

# Data Science

Lecture 1-1: Course Introduction and Organization



Lecturer: Yen-Chia Hsu

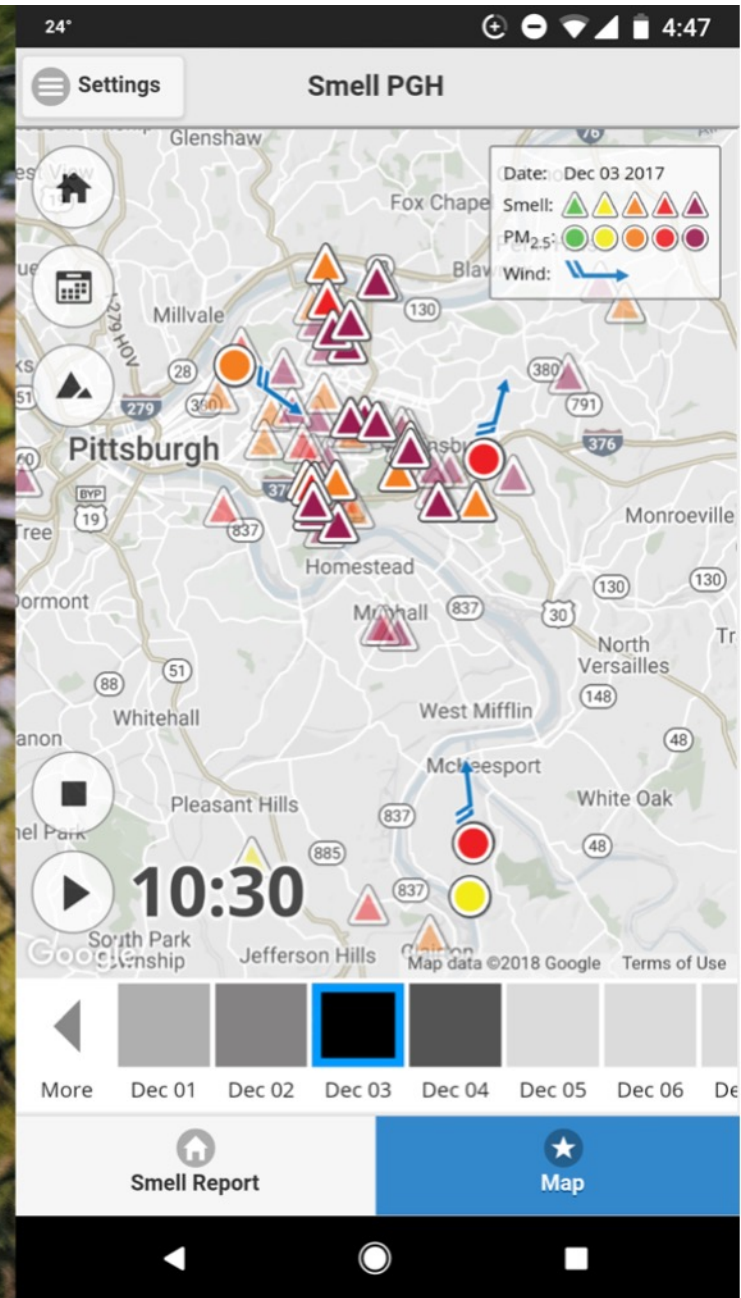
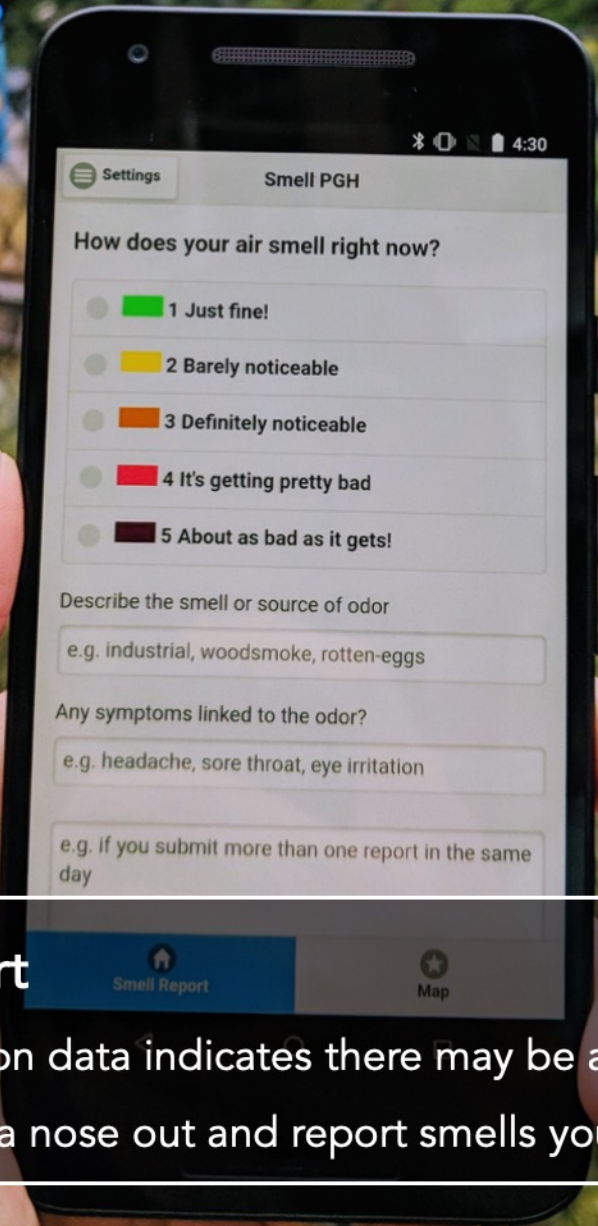
Date: Feb 2024

