

End of semester:

- ▶ Ethics day 1 (Monday)
- ▶ Ethics day 2 (Wednesday)
- ▶ Ethics day 3 (Friday)
- ▶ Presentations day 1 (next week Monday)
- ▶ Presentations day 2 (next week Wednesday)
- ▶ Review for the final exam (next week Friday)
- ▶ Final exam (Wednesday, May 7, 10:30-12:30)

Today:

- ▶ Reminders and clarification about project
- ▶ Discussion of Boston Housing dataset

1. The author says, “I’ve looked, and I can’t find anyone that will talk about that column [regarding African-American population] *in a technical fashion*. There is plenty of commentary about the column, but not about how the already-questionably-included data has effectively been *destroyed*, and so in its current form is unusable and an unfortunate artifact of past assumptions about society.” To what extent does the *technical* analysis of the data contribute to thinking about the *ethical* aspects?

2. The author says, “If this [the dataset] is used to predict, or even *influence future models by its very existence*, the systemic racism will continue to be a pricing factor.” To what extent does the inclusion of a dataset like this in widely-used libraries and pedagogical sources (textbooks and tutorials) influence future models? Does the pedagogical use of the Boston Dataset compromise students who learn data science, statistics, or machine learning from its use?

3. The author suggests that there are—or should be—some ethics/norms/standards in the data science community. For example, “There’s a rule to be accepted here about the publishing of data: *Publish only original data*. Any transformation should only be applied inside a model, and noted.” and “In addition to the social concerns here, there is the abstract issue mentioned earlier that should always be on the mind of a data scientist: Are you sure your data makes sense?” To what extent can an ethical system like this help machine learning (and related fields) be more likely to be used for good than for harm?

## Coming up:

### **Due Thurs, Apr 24:**

*Read and respond to ethics readings on AI and human learning  
(See Canvas)*

**Apr 28 and 30:** *Project presentations*

**Due Apr 30:** *Final submission of notebooks  
(That's all.)*