2024 Collaborative Multiracial Post-Election Survey (CMPS)

Principal Investigators: Lorrie Frasure, Natalie Masuoka, Angela Ocampo

Data Extract II (November 24 Release)

Methodology Statement II (replaces September 10th version):

Combined 2024 CMPS dataset for adult Primary Samples (PS) and Oversamples (OS) (November 24, 2025 Release)

Note: Scholars should use this November 24, 2025 version of the CMPS dataset, which includes fresh adult Primary Samples data for their analyses.

On September 10, 2025 the CMPS released a total of 15,055 completed interviews in the adult Primary Sample (PS). During the data review process, an additional 155 cases that are part of the adult Primary Sample (PS) were found to be valid and available for analysis, and those have been added to further increase the sample size of the adult Primary Sample (PS). The Data Extract II release (November 24, 2025) includes these additional 155 completed interviews, for a total of 15,210 adult Primary Sample (PS) respondents.

This new release also includes 4,424 completed interviews with adult Oversample (OS) groups including hard-to-reach populations.

Both the adult Primary Sample (PS) and Oversamples (OS) are now combined in a single dataset for a total of 19,634 interviews.

The combined adult dataset was collected online in a respondent self-administered format from May 15, 2025 to October 6, 2025.

The survey (and invitation) was available to respondents in 11 languages: English, Spanish, Chinese (simplified), Chinese (traditional), Korean, Vietnamese, Tagalog, Arabic, Urdu, Farsi, and Haitian Creole. Due to the scholarly interest in the 2024 election, the project started with a large sample of registered voters from online sources that verified voter status. In addition, the data include a sample of non-registered adults, including non-citizens.

For purposes of sample management, we have assigned respondents in the adult Primary Sample to a single group in the table below. However, the instrument allowed respondents to self-identify with multiple categories as we explain in detail below.

2024 CMPS Adult Primary Samples (November 24, 2025 Release)

	PS Total	Latino	Black	Asian	White
Total	15,210	4,094	4,989	3,697*	2,117

*Note regarding adult Primary Sample release: AAPI combined sample of 4,010 includes 313 Native Hawaiian and Pacific Islander (NHPI)

Designing Question Content. In spring 2024, scholars from academic institutions were invited to collaborate on the 2024 Collaborative Multiracial Post-Election Survey (CMPS). The 2024 CMPS research design builds upon the 2016 and 2020 CMPS, which were also cooperative, user content driven, multi-racial/ethnic/lingual, post-election online surveys in the United States. This 2024 CMPS includes nearly 200 scholars from colleges/universities, including "Contributors" who purchased survey content and "Collaborators" who purchased data access, prior to the release of the individual level data extract. Survey questions were user generated. Users who contributed survey content could submit questions for just one single racial group, or common questions across multiple racial/ethnic, or oversample groups, depending on their interest. In cases where two different users submitted very similar questions, the PIs worked to create a single common question.

Respondent Sample Source. Data for this project came from two primary sources: first, large national online sample vendors were utilized due to their geographic coverage across all 50 states and capacity to deliver samples across race and ethnicity. Second, data came from experienced specialized sample vendors who have recruited targeted samples including: Latino, African American, Asian American, Native American, MENA, Muslim, Jewish, as well as immigrant communities and other hard-to-reach populations. Survey samples were collected from over 15 online sample vendors that have reliable and demographically accurate samples of the target populations. This allowed for broader and more diverse panels of target populations, including non-English speakers and other hard-to-reach populations.

All sets of sample sources were cross-checked and deduplicated using unique online respondent ID numbers to ensure any invited participant was only included once. Regardless of the sample vendor source, participants were randomly selected to participate in the study, receiving up to five email, text message, or app notification requests to participate. Invitations were bilingual for selected samples. For all participants, voter registration status was requested and confirmed in the screening process prior to their participation.

Respondent Recruitment: Respondents were offered a gift card or reward points of up to $100 as compensation for their participation. Respondents were able to pause the survey at any point and return for completion. Respondents who paused the survey were sent reminders to complete the survey. In

addition, respondents who suspended the survey for more than four days were sent new invitations to continue where they left off and complete the survey to encourage higher response rate of initially selected sample records.

