

category¹⁴. Some states like North Carolina do not disaggregate their race/ethnicity categories by gender or age but do disaggregate them by week. This is extremely helpful for analyzing the efficacy of a state's various rollout strategies over time. Many state data dashboards only show cumulative demographics.

What the missing data hides

Without a stable source of race and ethnicity information, governments are unable to prioritize the segments and geographics of the population that need the vaccine the most. I set out to attempt to construct a more accurate picture of the demographics of people getting the vaccine in the hopes that I help could fill that knowledge gap.

I used location data from [SafeGraph](#), a company that aggregates anonymized location data from numerous applications. The applications, which include weather, navigation, and social media apps, report the data to SafeGraph when users have opted in to location tracking. The apps span nearly 47 million mobile devices in the United States. The location data is aggregated by certain points-of-interest (POIs) such as schools, airports, and parks.

SafeGraph provides free access to its data for academic researchers. I applied to the program and was granted access. I then downloaded all the data from 2021 in its Weekly Patterns dataset. Weekly Patterns tracks the number of people that visit a POI each week. It breaks down that count by the home census block group of the visitors. (A census block group (CBG) is a subdivision of a census tract; a home census block group is the CBG that SafeGraph [thinks](#) an individual's home is in.) For example, the data is able to tell me how many people who live in a certain CBG traveled to a certain POI (e.g. CVS pharmacy) in a given week. However, the data does not tell me who any of the people are, what time they went, or where precisely they live.

¹⁴ *COVID-19: vaccine summary data through Apr 13, 2021*, http://www11.doh.state.fl.us/comm/_partners/covid19_report_archive/vaccine/vaccine_report_latest.pdf

In summary, SafeGraph gives me an estimate of, for each point-of-interest that SafeGraph tracks (roughly 4.5 million in the United States), how many people visit it in a given week and from which census block group. If I can find which POIs distribute vaccines, and if I can find the demographic information of the census block groups of the sites' visitors, then I can construct an estimate of the demographic makeup of vaccine recipients.

To lessen the computation time and policy research required to make a reliable estimate, I limit this analysis geographically to Richmond City, Virginia. I picked Richmond because it is a densely populated urban area and, as of April 10, 37.4% of vaccine data records did not include race or ethnicity information (Dottle et al., 2021). This measure ranks Richmond relatively low compared to the fifty states, according to data from the Kaiser Family Foundation (KFF, 2021). Richmond is also worthy of investigation given that that percentage of missing data has grown in the last month, as evidenced by the blue line (the second highest line) in Fig. 1.

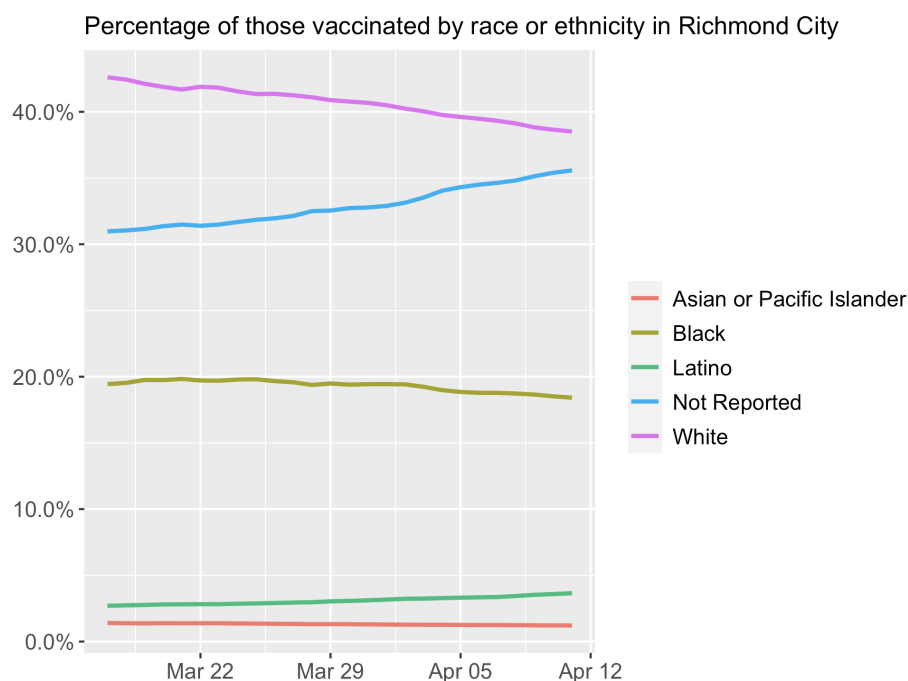


Fig. 1. Data from the [Virginia Department of Health](#).