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 (e.g., processing structured and unstructured data).

**Smell PGH** **Smell Event Alert**

Local weather and pollution data indicates there may be a Pittsburgh smell event in the next few hours. Keep a nose out and report smells you notice!





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

**LEGEND** | ANALYSIS

Tree cover gain - 2000-2020  
 ● Tree cover gain

Tree cover loss - 2001-2021  
 ● Tree cover loss

Displaying Tree cover loss with canopy density > 30% ▾

2001 2004 2008 2011 2015 2018 2021

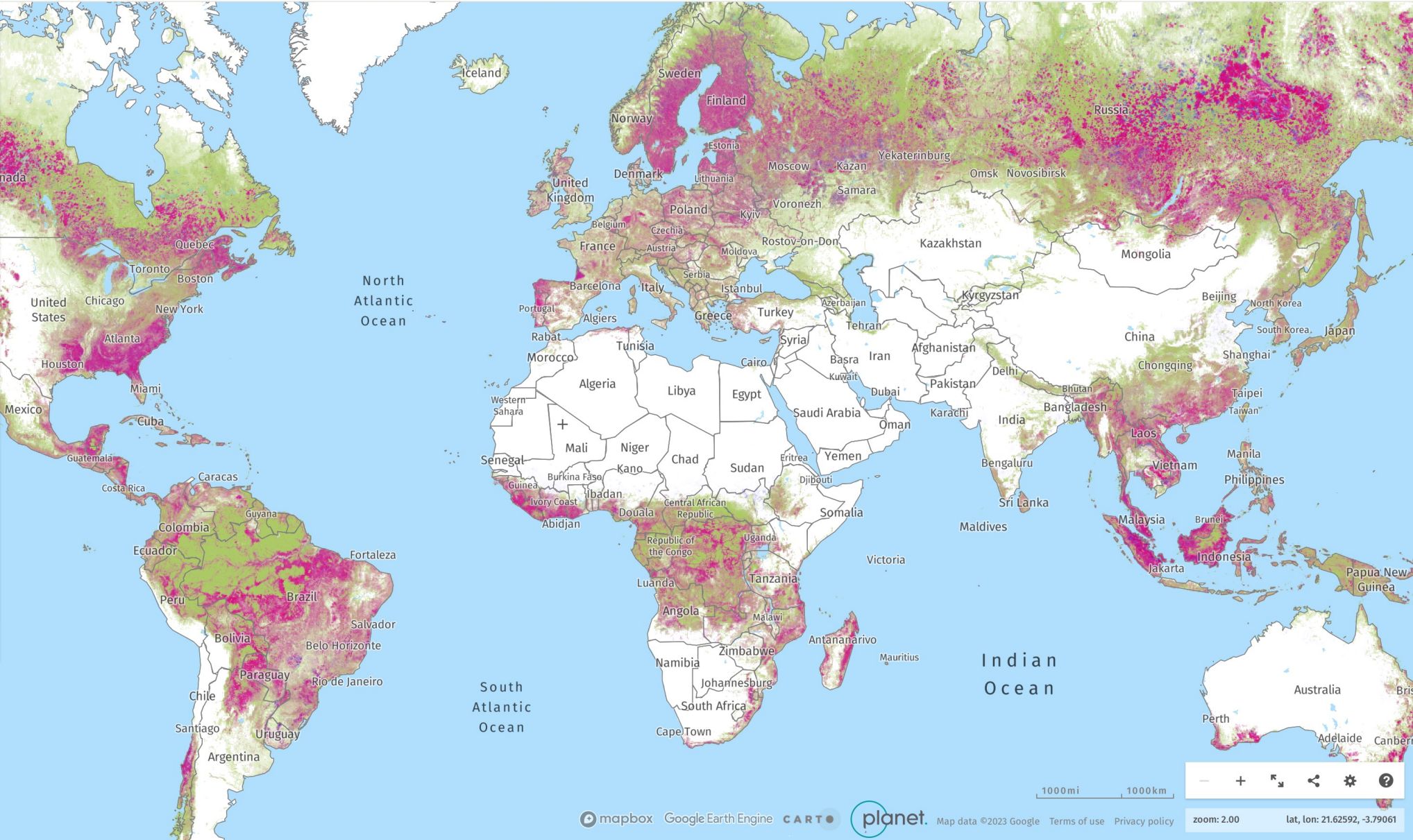
Tree cover loss is not always deforestation.

Tree cover - 2010  
 ● Tree cover

Displaying Tree cover with canopy density > 30% ▾

Displaying Tree cover for 2010 ▾

PLANET SATELLITE IMAGERY (TROPICS) ▾



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

Electronics Deals Week Last day

Back to results



**Nintendo Switch Console, Grijs (Nintendo Switch)**  
 Visit the Nintendo Store  
 Platform : Nintendo Switch  
 ★★★★★ 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?