Adult Sample response rate: The adult Primary Sample completed data collection on September 4, 2025 and the adult Oversample groups completed data collection on October 6, 2025. In total, 116,944 unique participants were sent invitations to participate in the survey and 106,239 clicked the initial invitation (90.8%). Among this group, 8,180 (7.7%) people dropped out immediately after the first item asking language preference, and 7,553 did not pass the CAPTCHA/bot detection check box, resulting in 90,506 who saw the first question asking year of birth. From here, 59,266 participants did not proceed to the first substantive question, meaning they either dropped out, they were screened out and disqualified, they were terminated because quotas were already full, or they failed attention checks during the screening process. This left 31,240 live participants who qualified for the survey and started Question 1, and 19,634 who completed the full survey as of October 6, 2025 (62.8% cooperation rate), which yields an overall online-survey participation rate of 62.8% and a response rate[1] of 19.9%. Out of the 19,634 completed interviews, 9,398 can be classified as part of Oversample Groups and hard-to-reach populations, as well as 15,210 are included in the Primary Sample (PS -Black, White, Latino, and Asian American). Please see section below **How respondents self-identify on the 2024 CMPS** on page 4. Among the full adult sample, a total of 1,471 surveys were completed in Spanish, and 431 were completed in an Asian or other languages.

Respondent Quality Control: Quality control efforts included the elimination of respondents who demonstrated clear non-serious responses (i.e., speeding or straight-line responses). Respondents were required to complete CAPTCHA checks to eliminate bots, and multiple attention checks and red-herring questions were included at the front end to eliminate bot, AI-driven, or extremely low-attention respondents. Time checks were also instituted to prevent respondents from answering the entire 800 item survey in less than 20 minutes. Finally, an open-ended response option was included and monitored daily to eliminate bots, AI-driven responses, and non-attention respondents.

Preliminary analysis related to inconsistent/illogical responses (i.e., marriage year less recent than birth year) suggests that the vast majority of the sample is of high quality. Of course, no research study can force a participant to pay close attention and be consistent. Indeed, the real world of opinion formation is messy, busy, noisy, and filled with distractions. However, our participants were given multiple reminders to proceed slowly, read questions carefully, and give their best honest answers. Still, each researcher needs to make their own judgement and decisions about whether or not they want to flag or isolate particular respondents from the sample due to inconsistent response patterns, and the threshold will vary between researchers. This is the case in all surveys.

[1] 106,239 clicked the invite link and we removed 7,553 bots, leaving 98,686 potential respondents and 19,634 completed the survey for an online-survey response rate of 19.9% (as of October 6, 2025).

Data Collection: Overall, the survey contained more than 800 unique questions including split samples, branch-items, and group-specific questions, and the average respondent completed over 400 items. The overall median completion time was 55.6 minutes and three-quarters of respondents completed in 81.3 minutes or less (see variable 'qtime'). Please note that completion times varied by sample group because some respondents had different numbers of questions to answer.

In order to balance any order effects or length effects the survey was divided into three segments:, questions 1-127, 128-260, and 261-404 (note there are 404 questions, but over 800 items, because a single question could be a grid that asks views on five separate items). These segments were then rotated as to which were presented first, second, or third, so that items that appear on the final third of the questionnaire are not always presented last to respondents and items that appear first on the questionnaire were not always presented first to respondents, and so on. This eliminates any concerns over respondents being fatigued by a longer survey and putting less effort into questions in the final third compared to the first third. These three segments were randomized, and a variable named "blockassignment" is available in the survey for users to identify the order of questions for each respondent.

2024 CMPS Adult Oversamples (OS) and Hard-to-Reach Populations (November 21, 2025 Release)

	Sample Size
Afro-Latino	1,991
Black Immigrant	1,304
Jewish	772
LGBTQ	2,258
Native American	1,293
NHPI	430
MENA	432
Muslim	918

2024 CMPS Oversample Categories. The CMPS allows respondents to self-identify and mark "all that apply." This is because adult residents in the United States may have multiple identities, and the Census has recognized multiple racial/ethnic identities since 2000. As such, oversample categories are NOT mutually exclusive. When considering factors such as religion, sexual orientation, immigrant status and more, we should expect to see diversity in overlapping identities. The oversample categories and final sample sizes are based on respondent self-categorization and self-reporting.

