



What is ML?

Fairness

Read & Discuss

Types of Bias

- Reporting
- Automation
- Selection (coverage, non-response, sampling)
- Group attribution (in-group, out-group)
- Implicit (confirmation, experimenters)

Ethical Considerations

- What ML fairness risks exist?
- How would a “fair” model behave?
- How can we mitigate risks?
 - model
 - dataset
 - context

