

Data Science Clinic

Practical experience applying Data
Science and technology-based solutions

Data Science Clinic DATA 27100 and 27200

Civic & Tech Data Clinic CAPP 30300, PPHA 30581, MACS 3030



THE UNIVERSITY OF CHICAGO

**DATA SCIENCE
INSTITUTE**

Outline

- Introductions – Who
- Logistics – How the clinic runs
- Grade Breakdown
- Break up into groups
 - First “week” Schedule
 - Meet and Greet
 - Scheduling Meetings

Clinic People

Who: Nick Ross

Jobs

- Data Science Clinic Director (2022-)
- Director of Backend Engineering & Data Science at The Meta (Kovaak) (2020-2022)
- Assistant Professor of Data Science at USF (2014-2020)
 - Director of Industry/Academic Partnerships
- Direct of Analytics Sega (2014-2015)
- Director of User Analytics and Acquisition at TinyCo (2011-2014)
- Senior Consultant Bates White (2002-2006)

Education

- PhD at UCLA in Management (2012)
- Masters in Economics at UC Davis (2007)
- BA in Applied Math/Statistics at UC Berkeley (2002)



Jobs

- Assistant Clinic Director, Data Science Institute (2021-present)
- Data Science Fellow, Mathematica Policy Research (2019)
- Data Science Consultant, The World Bank (2016-2018)
- Owner, full stack developer, TJH Media (2013-2018)
- Associate, Quantitative Rates Trading, Credit Suisse (2008-2013)

Education

- MS Computational Analysis and Public Policy, UChicago (2020)
- BA Economics, UChicago (2008)



THE UNIVERSITY OF
CHICAGO

People Organization

- Each Project has an **External Mentor**
 - Each Project has an **Internal Mentor**
 - Each Project has a TA
 - Each Project has students
 - **Bennett:** System Administrator (if you use the cluster they might jump in to help)
-
- Nick & Tim: Internal Mentor as well as Admins



Your TA was chosen based on their knowledge and experience.

Their primary purpose is to keep the project running. They are technical resources.

TA's will:

- Take attendance during their working sessions.
- Provide feedback on your communication, professionalism and coding during the sessions.

Being unprofessional, rude, condescending or behaving generally unprofessional to the TA is grounds for failing this course.

Your Internal Mentor

Your clinic mentor was chosen based on the mentor's skillset match to the project requirements.

Your mentor provides general project oversight and is a technical resource for the team.

Your External Partners

Clinic students gain experience working directly with the external mentor during the quarter.

The external partner is the domain expert for the project providing background and guidance in the application of data science to the domain.

Some partners may require NDAs, background checks or additional training specific to that organization.

Do not share details about your project or data with anyone outside of your project team, your mentor, clinic staff, and TAs.

You will need to come up to speed on the technologies they are using. Part of this process is figuring it out.

Logistics

Week 1-2: Domain and Technical Ramp-Up

- Access, Data and Problem Statement
- Setting up meeting times

Weeks 2-8: Iterative Project Development

- Expected to work 10-12 (UG), 15 (G) hours per week
- Writing code, doing analysis, etc.

Weeks 9-10: Documentation, Production Planning, and Video

- Document all progress and approaches for future teams
- Transition codebase to be ready for production
- Final Deliverables:
 - 10 - 12 minute video presenting your work
 - One-pager and Email
 - Possible Others...



- Everything is conditioned off of your **weekly mentor meeting times**
- Meeting times are set by discussion with your TA and your Mentor
 - This is done the first week of the course via **Org Report**.
- Each week you will be required to:
 - Attend the two TA sessions (in-person)
 - Attend the mentor meeting (possibly in-person)
 - Fill out your Weekly Progress Report (due midnight *before* mentor meeting)
 - Fill out your Weekly Planning Report (due midnight day of mentor meeting)

Mentor Meetings

Every mentor runs their mentorship differently, but some things for everyone to keep in mind:

- Mentors are here to help
- Mentors will report back on any issues
- Show up ready to discuss the project for your weekly meeting
- **No new issues should be raised in your weekly meetings. Slack is the place to report issues as they arise.**
- **Mentor meetings may be in person or remote at the discretion of the mentor**

- TA Sessions are two, **in-person**, hour-long sessions per week.
- Attendance is taken.
- These are working sessions to diagnose issues, etc.
- The primary way that you will interact with the TA.
- Location will be in Ryerson Annex and will be given to you over slack at the start of week #2.

Weekly Logistics

- **Assume** that your TA sessions are Mon and Fri and your Mentor meeting is Wed

Weekly Schedule	Monday	Tuesday	Wednesday	Thursday	Friday
Meeting	TA Session (in-person)		Mentor Meeting		TA Session (in-person)
	TA takes attendance	Weekly Progress Report due midnight (Graded)	Weekly Planning Report due midnight (Graded)		TA takes attendance

- Each week you will be graded (individually) on:
 - **Planning report:**
 - Did you describe the tasks that were supposed to do?
 - **Progress report:**
 - Did you complete the tasks outlined?
 - Attendance at working sessions
- Since you wrote your planning document it is important to be complete and specific around the tasks.
 - If there is a misunderstanding about “what” was supposed to be completed it is on you... you wrote the task description!
- If you procrastinate or get stuck you may be docked points and credit.

In general...

