

# Democracy Fund + UCLA Nationscape User Guide

FEBRUARY 2020

## TABLE OF CONTENTS

Acknowledgements.....	3
Release .....	4
Methodology and Representativeness Assessment.....	5
Data Collection .....	6
Data Analysis, Weights, and Variance Estimation.....	7
Orientation to Data Files.....	8
Terms of Use .....	9
Appendices: Change Log, Weighting procedure.....	10

## **Acknowledgements**

Nationscape is funded by Democracy Fund Voter Study Group and UCLA with additional support from the Marvin Hoffenberg Chair in American Politics and Public Policy at UCLA through a partnership agreement between these parties.

The study was designed and data collection was managed by Nationscape Principal Investigators, Chris Tausanovitch (Associate Professor, Department of Political Science, UCLA) and Lynn Vavreck (Marvin Hoffenberg Professor of American Politics and Public Policy, UCLA) with design input from the Democracy Fund Voter Study Group, including former research director, John Sides (Professor, Department of Political Science, Vanderbilt University) and Robert Griffin, who currently holds that position.

UCLA graduate students staff Nationscape and execute the fielding of the survey each week. They are: Alex Rossell Hayes, Project Manager; Tyler Reny, Project Coordinator; and Aaron Rudkin, Project Developer.

Data collection is performed by LUCID, Inc., which collaborated with Nationscape personnel on aspects of the study's design, sampling plan, and feasibility. Lead personnel on the project at LUCID include: Lauren Astrachan, Analyst, Project Success; Ashley Burkardt, Lead, Project Success; Kyle Cousans, Senior Manager, Sales Engineering; and Allison St. Martins, Associate Account Manager.

## Release

This document is a brief guide to Democracy Fund + UCLA Nationscape — accompanying the release of Phase 1 data in January 2020.

Nationscape is a 16-month election study conducted by researchers at UCLA. This is the first year the project has existed. The project completes roughly 6,250 interviews each week. It started in July of 2019 and will conclude in January of 2021. Phase 1 of the data, released in January of 2020, includes nearly 156,000 cases collected over 24 weeks. Data collection began with the week of July 18, 2019 and concluding with the week of December 26, 2019 (last interview on January 1, 2020).

Each weekly survey is released as its own dataset. There is no cumulative file release.

## Current Release

This Winter 2020 data release (Phase 1, version 20200131) is the first public release of Nationscape data, as such there are no additions or changes to prior releases at this time.

If you find errors or have comments or question about the data, please communicate with Nationscape Project staff by writing to Chris Tausanovitch at [ctausanovitch@ucla.edu](mailto:ctausanovitch@ucla.edu).

## Citation

Citations to the dataset should be to the weekly studies as follows:

Tausanovitch, Chris and Lynn Vavreck. 2020. *Democracy Fund + UCLA Nationscape, October 10-17, 2019* (version 20200131). Retrieved from [URL].

## **Methodology and Representativeness Assessment**

A detailed discussion of the sampling methodology, survey execution, and assessment of the representativeness of Nationscape can be found in:

Tausanovitch, Chris and Lynn Vavreck, Tyler Reny, Alex Rossell Hayes, and Aaron Rudkin. 2019. "Democracy Fund + UCLA Nationscape Methodology and Representativeness Assessment."

This document can be downloaded [here](#) by clicking on the "Nationscape Methodology and Representativeness Assessment" link under "Data Resources." These analyses were performed in July of 2019 and use an initial set of weights that have since been updated to allow for more specificity.

A description of the current weighting methodology can be found at the end of this document.

## Data Collection

**Field dates:** Nationscape conducts weekly surveys. The first wave went in to the field on July 18, 2019. For this release the last wave went in to the field on December 26, 2019. The name of each dataset reflects its first field date. Each survey is in the field one week.

**Mode of interview:** Interviews are conducted online anywhere the respondent has access to a networked computer.

**Interview length:** The weekly questionnaires are designed for a 15-minute median administration time, not counting screening time by the sample provider. Observed median times to completion vary from 13.5 to 16 minutes.

**Languages:** Nationscape pilot surveys were offered in Spanish and English with respondents able to choose their language question by question. Due to lack of use during the pilot phase, the Spanish option was discontinued upon official launch of wave 1.

**Participation rates:** On average, across all waves in this release, of those selected to be interviewed, 12.2 percent declined immediately. Another 5.1 percent dropped off elsewhere in the survey without completing. We removed 7.9 percent for speeding or straight-lining through the survey. This results in an average yield of 74.8 percent of the original invited sample. We define speeding as completing the survey in fewer than six minutes and straight-lining as selecting the same response for every question in the three-policy question batteries.

