Summary of Lessons from the Knowledge Series:
Using Twitter Data for Development Research and Evaluation

March 10, 2021

This is a summary of the webinar where World Bank and IEG staff shared experiences working with sentiment analysis, including analyzing Twitter data, as a tool for collecting secondary data for development research and evaluation. The conversation focused on the main lessons learned from conducting exploratory research using Twitter data from Colombia and the former Federally Administered Tribal Areas (FATA) of Pakistan.*

Sentiment Analysis is the process of applying various kinds of Artificial Intelligence algorithm, mostly natural language processing and other types of text analysis to identify, extract, quantify and study opinions and subjective statements of people **

Overall, the session clearly demonstrated that there is clear potential to use social media data for sentiment analysis.

CHOICE OF INSTRUMENT AND RESEARCH CASES
Discussants highlighted that both cases used Twitter for data collection efforts because of the following features:

- Twitter has high penetration in the areas where the research was conducted.

- For certain topics and research questions, Twitter could be a good choice to collect feedback because regular users/citizens tend to use it to discuss socially critical and contentious issues. For example, this type of analysis may be useful for evaluations and research concerning citizen engagement, social contracts, Community-Driven Developments and service delivery issues, where researchers want to measure public buy-in.

- Twitter provides options for simple data extraction at a reasonable cost and with fairly high accuracy.

Discussants also noted that sentiment analysis can provide additional contextual information to support the analysis around the question of whether the World Bank Group is supporting the right issues and providing adequate support to critical issues.

Case of former FATA/Pakistan – World Bank
Some of the emerging research was conducted by the World Bank's DEC team to compliment sources of data in data poor countries i.e. places where there was no existent government data infrastructure, administrative data, or where running a survey was not possible. The team used Tweets from regular citizens in the former FATA region on the various service delivery topics.

- To ensure that the tweets are from the people in FATA, the team used geolocation of user accounts to ensure that the self-reported data is accurate and to avoid bot accounts.

- The team used word frequency analysis with a cascading structure by finding words co-occurring with the main topics they were interested in. Co-occurinng words were then grouped into four main categories: violent conflict, service delivery, rights and activism, and rule of law.

Case of Colombia - IEG

The research was conducted as part of IEG's country case study for the evaluation of World Bank Group's Support to Private Capital Mobilization. The analysis focused on Tweets focusing on the reactions from Colombian citizens to the tax reforms implemented between 2015-2019. IEG was interested in understanding the effects of the World Bank's Development Policy Loans to the Government of Colombia to conduct the tax reforms. Key research steps included:

- The question of how much interest there was from the public was verified by using Google search intensity data for keywords such as "reforma tributaria" or tax reform. The research found that citizens followed the governments tax reform efforts closely. Citizens provided feedback precisely at the peaks of the reform, which allowed collection of relevant data.
- The team created search strings by using search terms, which were then fed into a program to analyze the data. Tweets were classified using a tool as positive or negative. Noteworthy, research found that a positive score may not necessarily mean positive development. For example, if one uses hashtags only, there may be a skewed perception of reality on the ground and incorrect data.

RISKS AND LIMITATIONS

- There are biases that come with contentious issues such as tax reforms or service delivery. For instance, people maybe more likely to share opinion if they are not satisfied with a change rather than if they have a positive experience.
- Using twitter data may not lend itself to cross country generalizability i.e. it may be difficult to find another country that can be comparable in terms of people's perception of government programs.
- There might be self-selection bias, given the various layers of selecting the tweets and accounts. For general population it is hard to get to representativeness just looking at Twitter data, so it should be used as supplementary along with other tools like surveys.
- There is a gap between the reference population (those whose opinion we are trying to understand) and the actual (active) twitter user community (in general and around a particular issue). The gap between the reference population and the actual user community can be relatively small but can also be very large. It is important to reflect on the nature and extent of this gap for particular contexts.

BENEFITS & OPPORTUNITIES FOR FUTURE EXPLORATION

- There are benefits to using social media data, especially under COVID-19 travel restrictions, to find additional ways to reach out to stakeholders and beneficiaries. For instance, social media data may allow comparing information shared by government representatives and formal sources with what citizens say directly about a certain policy or issue. This research tool can also complement the other formal sources of data like surveys while remaining low cost and fast.
- Using social media data and sentiment analysis should be understood as very context specific analysis. For example, the case of Colombia was chosen because Twitter penetration in the country is around 60% and citizens are directly impacted by the tax reforms. This is not a globally applicable analysis and need to be used for issues and conditions that lend themselves favorable to such analysis.
- This type of analysis allows experimentation and quick returns, specifically when there is lack of access to data. Sentiment analysis can be thought of as the "second best thing" understanding that it is sometimes better to have some data than to not have any at all.
- Social media data can be useful if research is focused on understanding social reactions or on making broader predictions on an issue (instead of measuring direct impact), because then bias may matter less.

- This tool may be useful for private sector especially for project monitoring in real time. For example, it could be used, depending on context, to monitor how some corporate sponsors and big financial firms are reacting to policies the World Bank Group supports.
- In countries where World Bank Group interventions have a certain level of visibility or prominence, sentiment analysis can be used to pick up elements of consequences of the intervention through the public's views about these consequences. For evaluation more specifically, this type of analysis could support the assessment of relevance and effectiveness.