



## Bias in AI

Study online at [https://quizlet.com/\\_a8mxvt](https://quizlet.com/_a8mxvt)

Prejudice in favor of or against something compared to another, usually in a way considered to be unfair	Human Bias
A statistic is biased if it is calculated in such a way that it is systematically different from the population parameter being estimated	Statistical Bias
race/color, religion/creed, national origin or ancestry, gender, sexual orientation, age, physical or mental disability, veteran status, genetic information, and citizenship	Protected Classes
Bias in a model introduced intentionally, by design. When someone creates an algorithm, they introduce a specific outcome they want to achieve, and the machine learning software creates models which generate the intended result.	Intentional Bias
Bias that stems from inaccurate data being fed to the machine learning model. Data from a biased source must be cleaned and adjusted if it's to be used in machine learning. Data not representative of whole population.	Source/sample Bias
Bias caused by macro conditions which all data about a particular target population unreliable.	Target Bias
Bias caused by limitations in our measurement tools and instruments. When a tool or instrument is unable to capture all the data needed to build a model, it can create unintentional biases or distortions.	Tool or Instrument Bias
people who built the model might be biased against black or colored people	Bias People



## Bias in AI

Study online at [https://quizlet.com/\\_a8mxvt](https://quizlet.com/_a8mxvt)

the strategy the company pursued to build their software might have been biased from the start, intending to focus only on economics but inadvertently also including race	Bias Strategy
the data sources themselves might have been biased, depending on what data was used to construct the model	Bias Data Sources
The algorithm chosen for the model might have been incorrect, choosing a clustering algorithm that used data it shouldn't have, or created correlates inadvertently.	Bias Algorithm
the model might have drifted over time, possibly by incorporating feedback from the client, and eventually ended up redlining as the default	Bias Model
what we do with a model's results - we attempt to explain them after the model has run.	Explainability
when we take a model apart, reducing it down to its source code, down to its individual components, and then walk through how the model function at each step, so that we can show why a particular step is or is not biased.	Interpretability