

1.

Data: The Monthly income and spending over 100 salaryman people

Outliers: A billionaire income and spending

Problem: Regression line will be pulled toward the billionaire's huge spending which the line will go up a lot. Then this makes seem like typical people spend much more per dollar earned than they actually do

2.

The basis is that statistical techniques are value-laden decisions about whose data is most important, rather than being pure mathematics. Prioritizing some cases over others is an ethical choice made by a method like linear regression, which heavily considers outliers.

I agree. Every model simplifies reality, which always helps some viewpoints while harming others, so no approach is neutral. Using any tool without recognizing the values and trade-offs it imposes on your data is an ethical failure, not just using a particular tool.

3.

- a. machine learning researcher and The Companies Using the Software
- b. machine learning researcher and The Companies Using the Software