Academic Partnerships and Data for Good at Meta:

An introduction to data and products for Researchers across Meta

April 5th, 2022

∞ Meta

Agenda

1. **Introduction to Meta's Data & Transparency Team**
   a. Academic Partnerships
   b. FORT (Facebook Open Research & Transparency)
   c. Data for Good

2. **FORT Products and Datasets, and How to Apply**

3. **Data for Good Products and Datasets, and How to Apply**

4. **Q&A**

Today's Speakers

Christina Fan
Academic Partnerships


Annique Wong
Academic Partnerships

Carlos Ahumada
Data for Good
Public Policy Manager

Data & Transparency Team

- Meta technologies are used by billions of people around the world, and we want to continue to empower people to make an impact using those products in their communities.

- We are bringing together multiple efforts across the company into one team with a cohesive approach and vision for data sharing and transparency, including FORT, Data for Good, and Academic Partnerships.

- Part of the focus of this team will be to ensure stakeholders have access to the right data and tools to enable independent study into the impact of Meta's technologies and share data to advance the world's understanding of key social issues.

Our transparency efforts are an ongoing work in progress — we welcome feedback as our teams come together!
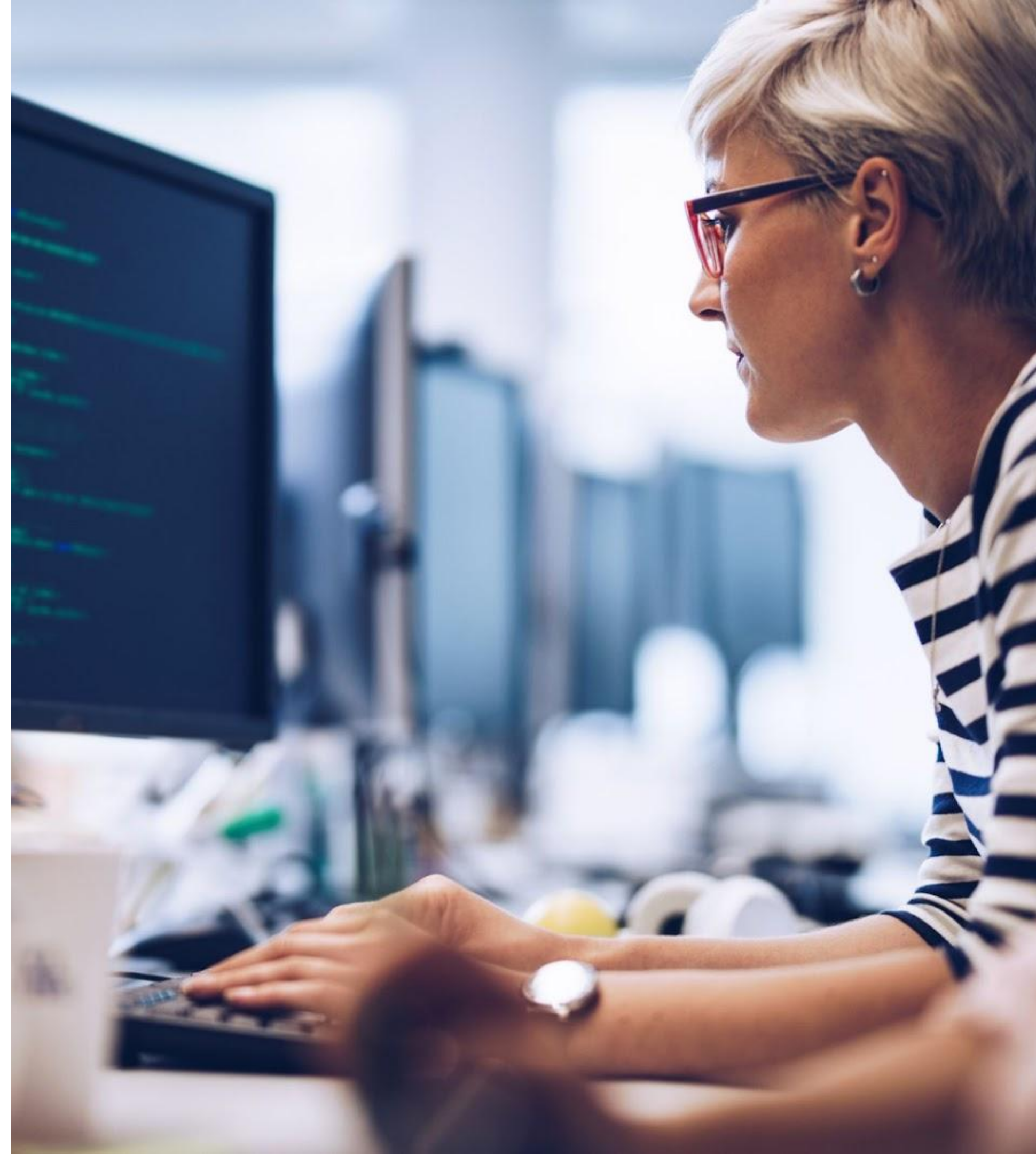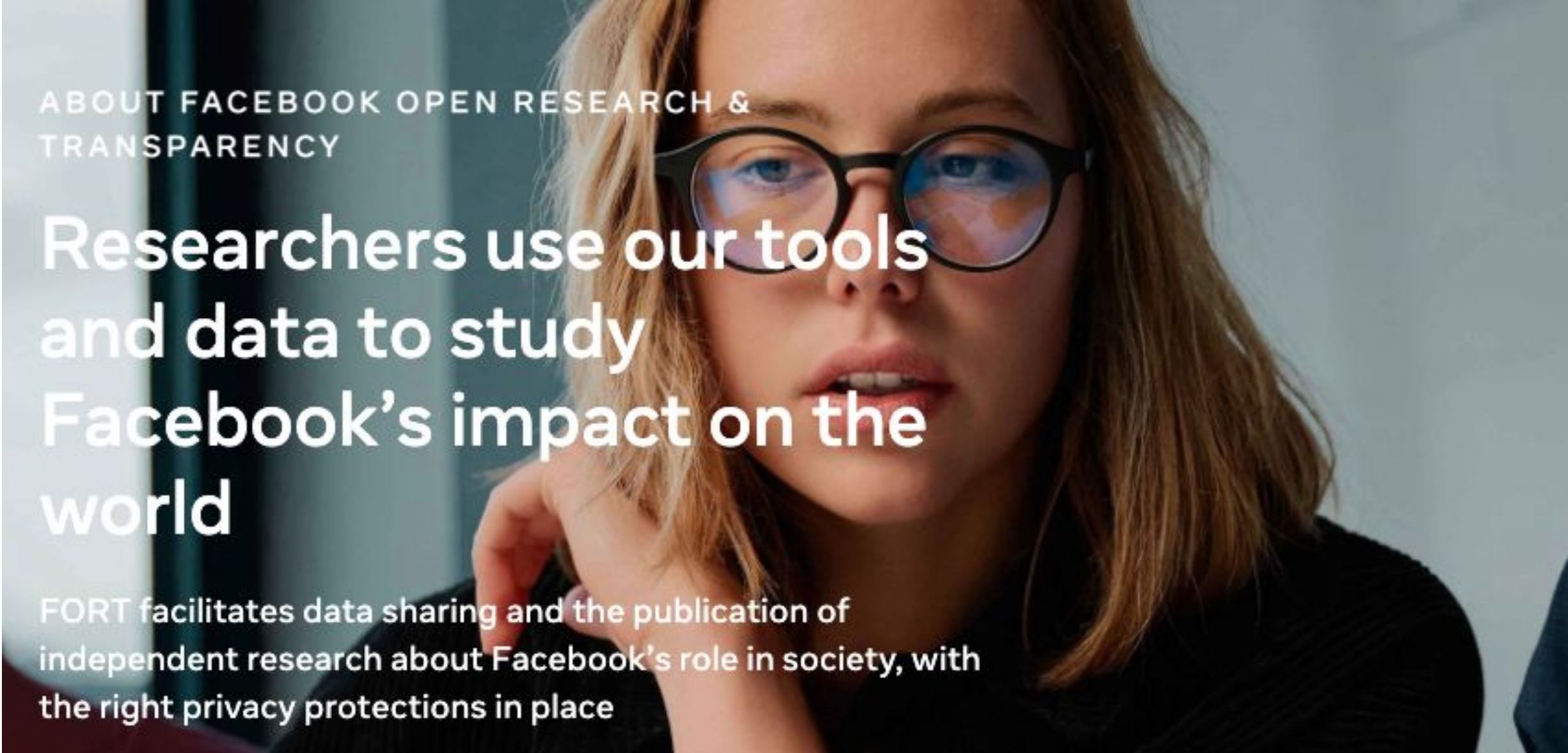
Academic Partnerships Team

