2025 Philadelphia Residents Survey

Methodology Report

Prepared for The Pew Charitable Trusts

Pew

June 2025

Prepared by:

Robyn Rapoport, Rob Manley, and Paula Armendariz
SSRS

Project Number: X1069

ssrs
redefining research

ssrs redefining research

Table of Contents

Project Overview

SSRS conducted The Pew Charitable Trusts' 2025 Philadelphia Residents Survey. The goal of this survey was to better understand what residents think about Philadelphia, their perceptions of the job the current mayor is doing, and how they feel about important public policy issues and current events facing the city.

The 2025 Philadelphia Residents Survey obtained surveys via a mixed-mode online, phone (via call-ins), and mail survey design. N=2,289 Philadelphia County residents completed the survey, with n=1,776 completing via web, n=111 calling in and completing via phone, and n=402 completing via mail. Data collection was conducted in English (n=2,207) and Spanish (n=82) from Jan. 2, 2025, to March 7, 2025.

The data for this survey was weighted to represent the adult population in Philadelphia. The total sample design effect for this survey is 1.79 and the margin of sampling error for the complete set of weighted data is plus or minus 2.7 percentage points.

This report provides additional information about the methods used to collect the data and report the survey results.

Sampling Methods

Overview

The target population for this survey was adults age 18 or older living in Philadelphia County, Pennsylvania. Additionally, Pew was interested in three key subgroups: African American/Black, Hispanic, and Asian Philadelphians. Given the population distributions of these subgroups in Philadelphia County, SSRS oversampled Asian and Hispanic Philadelphians and closely monitored the number of completed interviews with Black Philadelphians to ensure that Pew would be able to conduct analyses of these subgroups.

Sample Frame

The sample was drawn from the Delivery Sequence File (DSF), a database maintained by the U.S. Postal Service (USPS). The DSF is an authoritative repository containing a detailed listing of all known delivery points in the United States. This includes residential and business addresses, post office boxes, and drop points, as well as information on seasonal and vacant properties. The DSF is updated regularly to maintain accuracy and completeness, and reflect changes such as new construction, property demolitions, or address reassignments.

Each delivery point in the DSF is accompanied by essential metadata, including status codes that indicate whether an address is active, vacant, or seasonal. The database also provides details about delivery characteristics, such as the type of address (e.g., single-family residence, apartment, or business) and the preferred delivery sequence for postal carriers.

Our sample provider, Marketing Systems Group (MSG), is licensed to draw samples from the DSF. Moreover, MSG has geo-coded each address to its unique Census Block Group, enabling precise geographic targeting for use in sample designs. Additionally, MSG can append a wide range of consumer and surname flags to samples pulled from the DSF.

Sample Design

The sample was drawn from all residential addresses in Philadelphia County. Records marked as seasonal, educational, vacant or "no stat" were excluded from selection.[1] The sample design used two sets of geographic strata.

Geographic Stratification

The first set of geographic strata divided Philadelphia County into seven neighborhoods defined by ZIP code. Table 1 lists the neighborhoods and their corresponding ZIP codes.

Table 1: Neighborhood Strata Definitions

NEIGHBORHOOD	ZIP CODES
South/Center	19102, 19103, 19106, 19107, 19145, 19146, 19147, 19148
Southwest	19142, 19143, 19153
West	19104, 19131, 19139, 19151
Lower Northeast	19122, 19124, 19125, 19133, 19134, 19135, 19137, 19149
Upper Northeast	19111, 19114, 19115, 19116, 19136, 19152, 19154
North	19120, 19121, 19123, 19126, 19130, 19132, 19138, 19140, 19141, 19150
Northwest	19118, 19119, 19127, 19128, 19129, 19144

The second set of geographic strata divided each neighborhood into the five race density strata, based on the incidence of Asian and Hispanic residents at the block group level. The data for this stratification was sourced from the 2023 block group level Census Bureau's Planning Database (PDB), which includes selected operational, housing, demographic, and socioeconomic statistics derived from the Decennial Census and the 2017-21 American Community Survey (ACS) five-year files.[2]

[1] "No stat" records in the DSF refer to addresses that the USPS does not actively deliver to. These records typically include buildings under construction, vacant properties, and special delivery situations where logistics make delivery impractical (e.g., extremely rural areas).

[2] U.S. Census Bureau. (2023). *2023 Planning Database*. Retrieved from https://www.census.gov.

