

# Causation

Readings on causation as it applies to data science, and more specifically data science ethics. A primary focus of these readings is causal inference and socially sensitive attributes (e.g., race and gender).

Title	Citation
The Metaphysics of Causation (Stanford Encyclopedia of Philosophy)	Gallow ( <a href="#">2022</a> )
The Problem of Induction (Stanford Encyclopedia of Philosophy)	Henderson ( <a href="#">2022</a> )
Causation	Scheines ( <a href="#">n.d.</a> )
Causal Diagrams for Empirical Research	Pearl ( <a href="#">1995</a> )
Disparate Causes, Pt. I	Hu ( <a href="#">2019</a> )
Variation Semantics: When Counterfactuals in Explanations of Algorithmic Decisions are True	Hudetz & Crawford ( <a href="#">2022</a> )
Causal Discovery Algorithms: A Practical Guide	Malinsky & Danks ( <a href="#">2018</a> )
On the Explanatory Depth and Pragmatic Value of Coarse-Grained, Probabilistic, Causal Explanations	Kinney ( <a href="#">2018</a> )
Fairness in Decision-making — the Causal Explanation Formula	Zhang & Bareinboim ( <a href="#">2018</a> )
Evaluations of Causal Claims Reflect a Trade-Off Between Informativeness and Compression	Kinney & Lombrozo ( <a href="#">2022</a> )
The Use and Misuse of Counterfactuals in Ethical Machine Learning	Kasirzadeh & Smart ( <a href="#">2021</a> )
Eddie Murphy and the Dangers of Counterfactual Causal Thinking About Detecting Racial Discrimination	Kohler-Hausmann ( <a href="#">2017</a> )
"But What Are You <i>Really</i> ?": The Metaphysics of Race	Mills ( <a href="#">1998</a> )
What is "Race" in Algorithmic Discrimination on the Basis of Race?	Hu ( <a href="#">Forthcoming</a> )

# References

- Gallow, J. D. (2022). The Metaphysics of Causation. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Fall 2022). <https://plato.stanford.edu/archives/fall2022/entries/causation-metaphysics/>; Metaphysics Research Lab, Stanford University.
- Henderson, L. (2022). The Problem of Induction. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Winter 2022). <https://plato.stanford.edu/archives/win2022/entries/induction-problem/>; Metaphysics Research Lab, Stanford University.
- Hu, L. (2019). *Disparate causes, pt. i*. Retrieved from <https://www.phenomenalworld.org/analysis/disparate-causes-i/>
- Hu, L. (Forthcoming). What's 'race' in algorithmic discrimination on the basis of race? *Journal of Moral Philosophy*.
- Hudetz, L., & Crawford, N. (2022). *Variation semantics: When counterfactuals in explanations of algorithmic decisions are true*. Retrieved from <https://philsci-archive.pitt.edu/20626/>
- Kasirzadeh, A., & Smart, A. (2021). The use and misuse of counterfactuals in ethical machine learning. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 228–236. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445886>
- Kinney, D. (2018). On the explanatory depth and pragmatic value of coarse-grained, probabilistic, causal explanations. *Philosophy of Science*, (1), 145–167. <https://doi.org/10.1086/701072>
- Kinney, D., & Lombrozo, T. (2022). Evaluations of causal claims reflect a trade-off between informativeness and compression. *Annual Meeting of the Cognitive Science Society*. Retrieved from <https://api.semanticscholar.org/CorpusID:269447691>
- Kohler-Hausmann, I. (2017). The dangers of counterfactual causal thinking about detecting racial discrimination. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3050650>
- Malinsky, D., & Danks, D. (2018). Causal discovery algorithms: A practical guide. *Philosophy Compass*, 13(1), e12470. <https://doi.org/10.1111/phc3.12470>
- Mills, C. W. (1998). "But what are you *really*?": The metaphysics of race. In C. W. Mills (Ed.), *Blackness visible: Essays on philosophy and race* (pp. 41–66). Cornell University Press.
- Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82, 669–688. Retrieved from <https://api.semanticscholar.org/CorpusID:10023329>
- Scheines, R. (n.d.). *Causation*. Retrieved from <https://www.cmu.edu/dietrich/philosophy/docs/scheines/causation.pdf>
- Zhang, J., & Bareinboim, E. (2018). Fairness in decision-making — the causal explanation formula. *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI*

*Symposium on Educational Advances in Artificial Intelligence*. New Orleans, Louisiana, USA: AAAI Press.