Who we are
- Launched in March 2021 to serve the independent academic researcher community globally

How we can support
- Supporting partners through partner onboarding and data access setup
- Promote researcher & product dialogues, feedback loops, and supported data solutions long-term
- Support partners through product training and escalations
- Promote researcher community and shared learnings across our ecosystem of partners

ABOUT FACEBOOK OPEN RESEARCH & TRANSPARENCY

Researchers use our tools and data to study Facebook's impact on the world

FORT facilitates data sharing and the publication of independent research about Facebook's role in society, with the right privacy protections in place

WHAT WE OFFER

We support research by sharing privacy-protected data, releasing analysis tools for the academic community, and collaborating on long-term projects.

Data for Good

Empowers partners (governments, academics, researchers, NGOs) with privacy-preserving data that strengthens communities and advances social issues.

Facts

- Today 500+ partners in 70 countries and dozens of projects.

- Positive impact on health issues, crisis response, economic development and climate change.

FORT: Datasets and Products

Meta

What are our Ad Transparency tools & products?

Ad Library

The Ad Library provides advertising transparency by offering a comprehensive, searchable collection of all ads currently running from across Meta apps and services, including Instagram.

Ad Library Report

The Ad Library Report allows users to explore, filter and download data for ads about social issues, elections or politics. Users can see overall spending totals, spending by specific advertisers and spend data by geographic location.

Ad Library API

The Ad Library API allows you to perform customized keyword searches of ads stored in the Ad Library. You can search data for all active and inactive ads about social issues, elections or politics.

How to get access?

- **Anyone can explore the Ad Library and Ad Library Report, with or without a Facebook account.**
- **The Ad Library API requires a Meta Developer account and identity/location confirmation.**

What is FORT's Researcher Platform and how do I get access to it?

The Ad Transparency Datasets, URL Shares Dataset, and Researcher API can only be accessed on the Researcher Platform. You must be an academic researcher in order to apply for access to these datasets and the Researcher Platform

- The Researcher Platform runs on a modified version of JupyterHub and is hosted on Amazon Web Services.

- Researchers can upload their own data and join it with Facebook data in a privacy protected environment.

- The Researcher Platform is designed such that data cannot be downloaded (barring a subset of URL Shares data) to ensure that it is being used for research purposes.

The Platform is governed by legal agreements that researchers are required to sign (dependant on the data):
- Research Participation Agreement (RPA)
- Research Data Agreement (RDA) – requires a university signatory

What are the Ad Transparency Data Sets?

In February 2021, Meta announced it would make targeting information for more than 1.65 million social issues, electoral, and political ads available to academic researchers for the first time.

Besides the Ad Library Tools, we also provide election-related ad datasets to study the impact of Facebook's products on elections:

Ad Targeting Dataset

Includes the **targeting information selected by advertisers running** social issues, election, and political ads that ran between August 3, 2020, and November 3, 2020. Ads with fewer than 100 impressions are excluded.

Ad Library Dataset

Includes social issues, election, and political ads that are part of the **Ad Library product.**

How to get access?
- **Academic researchers must first apply for access.**
- **If approved, Meta will then contact you about the RDA and verify your identification.**

What is the URL Shares Data set?

The data set summarizes the demographics of people who viewed, shared, and otherwise interacted with web pages (URLs) shared on Facebook starting January 1, 2017 up to and including February 28, 2021.

- The URL Shares Data set includes differentially private individual-level counts of the number of people who viewed, clicked, liked, commented, shared, or reacted to any URL, for any URL with at least 100 public shares.

- Counts are aggregated at the level of country, year-month, age bracket, gender, and for U.S. users, political page affinity.

How to get access?
- **Academic researchers must first apply for access through Social Science One, a third party channel partner used to maintain the independence of researchers.**
- **If you are approved by Social Science One, Meta will then contact you about the RDA and verify your identification. Researchers are onboarded once a quarter.**

What is the Researcher API?

The API provides **near real-time data, as well as billions of historical data points,** on public discussions on Facebook, which equips researchers to study existing topics of interest, as well as gain a window into evolving or emerging topics of interest within the Platform.

The Researcher API Endpoints provides researchers with:

- Aggregated insights on public forums (Pages, Groups, Events) on the Facebook Platform

- Access to raw, anonymized data from public forums on the Facebook Platform

- Ability to keyword search across the public forums on the Facebook Platform

How to get access?

- **We have invited a small group of qualified academics to test this product and provide feedback so we can iterate and improve it, before launching to a broader group of researchers.**
- **To sign up for the latest updates, please visit:** https://fort.fb.com/intake

Future data/tooling: Survey Insights

- We've built a process that allows survey researchers to connect individual-level FB & IG data to their existing surveys via explicit consent, without ever exposing those identities to Meta or to the researchers

- Our goal: enable social scientists to conduct deeper survey research by leveraging panelists' ground-truth social media engagement data, and rely less on self-reported responses

- Currently in "pilot" mode, with a few early partners and projects to date: ANES, Annenberg, and an interventional study of the US 2020 elections

- More soon…we're spending 2022 "productizing" this process: prior to making it available more broadly, ensuring it's repeatable and scalable with a good "user experience" for researchers and proper governance

What kind of research does FORT data support?