Since neighborhoods are defined by ZIP code and race density strata are determined at the block group level, we created a crosswalk linking block groups to neighborhoods. Each block group was assigned to a single neighborhood based on the plurality of its population. For example, if 60% of a block group's population resided in Neighborhood 1 and 40% in Neighborhood 2, the block group was assigned to Neighborhood 1 for sampling purposes.

The five race density strata listed in Table 2 were defined within each neighborhood using the following process. Block groups were divided into two sets of deciles: one based on incidence of non-Hispanic Asian residents and the other based on incidence of Hispanic residents. These two sets of deciles were combined, resulting in up to 100 possible race density strata. A target incidence was created for each race density stratum that combined the Hispanic and Asian incidence figures.

Table 2: Race Density Stratification (Within Neighborhood)

STRATA	RACE DENSITY
One	Highest density Asian/Hispanic block groups
Two	High density Asian/Hispanic block groups
Three	Middle density Asian/Hispanic block groups
Four	Low density Asian/Hispanic block groups
Five	Lowest density Asian/Hispanic block groups

The target incidence, $I(T)_i$ in block group i was defined as the sum of the block group's Hispanic incidence, $I(H)_i$, and an adjusted Asian incidence. The Asian incidence $I(A)_i$ was adjusted by multiplying it by the ratio of the overall Hispanic incidence to the overall Asian incidence. $I(T)_i = I(H)_i + \left((I(H)/I(A)) \times I(A)_i\right)$. This target incidence was created to enable simultaneous and equal targeting of both groups.

The final step was to sort each neighborhood's race density strata by the target incidence and divide into quintiles, creating the five race density strata. Table 3 shows the address-based sample survey (ABS) frame distributions and Asian and Hispanic incidences from the PDB across all strata.

Table 3: Frame Counts and Estimated Incidences Across Geographic Strata

REGION	RACE DENSITY STRATA	FRAME	ASIAN INCIDENTS	HISPANIC INCIDENTS	ASIAN or HISPANIC INCIDENTS
1 South/Center	1 Highest density Asian/Hispanic	25,549	32.9%	17.7%	50.6%
1 South/Center	2 High density Asian/Hispanic	26,601	16.6%	9.0%	25.6%
1 South/Center	3 Middle density Asian/ Hispanic	27,798	10.7%	7.2%	17.8%
1 South/Center	4 Low density Asian/Hispanic	27,399	5.9%	5.2%	11.1%
1 South/Center	5 Lowest density Asian/Hispanic	27,976	1.3%	2.9%	4.2%
2 Southwest	1 Highest density Asian/Hispanic	7,958	13.6%	7.1%	20.7%
2 Southwest	2 High density Asian/Hispanic	8,671	2.4%	8.2%	10.6%
2 Southwest	3 Middle density Asian/Hispanic	9,292	1.4%	4.4%	5.9%
2 Southwest	4 Low density Asian/Hispanic	8,410	0.4%	1.4%	1.8%
2 Southwest	5 Lowest density Asian/Hispanic	8,963	0.0%	0.0%	0.0%
3 West	1 Highest density Asian/Hispanic	16,882	26.3%	7.3%	33.6%
3 West	2 High density Asian/Hispanic	16,500	8.3%	8.9%	17.2%
3 West	3 Middle density Asian/Hispanic	16,617	2.8%	3.2%	6.0%
3 West	4 Low density Asian/Hispanic	17,604	0.1%	1.7%	1.8%
3 West	5 Lowest density Asian/Hispanic	16,905	0.0%	0.0%	0.0%
4 Lower Northeast	1 Highest density Asian/Hispanic	21,093	12.2%	67.0%	79.2%
4 Lower Northeast	2 High density Asian/Hispanic	21,428	12.1%	48.4%	60.5%
4 Lower Northeast	3 Middle density Asian/Hispanic	27,406	4.7%	36.0%	40.7%
4 Lower Northeast	4 Low density Asian/Hispanic	22,959	3.5%	21.3%	24.8%
4 Lower Northeast	5 Lowest density Asian/Hispanic	26,047	1.8%	6.7%	8.6%
5 Upper Northeast	1 Highest density Asian/Hispanic	19,779	26.1%	16.8%	42.9%
5 Upper Northeast	2 High density Asian/Hispanic	22,615	14.9%	16.1%	31.0%
5 Upper Northeast	3 Middle density Asian/Hispanic	21,816	9.1%	13.5%	22.6%
5 Upper Northeast	4 Low density Asian/Hispanic	21,660	4.4%	9.5%	13.9%
5 Upper Northeast	5 Lowest density Asian/Hispanic	21,469	1.3%	5.2%	6.5%