Page 1 of 6


<




**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




**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



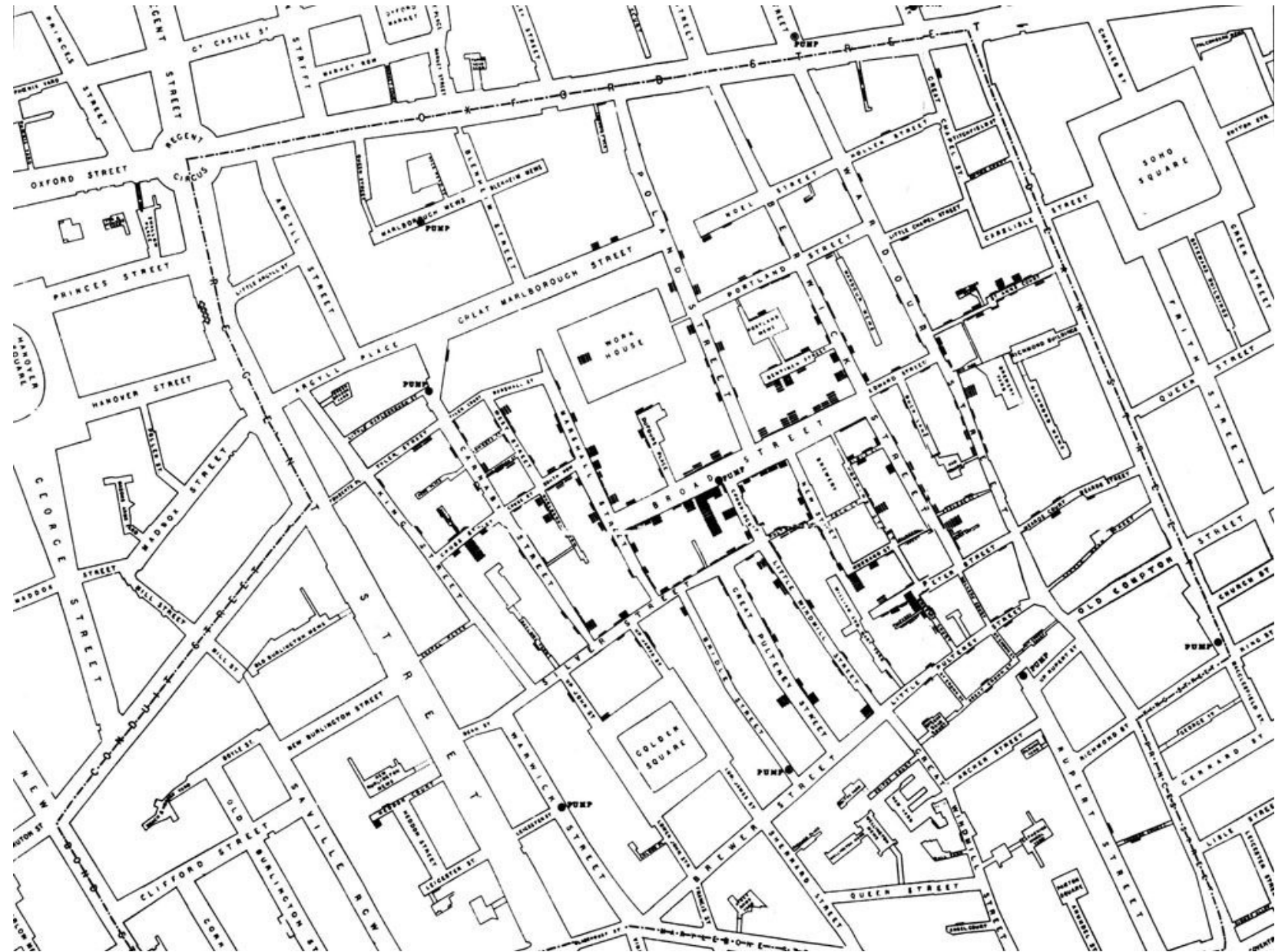
**Mario Kart 8 : Deluxe (Nintendo Switch)**  
 Nintendo  
 ★★★★★ 39,054  
 Nintendo Switch  
 €49.95