- This is not a p-set where you simply complete the task the night before
- This is iterative analysis:
 - You will get stuck
 - Waiting until the night before a report is due is a recipe for failure and poor grades
- Even if a task seems relatively straightforward:
 - No one has done it before, on this data set, using this code.
 - We (mentors) don't know exactly what will happen.
- We have very little sympathy for getting stuck at the last minute

Canvas

- Grade updates
- All graded submissions
- Major course announcements

Github

- Code submission/review
- Clinic Repo:
 - <https://github.com/dsi-clinic/the-clinic>
 - <https://dsi-clinic.github.io/the-clinic>

Slack

- Quick access to mentors and course staff
- Questions here instead of in meetings
- **Keep conversation in team channel for visibility**
- **Use threads**



- Your default position should be that:
 - All data used
 - All code produced
 - All conversations
 - **Are confidential.**
- If your laptop is stolen or compromised in some way
 - **Tell us ASAP.**
- Obvious question: What if I want to talk about this on my resume or in an interview?
 - There is a doc on the clinic github page that has some good rules of thumb about this
 - If you have specific questions ask Tim/Nick
- SLACK with Nick/Tim:
 - Nick and Tim are in every slack channel, but we don't grade off of them for *content*. Obviously if someone is being unprofessional, etc. that will affect their grade.
 - There are some administrative / pedagogy reasons we need to be in each channel, but we aren't monitoring them day-to-day.
 - You can ignore us in the channel or call out (by tagging) if you are having issues.

Grading and Assessment

Grade Break down

- 10% Peer Evaluation
- 20% Weekly Reports
- 5% Org report, computer setup
- 10% Professionalism
- 15% Mid Quarter presentation
- 40% Final Deliverables

- “BIG” Deliverables:
 - a. Mid-Quarter Presentation
 - b. Final Deliverables:
 - Video Presentation
 - One Pager
 - Client Deliverables
 - Technical Deliverable
- All grades are posted and all assignments are to be submitted via Canvas.
 - a. You should have received an invite to the DATA 27100 Canvas Page
 - b. NO make up peer reviews
- Where can I find information about grading standards, rubrics, etc.?
 - On the github clinic page!

First few days

- Org Report is due **FRIDAY Night** at midnight.
 - Submit as a *group* via Canvas
 - You need to coordinate with your mentor and TA.
 - *If you have any issues getting hold of someone, please let me know ASAP.*
- Thursday: Technical Demonstration & Verification
 - Your computer needs to be set up (see Computer setup assignment on Canvas) before class.
 - If you are returning to the **same project**, no need to attend.

Week #2 Logistics

- You will get your meeting times and locations early in the week (before your first meeting)
- No TA sessions before your first mentor meeting
- No Week #2 progress report
- Week #2 Planning Report due at midnight on the day of your mentor meeting

External Mentor Meetings

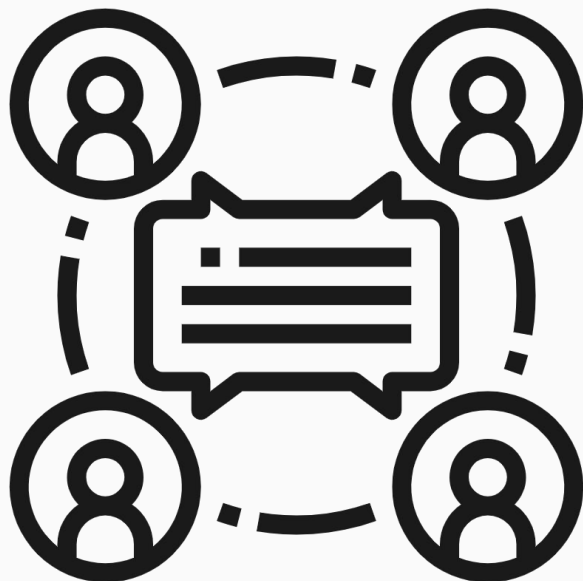
- When you meet with your external mentors you need to fill out a mentor meeting notes template
- This will help you stay organized and identify specific action items required.
- You can find more information about this in the github repo.
- Note that this is graded as a group, so make sure that you have a designated note-taker.

Break OUT

Mentor Information

	A	B	C	D	E
		Project	Mentor	TA	Cluster
	1	Argonne	Bill Trok	Yushan (Cassie) Tang	N
	2	Center for Living Systems	Liya	Justin Wang	Y
	3	CMAP	Anna Woodard	Yushan (Cassie) Tang	Y
	4	College Financial Health	Alexander Bogatskiy	Polly Ren	N
	5	Data and Democracy	Julia Mendhelson	Harper Schwab	N
	6	DSI - Agents	Tim Hannifan	Dongwei (Harper) Lyu	N
	7	DSI - HPC	Nick	Justin Wang	N
	8	Food System 6	Meghan Hutch	Gayathri Jayaraman	N
	9	Groundwork Bridgeport	Jie Jian	Dongwei (Harper) Lyu	Y
	10	IDI - Grievences	Jingchao Fang, Ganghua	Harper Schwab	N
	11	Internet Equity Initiative	Jonatas Marques	Gayathri Jayaraman	N
	12	Kids First Chicago	Tim Hannifan	Jack Sanderson	N
	13	Pesticide Action Network	Kriti Sehgal	Cissy Choy	N
	14	RAFI - Grocery	Kelly Smalenberger	Jack Sanderson	N
	15	UChicago Library	Susanna Lange	Cissy Choy	N
	16	UChicago Transportation	Nick	Polly Ren	N

Breakout: Coordinating with your Team



Meet your team:

Name / Program / Year

Identify your mentor(s) / TA in Slack

Weekly team meeting time

Start filling out Org Report

- If you ping someone and do not hear by Wednesday at 5pm, please slack me directly.

Verify Github, Canvas and Slack

Begin Installing Software for Thursday



The End