6 North	1 Highest density Asian/Hispanic	31,560	12.8%	43.5%	56.3%
6 North	2 High density Asian/Hispanic	29,125	9.1%	13.1%	22.3%
6 North	3 Middle density Asian/Hispanic	35,994	3.5%	8.6%	12.1%
6 North	4 Low density Asian/Hispanic	32,348	0.5%	2.5%	3.0%
6 North	5 Lowest density Asian/Hispanic	32,336	0.0%	0.0%	0.0%
7 Northwest	1 Highest density Asian/Hispanic	13,375	6.0%	11.7%	17.7%
7 Northwest	2 High density Asian/Hispanic	12,792	4.6%	4.5%	9.2%
7 Northwest	3 Middle density Asian/Hispanic	14,565	1.4%	4.7%	6.1%
7 Northwest	4 Low density Asian/Hispanic	13,781	0.9%	2.2%	3.1%
7 Northwest	5 Lowest density Asian/Hispanic	14,240	0.1%	0.9%	1.1%
Total		**719,513**	**7.6%**	**13.3%**	**20.9%**

Within each neighborhood, race density strata were sampled disproportionately to increase the incidence of Hispanic and Asian respondents. Table 4 shows subgroup incidences from the PDB across the race density strata aggregated across all neighborhoods.

Table 4: Frame Counts and Estimated Incidences Across Race Density Strata

			PDB INCIDENCES		
RACE DENSITY STRATA	FRAME	DIST'N	ASIAN, NON-HISPANIC	HISPANIC	ASIAN OR HISPANIC
Highest density Asian/Hispanic	136,196	18.9%	19.5%	28.7%	48.2%
High density Asian/Hispanic	137,732	19.1%	11.1%	16.7%	27.7%
Middle density Asian/Hispanic	153,488	21.3%	5.4%	12.7%	18.1%
Low density Asian/Hispanic	144,161	20.0%	2.6%	6.9%	9.4%
Lowest density Asian/Hispanic	147,936	20.6%	0.8%	2.6%	3.3%
Total	**719,513**	**100.0%**	**7.6%**	**13.3%**	**20.9%**

The higher density race strata were oversampled relative to the lower density strata thus increasing the incidence of Asians and Hispanics. Table 5 shows the resulting sample distribution and the adjusted PDB incidences of key subgroups.

Table 5: Frame and Sample Distribution Across Race Density Strata

RACE DENSITY STRATA	FRAME DIST'N	SAMPLE DIST'N
Highest density Asian/Hispanic	18.9%	30.4%
High density Asian/Hispanic	19.1%	22.0%
Middle density Asian/Hispanic	21.3%	19.6%
Low density Asian/Hispanic	20.0%	13.8%
Lowest density Asian/Hispanic	20.6%	14.2%
Total	**100.0%**	**100.0%**

	INCIDENTS	INCIDENTS
Asian, non-Hispanic	7.6%	9.9%
Hispanic	13.3%	16.2%
Asian or Hispanic	20.9%	26.1%

Asian Surname Stratification

In addition to the geographic stratification, Asian surname flags were also used to target Asian respondents. To do this, the sample was released in two phases. Asian and Hispanic surname flags were appended to the first sample release though they were not used in the sampling process. The flags were used to determine what oversampling would be needed in the second sample release to get a sufficient number of Asian and Hispanic interviews.

After analyzing the performance of the first sample release, it was clear that we would not need to oversample any Hispanic surname records in the second sample release, but we would need to oversample Asian surname flagged records. Table 6 shows how much of the sample was released in each wave and how many of the records were flagged as Asian surname.

Table 6: Asian Surname Flagged Across Sample Waves

	WAVE 1	WAVE 2	TOTAL
Total	**11,700**	**7,801**	**19,501**
Asian surname flagged	1,030	5,241	6,271

The 35 geographic strata were combined with the two Asian surname strata, resulting in a total of 70 sample strata. The number of Asian surname flagged cases within each geographic stratum was estimated by applying the proportion of flagged cases from the initial sample order to the frame total. For example, the initial sample order included 668 records from the highest race

density stratum from the South/Center region, of which 87 were flagged as Asian surname. The total frame size of that stratum (25,549) was then multiplied by 87/668 to estimate the number of Asian flagged cases in the frame (3,327). Table 7 shows how the frame is distributed across all of the sample strata.

