RESCHEDULED CLASS MEETING

- **Thursday, November 10   12:30-1:20 pm (San Diego)**
  - Same Zoom link
  - Recording will be provided

RUNNING DATA SCIENCE PROJECTS

Q1 PROJECT TEAMS

Team X

- James Dai
- Vedan Desai
- Jun Oh
- Costin Smilovici
- Tyler Tran

Team Y

- Qixi Huang
- Joseph Perez
- Garrick Su
- Edwin Tse
- Maricela Vasquez

PRESENTATION

- "Improving Fairness in Machine Learning Systems…" by Costin

SKILLS NEEDED

- Strategic planning about the project: what's needed, what it will produce, what questions will be answered, …

- Shared understanding (within team, and with leadership_ around the scope of the analysis

- Accountability to one another

- Ability to jointly assess if you're reaching your goals

- Programming, statistics, communication, narratives, sociology of healthcare, ethics of healthcare, …

WORKING TOGETHER

- Have a shared understanding of tasks, schedules, etc.
    - Have a "primary" *and* a "secondary" person for each task
- Communicate about your strengths & weaknesses

WORKING TOGETHER

- Have a shared understanding of tasks, schedules, etc.
  - Have a "primary" *and* a "secondary" person for each task
- Communicate about your strengths & weaknesses

- Follow through on your commitments
- **Be honest with each other**

BREAKOUT GROUPS (BY TEAMS)

- Talk about roles, expectations, preferences, etc.

- Determine communication, shared understanding, etc.

- Start to plan next steps about Q1 replication project

EXPLORATORY DATA ANALYSIS

- *For the next week:* **Start to understand the MEPS data**
  - Identify questions that seem like they would be useful for understanding dataset
    - Aim for comprehensiveness, even if it's not possible…
  - Do the (initial) analyses to address those questions
    - If there is a script, then please run it (since that's what a replication is)…but make sure that you understand what the script does
    - If there isn't a script, then you'll need to write it
  - *Make sure to keep good records of what you have done!*