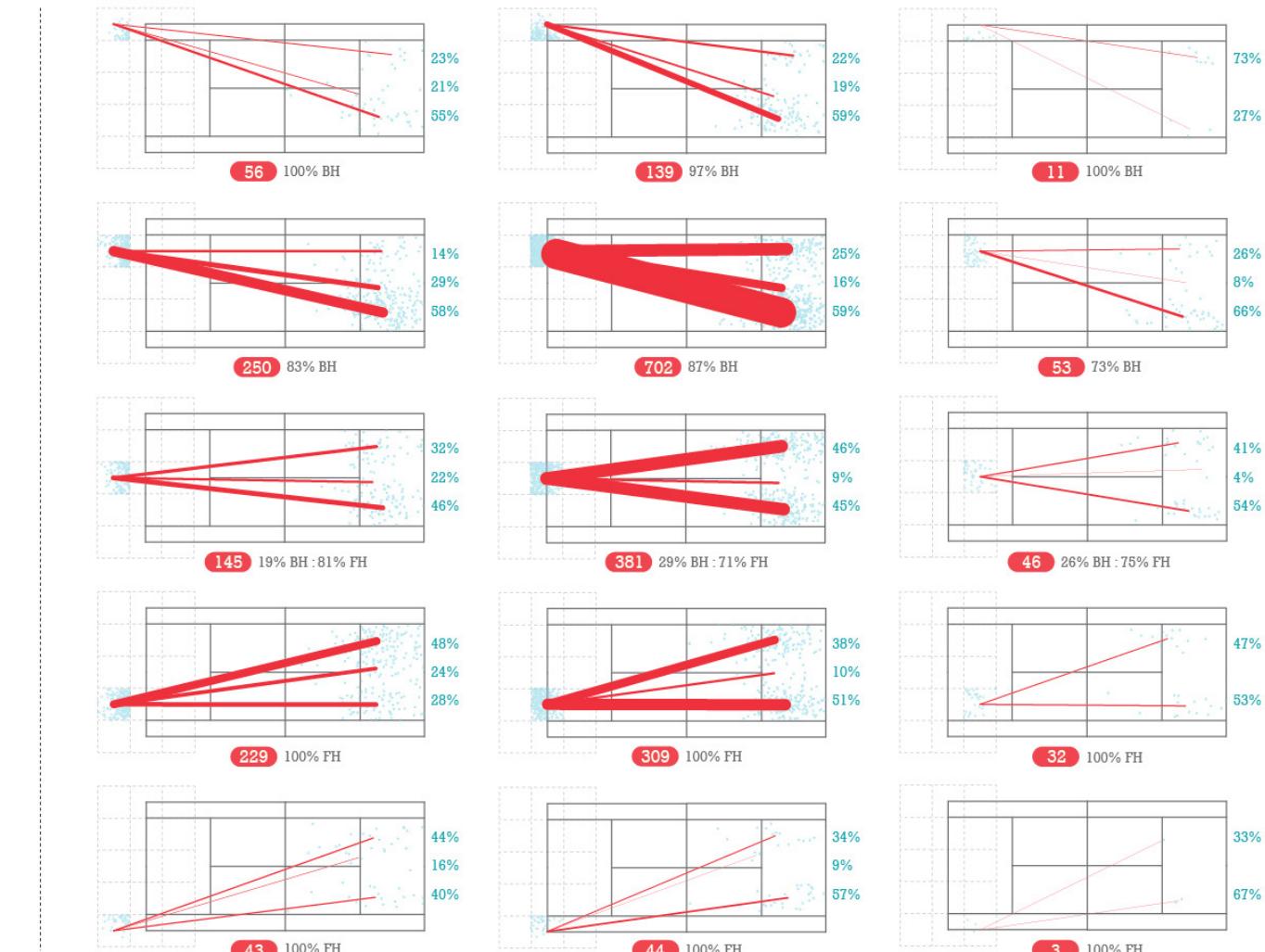
錦織圭のショットチャート

Shot charts are critical in understanding a player's on court behaviour. They are frequently used to map shot patterns from particular areas of the court. These patterns are of particular interest to coaches and players for pre and post match tactical analysis. Here we present 2,443\* shots from Kei Nishikori that were played over a period of 6 months in 2014–15 against opponents like Federer, Djokovic, Murray and Wawrinka.



Analysis: Damien Saundier (gamesetmap.com)  
Source: Hawk-Eye  
Photo: Getty Images

\* Shots that landed out are not included.  
Serve and serve returns are not included.  
Only shots that occurred within the grid are shown.