Table 7: Frame Counts Across Detailed Sample Strata

REGION	RACE DENSITY STRATA	FRAME	ASIAN FLAGGED	NOT ASIAN FLAGGED
1 South/Center	1 Highest density Asian/Hispanic	25,549	3,327	22,222
1 South/Center	2 High density Asian/Hispanic	26,601	1,980	24,621
1 South/Center	3 Middle density Asian/Hispanic	27,798	1,943	25,855
1 South/Center	4 Low density Asian/Hispanic	27,399	1,874	25,525
1 South/Center	5 Lowest density Asian/Hispanic	27,976	1,251	26,725
2 Southwest	1 Highest density Asian/Hispanic	7,958	918	7,040
2 Southwest	2 High density Asian/Hispanic	8,671	535	8,136
2 Southwest	3 Middle density Asian/Hispanic	9,292	535	8,757
2 Southwest	4 Low density Asian/Hispanic	8,410	177	8,233
2 Southwest	5 Lowest density Asian/Hispanic	8,963	179	8,784
3 West	1 Highest density Asian/Hispanic	16,882	2,067	14,815
3 West	2 High density Asian/Hispanic	16,500	911	15,589
3 West	3 Middle density Asian/Hispanic	16,617	134	16,483
3 West	4 Low density Asian/Hispanic	17,604	357	17,247
3 West	5 Lowest density Asian/Hispanic	16,905	89	16,816
4 Lower Northeast	1 Highest density Asian/Hispanic	21,093	5,541	15,552
4 Lower Northeast	2 High density Asian/Hispanic	21,428	4,125	17,303
4 Lower Northeast	3 Middle density Asian/Hispanic	27,406	3,551	23,855
4 Lower Northeast	4 Low density Asian/Hispanic	22,959	1,519	21,440
4 Lower Northeast	5 Lowest density Asian/Hispanic	26,047	1,873	24,174
5 Upper Northeast	1 Highest density Asian/Hispanic	19,779	2,755	17,024
5 Upper Northeast	2 High density Asian/Hispanic	22,615	2,144	20,471
5 Upper Northeast	3 Middle density Asian/Hispanic	21,816	1,673	20,143
5 Upper Northeast	4 Low density Asian/Hispanic	21,660	1,515	20,145
5 Upper Northeast	5 Lowest density Asian/Hispanic	21,469	1,069	20,400
6 North	1 Highest density Asian/Hispanic	31,560	4,897	26,663
6 North	2 High density Asian/Hispanic	29,125	1,767	27,358
6 North	3 Middle density Asian/Hispanic	35,994	1,873	34,121
6 North	4 Low density Asian/Hispanic	32,348	713	31,635
6 North	5 Lowest density Asian/Hispanic	32,336	713	31,623
7 Northwest	1 Highest density Asian/Hispanic	13,375	382	12,993
7 Northwest	2 High density Asian/Hispanic	12,792	375	12,417
7 Northwest	3 Middle density Asian/Hispanic	14,565	468	14,097
7 Northwest	4 Low density Asian/Hispanic	13,781	537	13,244
7 Northwest	5 Lowest density Asian/Hispanic	14,240	356	13,884

Fieldwork Procedures

Questionnaire Design

The questionnaire was initially developed by The Pew Charitable Trusts' staff members. SSRS provided feedback regarding new question wording, order, clarity, and other issues pertaining to questionnaire quality. Together, the SSRS and Pew teams worked to finalize the questionnaire.

Paper Survey Formatting

SSRS was responsible for formatting the questionnaire into a self-administered paper instrument. SSRS focused on clarity of format for any skip logic and for overall comprehension of the questionnaire. Efforts were made in the design to (1) encourage cooperation by offering easy-to-read, easy-to-maneuver hard copy; and (2) reduce the potential for confusion and thereby produce the most accurate data. SSRS formatted the survey in Microsoft Word and then worked with our professional printing service for execution and printing. Paper surveys were printed in both English and Spanish and were sent to Pew for approval before printing and mailing materials to contacts.

Letter and Postcard Design

The text for the survey invitation letter, reminder postcard, and follow-up reminder letter were developed by SSRS using the 2022 materials as a reference. SSRS translated them into Spanish and formatted the letters and postcards to prepare them for mailing. SSRS sent the Pew team the final postcard and letters for approval prior to printing and mailing the materials to contacts.