- **NYU CSMaP**: Cracking Open the News Feed: Exploring What U.S. Facebook Users See and Share with Large-Scale Platform Data (2021)

- **Arizona State University**: Mobilizing Consensus on Facebook: Networked Framing of the U.S. Gun Control Movement on Facebook (2022)

- **Anatoliy Gruzd's team at the Ryerson University Social Media Lab created a Dashboard to track online misinformation and disinformation about the 2022 Russian invasion of Ukraine using data from the URL Shares dataset**

We also host First Friday, a monthly community event for URL Shares Partners to share their research and collaborate!

How do I apply for access to FORT datasets and products?

You must be a university affiliated academic researcher in order to get access to FORT datasets and products (excluding Ad Library Tools).

- **Ad Library and Ad Library Report** (Open for anyone)
- **Ad Library API** (Requires Developer Account)
- **Ad Transparency Datasets** (Application via FB website; RDA and ID verification)
- **URL Shares** (Application via SS1, RDA and ID verification)
- **Researcher API** (General Access coming soon)

If you have any other questions, please reach out to openresearch@fb.com.

Data for Good: Datasets and Products

∞ Meta

What we offer

We build privacy-preserving stools to assist during humanitarian emergencies and to support global development by informing policy-making processes around the world.

Maps   Surveys   Insights

∞ Meta

How do we offer them?

Public

Privacy-preserving
Free to use
No License Agreement required
Shared through UN's HDX Portal
Open to the public
Governments Ready

Two types of tools &
access

Controlled

Privacy-preserving
Free to use
License Agreement required
Shared through Meta's Partner Portal
Only NGOs, academics, researchers
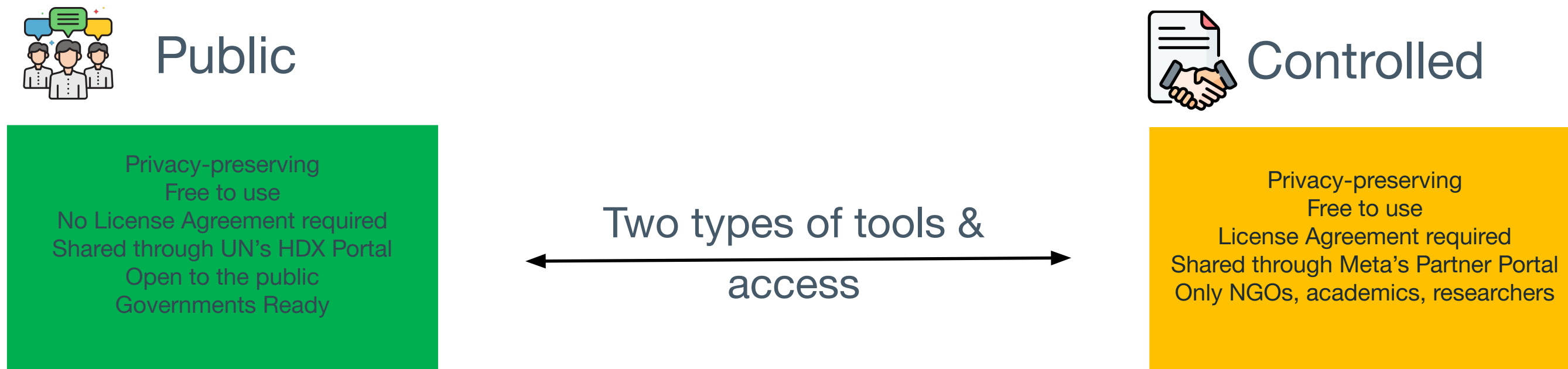
All Data for Good tools are **privacy-preserving** and comply with data protection legislation and regulation from governments around the world.