## Data Analysis, Weights, and Variance Estimation

Nationscape data can be analyzed using individual waves of the survey or any combination of waves. The number of completed interviews each week is roughly constant, so no week should dominate analysis in combined datasets, but researchers should be aware that some weeks have more completions than others.

Analyses should be weighted to accurately represent the population of interest. Sampling error calculations should account for the sample design and the effects of weighting on variance.

### Weights

We provide weights in the “weight” variable that can be used for a national sample.

The sample is not a simple random sample nor is it a random sample of any kind. Please see the “Democracy Fund + UCLA Nationscape Methodology and Representativeness Assessment” (Representativeness Assessment linked earlier in this document; description of the weighting procedure included at the end) for details on how weights are constructed to hit national targets and which national targets are used.

Instead, we use purposive sampling (selecting respondents based upon their characteristics) to obtain a sample that is constructed to be representative of the population in terms of a specified set of characteristics.

Researchers should take care in calculating margins of error or standard errors due to the non-random nature of the sample. Standard calculations based on random sampling will underestimate the true magnitude of random error in our sample. As a starting place, we recommend the adjustment proposed by Rivers and Bailey (2009).<sup>1</sup>

---

1 Rivers, Douglas, and Delia Bailey. “Inference from matched samples in the 2008 US national elections.” *Proceedings of the joint statistical meetings*. Vol. 1. Palo Alto, CA: YouGov/Polimetrix, 2009.

## Orientation to Data Files

### Data files

The data files being released in Phase 1 are constructed for delivery in Stata (.dta) format and we recommend this download. Each file contains a “readme” file to describe special features of that week’s data collection, if any.

### Variables

There are approximately 230 variables in each weekly file.<sup>2</sup> They are named to reflect the topic they measure. For example, party identification in three categories is named “pid3.” Each file contains variable labels and value labels for each variable.

Attached to this document is a spreadsheet of when each question was asked, wave by wave.

### Missing data

Missing data are mainly coded to indicate the reason they are missing using the following codes and categories:

888 Not asked of this respondent  
999 Not sure, don’t know

### Codebook

Each weekly survey has its own codebook and banner book of results. Nationscape staff have made every effort to maintain consistency across the survey waves on variable names, outcome categories, outcome codes and labels, and question wording. Any deviations from consistency are noted in the “readme” file associated with each week’s data collection and in the digest at the end of this document.

---

<sup>2</sup> There is a rotating set of questions that vary over weeks of fielding, resulting in different numbers of variables in different weeks.



## Terms of Use

Researchers using these data agree to:

1. Use them only for statistical purposes and not for investigation of specific survey respondents.
2. Make no use of the identity of any survey respondents discovered intentionally or inadvertently, and to advise Nationscape personnel of any such discovery immediately by emailing Chris Tausanovitch at [ctausanovitch@ucla.edu](mailto:ctausanovitch@ucla.edu).

Cite Nationscape data and documentation in work that uses Nationscape data and documentation. Citations to the data sets should be as follows:

Tausanovitch, Chris and Lynn Vavreck. 2020. *Democracy Fund + UCLA Nationscape, October 10-17, 2019* (version 20200131). Retrieved from [URL].

3. Acknowledge that the original collectors of the data — UCLA, LUCID, and Democracy Fund Voter Study Group — and all funding agencies, bear no responsibility for the use of the data or for interpretations or inferences based upon such issues.

## Appendix 1: Weekly Change Log

The following changes were made to the survey at the points noted below. All changes will persist through subsequent waves unless otherwise noted.

**Wave 2:** Disabled RealAnswer and ReCAPTCHA for last 700 completes in this week's data collection and in future waves.

**Wave 3:** Launched one day late.

**Wave 5:** Began asking gubernatorial and Senate general election ballot items only to respondents who have gubernatorial or senate elections in 2019 or 2020 based on their state, as provided by Lucid.

**Wave 10:** New display logic was incorporated to show Democratic primary ranking questions to just those who indicated that they were voting in the Democratic primary or were not sure. These variables include rank\_dems\_1, rank\_dems\_2, and rank\_dems\_3.

**Wave 12:** We made a small edit to the gun ownership question. Categories are now: (1) Yes, I personally own a gun; (2) I don't, but a member of my household owns a gun; (3) No one in my household owns a gun; and (4) Not sure. The change is to the underlined text above.