Example: analyzing tennis player behavior -- <https://tennismash.com/2016/01/18/kei-nishikori-shot-charts/>

**Explore Outcomes**

### Workspace Layout 101

Charts  Details [?](#) [X](#)

**Studies**

Sort by Average distance to exits ↑

1 2

Enable filters  Click and drag over axes to add filters

Desk row rotation Spacing between rows (ft.) Average distance to exits Views to outside Number of desks

10 of 10 [Create Revit Elements](#)

**Details**

**Outputs**

Average distance to exits	7.0
Views to outside	0.573
Number of desks	17.0

**Inputs**

Desk row rotation	-35
Spacing between rows (ft.)	12

Example: exploring automatically generated design options -- <https://www.autodesk.com/autodesk-university/de/node/132606>

This course has five teaching team members, with their first names underlined below.

Please refer to Canvas for the contact information.



Yen-Chia Hsu



Robin Langerak



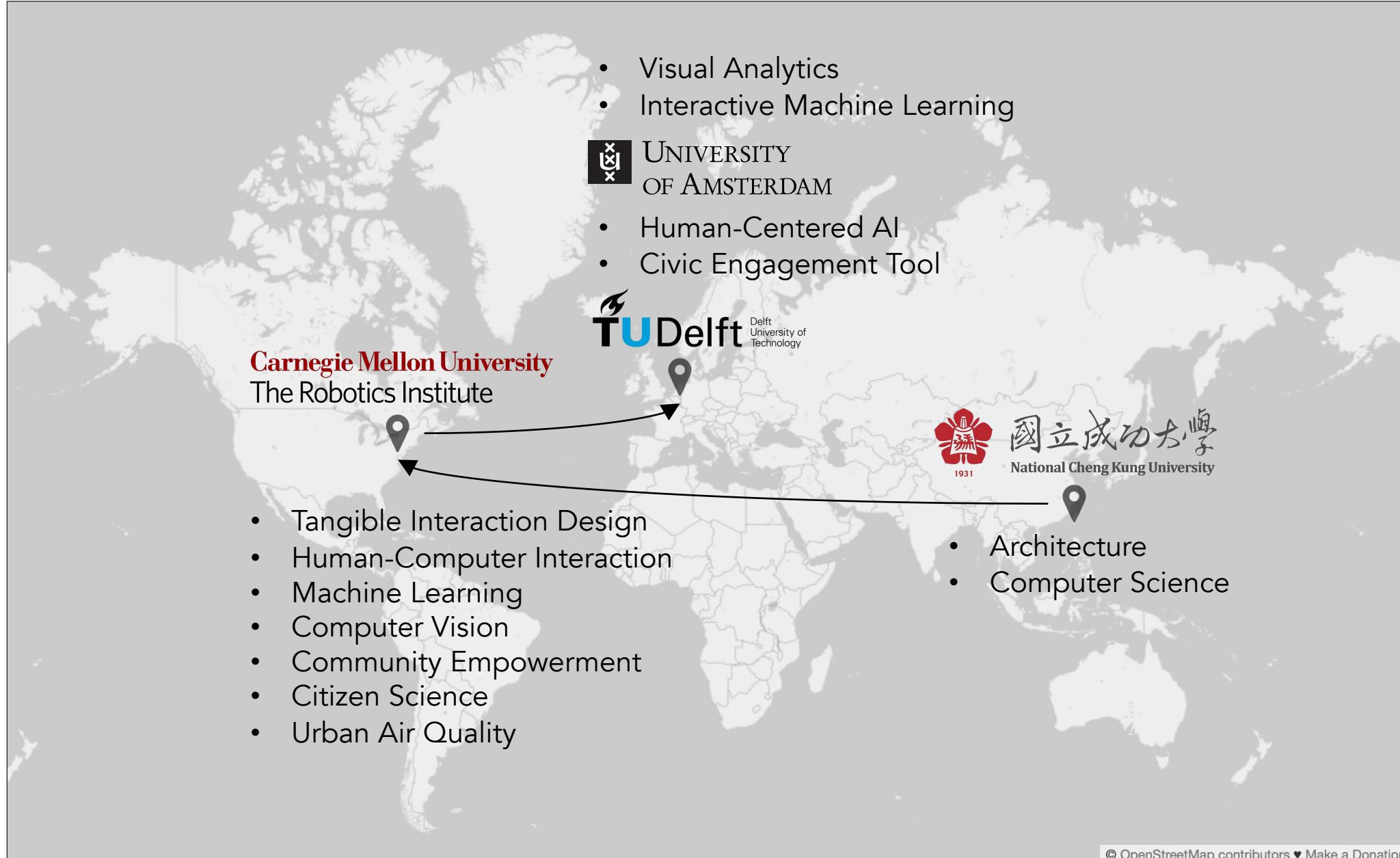
Rachel de Haan



Robert van Straten



Bryan Fleming



The screenshot shows a web browser window with the title "Course Overview — Data Scien" and the URL "multix.io/data-science-book-uva/docs/home.html". The page content is as follows:

**Data Science UvA**

Search this book...