>

# 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



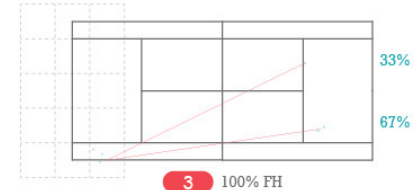
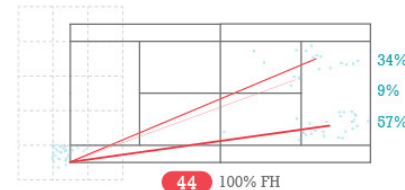
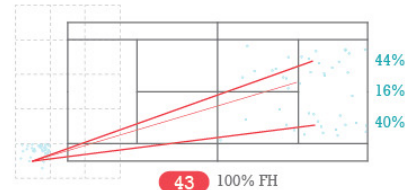
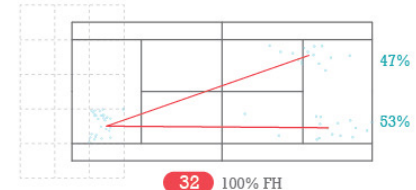
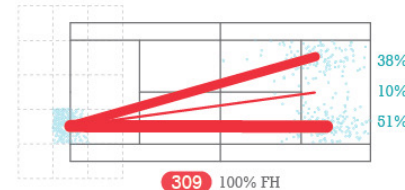
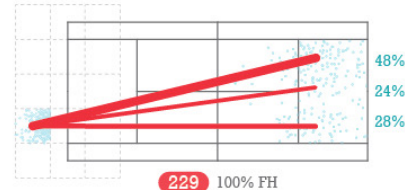
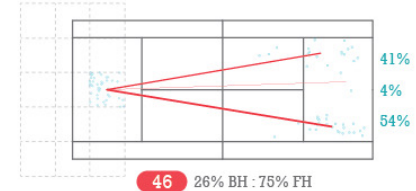
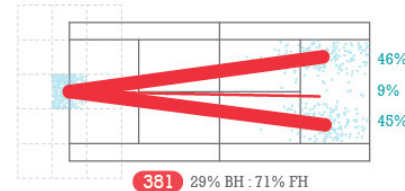
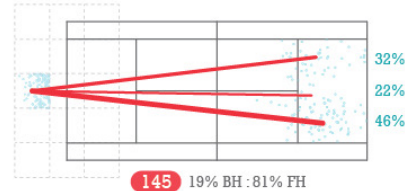
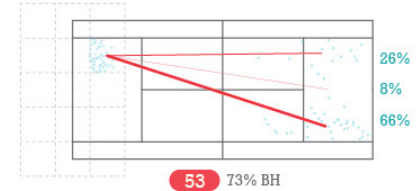
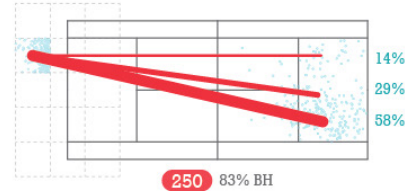
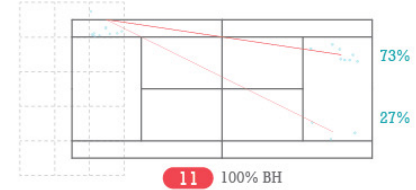
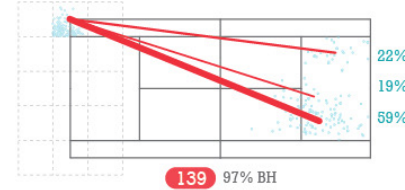
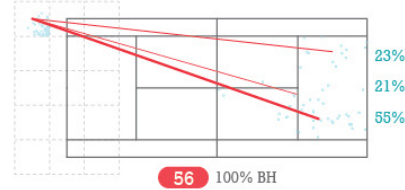
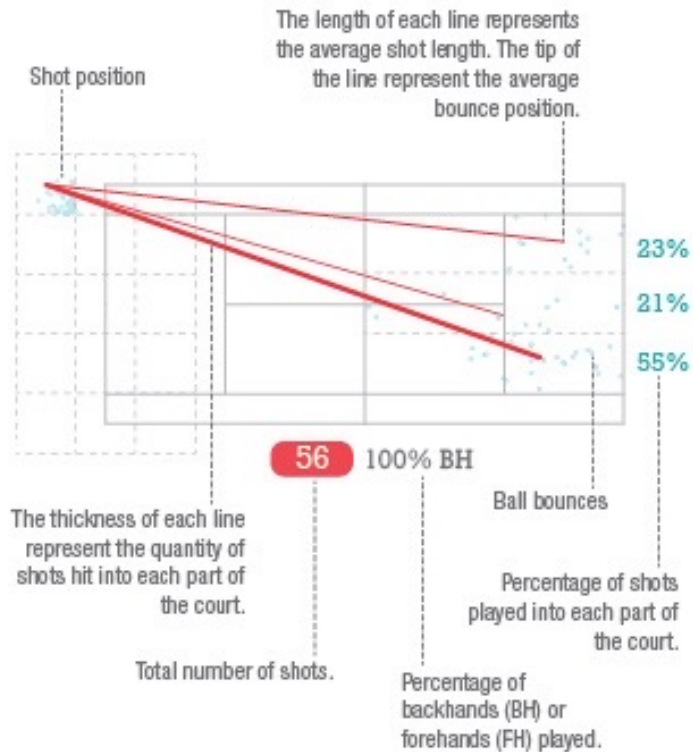


