July 5, 2022

The Honorable Ron Johnson
United States Senate
Washington, DC 20510

Dear Senator Johnson:

Thank you for your letter regarding mobility data and its use by the Centers for Disease Control
and Prevention (CDC) as part of the coronavirus disease 2019 (COVID-19) response.

CDC endeavors to use the best science available to inform our understanding of the public health
impacts of interventions and to inform recommendations. Responsible use of data science is a
key part of this understanding, and it allows the agency to leverage informative data assets in a
timely fashion while acknowledging, respecting, and safeguarding the sensitivities of these data.

In accordance with the agency's commitment to responsible stewardship of data, over the last
two years, CDC has used aggregated and anonymous population mobility data, derived from
mobile devices, for public health research on COVID-19 and other public health issues of
importance. These data are not tied to an individual and have multiple layers of privacy
protections to help ensure they cannot be misused or reidentified. These protections are crucial
for our agency, and the mobility data CDC uses meet these standards for aggregation and
anonymity.

As part of the COVID-19 response, CDC obtained aggregated data from commercial vendors to
better understand population-level impacts of COVID-19 policies and to shed important light on
other pressing public health problems. Given the aggregated level of these data, CDC does not
and could not use these data for monitoring compliance with COVID-19 orders or to track
individuals. Please see the attached enclosure for further information.

I appreciate your letter and support, and that of Congress overall, as we work together to fight
COVID-19. CDC remains committed to leading with science, promoting equity, and protecting
the American public during this pandemic. If you have further questions, please have your staff
contact Jeff Reczek in our CDC Washington Office at (202) 245-0600 or JReczek@cdc.gov.

Sincerely,

Rochelle P. Walensky, MD, MPH
Director, CDC

Enclosure

Enclosure

CDC Use of Population Mobility Data for Public Health Action

CDC's Legal Authority to Collect Mobility Data
Under the Public Health Service Act, CDC has general legal authority to request and collect data that can assist in its public health mission, including the ability to collect and make available information supporting public health research. Much of these data are collected through voluntary submissions, but CDC also has the authority to enter into contracts to procure data or services related to data, including contracts for research (See, e.g., 42 USC § 241). Additionally, the Department of Health and Human Services may take action in response to a public health emergency, including by entering into contracts and supporting investigations into the prevention of disease (42 USC § 247d).

The acquisition of mobility data is material to the investigation of population impacts of public health incidents of concern and associated mitigation measures. CDC appreciates the importance of responsible data stewardship and takes seriously the need to protect privacy and confidentiality. Mobility data obtained are aggregated and anonymized for public health research.

About the Data: Aggregated, Anonymized Mobility Data
Over the last two years, CDC conducted limited public health research related to the COVID-19 emergency and other public health problems using aggregated population mobility data— anonymized, aggregate data derived from mobile devices. These data have proven valuable for understanding tough public health problems that other data sources shed little light on, both for COVID-19 and other public health issues such as natural disaster response, chronic disease prevention, and toxic environmental exposures.

Since 2020, CDC has obtained mobility data from three commercial sources: SafeGraph, Cuebiq, and Google. For the first year of the pandemic, aggregated and anonymized location data from mobile devices were available free-of-charge to governmental and non-governmental agencies from these three sources. CDC used these data and publicly shared high-level mobility trends on CDC's COVID Data Tracker.[1] These data were available to CDC at no-cost until mid-2021. At that time, CDC, through the COVID-19 emergency response, purchased datasets from SafeGraph (April 2021) and Cuebiq (June 2021) to continue to understand how populations were returning to pre-pandemic activity levels. The county-level data from Google has always been, and remains, free and publicly available.

Privacy Protection: Multiple, Overlapping Layers
CDC did not seek, nor did it receive, individually identifiable data. CDC took measures to ensure that population mobility data it received were anonymized, contained no personally identifiable information (PII), and had robust privacy protections to help prevent against misuse or reidentification. These privacy protections fall into two general categories: (1) privacy controls and (2) procedural safeguards. Controls and safeguards are tailored to each dataset, and each source of data has a combination of protections designed to preserve privacy and prevent individual reidentification for that particular source.

