**Fairness: Types of Bias**



**Estimated Time:** 5 minutes

Machine learning models are not inherently objective. Engineers train models by feeding them a data set of training examples, and human involvement in the provision and curation of this data can make a model's predictions susceptible to bias.

When building models, it's important to be aware of common human biases that can manifest in your data, so you can take proactive steps to mitigate their effects.

## Reporting Bias

**Reporting bias** occurs when the frequency of events, properties, and/or outcomes captured in a data set does not accurately reflect their real-world frequency. This bias can arise because people tend to focus on documenting circumstances that are unusual or especially memorable, assuming that the ordinary can "go without saying."

## Automation Bias

**Automation bias** is a tendency to favor results generated by automated systems over those generated by non-automated systems, irrespective of the error rates of each.

## Selection Bias

**Selection bias** occurs if a data set's examples are chosen in a way that is not reflective of their real-world distribution. Selection bias can take many different forms:

* **Coverage bias**: Data is not selected in a representative fashion.
* **Non-response bias** (or **participation bias**): Data ends up being unrepresentative due to participation gaps in the data-collection process.
* **Sampling bias**: Proper randomization is not used during data collection.