To: Sarah Holland, Public Policy Manager

From: Douglas Hummel-Price, Technical Advisor

Date: 1/30/20

RE: Assessing Bias: An Open Data Tool Created by the Urban Institute

**Executive Summary**

Executing data science ethically requires careful consideration of a variety of factors such as transparency, accountability, and bias. Bias assessment poses challenges that necessitate proactive consideration to avoid creating analyses and algorithms that treat certain populations unfairly. A lot of data ethics work focuses on the technical side, e.g., which model specifications are appropriate, but relatively little has been done to help researchers evaluate bias within their actual data. The Urban Institute has developed a prototype of an open data bias assessment tool that examines data sets from individual cities in the US and provides numerical assessments of how demographically representative the dataset is.[[1]](#footnote-1)

**Background**

Roughly speaking, the relevant ethical discussions about bias revolve around two types of bias: algorithmic/model and data inherent. The former deals with the fact that different model types can introduce bias into an algorithm implementation; this bias can be minimized through a variety of methods not discussed here.[[2]](#footnote-2) However, these tools and methods are generally only as good as data fed into them. Datasets themselves often contain bias that stems from various sources. No data collection mechanism is perfect, so bias might creep in simply as a result of flaws in the collection method. For example, consider telephone-based surveys. Many of these surveys only sample households that have a land-line, so the resulting data skews towards characteristics associated with owning a land-line, such as being older or holding certain views towards new technology.[[3]](#footnote-3) Bias can also creep in when the underlying data are generated by a biased socio-cultural process. Historical crime data demonstrates this type of bias well.[[4]](#footnote-4) If an algorithm is taking in crime data from the 1950s and 1960s, for example, those data will disproportionately characterize people of color as violent and criminal, when the disproportionality originates in racist policing policies of those times.[[5]](#footnote-5) These data inherent biases can severely distort modeling and algorithmic performance, so researchers must proactively consider them in the course of their work.

**The Urban Institute’s Solution**

To help policymakers and researchers combat bias inherent in data, the Urban Institute has created a prototype for a tool that allows them to feed in their data sets and receive quantitative assessments of where the data set contains demographic bias. Led by Chief Data Scientist Graham MacDonald and Data Science Analyst Ajjit Narayanan, the Urban team started by narrowing the scope from all potential data sets to only examine data sets related to city-level open data. The tool assumes US Census data as the best true representation of underlying population distributions within a metro area.[[6]](#footnote-6) The resulting statistics can be used by the researcher to both evaluate the data set and provide greater transparency and accountability.

**Six Steps to Generate Bias Statistics[[7]](#footnote-7)**

The tool takes six steps to create bias assessment statistics. First, the tool automatically detects a dataset’s source city by sampling the data set and geocoding these data to a specific city, picking the modal city if the sample points to multiple cities. Next, the tool reads in the data set with geographic data with the spatial boundaries of the census tracts in the city and appends demographic data from the American Community Survey’s Data profile, such as race, education, unemployment, and poverty distributions. Third, the tool maps each data point to its corresponding Census tract; a significant amount of legwork for the tool involved handling notoriously challenging geospatial data for this step.[[8]](#footnote-8) Fourth, the tool calculates “tract reporting bias,” a numerical estimate of how many possible data points within a tract are represented in the actual data, i.e., how much data is available for a given neighborhood. Fifth, the tool uses the tract reporting bias along with the Census data to provide a percentage point difference between the input data and the Census data – if the Census data says that the Hispanic population makes up 30% of a city, but the input data suggests that population only comprises 20% of the city, the resulting metric would be (30% - 20%) = -0.10. Lastly, the tool calculates statistical significance of the bias metrics so researchers know whether the discrepancies are likely real or just an artifact of random sampling. The tool is built in a Python script and is run on Amazon Web Services cloud infrastructure. This allows Urban to offer the tool for free, as the AWS setup provides for minimal cost. Analyzing 1000 datasets for month, each with 500,000 rows, is estimated to cost $20 per year total.