Maps Example

High Resolution Settlement Layer

Milan
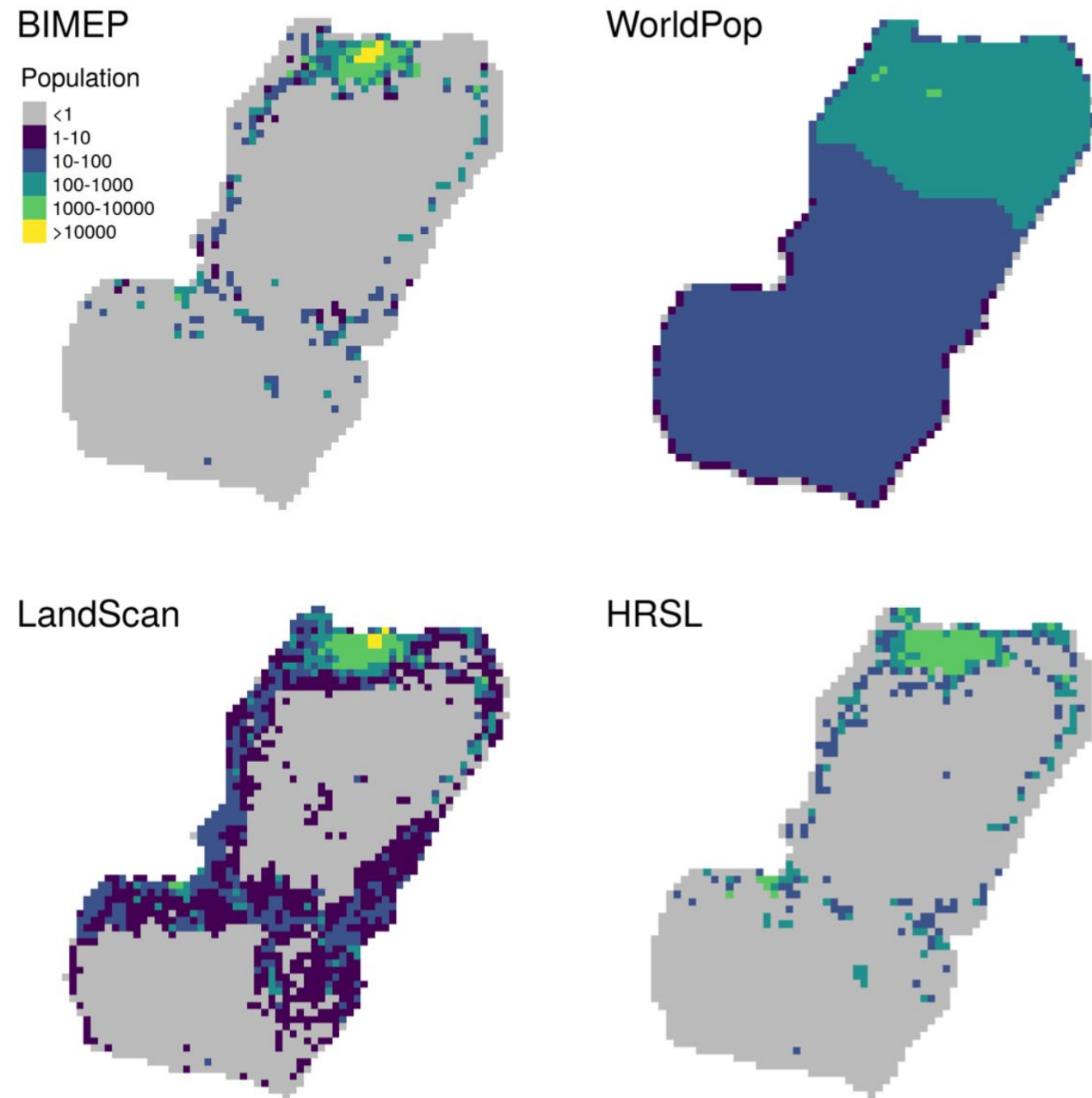
Fig. 2. Bioko population rendered at 1x1 km resolution. A. BIMEP; B. WP; C. LS; D. HRSL. Grey pixels represent uninhabited areas (population = 0).
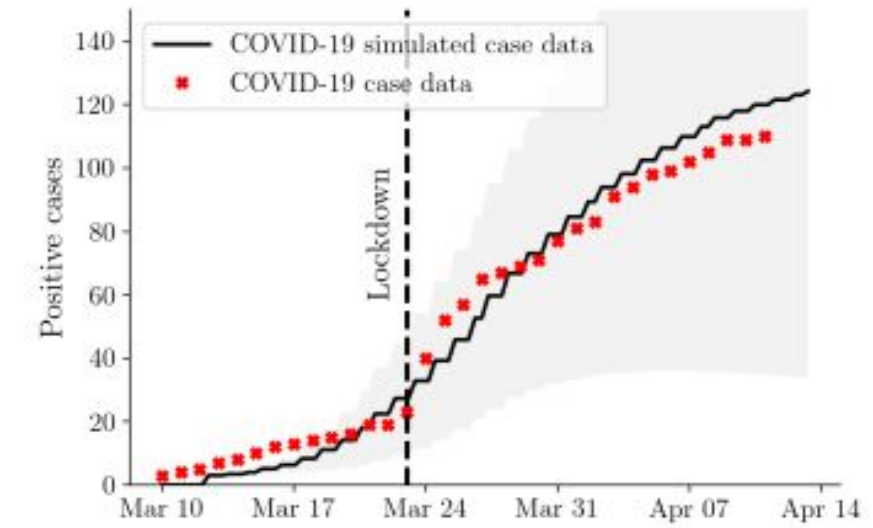
(a) Cumulative number of positively tested individuals
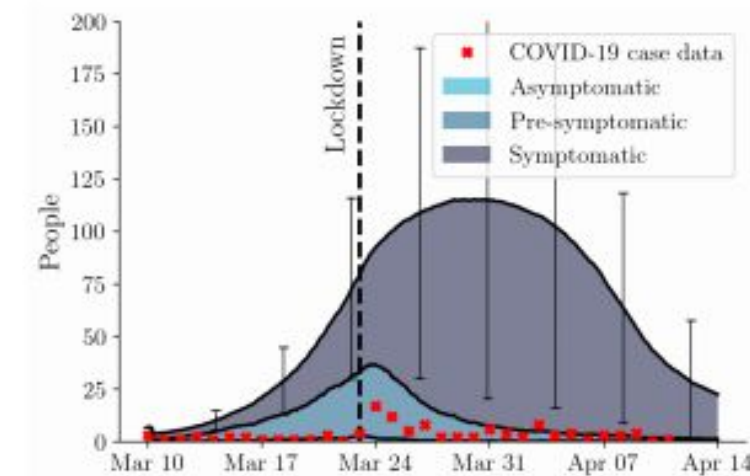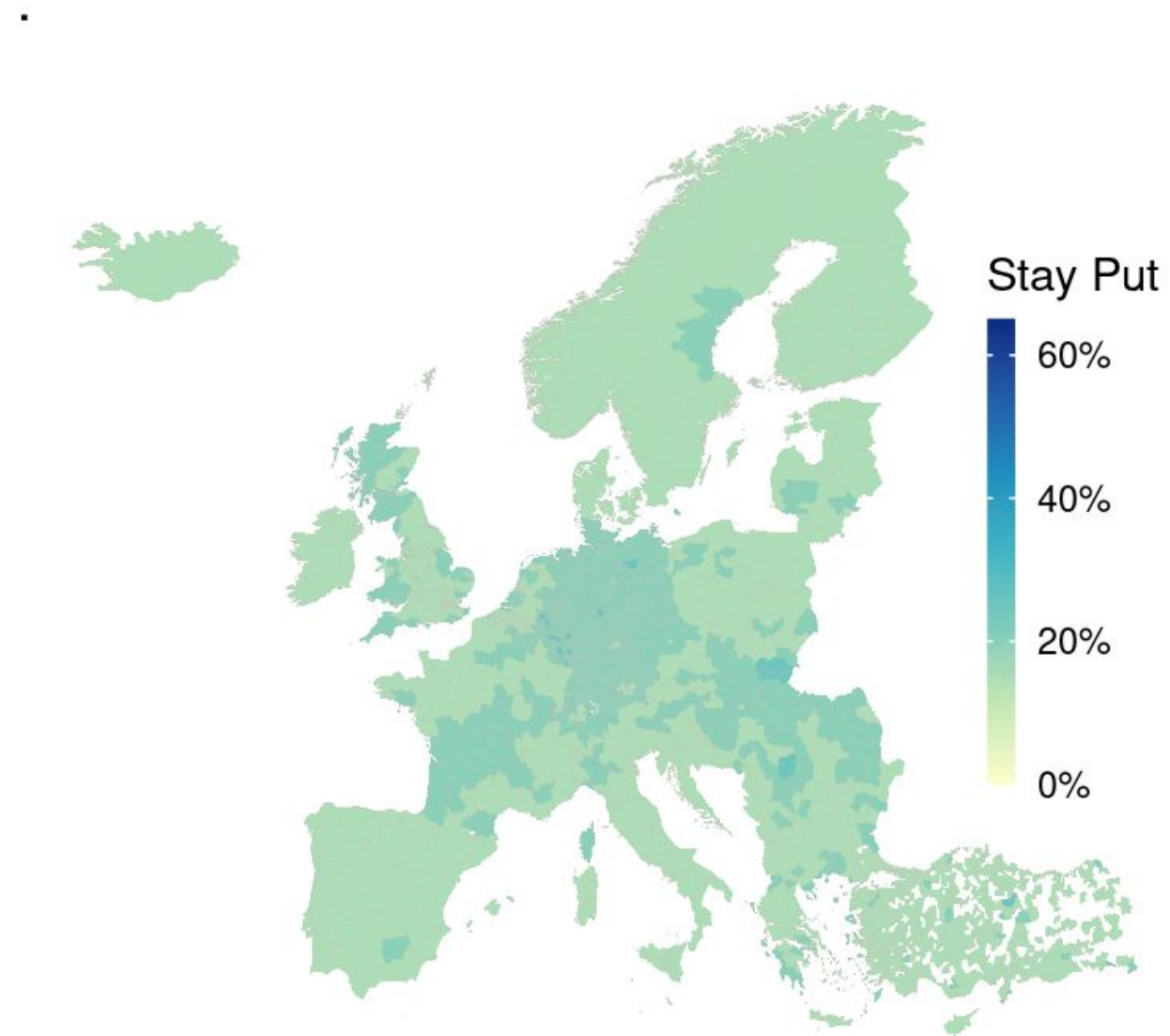