---

[1] https://covid.cdc.gov/covid-data-tracker/#mobility

Privacy Controls
Privacy controls were integrated into each of the mobility datasets that CDC received and were inextricably tied to the data. Examples of these privacy controls included:

- **Anonymity**: Device mobility data were anonymous. Demographic characteristics or specific information about device user were unknown. There were no PII in the data. CDC did not have race, ethnicity, gender, age, or other demographic information for individuals related to these data.
- **Aggregation**: The mobility data CDC received were aggregated in one or more ways, meaning that the data were summarized by time (e.g., SafeGraph data were aggregated over a week) and/or location (e.g., in datasets, location data were aggregated to either census block group[2] or county).
- **Differential privacy** was applied to mobility datasets, meaning data providers added "noise" to the datasets to fuzz out individual data. Examples of differential privacy include adding additional data points or moving the true location of a point of interest to another place in the data to protect privacy in rural or low population areas. This same technique is applied to U.S. Census data as a best practice for protecting privacy.
- **Geomasking** is a method of protecting confidentiality while preserving the usability of data that includes changing the geographic location of an individual in an unpredictable way. For example, Cuebiq utilizes geomasking for inferred home and work locations, and then aggregates these data by census block group.
- **Suppression** of geographies and locations with small numbers. For example, geographic areas with fewer data points were not provided.
- **Sensitive Points of Interest (SPOI)**: For data acquired through Cuebiq, the company omitted data on SPOI. SPOI include, but are not limited to, churches, religious facilities, sensitive businesses, military and correctional facilities, social demonstrations, locations with firearms, and more.

Procedural Safeguards
Procedural safeguards were built into the procurement process by which CDC accesses and utilizes these mobility datasets. Examples of procedural safeguards include:

- **Privacy Controls**: In the procurement documents for these mobility data, CDC required that mobility data sent to CDC must be anonymous and exclude PII, while still applying standard contractual information security requirements, privacy review, and risk assessments of the data.
- **Controls on CDC access**: Access to these mobility datasets has been limited to a small number of CDC users.
- **Datasets hosted within the provider's platform**: For these mobility datasets, the data is hosted on the providers' platform, rather than within CDC. This limits what data CDC staff can download and allows data access to be audited by the providers.

Contracts with SafeGraph and Cuebiq
For both SafeGraph and Cuebiq, CDC and other public health entities received the aggregated, anonymized data under no-cost programs from the emergence of COVID-19 until mid-2021, at which time CDC entered into one-year contracts with SafeGraph (April 2021; $420,000) and

---

[2] A census block group is defined by the U.S. Census Bureau to contain between 600 and 3,000 people. For more information, see https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_4.

Cuebiq (June 2021; $208,000) through the COVID-19 response using funds from the Coronavirus Aid, Relief, and Economic Security (CARES) Act.[3] Both contracts have expired. CDC did not share data purchased under these contracts with other federal, state, or local agency or private companies, and such sharing is prohibited by these contracts.

Mobility Data: Storage, Retention, and Linkage
Storage of Mobility Data
As a privacy safeguard, datasets from both Cuebiq and SafeGraph were accessed through provider platforms, which are discussed in detail below. As both contracts have expired, CDC no longer has access to data through the provider platforms.

- For SafeGraph, aggregate Weekly Patterns and Neighborhood Patterns data were made available through SafeGraph's Amazon Web Services S3 Bucket.
- For Cuebiq, data were able to be queried via Cuebiq's proprietary Workbench Platform, which is a Platform as a Service (PaaS) tool. User-authenticated access to the platform was limited to five CDC staff, and aggregate data products were able to be downloaded only after review by Cuebiq staff to ensure privacy control standards were met.

Retention of Mobility Data at CDC
CDC currently retains the analytic outputs and aggregate data derivatives from the Cuebiq and SafeGraph data. These aggregated outputs and derivatives may be in use for active projects and may be retained per federal and agency requirements for records retention.[4]

With regard to data security, data stored in CDC servers are governed by privacy and security standards specified by the CDC Office of the Chief Information Security Officer. In addition, as part of the acquisition process, CDC determined that both the overall information security risk and the PII confidentiality impact level were low, as the datasets included no personally identifiable information.