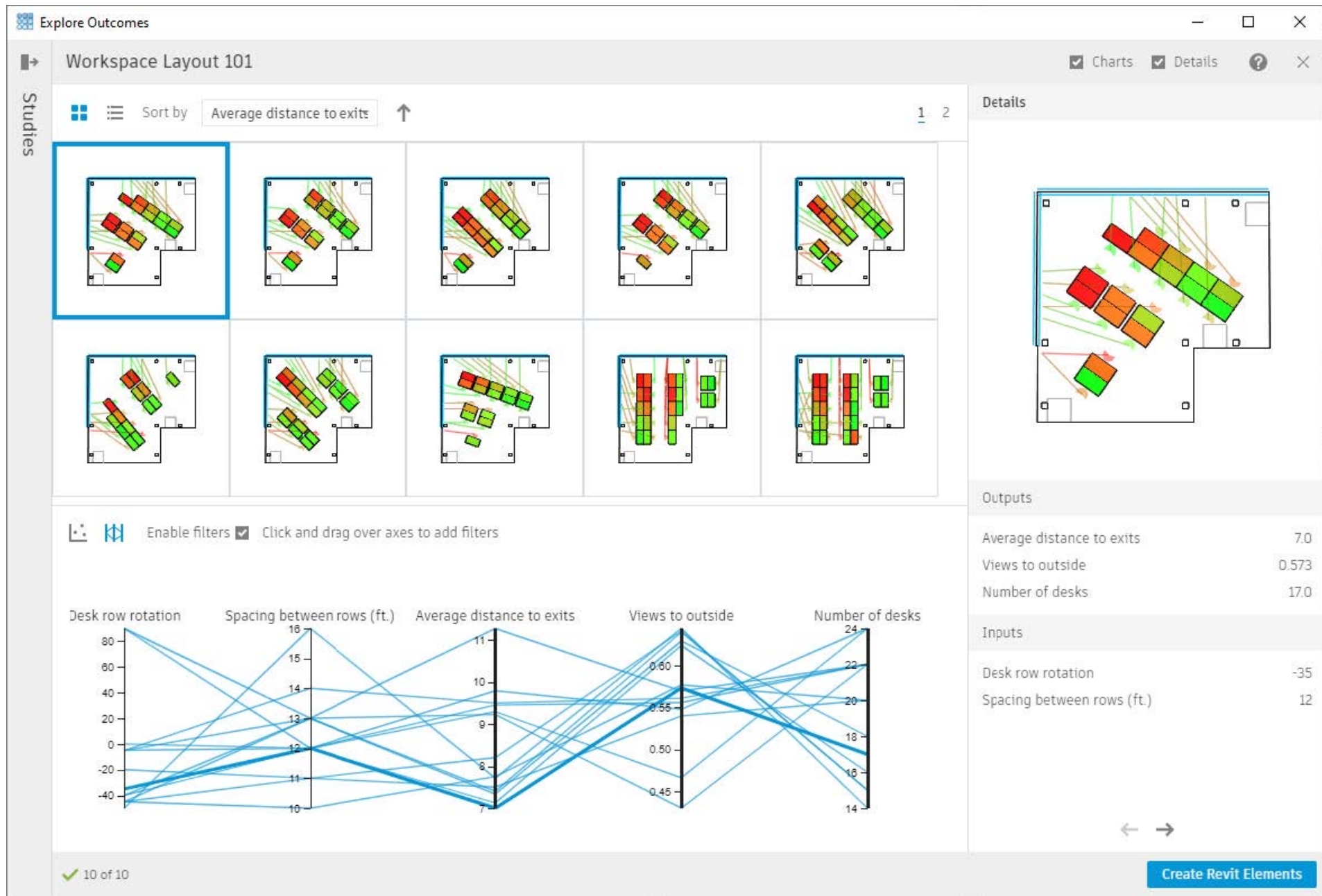


# Kei Nishikori Shot Charts

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.