(b) Daily number of infected individuals
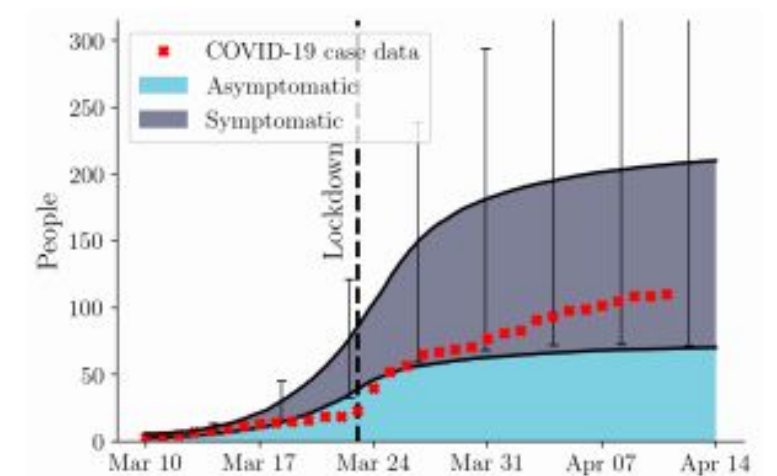
(c) Cumulative number of infected individuals

Figure 3: Predictions by our model using the parameters optimized to match COVID-19 case data before the introduction of restrictive measures by the government. Starting March 23, 2020, the model simulation implemented: social places and educational institutions close; public transport, supermarkets, offices reduce the rate of contamination by 50%; individuals reduce their check-in intensities by 20%. In panel (a), line and shading represent mean and standard deviation across 40 independent realizations of the model. In panels (b) and (c), lines represent means and error bars represent the standard deviation of total infections as given by the top line.
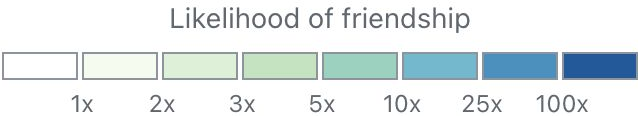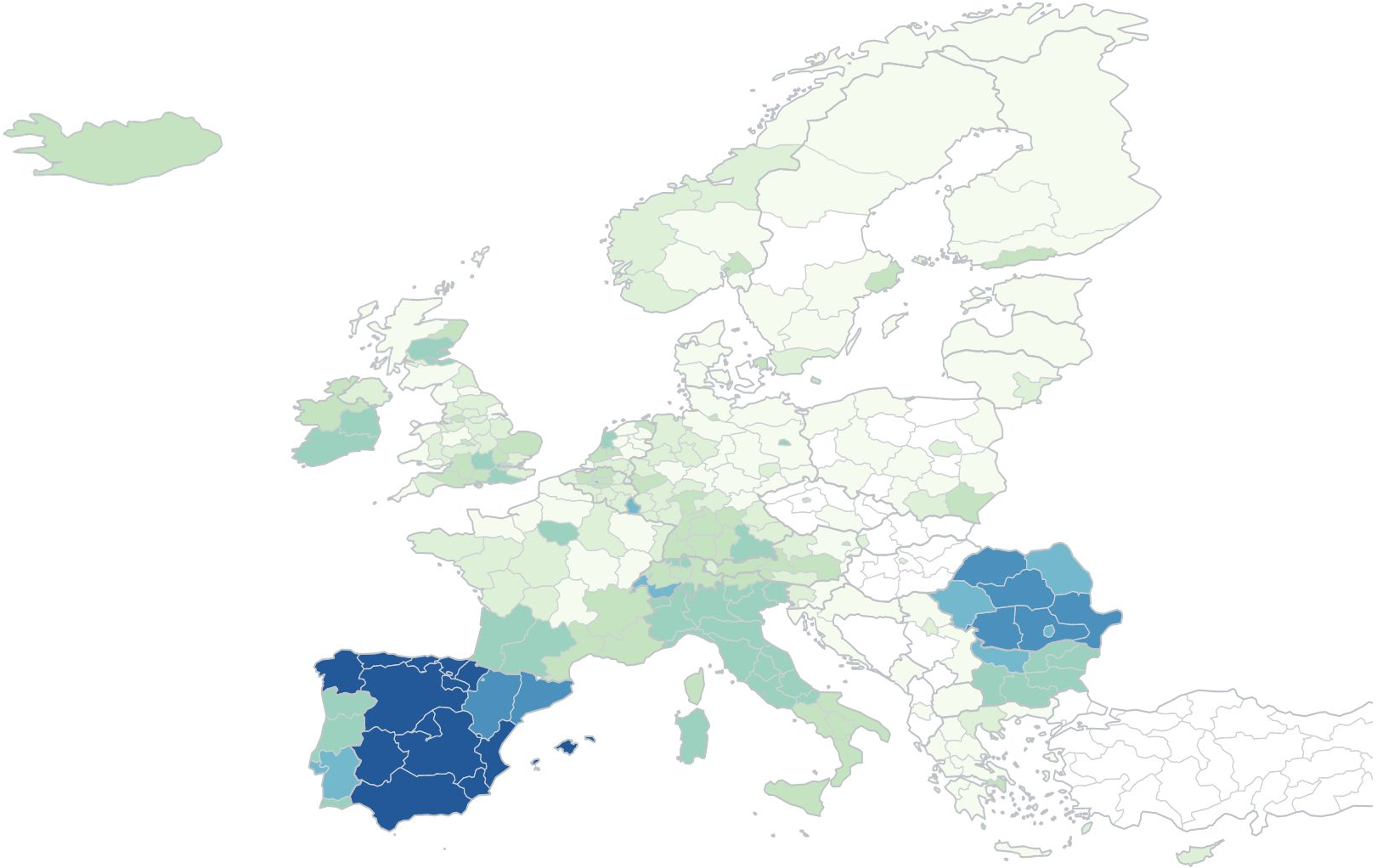
Movement Range Maps

