



Fraym Data Sources & Methods

A Primer on Fraym's Data Sources & Methods

Data Sources & Methods Overview

A large, abstract teal graphic on the right side of the slide, consisting of several overlapping, curved shapes that resemble a stylized letter 'L' or a series of connected arcs.

Fraym analysis utilizes both survey microdata and spatial data.

First, we summarize the process and protocols taken to ensure quality training data – which is the foundation for spatial and traditional analysis.

Then we show how Fraym transforms training data into spatial data using our technology for subnational analysis.

Data Production Process Review

To measure indicators around vaccination beliefs and perceptions, Fraym collected and processed data through the following steps:

Survey

1

Comprehensive yet targeted questionnaire developed for cover key analytical concepts and relevant socio-demographic information for interpolation—all while incorporating contextual idiosyncrasies and survey methodological best practices. This novel data fills gaps in existing data.

Sample Design

2

Survey samples are designed to be nationally representative from a demographic and socioeconomic perspective. This ensures that Fraym can accurately predict across the target geography for every indicator of interest.

Spatial Data Production

3

Fraym's proprietary machine learning technology takes variables from the novel survey, combines them with hundreds of spatial, demographic, and socioeconomic datasets as covariates and outputs an interpolated raster layer with zip code level estimations for proportions and populations for all key indicators.

Analysis

4

Using the interpolated output, Fraym runs analyses on the indicators of interest, including how they correlate and vary spatially, as well as aggregates the data to the county and state level.

Sample Design

Quality training data is critical to producing valid and actionable insights.

- Fraym worked with multiple vendors to conduct an online survey with zip code level targeting to collect a large, nationally representative sample of American adults.
- Fraym set quotas based on 4 categories: **demographic, education, race / ethnicity, and geographic regions** to ensure representativeness across the data inputs to our models.
- Wave 2: Data collection took place in **September and October of 2024** in both **Spanish and English**. The total sample size was **16,737 respondents ages 18+¹**.

QUOTA CATEGORIES



Demographic



Geographic



Race / Ethnicity



Education

Training Data Overview

Data Collection:

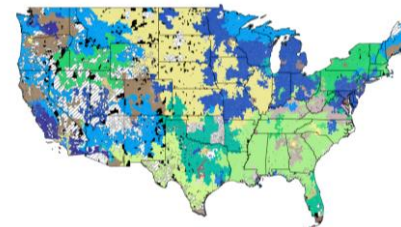
Initial Covariates: 683
 Sampling Units: 32,931
 Sampling Unit: zip code
 Data Sources: US Census Bureau

Covariate Types:

1. 2020 Decennial Census
2. Land Area / Water Area
3. Educational Attainment
4. Income
5. Urban/Rural Land use

Stratification

Fraym uses a novel approach to subdivide the country into near homogeneous groups at the zip code level.¹



Random Sampling

Within each stratum, random sampling occurred twice at two different hierarchical levels.

Zip Code Selection

Level 1: Within each stratum, zip codes are randomly selected.



Respondent Selection

Level 2: Within each zip code, respondents are randomly selected.



Survey Ethics: Risks, Privacy and Confidentiality

Fraym and our vendors take deliberate and thoughtful steps to safeguard the privacy and anonymity of respondents, and any data collected from them.

Regarding privacy, **Fraym does not collect or include personally identifiable information (PII)** such as names, home addresses, email addresses, or mobile phone numbers. Our vendors adhere to industry-standard best practices on data encryption to safeguard the privacy and anonymity of survey respondents and any data collected.

Access & Motivation Scores

The Access and Motivation Scores were constructed with Multiple Correspondence Analysis (MCA) using categorical indicators covering themes around vaccine concerns, considerations, beliefs, worries, and medical care.

Indicators for Access

Accessibility is defined by the individual's access to medical care combined with knowledge and concerns around cost and convenience that may act as barriers to getting vaccinated.

INDICATOR	DESCRIPTION
Medical Care: Resources and costs associated with health care	
Primary care physician	Individual currently has a primary care physician
Health insurance	Individual currently has health insurance coverage
Healthcare costs	Individual reports not having to delay medical care due to costs in the past year
Concerns: Perceived barriers when deciding to get vaccinated	
Online appointment	Individual is not concerned about getting an appointment online
Vaccine location knowledge	Individual is not concerned about knowing where to get a vaccine
Vaccine site difficulty	Individual is not concerned about getting to the vaccination site
Separate appointments	Individual is not concerned about having to schedule separate appointments for different vaccines
Eligibility	Individual is not concerned about being eligible for vaccine
Cost concerns	Individual is not concerned about the costs associated with getting the vaccine, including clinic payments, transport and time costs
Open times	Individual is not concerned about vaccination sites not being open at convenient times

Indicators for Motivation

Motivation is defined as the individual's concerns, confidence, and beliefs around infectious diseases and vaccines.

INDICATOR	DESCRIPTION
Concerns: Worries about the disease and/or about the vaccine	
Worries about disease	Individual is worried about being infected by disease
Vaccine safety	Individual is not concerned about vaccine safety
Vaccine effectiveness	Individual is not concerned about vaccine effectiveness
Vaccine side effects	Individual is not concerned about vaccine side effects
Vaccine disease	Individual is not concerned that the vaccine may cause the disease it is supposed to prevent
Trusted peers	Individual is not concerned that their trusted peers are not getting the vaccine
Confidence: Factors that related to individual's confidence in the vaccine	
Recommended by medical professional	Individual has been advised by a medical professional to get vaccinated
Beliefs: Views regarding the nature of the disease	
Community spread	Individual believes that the disease is spreading in their community
Healthy infection	Individual believes that despite being healthy, getting infected matters
Friends and family risk	Individual believes their friends and family are at risk of the disease
Disease is real	Individual believes the disease is real



Have any questions? Contact us:

s.walsh@fraym.io