For example, a respondent can be in the LGBTQ oversample as well as in the Afro-Latino oversample. A respondent can be in the Black immigrant oversample and also in the Muslim oversample. A respondent can be in the Native American oversample and also identify as Hispanic-Latino in the primary adult sample, and so on.

How respondents self-identify on the 2024 CMPS. Throughout the history of this project, it has been the intention of the CMPS team to collect the most diverse sample of respondents in the United States as possible, and to allow respondents to self-identify their race, ethnicity, religion and immigration status.

The 2024 Census ACS revealed that the multi-racial population (respondents who marked two or more race categories) increased to a record high of 13.2%, or nearly 45 million Americans, up from just 2.4% or 6.8 million in 2000. From 2010 to 2020 alone, the multi-racial category increased by 127%. For 50 years of the Census, Hispanic or Latino respondents are almost by definition "multi-ethnic" as they answer both race and ethnicity questions. The outcome is that in both the real world and in our data, respondents do not necessarily fit easily into one box. As a general trend, most respondents do identify as single-race. In the 2024 CMPS, 17,414 (88.7%) selected only one race option, while 2,220 (11.3%) of respondents chose two or race/ethnic categories.

Moving from the adult Primary Sample to the adult Oversample and hard-to-reach communities, there is a possibility of a greater amount of overlapping identities among respondents. For example, in the Black Immigrant oversample of 1,304, there are 105 respondents who identify their religion as Muslim and as such, they are also part of the Muslim oversample. In the adult Black Primary Sample, end-users will find that there are 546 immigrants, which aligns with their proportion in the national population, given that about 10% of the adult Black population in the U.S. is foreign born today. Within the adult Latino Primary Sample, there are 690 people who identify as Afro-Latino, which one would expect in a representative sample of all Latino adults. These overlapping identities are what make each independent adult Primary Sample or Oversample accurate and representative.

However, scholars who decide to *only* examine Black immigrants, or only Afro-Latinos, will find that the additional oversamples being released as part of this data extract also increase their sample sizes, which allows for a much larger and richer set of observations to work with. For instance, while there are 690 Afro-Latinos in the adult Latino Primary Sample, in full, the OS includes a total of 1,991 Afro-Latino respondents.

There is a reasonable and expected amount of overlap between the adult Primary Sample and Oversamples. Yet, there is considerably less overlap across the Oversamples, as they are more unique. For instance, there are only 15 respondents who are *both* Native American and Jewish. There are only 10 respondents who identify as *both* MENA and Afro-Latino.

Finally, with respect to sexual orientation, the LGBTQ oversample stratifies all race, ethnicity and religious categories, as we would expect. For this reason, we have created a unique and independent weight **(wt_adult_lgbt)** to use for analysis of this population.

The CMPS offers variables "qraceid" and individual dummy variables for race/ethnicity and religion, which all may be relevant to understanding a respondent's race or ethnicity. An administrative variable to track quota completes is called "main_eth" but is not necessary for analysis. The CMPS data also offers other variables that identify respondents in the adult Oversamples and hard-to-reach populations. End-users can choose how to subset the data for analysis of different racial or ethnic groups, as well as of any other oversample or hard-to-reach populations. We urge end-users to use their own judgement and theoretical priors to determine how they will classify respondents, and to disclose these decisions within their analysis.

California sample with oversample of Black respondents

As the largest state, we expected to complete a large number of interviews in California, especially among Asian Americans and Latinos, where California is their largest population state. While there are more than 2 million Black Californians, their population is comparatively smaller than Asian Americans (6M) and Latinos (15M) in California, and in most surveys in California, the Black sample is too small to analyze. This year the CMPS purposefully included an oversample of Black Californians to bring their sample up to 850, which marks the largest ever sample of Black Californians in a comparative multiracial academic study. Because of this oversample, in the adult Primary Samples, hard-to-reach samples, and the California sample, weights are used to adjust the Black California sample to match the Census estimates and bring all samples into balance. Yet the adult oversample of Black Californians at 850 provides end users with an opportunity to do both comparative and standalone analysis of this population.