(Used by governments in the context of COVID-19)
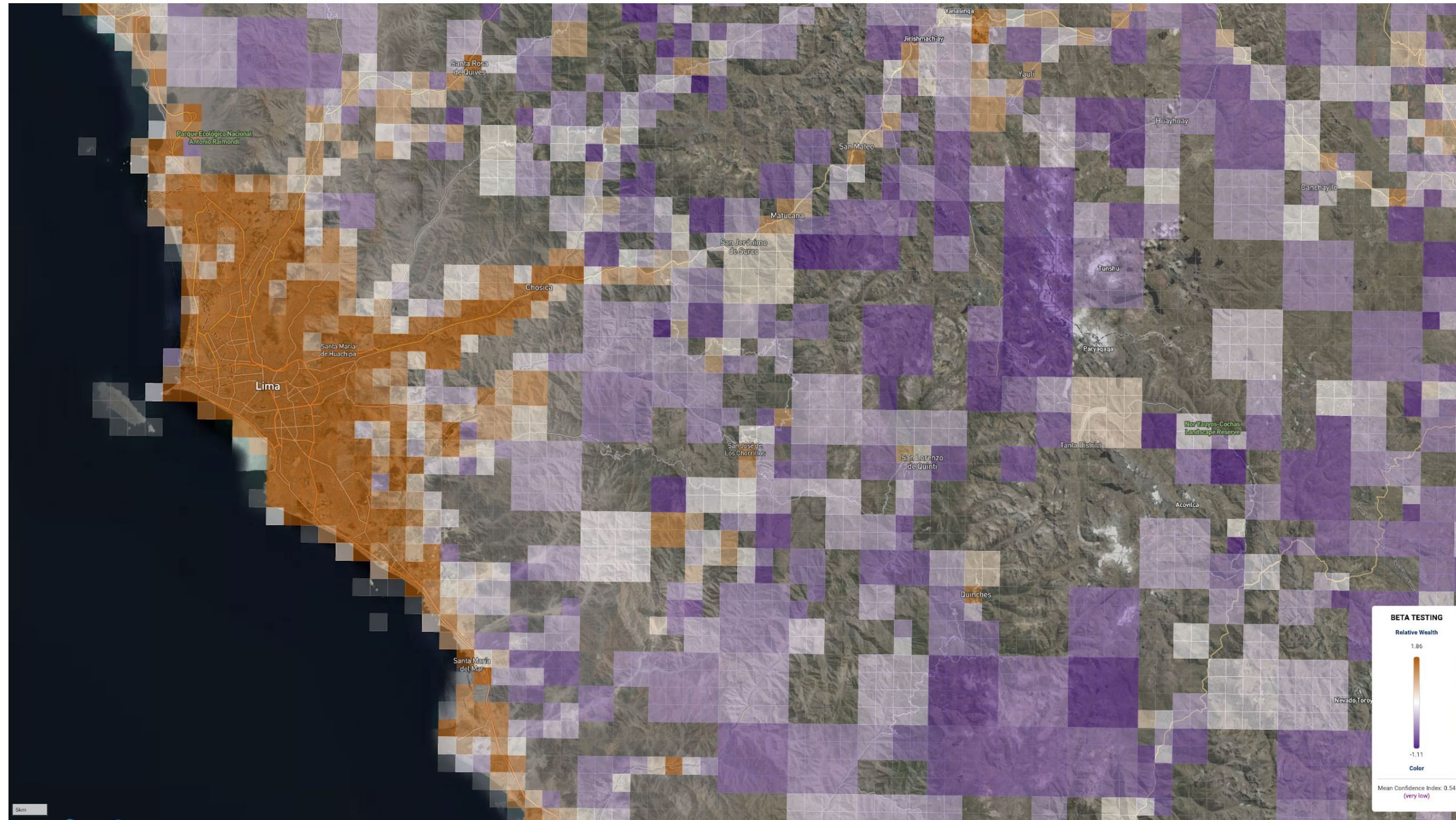

Date: 2020-03-04

Relative Wealth Index

Peru

|	Resolution	Source	Min	Mean	Median	Max
INPUT DATA						
Road density*	lat/lon	Open Street Map[1]	0	0.0007	0	0.08
Urban or built up*	15 arc-seconds	NASA (MODIS)[2]				
Elevation*	3 arc-seconds (~ 90 meters)	USGS[3]	-24	612	294	7643
Slope*	3 arc-seconds (~ 90 meters)	USGS[3]	0.0	0.024	0.008	2.108
Precipitation*	0.25-degree	NASA/Japan Aerospace Exploration Agency[4]	-48.7	1.7	0.0	2233.6
Population*	1 arc-second (~ 30m)	Humanitarian Data Exchange[5]	10	608	73	516163
# Cell towers[+]	2.4km tiles	Facebook[6]	0	13	0	71004
# WiFi access points[+]	2.4km tiles	Facebook[6]	0	369	1	1949963
# Mobile devices[+]	2.4km tiles	Facebook[6]	0	217	0	454962
# Android devices[+]	2.4km tiles	Facebook[6]	0	168	0	291831
# iOS devices[+]	2.4km tiles	Facebook[6]	0	49	0	204058
Nightlights / Radiance (VIIRS)*	15 arc-seconds	National Centers for Environmental Information Earth Observation Group[7]	0.0	1.9	0.3	58843.3
Satellite Imagery[+]	0.58 m/pixel	Digital Globe[8] (Bing tile 15)				

Survey Examples

Global State of Small Business Surveys (2021)
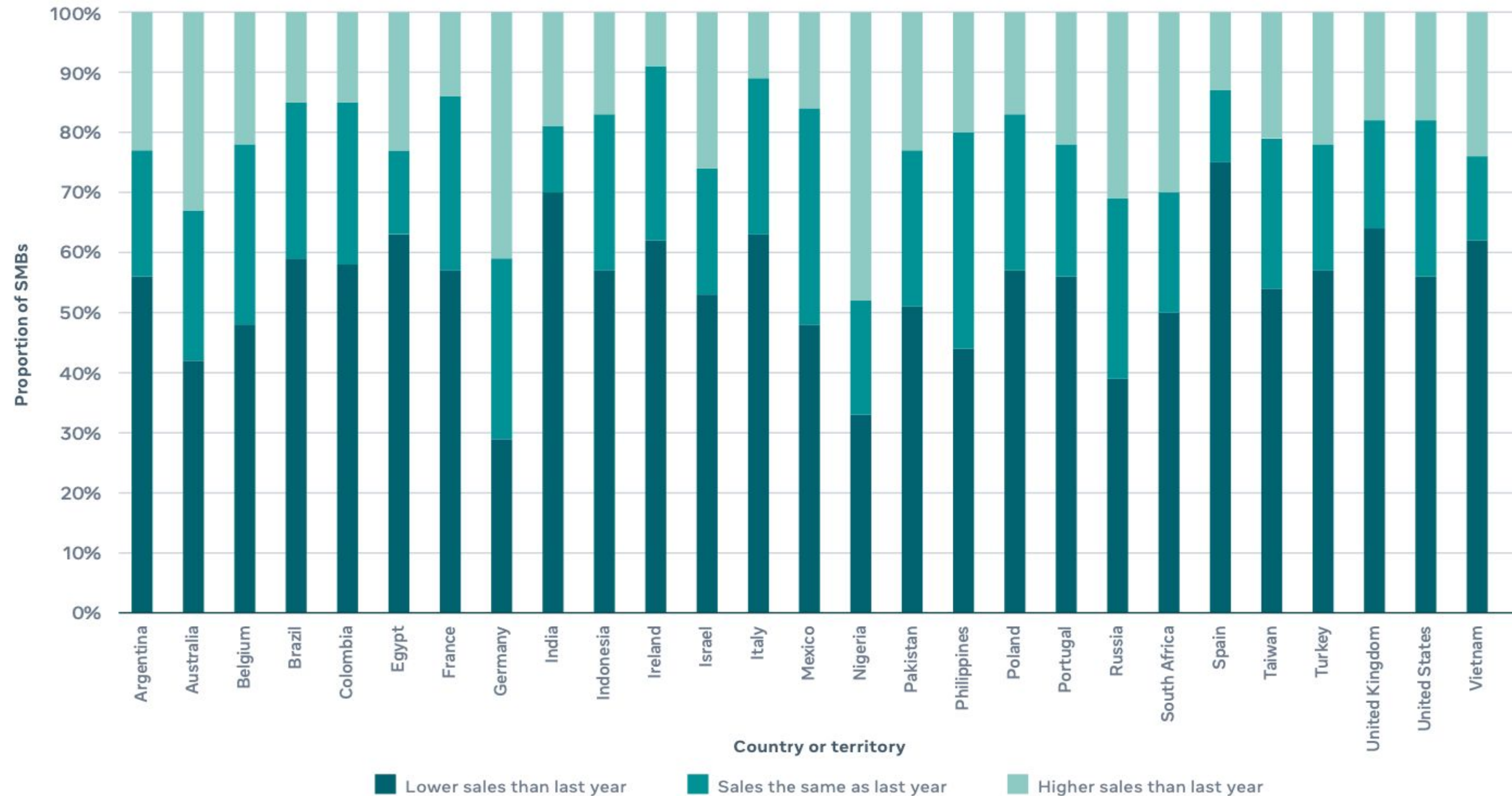
Covid-19 Trends & Impact Survey (2021)