**Wave 16:** Beto O'Rourke removed from democratic vote intent and ranking questions at 3:50 p.m. PST on November 1 when he dropped out of the race.

**Wave 17:** Senate primary question wording changed from, "Someone more critical/supportive of the President" to "Someone more critical/supportive of President Trump."

## Appendix 2: Weighting Procedure

All targets are derived from the [2018 ACS](#) population over the age of 18, except for 2016 vote.

Weights for gender, census region, Hispanic ethnicity, race, education, age, household language, and country of birth are derived by directly comparing responses in the Nationscape survey with responses to the ACS. Targets appear in table below.

Nationscape respondents may choose not to provide household income. The target for non-response is constructed as the proportion of Nationscape respondents who did not provide an answer. Targets for response categories are based on ACS responses multiplied by the proportion of respondents who chose to answer the Nationscape income question. In effect, this means non-respondents are not weighted for income.

Metropolitan status is derived by crosswalking Nationscape respondents' ZIP codes to Census Bureau Public Use Microdata Areas (PUMAs) using the [2014 MABLE Geocorr](#) based on population percentage. PUMAs are assigned a metropolitan status based on the average status of ACS respondents within them. Groupings are below. Nationscape respondents' PUMA-level metropolitan statuses are weighted to ACS respondents' PUMA-level metropolitan statuses.

Interactions are calculated between several demographic variables. Because some interactions produce very small target populations, some variables are coarsened in the interactions. Interactions and coarsened categories are listed below.

The 2016 vote is derived from the [United States Elections Project](#) and [MIT Election Lab](#). We start with the final vote distribution and estimated turnout. Then, we adjust turnout downwards to account for the 2.7 percent of voters who were not eligible to vote in 2016 but were 18 at the time that our survey was fielded. We then adjust turnout upwards to account for the fact that in the target population we expect 6 percent of voters will be nonvoters in 2016 who lie about voting on our survey. This number comes from the best available estimate.<sup>3</sup>

### Weighting Targets

#### Age

18–23	24–29	30–39	40–49	50–59	60–69	70+
9.4%	10.9%	17.4%	16.3%	17.0%	15.2%	13.8%

#### Birthplace

The United States	Another country
82.1%	17.9%

---

<sup>3</sup> Jackman, Simon, and Bradley Spahn. “Why does the American national election study overestimate voter turnout?.” *Political Analysis* 27.2 (2019): 193–207.

## Education

No high school diploma	High school diploma	Some college	Associate's degree	Bachelor's degree	Graduate degree
11.6%	27.3%	21.9%	8.4%	19.3%	11.5%

## Gender

Female	Male
51.7%	48.3%

## Hispanic

Mexican	Other Hispanic	Not Hispanic
9.8%	6.5%	83.7%

## Household Income

\$19,999 or less	\$20k-\$34,999	\$35k-\$49,999	\$50k-\$64,999	\$65k-\$79,999	\$80k-\$99,999	\$100k-\$124,999	\$125k-\$199,999	\$200k or more	Unanswered
9.7%	10.6%	10.9%	10.5%	9.2%	10.4%	10.1%	14.6%	9.1%	4.9%

## Language

English only	Spanish	Other language
78.1%	13.0%	8.8%

## Race

AAPI	Black	White	Other race
6.9%	12.1%	74.1%	7.0%

## Region

Midwest	Northeast	South	West
20.8%	17.4%	37.9%	23.8%

## Urban

Rural	Suburban	Urban-suburban	Urban
24.1%	24.3%	42.2%	9.4%

## Vote 2016

Clinton	Trump	Other vote	No vote
27.7%	26.5%	3.3%	42.5%

## Metropolitan status grouping

	Rural	Suburban	Urban-suburban	Pure urban
Individual	0-1	2	3	4
PUMA average	0.0-1.6	1.6-2.4	2.4-3.2	3.2-4.0

## Interactions

Interaction	Coarsening
Age × Hispanic	Combine 60–69 and 70+
Age × Race	Combine 60–69 and 70+
Education × Gender	None
Education × Hispanic	Combine Bachelor's degree and Graduate degree Combine Some college and Associate's degree
Education × Race	Combine Bachelor's degree and Graduate degree Combine Some college and Associate's degree
Gender × Race	None
Hispanic × Language	Combine English only and Other language
Hispanic × Race	Combine AAPI, Black, and Other race