
2022 National Household Travel Survey

National ABS Data User Guide

Date: 7/21/2023

Submitted by: Ipsos Public Affairs LLC

2020 K Street, NW, Suite 410, Washington, DC 20006

Submitted to: Daniel E. Jenkins, PE

Senior Transportation Specialist, National Travel Behavior Data Program Manager

Office of Highway Policy Information



US Department of
Transportation

Federal Highway
Administration





Table of Contents

1. Introduction to the 2022 NHTS	1
1.1 Uses for the NHTS.....	2
1.2 Overview of Survey Content.....	4
2. The NHTS Program Evolution	5
2.1 2022 NHTS Survey Changes.....	5
2.2 Evolution of NHTS.....	5
3. Survey Procedures and Methodology	8
3.1 Sampling Design and Selection.....	8
3.1.1 Address Based Sample (ABS).....	8
3.1.2 Spanish Surveys.....	9
3.2 The Survey Process.....	10
3.3 Reducing Respondent Burden.....	11
3.4 Data Collection Method.....	11
3.5 Household Reporting.....	13
4. Data Validation & Edits	14
4.1 Clerical Review.....	15
4.2 Recontact Process.....	17
5. Description of Data Files	18
5.1 Basic Data File Structure.....	18
5.2 Data File Relationships.....	18
5.3 Records Included in Data Files.....	19
5.4 Data Documentation.....	20
5.5 Data File Conventions and Special Codes.....	20
5.5 Trip Purpose Variables.....	21
6. Survey Response Rates	21
6.1 Classification of all Sampled Addresses.....	22
6.2 Profile and Overall Survey Response Rate.....	23
7. Weighting	23
7.1 Weighting the Data.....	23
7.2 Which Weight to Use?.....	24
7.3 Weighting Targets and Benchmarks.....	24



Appendix 1:	25
Resources on the NHTS Website	25
Travel Concepts	25
Trip	25
Person Trip	25
Person Miles of Travel (PMT)	25
Vehicle Trips	26
Vehicle Miles of Travel (VMT)	26
Vehicle Occupancy	26
Merging Data from Multiple Files	26
<i>ID Numbers</i>	26
ID Variables Not Always Sequential	28
Sum of Weights / Control Totals	28
Non-sampling Errors	29
Sampling Errors	29
Appendix 2: Trip Purpose and Mode Changes, 2017 to 2022	30



Index of Tables and Figures

Tables:

Table 1: NHTS Evolution 2001-2022	5
Table 2: Daily Data Validation Steps	14
Table 3: Common Errors, in Order of Frequency	16
Table 4: Data Files	18
Table 5: Special Codes.....	20
Table 6: Trip Purpose Variables	21
Table 7: Household Dispositions.....	23
Table 8: Profile and Overall Response Rates	23
Table 9: Examples of Link Variables between the 2022 NHTS Data Files	26
Table 10: Number of Cases and Sum of Weights for Households, Persons and Trips.....	28
Table 11: Number of Cases and Projected Numbers for Key Travel-Related Results.....	29
Table 12: Trip Purpose Codes, 2022 and 2017 NHTS.....	30
Table 13: Means of Transportation Codes, 2022 and 2017 NHTS	31

Figures:

Figure 1: ABS Monthly Sample Selection	9
Figure 2: ABS Sample Data Flow	12
Figure 3: 2022 NHTS Data File Relationships	19
Figure 4: Example of a Full Travel Day	20
Figure 5: Merging Data Files	27



1. Introduction to the 2022 NHTS

The National Household Travel Survey (NHTS) is the U.S. Department of Transportation's (USDOT's) nationally representative data source for daily local and long-distance passenger travel. This inventory of travel behavior has been conducted every five to eight years since 1969 in an effort to capture travel mode and trip purpose among U.S. household residents.

For more than half a century, the NHTS has been a vital and reliable instrument for the collection and analysis of national transportation details. There have been nine iterations of the survey—in 1969, 1977, 1983, 1990, 1995, 2001, 2009, 2017, and now in 2022—that provide a significant record of how the nation's travel behavior has changed alongside the country's demographic, economic, and cultural composition.