Global COVID-19 vaccination acceptance
Weighted 7-day mean percent of respondents who said 'yes, definitely' or 'yes, probably'.
Responses from January 24 - January 30, 2021.
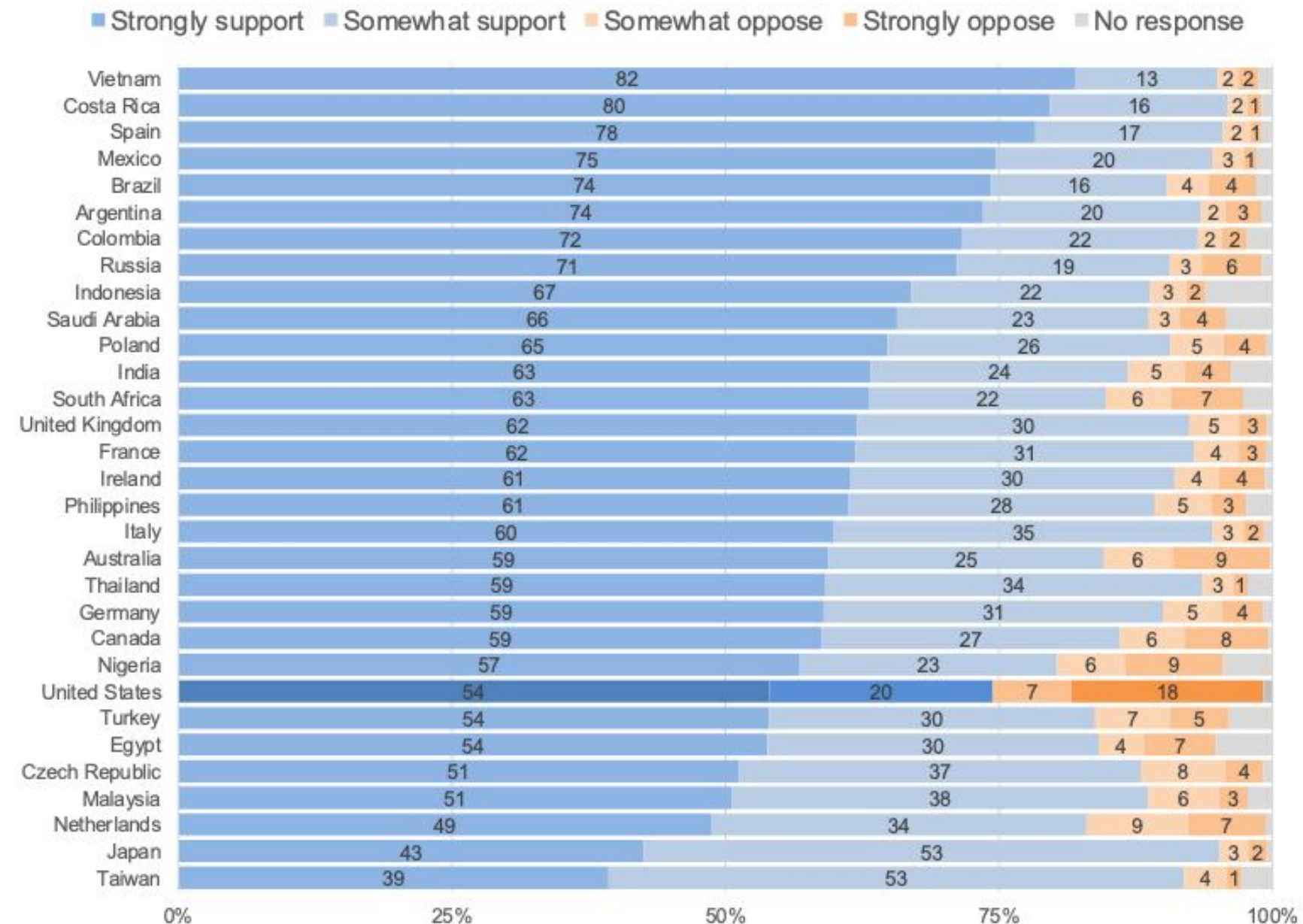
37%   74%
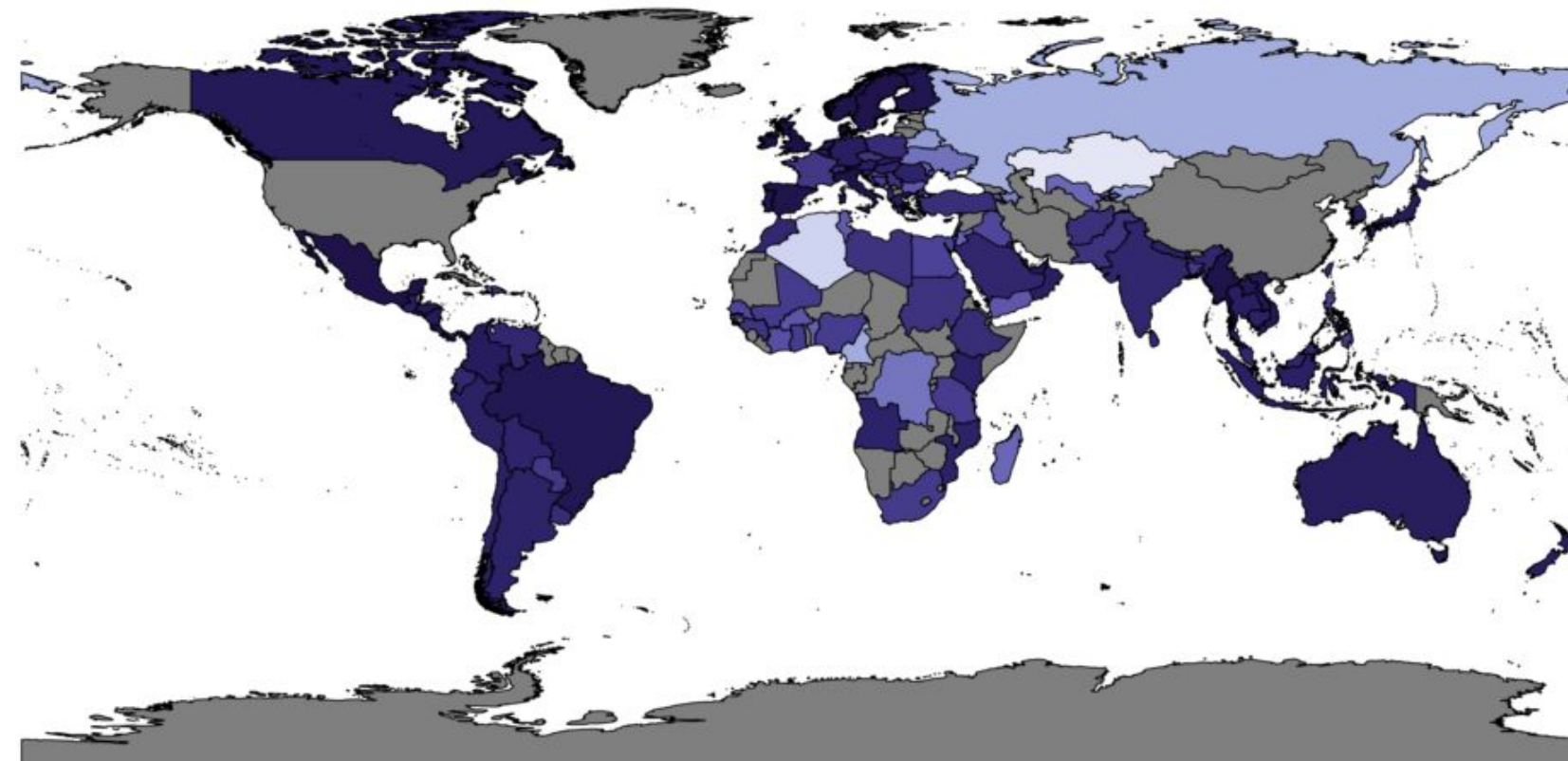
If a vaccine to prevent COVID-19 were offered to you today, would you choose to get vaccinated?
Asked of all survey respondents who had not already been vaccinated.
Data from COVID-19 Symptom Survey collected by the University of Maryland in partnership with Facebook, January 24-30, 2021.