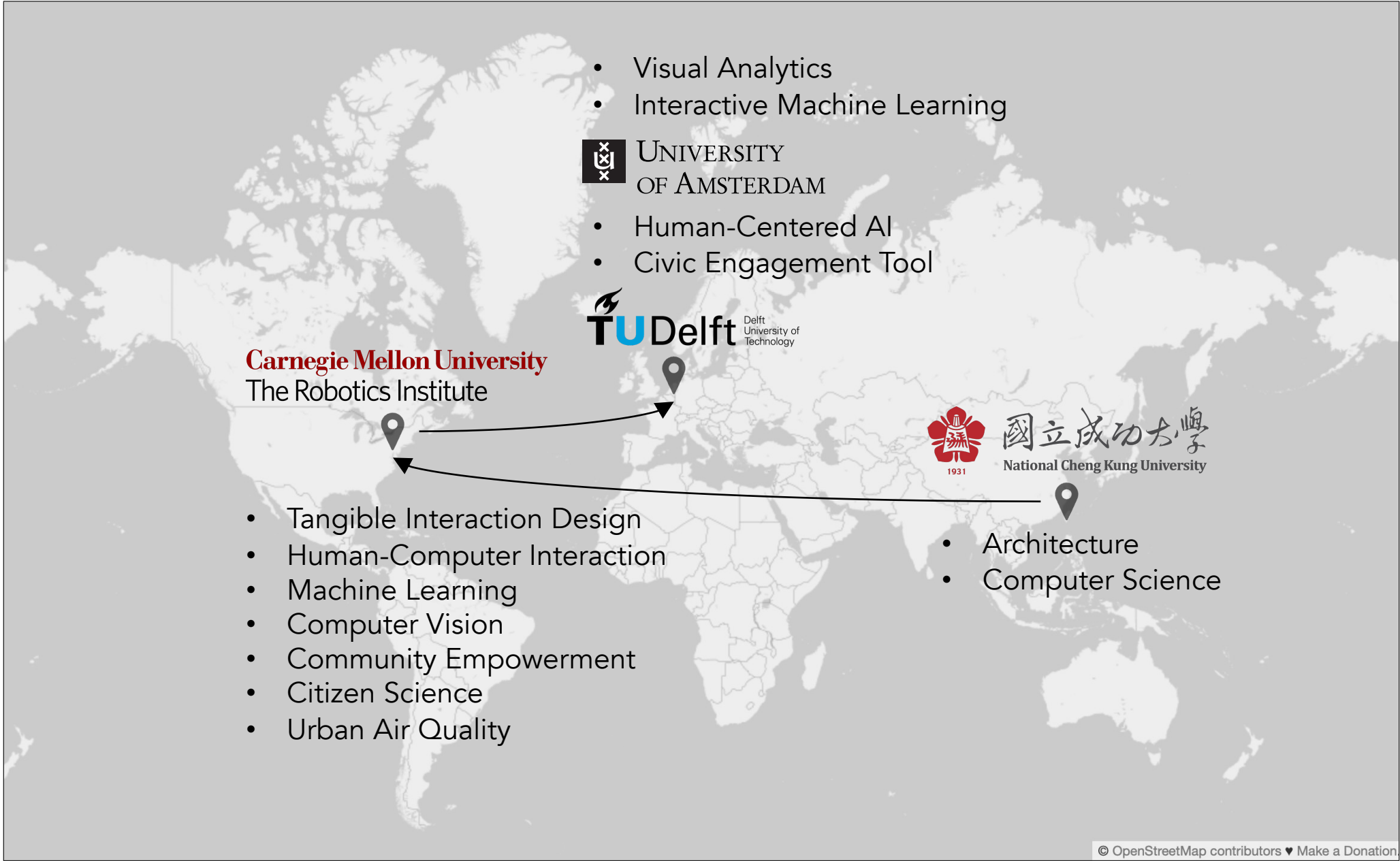




This course has five teaching team members.