**Course Overview**

Course Syllabus

**LECTURES**

L1: Introduction  
L2: Data Science Fundamentals

**MODULES**

Structured Data Processing

**OTHERS**

Warm-up Python Coding

**Course Overview**

• Course Name: Data Science (2022/2023)  
• Program: The third year of Bachelor Informatiekunde (i.e., Information Science)  
• Institution: Informatics Institute, University of Amsterdam  
• Instructor: Yen-Chia Hsu <[y.c.hsu@uva.nl](mailto:y.c.hsu@uva.nl)>  
• Refer to the [course syllabus](#) for details  
• Refer to [datanose](#) for the time table and classroom location  
• Refer to [Canvas](#) for announcements, the link to live lectures, and other members' emails in the teaching team  
• **Bring your laptop to the lectures that have the tutorials**

All the course content on this website is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#).

**Schedule Outline**

**⚠ Warning**

This schedule may be changed during the development of this course.

Below is the outline of weekly activities. The term "notebook" refers to the Jupyter Notebook script. The term "tutorial" refers to step-by-step guidances of a notebook script. We strongly recommend you to bring your laptop during the lectures with tutorials.

**ℹ Preparation for Lectures**

In the link for each lecture, there is a preparation section. You **have to** prepare them before going to the lectures or tutorials. If you come to the class without doing the preparation part,

# Course Outline

- Week 1: Introduction + Recap of Data Science Fundamentals
- Week 2: Structured Data Processing (Tutorial + Lecture)
- Week 3: Deep Learning Overview + Crowdsourcing
- Week 4: Mid-term Exam (no work sessions and only one lecture)
- Week 5: Text Data Processing (Tutorial + Lecture)
- Week 6: Image Data Processing (Tutorial + Lecture)
- Week 7: Human-Centered Approach in Data Science
- Week 8: Final Exam (no work sessions and no lecture)
- Check <https://multix.io/data-science-book-uva/docs/home.html#schedule-outline> for more details

# Administration

- Announcements will be made on Canvas, so please check it periodically.
- Lectures will be streamed and recorded (the link is on Canvas).
- Lectures may be given virtually if unexpected situations happen, same as work sessions.
- Use email or TicketVise on Canvas to ask questions.
- This course does not track attendance.
- You need to do the preparation part before coming to the lectures.
- You are expected to treat others with mutual respect and appreciation regardless of any differences.
- It is strongly recommended to stay home if you have symptoms associated with COVID-19.
- Check <https://multix.io/data-science-book-uva/docs/syllabus.html> for more details

# Grading

- Assessment is based on two exams (one midterm for 40% weight, and one final term for 60% weight).
- There is a resit, which is 100% weight (will override your original weighted sum of partial grades).
- Assignments are not graded, but you should do them to prepare for exams.
- Exams are based on coding questions (and may also contain multiple choice questions).
- We will provide mock-up exam notebooks and test cases for you to practice.
- You may bring an A4-size cheat sheet with two sides of content (written or printed) to the exams.
- No other materials are allowed during the exams (except the cheat sheet).
- You must hand in the cheat sheet immediately after the exam (not allowed to take it with you).
- Check <https://multix.io/data-science-book-uva/docs/syllabus.html> for more details

# Academic Code of Conduct

- **OK** to discuss assignments with classmates
- **OK** to use existing solutions as part of your projects or assignments (but you need to clarify your contributions and cite the source properly)
- **OK** to publish your project portfolio (e.g., source code) after the course is over
- **NOT OK** to ask someone to do assignments or projects for you
- **NOT OK** to copy solutions or written content from classmates
- **NOT OK** to pretend that someone's solution or idea is yours
- **NOT OK** to post solutions for your assignment or course exams online
- **ASK** the teaching team if unsure

# Work in Progress!

- The course is re-designed!
- Course materials are work-in-progress and may need tweaking along the way.
- This is the first time that we give this re-designed data science course.
- Several topics are still objects of research, and we may not have good answers to them.
- We appreciate your:
  - **enthusiasm** for embracing the complexity in data science
  - **patience** for the imperfect course logistics and materials
  - **feedback** for helping us improve the course