Wrap-Up

Today

- Please fill out *Course evaluations*
  - It's helpful *data* for me, my department, and the university.
  - Each comment matters. Each rating matters.
  - Be honest, be balanced.

Logistics

- Sign up for final project presentation slot (see Moodle)
  - Make the title a *link* to your slides
  - Lab is busy, so: **North Hall 253**
- All work should be in before Finals (or make specific arrangements with me)

Wrap-Up of Virtuous Data Science

Objectives

- Describe how Reformed concepts of justice and shalom apply to data collection, analysis, sharing, and use.
- Give examples of specific concerns around privacy, bias, accountability, and transparency
- Describe steps and dispositions that individual data scientists can take to act justly in their profession

Perspectives Debrief

1. What *case studies* do you want to remember?
2. What *principle* seemed most important?
3. If we're only considering those 12 principles, **what's missing**?

Discuss these with neighbors.

A bullet-point summary of biblical *justice*

- Community above individual (voluntarily)
- Equity: equal treatment, dignity
- Collective responsibility
- Individual responsibility

From: Tim Keller, A Biblical Critique of Secular Justice and Critical Theory

- Advocacy for poor and marginalized

Community

> The righteous are willing to disadvantage themselves to advantage the community; the wicked are willing to disadvantage the community to advantage themselves.

So:

- privacy: what *must* we share? what must we *not* share?
- integrity in data collection, analysis, reporting, communication

Thus:

- Data analysis process: *reproducible, transparent*

Equity: Everyone must be treated equally and with dignity.

- *direct* impact
  - fair risk assessment (see Discussion and COMPAS)
  - fair surveillance (don’t hyper-surveil the poor etc.)
  - fair resource allocation
- *indirect* impact:
  - don’t show ads for criminal background checks more often for Black names
  - don’t tolerate higher speech recognition error rates for minorities

Should we even be predicting peoples' lives?

- Risk assessment for criminality, loan approval, etc. requires predicting peoples' future actions and situations
- These predictions might be terribly inaccurate. *Should we be trying at all*?

> Despite using a rich dataset and applying machine-learning methods optimized for prediction, the best predictions were not very accurate and were only slightly better than those from a simple benchmark model.

Salganik et al, **Measuring the predictability of life outcomes with a scientific mass collaboration**. PNAS, April 2020

Corporate responsibility: I am sometimes responsible for and involved in other people's sins.

- Even if *I* intend no prejudice, my *algorithm* could be prejudiced because of training data.
- Even if my work is honest, I could be supporting a company that exploits other workers directly or rely on conflict minerals and child labor
- Environmental responsibility is both individual and collective

http://opiniojuris.org/2020/01/13/the-mighty-apple-google-tesla-dell-and-microsoft-in-the-dock-a-look-at-the-child-labour-lawsuit/

Individual responsibility: I am finally responsible for all my sins, but not for all my outcomes.

- I must do what’s right, whether or not my company’s policies require it.
- When something isn’t right, I need to say something even if it risks my job.

Advocacy: We must have special concern for the poor and the marginalized.

- By *exposing injustice* through visualization and modeling
- By **listening to** and **amplifying**, not **speaking for**.
- e.g., beware of doing “parachute research” or de-contextualized “Data for Good”

Data science, and data scientists, *are not saviors*.

Incarnation

In your relationships with one another, have the
same mindset as Christ Jesus:
Who, being in very nature God,
did not consider equality with God something
to be used to his own advantage;
rather, he made himself nothing
by taking the very nature of a servant,
being made in human likeness.
And being found in appearance as a man,
he humbled himself

by becoming obedient to death—
even death on a cross!

> Philippians 2:5-8, NIV

Learning More

Courses

- DATA 303: Applied Modeling and Visualization
- CS 344: Machine Learning (AI)
- STAT 245: Applied Data Analysis
- STAT 341: Computational Bayesian Statistics

Some further reading on data ethics

- The Oxford Handbook of Ethics of AI
- Coded Bias documentary
- Fast.AI Data Ethics course
- Ethics and Data Science by Mike Loukides, Hilary Mason, DJ Patil

Other resources on Data Ethics

- AI Now Institute
- Data and Society
- AlgorithmWatch
- Harvard BKC
- ACM Conference on Fairness, Accountability, and

Who/What I’m Reading / Following: Tech

- RStudio AI blog
- tidyverse blog
- RWeekly
- distill.pub
- Harvard Data Science Review

“What can I do?”

- Practice the “data dispositions”
  - Humility (cite sources, acknowledge limitations, validate results)
  - Integrity (check assumptions, reproduce analyses, evaluate others’ claims)