Programming

Once the questionnaire was finalized, the survey was programmed into SSRS's Confirmit platform for web administration in both English and Spanish. Extensive checking of the program was conducted to ensure that skip patterns followed the questionnaire's design. The program was also checked on multiple devices, including desktop computers, mobile devices (such as tablets and smartphones), and different web browsers to ensure consistent and optimized visualization across devices and web browsers.

SSRS generated unique survey passwords that were assigned and provided via mail to potential respondents. The web survey was accessed directly by respondents using their unique passwords. This also gave respondents the ability to return to their survey later if they chose to suspend their interview.

Pretest

Prior to the start of data collection, SSRS conducted 10 cognitive pretest interviews of the online and hard-copy survey instruments between Nov. 25, 2024, and Dec. 4, 2024. All cognitive pretest interview participants were recruited from the SSRS Opinion Panel. To qualify for a cognitive pretest interview, participants needed to currently reside in Philadelphia, be at least 18 years old, have access to an internet-enabled computer or smartphone, and be at least somewhat comfortable using Zoom.

SSRS provided Pew with recordings from the interviews and a detailed memo including feedback on the overall instrument, new questions, and survey usability. Following the pretest phase, adjustments were made to the questionnaire, survey program, and paper surveys to prepare for the full launch.

Data Collection

A sequential web/phone and mail mixed-mode methodology was used to conduct interviews for this survey. Data collection was conducted in English and Spanish. Eighty-two percent of the surveys were completed via web/phone, while 18% were completed via mail mode.

All selected sample records received a double-sided invitation letter with one side in English and the other side in Spanish. This letter included The Pew Charitable Trusts' logo and was addressed to "Philadelphia Area Resident." The letter's text was developed in collaboration with Pew and included a QR code, a short web link for the survey, and a personalized PIN to access the web survey. The invitation letter included a $1.25 cash pre-incentive and an offer of a $10 payment upon completion of the survey via a virtual gift card code. The $10 payment was disbursed immediately after completion of the web survey.

Four days after mailing the invitation letter, a reminder postcard in English and Spanish was sent to all contacts. The purpose of this mailing was simply to remind potential respondents to reply to the initial mailing. The postcard did not contain the survey web link or the target respondent's personal PIN.

Seven business days after the survey invitation letters were mailed, questionnaire packets were sent via first-class USPS mail. This mail option ensured that we were able to reach respondents who do not have internet access, are unable to complete the survey online, or simply prefer to complete it on paper. This mailing was sent in a 9-by-12-inch envelope and contained:

- A personalized reminder letter—double-sided with one side in English and the other side in Spanish—printed in color with The Pew Charitable Trusts' logo, explaining the nature of the survey.

- Two 8-page questionnaire booklets (one in English and one in Spanish).
- A postage-paid business reply envelope.

For respondents completing the survey via mail mode instead of web, a promised $10 post-paid honorarium was provided to respondents as a check mailed after receipt and verification of completion of the mail questionnaire. Table 8 shows the contact schedule for when each mailing was sent out.

Table 8: Contact Schedule

MAILING	DATE
1st release invitation letters mailed	Jan. 2, 2025
1st release reminder postcards mailed	Jan. 6, 2025
1st release reminder letters mailed	Jan. 13, 2025
2nd release invitation letters mailed	Feb. 11, 2025
2nd release reminder postcards mailed	Feb. 14, 2025
2nd release reminder letters mailed	Feb. 21, 2025

Data Processing and Quality Control

Data was checked after the first night of interviewing and throughout the field period to confirm that skip patterns were correctly followed. In addition, the back-end programmer created a program consisting of instructions derived from the skip patterns designated on the questionnaire to check the data. The program confirmed that data was consistent with the definitions of codes and ranges and matched the appropriate bases of all questions. The SSRS team also reviewed preliminary SPSS files and independently checked all the created variables to ensure that all variables were accurately constructed.

As a standard of practice, quality checks were incorporated into the survey. Quality control checks for this study included a review of "speeders," reviewing the internal response rate (number of questions answered divided by the number of questions asked) and open-ended questions.[3] Respondents who failed the quality checks employed were not included in the final dataset. Since the respondent set for this study only included addresses that were randomly selected and invited to the study, the risk of bots and fake profile respondents was greatly minimized. In addition, every address received a personalized, unique survey participation code that the respondent would need to enter to take the survey, and our system would not accept multiple web completions for the same participation code.

