

# 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**

- 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, contract 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
- **Tim** and **Nick**: Clinic Administration & Oversight
- **Bennett**: System Administrator (if you use the cluster he might jump in to help)



# Working with Your TA

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.

## Working with Your Clinic 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.



# Working with 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.

- 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.**
- Make sure that multi-factor authentication is set to on for your github account.
- 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

# 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 Planning Report (due midnight day of mentor meeting)
  - Fill out your Weekly Progress Report (due midnight *before* 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 <b>Progress</b> Report due midnight (Graded)	Weekly <b>Planning</b> 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

- The progress report is a weekly assessment of the work you did.
- 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
  - Get stuck
  - Do the bare minimum to complete the task
  - Complain that a task wasn't clear
    - Remember – it was your task description!
- You will get docked points.

## 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>

## Slack

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



# Grading and Assessment

# Grade Break down

- 10% Peer Evaluation
- 20% Weekly Reports
- 10% Surveys, initial org reports, computer setup
- 20% Professionalism
- 10% Mid Quarter presentation
- 30% Final Deliverables

- “BIG” Deliverables:
  - a. Mid-Quarter Presentation
  - b. Final Deliverables:
    - Video Presentation
    - One Pager
    - Client Deliverables
- 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

# Grade and Canvas

- All grades are on canvas
- The “total” column doesn’t weight the score correctly
- In order to avoid you looking at it, you have all been given a 0/10,000 on a “FAKE ASSIGNMENT”
- This obviously doesn’t effect your grade

Surveys and Peer Revi...	FAKE	Total
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%
-	0%	0%

- 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.*
- Tuesday: Technical Demonstration & Verification
  - Your computer needs to be set up (see Computer setup assignment on Canvas) before class on Tuesday.
  - If you are returning to the same project, no need to attend.



## Week #2 Logistics

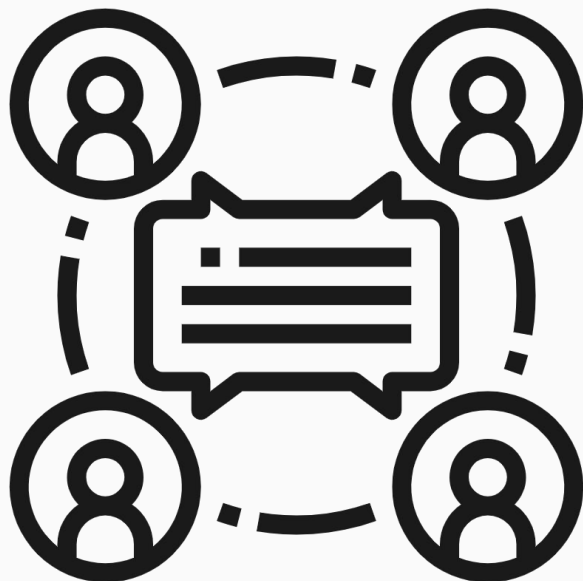
- You will get your TA meeting times and locations
- 1st Mentor Meeting date:
  - **Week #1 Progress Report due the night before.**
    - Template will be provided – basically “did you get your computer verified”
  - **Week #2 Planning Report due that night.**

# BreakOUT

# Mentor Information

	A	B	C	D	K
1		<b>Project</b>	<b>Mentor</b>	<b>TA</b>	<b>Github Link</b>
2	1	<b>AmFam</b>	Anna Woodard	Jessica Ji	<a href="https://github.com/dsi-clinic/2023-autumn-amfam">https://github.com/dsi-clinic/2023-autumn-amfam</a>
3	2	<b>Argonne</b>	Bill Trok/YJ Choe	Victor Perez Martin	<a href="https://github.com/dsi-clinic/2023-clinic-Argonne">https://github.com/dsi-clinic/2023-clinic-Argonne</a>
4	3	<b>Argonne/Fermi</b>	Isaac Mehlhaff	Jessica Ji	<a href="https://github.com/dsi-clinic/2023-autumn-argonne-fermi">https://github.com/dsi-clinic/2023-autumn-argonne-fermi</a>
5	4	<b>BankTrack</b>	Trevor Spreadbury	Fei Wang	<a href="https://github.com/dsi-clinic/2024-winter-clinic-banktrack">https://github.com/dsi-clinic/2024-winter-clinic-banktrack</a>
6	5	<b>CMAF</b>	Anna Woodard & Tim	Victor Perez Martin	<a href="https://github.com/dsi-clinic/2024-winter-cmaf">https://github.com/dsi-clinic/2024-winter-cmaf</a>
7	6	<b>CTC</b>	Patricia C. & David U.	Yuwei (Grant) Chen	<a href="https://github.com/dsi-clinic/2024-winter-ctc">https://github.com/dsi-clinic/2024-winter-ctc</a>
8	7	<b>Climate Cabinet</b>	Trevor Spreadbury	Avery Schoen	<a href="https://github.com/dsi-clinic/2024-winter-climate-cabinet-campaign-finance-tracker">https://github.com/dsi-clinic/2024-winter-climate-cabinet-campaign-finance-tracker</a>
9	8	<b>CREF</b>	Todd Nief	Rita Xu	<a href="https://github.com/dsi-clinic/2024-winter-compostable">https://github.com/dsi-clinic/2024-winter-compostable</a>
0	9	<b>Fermi - ML Simulations</b>	Peter Lu/Rituparno	Yuwei (Grant) Chen	<a href="https://github.com/dsi-clinic/2023-Autumn-Clinic-Fermi-CaloDiffusionPaper">https://github.com/dsi-clinic/2023-Autumn-Clinic-Fermi-CaloDiffusionPaper</a>
1	10	<b>Fermi - GNN</b>	Chong Liu	Yiran Hao	<a href="https://github.com/exatrkx/NuGraph">https://github.com/exatrkx/NuGraph</a>
2	11	<b>Internet Equity</b>	Tim	Ridhi Purohit	<a href="https://github.com/dsi-clinic/2023-autumn-internet-equity">https://github.com/dsi-clinic/2023-autumn-internet-equity</a>
3	12	<b>Invenergy</b>	Vasileios Charisopoulos	Avery Schoen	<a href="https://github.com/dsi-clinic/2024-winter-invenergy">https://github.com/dsi-clinic/2024-winter-invenergy</a>
4	13	<b>Perpetual</b>	Launa Greer	Sarah Walker	<a href="https://github.com/dsi-clinic/2024-winter-clinic-perpetual">https://github.com/dsi-clinic/2024-winter-clinic-perpetual</a>
5	14	<b>Prudential</b>	Nick	Yiran Hao	Sagemaker
6	15	<b>RAFI</b>	Todd Nief	Rita Xu	TWO OF THEM: <a href="https://github.com/dsi-clinic/2024-winter-rafi-poultry">https://github.com/dsi-clinic/2024-winter-rafi-poultry</a>
7	16	<b>University of Northern Iowa</b>	Satadisha Saha Bhowmick	Ridhi Purohit	<a href="https://github.com/dsi-clinic/2023-autumn-matloob-lab">https://github.com/dsi-clinic/2023-autumn-matloob-lab</a>
8	17	<b>WBEZ</b>	Susanna Lange	Sarah Walker	<a href="https://github.com/dsi-clinic/2024-winter-wbez">https://github.com/dsi-clinic/2024-winter-wbez</a>
9	18	<b>IRC</b>	Cristina Garbacea	Fei Wang	<a href="https://github.com/dsi-clinic/2023-autumn-irc">https://github.com/dsi-clinic/2023-autumn-irc</a>

# 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