Climate Change Opinion Survey (2021)

Strong International Support for the Paris Agreement

Legend: Strongly support · Somewhat support · Somewhat oppose · Strongly oppose · No response

Country	Strongly support	Somewhat support	Somewhat oppose	Strongly oppose	No response
Vietnam	82	13	2	2	
Costa Rica	80	16	2	1	
Spain	78	17	2	1	
Mexico	75	20	3	1	
Brazil	74	16	4	4	
Argentina	74	20	2	3	
Colombia	72	22	2	2	
Russia	71	19	3	6	
Indonesia	67	22	3	2	
Saudi Arabia	66	23	3	4	
Poland	65	26	5	4	
India	63	24	5	4	
South Africa	63	22	6	7	
United Kingdom	62	30	5	3	
France	62	31	4	3	
Ireland	61	30	4	4	
Philippines	61	28	5	3	
Italy	60	35	3	2	
Australia	59	25	6	9	
Thailand	59	34	3	1	
Germany	59	31	5	4	
Canada	59	27	6	8	
Nigeria	57	23	6	9	
United States	54	20	7	18	
Turkey	54	30	7	5	
Egypt	54	30	4	7	
Czech Republic	51	37	8	4	
Malaysia	51	38	6	3	
Netherlands	49	34	9	7	
Japan	43	53	3	2	
Taiwan	39	53	4	1	

In 2015, [name] signed an international agreement in Paris with 196 other countries to limit the pollution that causes climate change. Do you strongly support, somewhat support, somewhat oppose, or strongly oppose [name]'s participation in the Paris Agreement?
February 2021

YALE PROGRAM ON Climate Change Communication

Academic Collaboration in the EU

Academic collaboration in the Europe

Researchers from the **London School of Economics** used the Social Connectedness Index as part of a model to assess quality of life in Germany.

A research group from the **University of Catalunya** was the leading source of analysis on mobility data for the European Commission. Their reports were delivered weekly to the policy-makers for the EU.

Researchers from the **Mercator Research Center and others,** used **Facebook Movement Range** the impact on cycling traffic, health and mobility of provisional bicycle infrastructure in 106 European cities.

Researchers from the **Max Planck Institute and others** used Population Density Maps to develop a region-specific population model to estimate the impact that interventions like social distancing, contact tracing, and curfews.
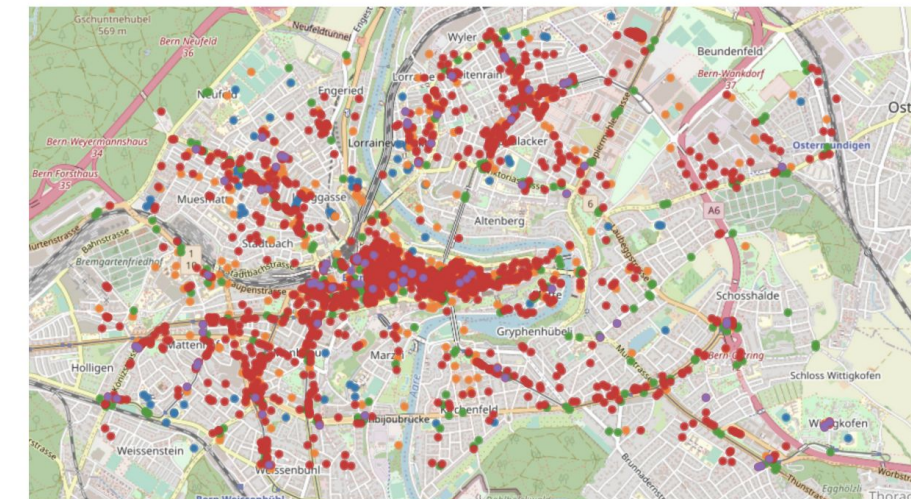
Figure 1: **Site locations by category in the mobility model of Bern, Switzerland.** Circles depict schools and research institutes (blue), social places (orange), bus stops (green), workplaces (red), supermarkets (purple).

Resources

Thank you

https://research.facebook.com/data/