Given the multimodal nature of this survey, the only exception to multiple completions were respondents completing via web and mail. As such, there were some duplicate cases (i.e., respondents who completed a paper and web survey) that needed to be addressed. To handle

[3] This quality control check was only for interviews completed via the web survey.

duplicate cases, SSRS employed the following rules: 1) the survey that was completed first was kept and 2) if completed on the same date, the survey with the highest internal response rate was kept.

Weighting

Weighting is generally used in survey analysis to compensate for sample designs and patterns of nonresponse that might bias results. Weighting ensures that survey estimates are unbiased, and the demographic profile of the sample matches the profile of the target population. The sample is weighted to be representative of Philadelphia County's residential adult population.

There are three steps in the weighting process: (1) a base weight to account for the disproportionately stratified sample design, (2) a probability of selection adjustment to account for sampling fractions within household, and (3) a calibration of sample demographics to target population parameters.

Base Weight

Design Weight

The design or sampling weight for each sample piece drawn from stratum i is given by $d_{0i} = N_i/n_i$, where N_i represents the number of records in the sample frame for stratum i, and n_i denotes the number of records released in stratum i. There are 70 sample strata overall, defined by the cross of the seven regions by the five race density strata, by the oversample flag.

Nonresponse and Unknown Eligibility Adjustment

The nonresponse and unknown eligibility (NRUE) adjustment distributes the design weights of (1) eligible nonrespondents among respondents and (2) records whose eligibility cannot be determined among records for whom eligibility is known. Starting with design weight, d_0, the NRUE adjustment can be written as:

$$f = \frac{\sum_{R,c} d_0 + \sum_{N,c} d_0 + e * \sum_{U,c} d_0}{\sum_{R,c} d_0}$$

where:

$$e = \frac{\sum_{R,c} d_0 + \sum_{N,c} d_0}{\sum_{R,c} d_0 + \sum_{N,c} d_0 + \sum_{I,c} d_0}$$

That is, the NRUE adjustment factor, f, is the sum of the design weights for respondents, eligible nonrespondents, and eligibility-adjusted unknown-if-eligible records, divided by the sum of the design weights for respondents. The eligibility factor, e, is the design-weighted percentage of records with known eligibility status that are, in fact, eligible. The cross of geographic strata by

race density strata were used to define 35 adjustment cells. The NRUE-adjusted design weight, d_1, is calculated as:

$$d_1 = \begin{cases} d_0 \times f, & for\ respondents \\ 0, & otherwise \end{cases}$$

Number of Adults Adjustment

The final adjustment applied to the design weight accounts for the sampling of one adult per household. Since only one interview was conducted in each household for the study, adults who live in households with more than one adult have a smaller probability of being selected to complete the survey. This adjustment can be expressed as A_i, where A is the number of eligible adults in household i, capped at three adults.

Final Base Weight

The final base weight is the product of the NRUE-adjusted design weight and the number of adults adjustment.

$$d_{1_{FINAL}} = d_1 \times A$$

The final base weight was trimmed at the second and 98th percentiles, and standardized overall, to sum to the number of interviews.

Calibration

With the base weight applied, the data was calibrated to balance the demographic profile of the sample to target population benchmark distributions. Benchmarks were derived from ACS data. Table 4 lists variables that were used in the calibration and the source of the benchmarks.

Weighting was accomplished by raking sample distributions to target population distributions using iterative proportional fitting. This procedure balances each calibration variable to target benchmarks individually and iteratively. The entire set of calibration variables is cycled through until the weights converge across all dimensions.

Weights were trimmed at the fourth and 96th percentiles to ensure that individual respondents do not have too much influence on survey-derived estimates.

Table 9. Calibration Variables and Sources

DIMENSIONS	SOURCE
Gender	ACS 2023 PUMS data[4]
Age	
Education	
Race/ethnicity by nativity	
Internet use	
Philadelphia neighborhood	ACS 2023 five-year estimates[5]

Missing data in the raking variables were imputed using hot decking. Hot deck imputation replaces the missing values of a respondent randomly with another similar respondent without missing data. Hot decking was done using an SPSS macro detailed in "Goodbye, Listwise Deletion: Presenting Hot Deck Imputation as an Easy and Effective Tool for Handing Missing Data" (Myers, 2011).