To find out the POIs that distribute vaccines, I downloaded a master list of vaccine providers¹⁵ compiled by URISA's GISCorps, a program that offers volunteer geographic information systems services. I also scraped the CDC's VaccineFinder website, which lists additional local pharmacies. I then filtered out the providers that weren't in Richmond. The final list contains the latitude-longitude coordinate of each vaccine provider. SafeGraph's POI dataset is a collection of polygons, each of which represents the footprint of the POI's building. I used Python to match each latitude-longitude point to a POI by finding a polygon that contained that point. In the case where the polygons of several POIs encompassed a point, I assigned the point to the most reasonable POI (e.g. if a point was inside both a CVS and the surrounding Mall of America, CVS was likely the true vaccine provider).

I was able to find the SafeGraph points of interest for 27 of the 29 vaccine providers I found. The Weekly Patterns data, for each POI in each week, details the home CBG for each visitor to that POI. For each home CBG, I downloaded its proportion of white, Black, Asian, and Hispanic residents, as well as its median income, from the United States Census.

The 27 providers fell under the following 7 corporations. Through press releases and news articles, I identified the opening date of each provider and only considered Weekly Patterns data for each provider from after their opening date: CVS has offered vaccines in Richmond since February 12, 2021 (North, 2021); Walgreens since February 24 (Fox, 2021); JenCare Senior Medical Center since February 23 (ChenMed, 2021); Kroger since February 24 (Fox, 2021); MinuteClinic since February 12 (North, 2021); Publix Super Markets since March 25 (Brous, 2021); Rite Aid since January 21 (WAVY.com, 2021); Walmart Pharmacy #10-2821 since February 24 (Fox, 2021); Westwood Pharmacy since January 19 (Fox, 2021).

¹⁵ https://covid-19-giscorps.hub.arcgis.com/datasets/c50a1a352e944a66aed98e61952051ef_0

I then took the weighted average of the proportion of each demographic category of the CBGs that had residents visit vaccine providers since February. I reweighted¹⁶ the data each week to match the vaccine allocations¹⁷ that were given to Virginia.

The results, as of April 11, are in the second column of the table below. The third column contains the racial/ ethnic makeup of vaccine recipients in Richmond in the same period. That makeup does not include the 37.4% of vaccine recipients who did not have their race/ethnicity reported. The fourth column contains the overall racial/ethnic makeup of Richmond City according to the 2015–2019 American Community Survey (U.S. Census Bureau, 2019).

Race or ethnicity	SafeGraph-based estimates	VA Department of Health data	Richmond City Demographics
White	59.7%	59.0%	45.5%
Black	31.2%	28.6%	46.9%
Asian	3.7%	1.8%	2.1%
Hispanic	5.5%		6.9%
Latino		6.4%	

As stated before, the Virginia Department of Health only reports race and ethnicity for 62.6% of Richmond City’s vaccination data. We see that the incomplete data reported by the state seems to overrepresent the share of White people getting vaccinated by about 2.7 percentage points. The SafeGraph analysis also indicates that Black people are being vaccinated at a share of 6 percentage points above what is being reported, and Asian people are being vaccinated at a share of 1.7 percentage points above what is being reported.

There are many nuances in Virginia’s vaccine rollout that are ignored in this preliminary analysis. For one, this model does not account for days when a pharmacy had no appointments

¹⁶ <https://www.safegraph.com/blog/measuring-and-correcting-sampling-bias-for-accurate-demographic-analysis>

¹⁷ See the following datasets: COVID-19 Vaccine Distribution Allocations by Jurisdiction for [Pfizer](#), [Moderna](#), and [Janssen](#).

available due to a lack of vaccine supply. Moreover, the SafeGraph data itself is nowhere near perfect. A bias audit of SafeGraph’s data published one month ago — the first of its kind — found that older and non-White people were less likely to be captured by mobility data (Coston et al., 2021). (The study did not break out “non-White” into more specific groups.)

As of April 10, just three states — Wyoming, New Hampshire, and Montana — neither report race nor ethnicity data, down from thirty-three in mid-January (Goldfarb & Schechtman, 2021). But though 47 states now report this data, 44% of race and ethnicity data is still shielded from federal view. At the state level, there is a large range on data quality, best shown in the following visualization by Rachael Dottle in Bloomberg News (Dottle et al., 2021):



Each circle represents a state¹⁸. The deeper the purple, the more reliable the data. We see that there is a large gap among the states, and a good number of states are struggling to make their way to the right side. This sparsity was caused by the lack of a standard data collection and presentation process from the federal level. For future and present public health crises, governments must regulate and bolster this critical piece of infrastructure to ever be able to claim to be prioritizing racial inequality.