**Recommended Use Cases**

The tool can be applied to most datasets from individual cities in the US that involve humans. It is specifically useful for smaller municipalities who may not have the budgets to hire full-scale data science teams, although it can certainly be applied in larger cities. By providing statistical significance statistics, the tool gives a measure of certainty that policymakers can use to decide whether the uncertainty is allowable for a given problem. For example, the tool is great for evaluating low- and medium-stakes programs like the locations of bikeshare stations or 311 requests. The tool probably is not ideal for scenarios in which certainty must be high, such as legal evaluations. The tool is not designed to provide a legal basis for whether a data set constitutes protected-class discrimination, or robust analysis for high-stakes health questions. However, in both instances, it does provide a starting point for such evaluation.

**Limitations and Other Considerations**

Graham MacDonald is quite open about the limitations of the tool, several of which revolve around the connection between the true population, the census data, and the input dataset. Since the tool can only use Census data to approximate underlying distributions, it is limited by the accuracy of that Census data. For example, Hispanics are often underrepresented in Census data, particularly in cities like LA, so the tool would not be appropriate for research questions that rely on accurate data about the Hispanic population. The input dataset might also not map to the census data in a substantively appropriate way. Take 311 calls – one would expect more 311 calls in more heavily populated areas, but more 311 calls would also come areas with more issues like uncollected garbage. Interestingly, the tool does not really present any difficult trade-offs besides the marginal risk that the Urban Institute could grow a collection of datasets for nefarious purposes if it desired; given the reputation and goals of the Institute, this seems like an unlikely worry. Lastly, researchers should be wary of their own automation bias; providing bias statistics could lead some researchers to become complacent about additional methods of reducing bias.

**Final Thoughts**

The open data bias assessment tool created by the Urban Institute is powerful in the right situations. It provides researchers and policymakers with the ability to identify blind spots in their data and resulting analyses, which in turn allows them to create better modeling, provide insight into data collection mechanism, and present an increased level of transparency and accountability. By presenting the bias statistics along with analysis, researchers demonstrate an openness that increases credibility and creates an additional avenue for accountability. The tool is currently limited to US cities and is only as good as the underlying Census data, but the tool itself is transparent about these limits and attempts to provide decision-makers with the information they need to make sound research and policy decisions.

1. Graham MacDonald and Ajjit Narayanan, “Toward an Open Data Bias Assessment Tool,” The Urban Institute, March 5, 2019,  [<https://www.urban.org/research/publication/toward-open-data-bias-assessment-tool>](https://www.urban.org/sites/default/files/publication/99844/toward_an_open_data_bias_assessment_tool_3.pdf). [↑](#footnote-ref-1)
2. Genie Barton, Paul Resnick, and Nicol Turner Lee, “Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms,” The Brookings Institution, May 22, 2019, <https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/>. [↑](#footnote-ref-2)
3. Kathleen Thiede Call, Michael Davern, Michel Boudreaux, Pamela Jo Johnson, Justine Nelson, “Bias in telephone surveys that do not sample cell phones: uses and limits of poststratification adjustments,” Medical Care 49, no. 4 (2011): 355-64, <https://www.ncbi.nlm.nih.gov/pubmed/21407032>. [↑](#footnote-ref-3)
4. Randy Rieland, “Artificial Intelligence Is Now Used to Predict Crime. But Is It Biased?” Smithsonianmag.com, March 5, 2018. <https://www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968337/> [↑](#footnote-ref-4)
5. There is evidence that such crime data bias exists even in recent years, but I have chosen a time frame in which policing policies were unambiguously racist so as to avoid distracting from the topic of this memo. [↑](#footnote-ref-5)
6. See Limitations and Other Considerations [↑](#footnote-ref-6)
7. Graham MacDonald and Ajjit Narayanan, “Toward an Open Data Bias Assessment Tool: Measuring Bias in Open Spatial Data,” Urban Institute, March 2019, <https://www.urban.org/sites/default/files/publication/99844/toward_an_open_data_bias_assessment_tool_3.pdf> [↑](#footnote-ref-7)
8. Graham MacDonald, untitled lecture given at Georgetown University for its Data Science for Public Policy Students, November 11, 2019. [↑](#footnote-ref-8)