The following table compares unweighted and weighted sample distributions to target population benchmark distributions for all variables used in the calibration.

Table 10. Sample Demographics

VARIABLE	CATEGORY	BENCHMARK	UNWEIGHTED	WEIGHTED
Gender	Male	46.2%	42.2%	45.0%
	Female	53.8%	57.8%	55.0%
Age	18-24	10.4%	7.6%	9.5%
	25-29	11.0%	10.9%	10.6%
	30-35	15.1%	12.8%	14.7%
	36-49	22.6%	23.2%	22.7%
	50-60	15.9%	14.3%	16.0%
	61-64	5.7%	6.0%	6.1%
	65 or older	19.3%	25.2%	20.4%
Education	Less than high school	12.0%	5.0%	9.5%
	High school graduate	31.1%	20.0%	30.2%
	Some college	22.9%	21.0%	23.8%
	College grad or more	34.0%	54.0%	36.5%

[4] U.S. Census Bureau. (2023). *American Community Survey Public Use Microdata Sample (PUMS)*. Retrieved from https://usa.ipums.org/usa/.
[5] U.S. Census Bureau. (2023). *American Community Survey 5-year estimates (2017-2023)*." https://www.census.gov/programs-surveys/acs/data.html. Accessed January 2025.

Race/ethnicity	White	35.9%	45.4%	37.8%
	Black	37.5%	24.1%	36.1%
	Hispanic	10.8%	14.4%	11.5%
	Asian	8.7%	12.1%	8.4%
	Another race	7.1%	4.1%	6.2%
Gender by race/ethnicity	Male, White/another race	20.7%	22.8%	21.2%
	Male, Black	16.1%	7.8%	14.4%
	Male, Hispanic	5.3%	5.6%	5.6%
	Male, Asian	4.1%	6.1%	3.8%
	Female, White/another race	22.3%	26.7%	22.8%
	Female, Black	21.4%	16.3%	21.7%
	Female, Hispanic	5.5%	8.7%	5.9%
	Female, Asian	4.6%	6.0%	4.6%
Age by race/ethnicity	18-29, White/another race	8.6%	7.7%	8.1%
	18-29, Black	7.4%	2.9%	6.2%
	18-29, Hispanic	3.4%	3.7%	3.6%
	18-29, Asian	2.0%	4.2%	2.1%
	30-49, White/another race	16.2%	17.5%	16.6%
	30-49, Black	13.4%	7.5%	12.5%
	30-49, Hispanic	4.6%	6.1%	4.9%
	30-49, Asian	3.4%	4.9%	3.4%
	50 or older, White/another race	18.2%	24.2%	19.3%
	50 or older, Black	16.7%	13.6%	17.3%
	50 or older, Hispanic	2.8%	4.6%	3.0%
	50 or older, Asian	3.3%	3.0%	2.8%

Education by race/ethnicity	HS or less, White/another race	13.7%	7.3%	12.7%
	HS or less, Black	19.5%	9.1%	17.4%
	HS or less, Hispanic	5.8%	6.5%	6.1%
	HS or less, Asian	4.0%	2.1%	3.4%
	Some college, White/another race	8.1%	8.5%	8.6%
	Some college, Black	11.3%	7.4%	11.4%
	Some college, Hispanic	2.6%	3.5%	2.8%
	Some college, Asian	0.9%	1.6%	1.0%
	College or more, White/another race	21.2%	33.6%	22.7%
	College or more, Black	6.7%	7.6%	7.2%
	College or more, Hispanic	2.4%	4.4%	2.6%
	College or more, Asian	3.7%	8.3%	4.0%
Home internet access	Yes	95.7%	95.2%	95.5%
	No	4.3%	4.8%	4.5%
Philadelphia neighborhood	South and Center	16.1%	25.6%	17.2%
	Southwest	6.4%	4.9%	5.9%
	West	11.4%	9.7%	11.1%
	Lower Northeast	17.9%	15.5%	18.4%
	Upper Northeast	16.9%	13.8%	16.6%
	North	21.9%	19.0%	21.2%
	Northwest	9.4%	11.4%	9.6%

Design Effect and Margin of Sampling Error

Specialized sampling designs and post-data collection statistical adjustments require analysis procedures that reflect departures from simple random sampling. SSRS calculates the effects of these design features so that an appropriate adjustment can be incorporated into tests of statistical significance when using these data. The so-called "design effect," or *deff*, represents the loss in statistical efficiency that results from a disproportionate sample design and systematic nonresponse. The total sample design effect for this survey is 1.79.