Linkage of Mobility Data by CDC
CDC does not have access to demographic characteristics or PII related to these data and, as such, CDC cannot link these data at an individual level to another dataset. Researchers can compare aggregated outputs at the census block group, tract, or county level to publicly available datasets, like the U.S. Census. These high-level comparisons allow CDC to make inferences about the demographic trends in population movement.

CDC's Use of Mobility Data to Advance Public Health Mission
CDC has used mobility data to assess the impacts of COVID-19; in addition, the insights derived from these mobility data provide policy makers with helpful information on the impact and effectiveness of policies that had profound effects on communities. This work has included

---

[3] P.L. 116-136. https://www.govinfo.gov/content/pkg/PLAW-116publ136/pdf/PLAW-116publ136.pdf
[4] For more information on CDC's Records Control Schedule, see:
Centers for Disease Control and Prevention, *CDC Unified Process Guide: MASO Records Control Schedule*, https://www2.cdc.gov/cdcup/library/process_guides/doc/CDC_UP_Process_Guide_MASO.doc.

Enclosure

publications related to the impacts of business closures,[5] stay-at-home orders,[6] and other community mitigation efforts[7,8] on the spread of COVID-19 in populations.

Additionally, data from SafeGraph, Cuebiq, and Google have informed the CDC COVID Data Tracker: Mobility dashboard.[9] This public-facing dashboard provides an option to view mobility trends by state and county. CDC also utilizes Google's publicly available mobility reports to aid in showing the timing and impact of COVID-19 mitigation policies. Data are available at the state[10] and county[11] policy levels.

To date, work with mobility data beyond COVID-19 has been limited though these data can be useful for other applications to challenging health concerns, like natural disaster response, chronic disease prevention, and toxic environmental exposures. These additional activities were included in the statement of work for the two mobility data acquisitions in 2021 as a good practice for efficiency and to reduce the costs of obtaining these data for public health research.

Aggregated, anonymized data provide unique insights into public health questions that are important for CDC to investigate. The privacy controls and procedural safeguards in place protect against risk of reidentification or misuse of the data.

---

[5] Borjas GJ. Business Closures, Stay-at-Home Restrictions, and COVID-19 Testing Outcomes in New York City. *Prev Chronic Dis* 2020; 17:200264. DOI: http://dx.doi.org/10.5888/pcd17.200264

[6] Moreland A, Herlihy C, Tynan MA, et al. Timing of State and Territorial COVID-19 Stay-at-Home Orders and Changes in Population Movement — United States, March 1–May 31, 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:1198–1203. DOI: http://dx.doi.org/10.15585/mmwr.mm6935a2

[7] Buchwald AG, Bayham J, Adams J, et al. Estimating the Impact of Statewide Policies to Reduce Spread of Severe Acute Respiratory Syndrome Coronavirus 2 in Real Time, Colorado, USA. *Emerging Infectious Diseases.* 2021;27(9):2312-2322. http://dx.doi.org/10.3201/eid2709.204167

[8] Lasry A, Kidder D, Hast M, et al. Timing of Community Mitigation and Changes in Reported COVID-19 and Community Mobility — Four U.S. Metropolitan Areas, February 26–April 1, 2020. *MMWR Morb Mortal Wkly Rep* 2020; 69:451–457. DOI: http://dx.doi.org/10.15585/mmwr.mm6915e2

[9] Centers for Disease Control and Prevention, *COVID Data Tracker: Explore Human Mobility and COVID-19 Transmission in Your Local Area,* https://covid.cdc.gov/covid-data-tracker/#mobility.

[10] Centers for Disease Control and Prevention, *COVID Data Tracker: State-Issued Prevention Measures at the State-Level,* https://covid.cdc.gov/covid-data-tracker/#state-level-covid-policy

[11] Centers for Disease Control and Prevention, *COVID Data Tracker: State-Issued Prevention Measures at the County-Level,* https://covid.cdc.gov/covid-data-tracker/#county-level-covid-policy