The continued collection and analysis of national transportation data allows agencies to monitor the performance and adequacy of current transportation facilities and infrastructure, and to plan for future needs. This data is used to inform program and project development, and is included in biennial reports to Congress on the nation's surface transportation system performance. The NHTS also serves as a base for further research and development at academic institutions and in private business. The following are examples of recent issues of interest that can be supplemented by the NHTS data:

- **Mapping Current Travel Patterns to Assess Infrastructure Issues and Community Needs:**
Data on travel-to-work patterns by mode and time of day can be used to identify areas for infrastructure improvements to reduce congestion. Data from the 2022 NHTS can also be used to assess access to public transit, paratransit, and rail services.
- **Assessing Travel by Various Demographic Groups:**
The 2022 NHTS includes measures of mobility and trip rates by different social and demographic data dimensions, including vehicle ownership and access to public transport at various income levels.
- **Modeling Developing Traffic Trends:**
Examination of 2022 NHTS data reveals the increase in online home delivery services for food, groceries, and household goods and their impact on travel behavior as well as the proliferation of internet-enabled ride-sharing apps. These developing traffic trends can be expected to influence the use of a community's roads and public transit systems.
- **Understanding New and Emerging Modes of Travel:**
An expanded data file for travel mode can be used to assess the impact of emerging travel modes, such as dockless bikeshares, e-scooters, and other urban mobility devices.
- **Understanding Energy Consumption, Environmental Concerns, and Community Interests:**
Household vehicle information is included in the 2022 NHTS and can help model energy consumption of the household vehicle fleet and the trend in vehicle type and usage in relation to gas prices.
- **Assessing Bicycle and Pedestrian Travel:**
The 2022 NHTS database includes measures of bike and walk travel for health measures and environmental concerns.



- **Analyzing Current Health and Safety Concerns:**
The 2022 NHTS provides measures of travel by mode to establish exposure rates for highway safety analyses.
- **Assessing Community Support for Projects and Policies:**
Comments from the 2022 NHTS respondents can be used to measure support for Federal, State, and local planning activities and policies.

The 2022 iteration of the NHTS continues this legacy of successfully fielding a comprehensive travel survey and creating a reliable database of America’s travel patterns and infrastructure needs. Over the course of a year, Ipsos developed the survey instruments and web-based platforms, pretested the surveys, conducted cognitive and usability tests of the instruments and electronic systems, developed the sample frames, fielded the surveys, cleaned, and weighed the data, and developed the weighted and unweighted data files. This process is described in the following sections and is intended to build understanding of the 2022 NHTS Data Files.

1.1 Uses for the NHTS

The data collected as a part of the NHTS serves to support a broad range of research needs for the USDOT, State and local agencies, academics, advocacy organizations, elected officials, and the media on pressing matters related to transportation, safety, and logistics. The following section contains recent (January 2022 – December 2022) examples of publications that used data from the NHTS as the basis for their findings. These examples cover a diverse range of topics in the areas of transportation, health, safety, environment, and engineering and were published in various journals including, but not limited to, the *Transportation Research Record*, the *Journal of Transport Geography*, and the *Journal of Transport & Health*.¹

Mapping Current Travel Patterns to Assess Infrastructure Issues and Community Needs: Anecdotal evidence taken from the NHTS suggests that the affordable housing crisis is forcing households to move to the outer reaches of major metropolitan areas, leading to an increase in commute distance for work. Using spatial regression, a study of the Orange County/Los Angeles area found that a higher ratio of jobs to affordable housing is associated with longer distance commutes, and there is substantial variation in jobs-housing fit across different neighborhoods in Los Angeles. The study recommended expanding and protecting the supply of long-term rental housing in job-rich neighborhoods to address the growing commute distances.²

Modeling Developing Traffic Trends: The publication, *Why Has Public Transit Ridership Declined in The United States?* Found that in the past decade, ridership in the United States declined 15 percent and rail ridership declined 3 percent. This decline was widespread and unlike trends in other developed countries. Using data from 215 Metropolitan Statistical Areas (MSAs) prior to the COVID-19 pandemic, they identified the factors responsible for this decline and were able to quantify the contribution of

¹ National Household Travel Survey - nhts.ornl.gov. (n.d.-a). https://nhts.ornl.gov/assets/2022_compendium.pdf

² Blumenberg, E., & Siddiq, F. (2022, February 18). Commute distance and jobs-housing fit - transportation. SpringerLink. <https://link.springer.com/article/10.1007/s11116-022-10264-1>



each. By providing a clear understanding of the causes of transit ridership decline, this research provides the foundation on which communities can craft an effective response to the problem.³