2024 CMPS California Adult Sample Size:

	CA Total	Latino	Black	Asian	White	Other
Total	3,545	1,000	850	1,121	324	250

Summary of Primary Sample and Oversample Variables

Variable name	Description
qraceid	adult PS or OS variable to identify respondents “primary race” using s3 and s4
asian	binary 0/1 variable for any respondent who selected Asian at s3r4
black	binary 0/1 variable for any respondent who selected Black at s3r3
latino	binary 0/1 variable for any respondent who selected Latino at s3r2
mena	binary 0/1 variable for any respondent who selected MENA at s3r6
nhpi	binary 0/1 variable for any respondent who selected NHPI at s3r7
white	binary 0/1 variable for any respondent who selected White at s3r1
os_native	binary 0/1 variable for respondent who selected Native American and said their ancestry is themself, their parents, or grandparents - excludes distant relatives
os_afrolat	binary 0/1 variable for respondent who selected Afro-Latino and said their ancestry is themself, their parents, or grandparents - excludes distant relatives
os_muslim	binary 0/1 variable for respondent who selected Muslim at s13r5
os_jewish	binary 0/1 variable for respondent who selected Jewish at s13r4, created proportionate to racial/ethnic estimate of Jewish American population
os_mena	binary 0/1 variable for any respondent who selected MENA at s3r6
os_lgbtq	binary 0/1 variable for any respondent who selected LGBTQ at s15 or s16
os_blkimm	binary 0/1 variable for respondent who selected Black at s3r3 AND was born outside the mainland U.S. at s25 OR foreign born parents at s25b and has ancestry to a country outside of the U.S.
california	binary 0/1 variable for any respondent from california at s21
main_eth	Administrative variable to track quotas across main racial groups

Methodology Used to Create Weights

We generated a series of weights based on the different samples within the CMPS 2024 project.

Primary Sample Weight

In the initial release of the adult Primary Sample, a weight variable named **"wt_adult_ps"** is based on the variable **"qraceid"** and attempts to weight each of the four main racial groups in the adult Primary Sample (Asian, Black, Latino, White) to their known Census demographic characteristics nationwide. Using the **"wt_adult_ps**" will only use the 15,210 completes in the adult Primary Sample, and will not utilize any of the Oversample or hard-to-reach sample completes.

For the variable **"wt_adult_ps:" We strongly recommend that this weight is *always* used in any analysis,** in particular, for scholars interested in running a "national" analysis when analyzing variation within a racial group, or for a single racial group using the Primary Sample. It employs a post-stratification raking algorithm that balances each demographic category within +/- 2 percent of its known demographics on the 2023 U.S Census American Community Survey (ACS). estimates which is estimated to be 60.1% White, not Hispanic, 18.2% Latino, 13.8% Black, 8.0% Asian American. The Primary Sample data has also been weighted to the 4-racial-group national combined average, which is useful for scholars interested in running a "nationally representative" analysis. The full data are weighted within each racial group to fall within the margin of error of the adult population in the 2023 ACS U.S. Census American Community Survey (ACS) 1-year data file for age, gender, education, income, state, nativity, and ancestry. This weight variable only applies to the Primary Samples' 15,210 cases.

Please note that the **"wt_adult_ps"** weight variable does not include data or values for any of the oversample groups.

Oversample Weight

For the newly released Oversample categories, we have also included a weight variable, here named **"wt_adult_os"** that we recommend all scholars employ in both simple frequency analysis or multivariate regression. This weight brings each of the eight ethnic, immigrant, or religious Oversample categories into balance with their best-known estimates of their national population.

For some groups such as Native American and MENA, which now have Census population estimates, we rely on Census data. For Muslims, we rely on the latest Pew[2] research. For other groups we rely on peer-reviewed scholarly estimates, combined with the best available Census data to estimate parameters of the population.

[2] https://www.pewresearch.org/short-reads/2025/06/18/how-us-muslims-compare-with-other-americans-religiously-and-demographically/

Because LGBTQ respondents cross all other racial, ethnic, immigrant and religious categories, we have defined a unique weight to balance this population called **"wt_adult_lgbt"** based on the latest data from the Williams Institute[3] and other scholarly research[4].