SSRS calculates the composite design effect for a sample of size *n*, with each case having a weight, *w*, as:[6]

$$deff = \frac{n\sum w^2}{(\sum w)^2}$$

The survey's margin of error is the largest 95% confidence interval for any estimated proportion based on the total sample—one around 50%. For example, the margin of error for the total sample is plus or minus 2.7 percentage points. This means that in 95 out of every 100 samples drawn using the same methodology, estimated proportions based on the entire sample will be no more than 2.7 percentage points away from their true values in the population. Margins of error for subgroups will be larger.

It is important to remember that sampling fluctuations are only one possible source of error in a survey estimate. Other sources, such as measurement error and reporting inaccuracy, may contribute additional error of greater or lesser magnitude.

Table 11. Design Effects and Maximum Margins of Sampling Error

GROUP	N	DESIGN EFFECT	MARGIN OF SAMPLING ERROR
Total sample	2,289	1.79	±2.7 percentage points
White/another race	1,133	1.73	±3.8 percentage points
Black	551	1.50	±5.1 percentage points
Hispanic	329	1.93	±7.5 percentage points
Asian	276	2.19	±8.7 percentage points

[6] Kish, L. (1992). Weighting for Unequal Pi. Journal of Official Statistics, Vol. 8, No.2, 1992, pp. 183-200.

Response Rate

Table 12 below reports the disposition of all sampled records that were contacted. The response rate estimates the fraction of all eligible sample that was ultimately interviewed. Response rates are computed according to American Association for Public Opinion Research standards.[7] The response rate for this survey was 18.3%.

Table 12: Response Rate

DISPOSITION	N
1. Complete (I)	2,289
2. Eligible, non-interview (R)	
Refusal & break-off	254
Deleted interview	1
3. Unknown eligibility, non-interview (UH)	
Nothing ever returned	15,528
Contacted, unknown if eligible	7
4. Not eligible, returned (IN)	
Ineligible	6
Undeliverable	1,415
Total records contacted	19,500
e=(I+R)/(I+R+IN)	64.2%
Response rate	18.3%

Deliverables

Preliminary

SSRS delivered a preliminary weighted SPSS on the first n=2,032 interviews obtained. This preliminary dataset was fully labeled and included created variables. SSRS also delivered a memo summarizing findings based on questions of interest to Pew.

Final

SSRS delivered to The Pew Charitable Trusts (1) a final weighted SPSS dataset, (2) two final weighted banners in Microsoft Word and Excel format, (3) a final version of the English questionnaire (for both web and mail administration), (4) a final banner specification memo, (5) a topline report, and (6) a final methodology report. As with the preliminary deliverables, SSRS delivered a memo summarizing findings based on questions of interest to the Pew team.

[7] The American Association for Public Opinion Research. (2023.) Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 10th edition. AAPOR.

Appendix A: About SSRS

SSRS is a division of AUS, a Mt. Laurel, New Jersey-based global market research and consulting firm. Through its affiliation with AUS, SSRS shares resources and experience with Marketing Systems Group.

SSRS is a full-service social science and market research firm managed by a core of dedicated professionals with advanced degrees in the social sciences. SSRS designs and implements solutions to complex strategic, tactical, public opinion, and policy issues in the U.S. and worldwide.

We partner with clients interested in conducting high-quality research. In the industry, SSRS is renowned for its sophisticated sample designs and its experience with all facets of data collection, including qualitative research, mixed methods, and multimodal formats.

The SSRS team specializes in creative problem-solving and informed analysis to meet its clients' research goals. SSRS provides the complete set of analytical, administrative and management capabilities needed for successful project execution.

SSRS is proud to be a charter member of the American Association of Public Opinion Research Transparency Initiative (www.aapor.org). The Transparency Initiative's goal is to encourage broader and more effective disclosure of research methods through proactively and routinely disclosing the critical research methods associated with publicly released studies.

SSRS is also a member of the Insights Association. Officially launched in January 2017, the Insights Association was formed through the merger of two organizations with long, respected histories of servicing the market research industry: CASRO and MRA. The result is a new, larger, and more-connected association with a unified, coordinated, and higher profile voice, aligned in mission and message, and ultimately more effective at advancing the industry and profession in which we all share an abiding passion. The Insights Association strives to effectively represent, advance, and grow the research profession and industry.