Check Canvas for their contact information.

- Yen-Chia Hsu, course coordinator
- Hua Chang Bakker, teaching assistant
- Lucas Belderink, teaching assistant
- Thomas Koene, teaching assistant
- Alejandro Monroy, teaching assistant



UvA Data Science

Search UvA Data Science [GitHub](#)

- Overview
- Syllabus
- Resources

LECTURES

- [L1: Introduction](#)
- [L2: Fundamentals](#)

MODULES

PRACTICES

- [P1: Python Warm-Up](#)

ASSIGNMENTS

OTHERS

## Course Overview

(Last updated: Jan 23, 2024)

- Course Name: Data Science (2023/2024)
- 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>
- Refer to the [course syllabus](#) for details.
- Refer to [DataNose](#) for the time table and classroom location.
- Refer to [Canvas](#) for announcements, links to live lectures, and teaching team members' emails.
- Previous editions: [2023 version](#)

All the content in this repository is licensed under [CC BY 4.0](#).

## Schedule Outline

**WARNING**  
This course is still under development (the second iteration), and the schedule may be changed.

This section provides the outline of weekly activities. We strongly recommend you bring your laptop during the lectures for classroom activities.

Below are the explanation of terms:

- `notebook`: a Jupyter Notebook script
- `tutorial`: step-by-step guidences of a notebook script
- `practice`: a practice (ungraded) for self-study
- `assignment`: an asssignment (graded)

This site uses [Just the Docs](#), a documentation theme for Jekyll.

# Schedule Outline

| Week | Content                                       | Deadline                               |
|------|---|--|
| 1    | Introduction + Data Science Fundamentals      |  |
| 2    | Structured Data Processing                    | Reflective writing due, Saturday 23:59 |
| 3    | Deep Learning Overview + PyTorch Tutorial     |  |
| 4    | Mid-term Exam (no seminars, only one lecture) |  |
| 5    | Text Data Processing                          | Reflective writing due, Saturday 23:59 |
| 6    | Image Data Processing                         | Reflective writing due, Saturday 23:59 |
| 7    | Multimodal Data Processing + Guest Lecture    |  |
| 8    | Final Exam (no seminars and no lecture)       |  |

# Administration

- Announcements will be made on Canvas.
- Lectures will be streamed and recorded with links on the Canvas home page.
- Lectures may be given virtually if unexpected situations happen, same as seminars.
- Use TicketVise (preferred) or email to ask questions.
- There is no attendance requirement.
- 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.

Assessment includes **two exams** (midterm 40%, final 50%) and **three reflective writing submissions** of assignments (10% total).



# Exams

- Check <https://multix.io/data-science-book-uva/#schedule-outline> about the coverage of exams.
- There is a resit, which is 90% weight (will override your original weighted sum of exam scores).
- Exams are based on multiple choice questions.
- We will provide one mock-up exam 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).
- Check the date and time of the exams carefully on UvA DataNose.

# Reflective Writing Submissions for Assignments

- We only grade your reflective writing submissions but not the assignments.
- Use the reflective writing template that we provide for submissions.
- We grade the reflective writing with pass/fail.
- You need to show that you have done the assignment by explaining what you have learned.
- We only accept submissions on Canvas (no email submissions).
- Submissions that are more than 24 hours late will receive a fail grade (no penalty within 24 hours).
- Assignment materials can appear in the mid-term or final exams.

For more details, check the course overview and syllabus pages below:

- <https://multix.io/data-science-book-uva/>
- <https://multix.io/data-science-book-uva/syllabus.html>



Questions?