Weights are applied to all demographic groups in the 2024 CMPS to attempt to balance the population characteristics in line with our best-known estimates. For many groups, this information is derived directly from the U.S. Census American Community Survey (ACS) annual estimates. Within hard-to-reach adult oversample populations, the ACS does not always have a defined estimate. Using ACS data from IPUMS[5] for Native Americans, Black immigrants, MENA[6] and NHPI researchers can access more information with fine-grained Census data to incorporate into weighting algorithms. For other groups such as Afro-Latinos, Muslim, Jewish individuals, the U.S. Census does not publish demographic statistics. Therefore, we consulted the extant literature and specialty research centers, such as Pew, Williams, Hutchins, Center for Migration Studies and others, for this information. Each group within both the adult Primary Sample and Oversample are balanced using weights to the best estimate of their population for gender, age, education, nativity, state, country of ancestry (where applicable) and partisanship for registered voters.

Weights are derived for each individual group, through a raking ratio algorithm[7], until convergence is achieved within the margin of error of each estimate. For example, first the Native American sample is balanced for gender, next for age, next for education, next for state, and so on. After each category has been balanced in comparison to the Census, a second round (rake) of weights are applied to each variable to further fine-tune the balance of each variable, where a variable is out of balance with the Census estimate.

As this is an academic research cooperative, if individual researchers have reason to believe the Oversample weights should be adjusted, we recommend you share your recommendations with the CMPS PI team. We also encourage end-users to help us explore additional weighting algorithms to improve our estimates.

California Sample Weight

For the California sample, a unique weight has been created called **"wt_adult_ca"** to balance each racial and ethnic group to match the latest Census ACS estimates for California adults.

---

[3] https://williamsinstitute.law.ucla.edu/visualization/lgbt-stats/?topic=LGBT#demographic
[4] https://www.pewresearch.org/short-reads/2023/06/23/5-key-findings-about-lgbtq-americans/
[5] https://usa.ipums.org/usa/index.shtml
[6] For more on deriving MENA estimates from ACS: https://www.cato.org/briefing-paper/middle-eastern-or-north-african-us-government-surveys-preview-mena-demographics
[7] Battaglia, Michael P, David C Hoaglin, and Martin R Frankel. 2009. "Practical Considerations in Raking Survey Data." *Survey Practice* 2 (5). https://doi.org/10.29115/SP-2009-0019.

Summary of CMPS Weight Variables

Group of Interest	Variable	Weight to use	Sample size
Adult PS: Asian	qraceid=4	wt_adult_ps	3,697
Adult PS: Black	qraceid=3	wt_adult_ps	4,989
Adult PS: Latino	qraceid=2	wt_adult_ps	4,094
Adult PS: White	qraceid=1	wt_adult_ps	2,117
OS: Afro Latino	os_afrolat	wt_adult_os	1,991
OS: Black Immigrant	os_blkimm	wt_adult_os	1,304
OS: Jewish	os_jewish	wt_adult_os	772
OS: LGBTQ	os_lgbtq	wt_adult_lgbt	2,258
OS: MENA	os_mena	wt_adult_os	423
OS: Muslim	os_muslim	wt_adult_os	918
OS: Native American	os_native	wt_adult_os	1,293
Adult NHPI	nhpi	wt_adult_os	430
California Sample	california	wt_adult_ca	3,545

*note: use qraceid for racial groups within California

****

Additional Data Support

Importing the Data

In this release, we provide three different data import options: SPSS (.sav), STATA (.dta) and CSV (.csv) format. Variable labels are available in both SPSS (.sav) and STATA (.dta) file formats. The CSV format includes only the data and does not include variable labels.

If you are using R and would like to import the data that includes variable labels, you will need to use the .dta format. To do this, please install and employ the `haven` package to import the data with variable labels in .dta format.

```
install.packages("haven")

library(haven)

cmps_2024 <- read_dta("[FILEPATH]/cmps2024_fulladult_112425.dta")
```

Alternatively, if you are using R and prefer to work with the data without variable labels, please employ the `read.csv` function to import the data:

```
cmps_2024 <- read.csv("[FILEPATH]/cmps2024_fulladult_112425.csv")
```

Suggested Stata and R Code to Employ Weights

Weights are recommended to be used with all analysis to balance the population. Depending on which group you are interested in analyzing, you should use different weights. As described above, the adult Primary Sample contains 15,210 completed interviews with the four main racial groups: Asian, Black, Latino, White and if you analyze this sample, you should focus your analysis only on these 15,210 cases.

