Beyond Numbers: Why Data Needs Human Interpretation

Junwei Chen

February 8, 2024

The field of data analysis has made progress in times and it plays even bigger role in aiding decision making processes across various industries. However there is a debate, about whether data should express their insights by themselves or if they require human intervention. In my personal opinion, I believe we should approach the idea of letting data speak on their own with caution and perhaps put it on hold for now. I state this for various reasons explained in the following parts.

The core of my argument lies on the necessity for expertise in various fields. Despite the emergence of data driven methods the intricacy of domains, it still requires judgment and intervention from human experts. Human experts play a role in interpreting data identifying patterns and grasping broader implications that quantitative approaches might overlook. As Jordan (2019) stated in his essay, while data analysis forms the foundation, qualitative insights from experts contribute depth and context to the analysis. An example of this, is that quantitative models can forecast market trends and investment opportunities. However, it is an experienced financial analysts who possess an intuitive understanding of market dynamics, geopolitical factors, and investor sentiments that defy quantification. With that being said, I assume combining human knowledge with data methods helps us make better decisions in lots of areas.

It is also important to take into account the ethical aspect when using data to make decisions. Sometimes, modern methods like Machine Learning ignore smaller groups and focus only on the majority. However, that is not always fair. Using only numbers might not capture the real-life complexities. My statement is backed by an essay written by D'Ignazio and Klein (2020, chap. 6). They argue that relying solely on numbers may not be fair or democratic as it can lead to misunderstandings and disregard insights. This highlights the importance of understanding the context because data alone might not capture the complexities of real-life situations. Imagine a city's transportation department, relying on machine learning algorithms and data analytics, prioritizing public transit routes based on high population density and commuter traffic, overlooking smaller neighborhoods and marginalized communities where

many low-income residents, the elderly, and people with disabilities depend on public transportation. In reality, we have to understand the situation behind the data to avoid making mistakes and to include everyone's perspective. It is important to be fair to everyone, not just the big groups.

The final reason, which is equally important revolves around the issue that although machines are great, at handling amounts of data they lack the ability to automatically filter out incorrect information. They just see everything as numbers. That means they might give us wrong results if the data is messy. We can't just rely on machines to handle data without humans checking it. Our computers don't really care about how messy our data looks; it's all a bunch of 1s and 0s to them. Regression models will happily provide coefficients for any set of numbers we feed into them (Au, 2020). The fact that computers can't distinguish and prepare data appropriately can greatly compromise the accuracy of results. This statement highlights a concern regarding our dependence on machines to handle data without the oversight and judgment that human involvement offers. For example, in pharmaceutical clinical trials, automated data analysis tools are crucial for processing extensive patient data. However, these tools often overlook inconsistencies and errors in data entry, potentially skewing outcomes and compromising the safety and efficacy assessments of new medications. Without human intervention to meticulously review and validate the data, the integrity of the entire research process is compromised. Therefore it remains essential to ensure procedures, for preparing data and maintain oversight in order to minimize potential inaccuracies and preserve the integrity of analyses based on data.

All in all, for all those aforementioned reasons, I am convinced that the combination of data analysis and human expertise in decision making processes is a must. While data driven methods offer insights, they must be complemented by judgment to overcome the complexities and nuances of real world situations. We have to be aware of the limits of data and use it wisely alongside human judgment. By merging data analysis, with human interpretation organizations can promote ethical and inclusive decision making frameworks that reflect a deeper understanding of both quantitative data and qualitative insights. This collaborative approach not only improves the accuracy and dependability of analyses, it is also ensures that decisions are based on a broader comprehension of the social, ethical and cultural aspects related to the issues being addressed.

References

- Jordan, Michael. 2019. "Artificial Intelligence–The Revolution Hasn't Happened Yet." *Harvard Data Science Review* 1 (1). DOI: 10.1162/99608f92.f06c6e61.
- D'Ignazio, Catherine, and Lauren Klein. 2020. *Data Feminism.* Massachusetts: The MIT Press. Website.
- Au, Randy. 2020. "Data Cleaning IS Analysis, Not Grunt Work," September. Article.