Descriptive Statistics

Ho Manh Toan

Hanoi, 20-11-2020

In my previous writing [1], I have shared my journey of learning to appreciate data. Talking about data analysis feels like a natural next step. However, let put the complicated analysis to one side, I want to focus on the power of the simplest form: descriptive statistics.

It is common to think that the more complicated, the better. Nonetheless, I believe that a descriptive analysis of a well-structured dataset can tell you many things. The public may have a hard time understanding a regression model, but a bar chart can grab their attention easily. Some of my most memorable papers used this simple method to explore an idea.

The first paper that I worked on by myself is an overview of the Vietnamese art market [2]. Case study was used to explore the issue. Due to its qualitative nature, the dataset consists of 35 cases of art fraud in the Vietnamese art market. The most prominent five cases were used for in-depth analysis. As one of the first works on the Vietnamese art market, the paper has attracted more attention than I hoped for: 12306 views, and 27 Google Scholar citations.

Recently, amid the COVID-19 pandemic, our research team has produced one of our most cited works: 66 Google Scholar citations in only seven months [3]. The paper collected news articles regarding the COVID-19 pandemic in Vietnam, and analyzed various social aspects of the pandemic: responses from the government, social media, and journalism. While the final dataset contains more than 14000 news articles, the analysis was relatively simple: content review and filtering number of articles based on keywords in different aspects. Nonetheless, the paper has covered most of the important aspects and provided useful information for the public.

Finally, utilizing a similar approach as the COVID-19 paper [3], our first paper on environmental issues focused on business and their roles in addressing environmental problems. The findings suggested a high involvement of government agencies in dealing with environmental issues, and a somewhat shallow focus on the environment from the business sectors.

Some of my favorite papers [5, 6] do not need any analysis at all. They focus entirely on concisely discussing an idea. In essence, data or statistics are tools that are helping us to explore an idea. Therefore, an elegant idea can flourish even with a simple approach such as descriptive statistics.

References

- [1] Ho, M. T. (2020, November 10). How I learn to love data. *OSF Preprints*. DOI: 10.31219/osf.io/pynjf
- [2] Vuong QH, Ho MT, Nguyen HKT, Vuong TT, Tran K, Ho MT. (2018). “Paintings Can Be Forged, But Not Feeling”: Vietnamese Art—Market, Fraud, and Value. *Arts* 7(4), 62. <https://doi.org/10.3390/arts7040062>.
- [3] La VP, Pham TH, Ho MT, Nguyen MH, Nguyen PKL, Vuong TT, Nguyen HKT, Tran T, Khuc Q, Ho MT, Vuong QH. (2020). Policy Response, Social Media and Science Journalism for the Sustainability of the Public Health System Amid the COVID-19 Outbreak: The Vietnam Lessons. *Sustainability*, 12(7):2931. <https://doi.org/10.3390/su12072931>.
- [4] Vuong QH, La VP, Nguyen THK, Ho MT, Vuong TT, Ho MT. (2020). Identifying the moral-practical gaps in corporate social responsibility missions of Vietnamese firms: an event-based analysis of sustainability feasibility. *Corporate Social Responsibility and Environmental Management*, doi:10.1002/csr.2029.
- [5] Vuong QH (2018). The (ir)rational consideration of the cost of science in transition economies. *Nature Human Behaviour* 2(1): 5, <https://doi.org/10.1038/s41562-017-0281-4>.
- [6] Vuong QH. (2020). Reform retractions to make them more transparent. *Nature* 582(7811):149. doi:10.1038/d41586-020-01694-x