¹⁸ For instance, the rightmost circle represents Maryland, which has race data for 95% of vaccine recipients. The leftmost circle represents Wyoming

Works Cited

- Brous, M. (2021, March 24). Publix Pharmacy Opens Appointments for COVID-19 Vaccinations in Three States. *Valdosta Daily Times*.
https://www.valdostadailytimes.com/news/business/publix-pharmacy-opens-appointments-for-covid-19-vaccinations-in-three-states/article_51388f06-ecf3-5a6a-a5af-d68a42bc1007.html
- ChenMed. (2021). *ChenMed Accelerates Safe Covid-19 Vaccine Administrations to Tens of Thousands of High-Risk Seniors Where They Receive Care*. Retrieved April 17, 2021, from
<https://www.prnewswire.com/news-releases/chenmed-accelerates-safe-covid-19-vaccine-administrations-to-tens-of-thousands-of-high-risk-seniors-where-they-receive-care-301233089>
- Cobler, N. (2021). Early state data point to higher vaccination rates among whites than other groups. *Austin American-Statesman*. Retrieved April 14, 2021, from
<https://www.statesman.com/story/news/coronavirus/2021/01/23/texas-provides-murky-covid-19-vaccine-data-race-ethnicity/6662987002/>
- Coston, A., Guha, N., Ouyang, D., Lu, L., Chouldechova, A., & Ho, D. E. (2021). Leveraging Administrative Data for Bias Audits: Assessing Disparate Coverage with Mobility Data for COVID-19 Policy. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 173–184. <https://doi.org/10.1145/3442188.3445881>
- Dottle, R., Gardner, A., & Green, J. (2021, April 7). U.S. South Is Fast Closing Country's Widest Black Vaccine Gaps. *Bloomberg News*.
<https://www.bloomberg.com/graphics/covid-vaccine-tracker-global-distribution/us-vaccine-demographics.html>
- Fox, A. (2021, February 23). Some Hampton Roads pharmacies beginning COVID-19 vaccine efforts this week. *WAVY.Com*.

<https://www.wavy.com/covid-19-vaccine/the-latest-some-hampton-roads-pharmacies-and-their-customers-await-covid-19-vaccine-delivery-this-week/>

Goldfarb, A., & Schechtman, K. W. (2021, January 15). *Analysis & updates | State-Level Vaccine Demographic Data is Messy and Incomplete—We Need Federal Data, Now*. The COVID Tracking Project. <https://covidtracking.com/analysis-updates/state-level-vaccine-demographic-data-is-messy-and-incomplete>

Harris, C. (2021, January 28). Texas to require providers to report race data on vaccine recipients. *Houston Chronicle*. <https://www.houstonchronicle.com/coronavirus/article/texas-to-report-race-data-covid-vaccine-patients-15905721.php>

KFF. (2021, April 14). *COVID-19 Vaccinations by Race/Ethnicity*. The Kaiser Family Foundation.

<https://www.kff.org/other/state-indicator/covid-19-vaccinations-by-race-ethnicity/>

North, E. (2021, February 4). CVS expanding Virginia vaccination locations from 28 to 36. *8News*. <https://www.wric.com/health/coronavirus/cvs-expanding-virginia-vaccination-locations-from-28-to-36/>

Painter, E. M., Ussery, E. N., & Patel, A. (2021). Demographic Characteristics of Persons Vaccinated During the First Month of the COVID-19 Vaccination Program—United States, December 14, 2020–January 14, 2021. *MMWR. Morbidity and Mortality Weekly Report*, 70. <https://doi.org/10.15585/mmwr.mm7005e1>

Porter, F. (2021, February 9). *NC Gov. Cooper: North Carolina Continues its Commitment to Equitably Distribute COVID-19 Vaccines*. <https://governor.nc.gov/news/north-carolina-continues-its-commitment-equitably-distribute-covid-19-vaccines>

U.S. Census Bureau. (2019). *American Community Survey 5-Year Estimates*.

<https://www.census.gov/programs-surveys/acs>

WAVY.com. (2021, January 21). *Here's when the COVID-19 vaccine will be available at your Hampton Roads pharmacy*.

<https://www.wavy.com/covid-19-vaccine/pharmacies-administering-covid-19-vaccine/>

Wells, K. (2021, February 23). *State releases race data for vaccinations, but nearly half of it is missing*. [https://www.michiganradio.org/post/state-releases-](https://www.michiganradio.org/post/state-releases-race-data-vaccinations-nearly-half-it-missing)

[race-data-vaccinations-nearly-half-it-missing](https://www.michiganradio.org/post/state-releases-race-data-vaccinations-nearly-half-it-missing)