Case Study 6.2
Measuring Gender Digital Inequality with Web Data

Author: Ridhi Kashyap (University of Oxford) and Ingmar Weber (Qatar Computing Research Institute)

Reducing gender inequalities in internet access and mobile phone ownership, along with improving digital literacy, have been recognised as important development targets within the UN Sustainable Development Goals (SDGs) framework. Tracking progress on gender digital inequalities is challenging, however, due to limited gender-disaggregated data, especially in less-developed country contexts. With support from Data2X (an initiative of the United Nations Foundation), and as a part of Data2X's "Big Data for Gender Challenge", we have been exploring the use of data obtained from social media advertising application programming interfaces (APIs) to generate real-time measures of gender digital inequality (http://data2x.org/big-data-challenge-awards/#digital).

As outlined in a recently published study in *World Development*, we leveraged Facebook's advertisement audience estimates (available from the platform's marketing API) to generate measures of gender gaps in internet and mobile phone access in a global perspective (Fatehkia, Kashyap, & Weber, 2018). These Facebook advertising audience estimates are publicly-accessible, allowing advertisers, or any user with a Facebook account, to query aggregate numbers of Facebook users by various geographic and demographic attributes such as age, gender, and device type. By providing aggregate data across different attributes for the platform's over 2 billion users, the data serve as a kind of digital census of Facebook users that can be valuably repurposed for social research.

We used the Facebook data to generate a "Facebook Gender Gap Index", an indicator of the ratio of female to male Facebook users in a given country. While the Facebook Gender Gap Index does not represent internet access per se, we found it to be highly correlated with official statistics on internet gender gaps (from the International Telecommunications Union or ITU) and mobile phone gender gaps (from the GSMA), for countries where data are available. The Facebook Gender Gap Index captured gender inequalities in internet access in less developed countries, where access to the internet is most unequal by gender.

We used these Facebook indicators to predict internet and mobile gender gaps found in official statistics, and then compared the performance of models using the Facebook indicators with two other types of models: 1) models using offline variables linked to a country's development (e.g., GDP per capita) or to the broader gender divide (e.g., gender gaps in literacy); and 2) models combining online Facebook variables with offline ones. For internet gender gaps, we found that models using Facebook data did better than those using offline indicators alone. As shown in Figure 6.3 panel (b), using Facebook data, we were able to significantly expand geographical coverage to of internet and mobile gender gap indicators compared to available statistics in the ITU database (as shown in Figure 1 panel (a), with the biggest gains for less developed countries. Higher values in the figure show greater levels of gender equality, with 1 indicating complete parity.

With help from our Data2X grant, our team has developed an online platform (www.digitalgendergaps.org) where we will release regularly updated measures of gender gaps in internet and mobile phone access across the world based on the approach described above. Ad audience estimates, like the ones we have described above, are available from most large web and social media platforms (e.g. Twitter, Google), and in ongoing work we are exploring the potential of applying our general approach to capture other forms of gender inequality, such as in education, digital literacy, and occupations.

Figure 6.3
The internet gender gap index: proportion of female population with internet access divided by proportion of male population with internet access