If you are interested in a different population from the Oversample, you will use a different weight and you should focus your analysis to only that portion of the larger sample. The weight variables are set to "missing" for respondents that do not fit that category. This is to ensure that the wrong weight is not accidentally applied to the wrong group. For example, the **wt_adult_ps** only has valid values for the 15,210 respondents in the Primary Sample and is set to missing for all other respondents which are NOT part of the Primary Sample. The same logic applies to all groups. The **wt_adult_ca** for the California oversample only has valid values for 3,545 cases of people who indicated their state was California and is set to missing for all other states.

Example STATA Code

If you are running analysis in Stata you might need to specify which sample you are interested in working with while applying the weight variables. For example if you want to assess presidential vote choice by gender, by race you would use the following, depending on the group of interest:

Latino Adult Primary Sample results by Gender:
tab q17 s14 if qraceid==2 [aw=wt_adult_ps], col

Black Adult Primary Sample results by Gender:
tab q17 s14 if qraceid==3 [aw=wt_adult_ps], col

Black Immigrant Oversample results by Gender:
tab q17 s14 if os_blkimm==1 [aw=wt_adult_os], col

Native American Oversample results by Gender:
tab q17 s14 if os_native==1 [aw=wt_adult_os], col

LGBTQ Oversample results by Gender:
tab q17 s14 if os_lgbtq==1 [aw=wt_adult_lgbt], col

Example R Code

If you are running analysis in R you will need to filter your data to the population of interest before applying the appropriate survey weight. We provide example code for the following scenarios:

Example 1: Analyzing the LGBTQ Adult Sample

```
library(survey)
library(tidyverse)

#filter dataset to only LGBTQ respondents
cmps_lgbtq <- cmps_2024 %>%
  filter(!is.na(wt_adult_lgbt))

#define survey design with weights
cmps_lgbtq_design <- svydesign(ids = ~1, data = cmps_lgbtq, weights = ~wt_adult_lgbt)

#create a weighted crosstab: LGBTQ by Gender
prop.table(svytable(~os_lgbtq + s14, cmps_lgbtq_design), margin = 1)
```

Example 2: Analyzing Oversample/Hard-to-Reach Adult Samples

```
library(survey)
library(tidyverse)

#filter dataset to oversample/hard-to-reach adult sample only
cmps_os  <- cmps_2024 %>%
  filter(!is.na(wt_adult_os))

#define survey design with weights
cmps_os_design <- svydesign(ids = ~1, data = cmps_os, weights = ~wt_adult_os)

#create a weighted crosstab: Black Immigrant by Gender
prop.table(svytable(~os_blkimm + s14, cmps_os_design), margin = 1)

#create a weighted crosstab: NHPI by Gender
prop.table(svytable(~nhpi + s14, cmps_os_design), margin = 1)
```

Example 3: Analyzing the Primary Adult Sample ONLY

```
library(survey)
library(tidyverse)

#filter dataset to primary adult sample only
cmps_ps <- cmps_2024 %>%
  filter(!is.na(wt_adult_ps))

#define survey design with weights
cmps_ps_design <- svydesign(ids = ~1, data = cmps_ps, weights = ~wt_adult_ps)

#create a weighted crosstab: Primary Race by Gender
prop.table(svytable(~qraceid + s14, cmps_ps_design), margin = 1)
```

Example 4: Analyzing the California Adult Sample ONLY

```
library(survey)
library(tidyverse)

#filter dataset to adult sample in California
cmps_ca <- cmps_2024 %>%
  filter(!is.na(wt_adult_ca))

#define survey design with weights
cmps_ca_design <- svydesign(ids = ~1, data = cmps_ca, weights = ~wt_adult_ca)

#create a weighted crosstab: Primary Race by Gender (among Californians)
prop.table(svytable(~qraceid + s14, cmps_ca_design), margin = 1)
```

Data Availability: In keeping with best practices and data transparency ethics in the social sciences, the data will be embargoed to contributors/collaborators and their co-authors for 1-year following the release of the combined adult individual level dataset. Following the data embargo, the adult individual level data will be deposited at Inter-University Consortium for Political and Social Research (ICPSR) and the Roper Center for Public Opinion Research for public use data deposit and archiving.

For publications, please cite:

Frasure, Lorrie, Natalie Masuoka, Angela Ocampo. 2025. The 2024 Collaborative Multiracial Post-election Survey (CMPS). Los Angeles, CA. November 2025 Release.