Figure 6.3a
Modeled using ITU ground truth data

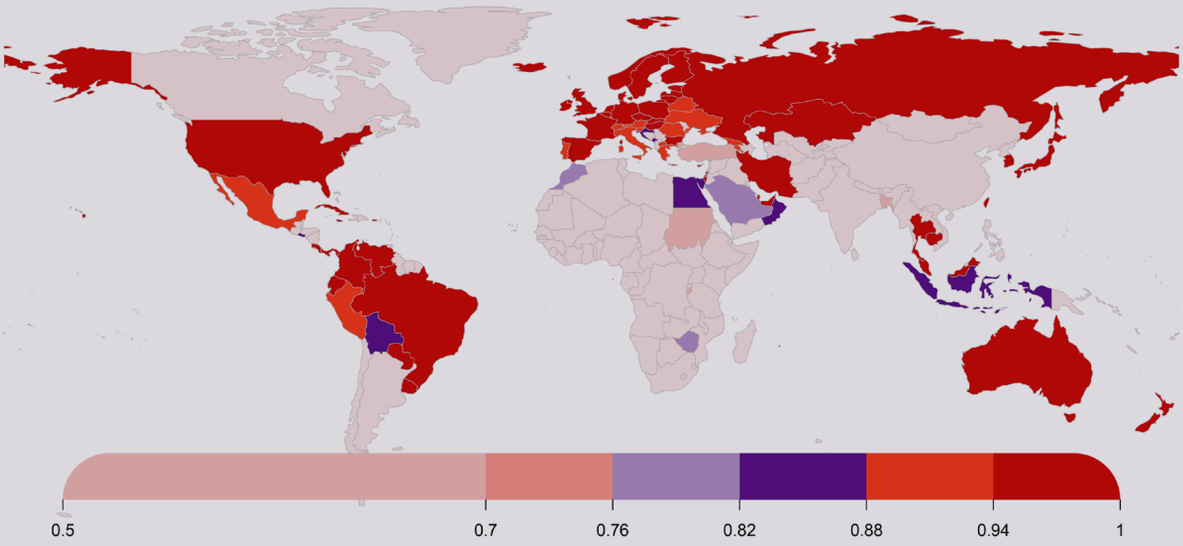
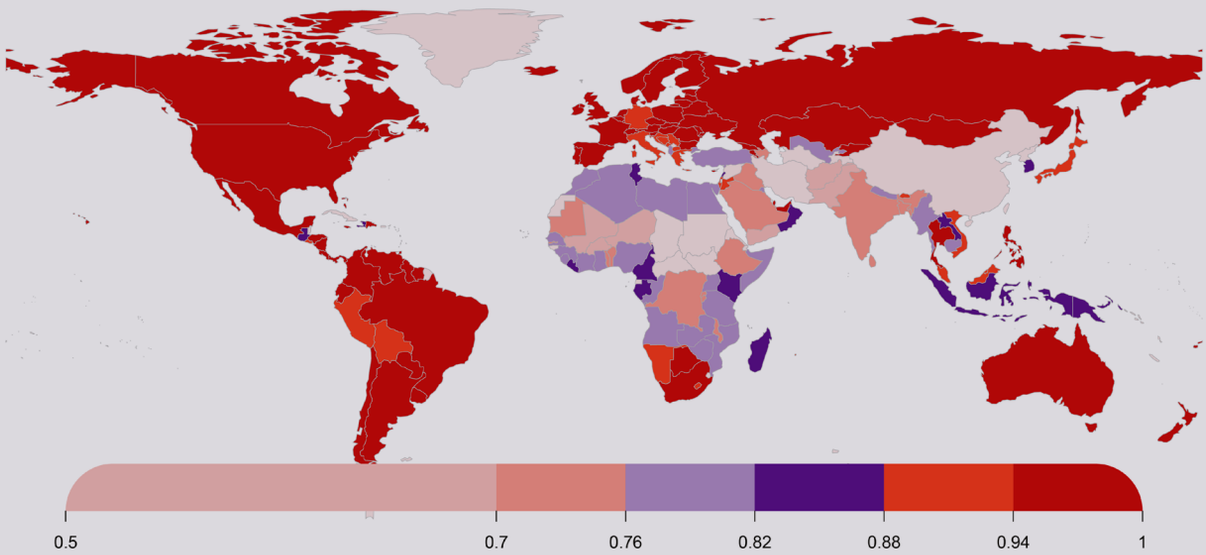


Figure 6.3b
Modeled using Facebook gender gap index



Source: Fatehkia, Kashyap and Weber (2018).
Note: Results were computed by (a) using ITU ground truth data, and (b) using Facebook18+ user gender gap index. Higher values indicate greater gender equality.