Assessing Bicycle and Pedestrian Travel: The California Department of Transportation had a goal of tripling walking and doubling biking and transit use in their most recent strategic plan. Using data from the 2012 California Household Travel Survey (CHTS) and the 2017 National Household Travel Survey (NHTS) suggest that biking, walking, and transit use in California decreased over this five-year period. In the publication, *The Mode Is Not the Methods: Assessing Changes in Biking, Walking and Transit in California Using the 2012 CHTS and 2017 NHTS*, an independent third party assessed each survey to see if the decline was likely to be real or the result of methodological differences. They found that overall, the use of biking, walking, and taking public transit had declined during this period. In this case, the NHTS was used as a ‘data check’ against another source, fulfilling an important role as a survey tool.⁴

Understanding Energy Consumption, Environmental Concerns, and Community Interests: The EcoDataLab created a consumption-based emissions inventory (CBEI) using data from the NHTS. Using the variable, household Vehicle Miles Travelled (VMT) from the NHTS and other sources, they were able to create a model that estimates the average consumption of goods and services and the associated emissions. This type of model can help policymakers understand the amount of GHG emissions that might be produced in their jurisdiction.⁵

Assessing Travel by Various Demographic Groups: The study, *Travel Behavior of Transportation-Disadvantaged Populations: Trends and Geographic Disparities* explored the travel behavior and mobility of transportation-disadvantaged groups, such as older adults, people with disabilities, low-income households, and rural residents, by analyzing data from the 2017 National Household Travel Survey (NHTS). It looked at variables such as trip rates, miles driven, travel mode, and other behaviors to understand the challenges faced by these groups in terms of transportation. Understanding trends in the travel behavior of transportation-disadvantaged populations is important for informing policy and transportation investment decisions.⁶

Understanding New and Emerging Modes of Travel: Shared-use mobility services such as ridesharing, bike sharing, and carsharing are becoming increasingly popular in large cities. One study, published in the article, *Interest of Shared Mobility and Emerging Vehicle Technologies in Rural America*, explored what would happen if these new transport modes were introduced into rural and small-urban communities. The study analyzed the interest and adoption patterns for shared mobility services in these areas using the 2017 NHTS data. Input about the interest and potential willingness to use these

³ Erhardt, G. D., Hoque, J. M., Goyal, V., Berrebi, S., Brakewood, C., & Watkins, K. E. (2022, July). Why has public transit ridership declined in the United States?. ScienceDirect. <https://linkinghub.elsevier.com/retrieve/pii/S0965856422000945>

⁴ Pike, S., & Handy, S. (2022, August 31). The mode is not the methods: Assessing changes in biking, walking and transit in California using the 2012 CHTS and 2017 NHTS: Published in findings. Findings. <https://findingspress.org/article/37777-the-mode-is-not-the-methods-assessing-changes-in-biking-walking-and-transit-in-california-using-the-2012-chts-and-2017-nhts>

⁵ EcoDataLab. (n.d.). Consumption-based emissions inventory methodology. CBEI Methodology. <https://www.ecodatalab.com/cbei/methodology>

⁶ Mattson, J., & Molina, A. (2022, February). Travel behavior of transportation-disadvantaged populations: Trends and geographic disparities. UGPTI. <https://www.ugpti.org/resources/reports/details.php?id=1057>



services, and adoption of various emerging vehicle technologies, could help improve understanding and planning for appropriate shared mobility services to meet the transportation needs in rural and small-urban communities.⁷

1.2 Overview of Survey Content

The 2022 NHTS data files form a thorough inventory of the travel habits and demographics of individuals and households. The daily travel data collected includes a record of all trips taken within a 24-hour period by all household members aged 5 or older. It is inclusive of US households of all types (urban, suburban, rural), regardless of household size or level of travel activity. Excluded from sampling are group homes, such as dormitories, prisons, rest homes, or other large occupancy residencies of 10 or more unrelated persons.

The 24-hour tour begins at 4 AM on the assigned travel date and ends at 4 AM on the following day (that is, if someone is assigned the 1st of March, their travel day begins at 4 AM on the 1st of March and ends at 4 AM on the 2nd of March). Respondents reported the trip purpose, transportation mode, travel time of day, travel day of the week, and the vehicle occupancy for each trip. This data was then linked with the household's vehicle characteristics, demographic characteristics, and socio-economic characteristics to create comprehensive travel records.

In addition to general person, household, vehicle, and daily travel data, the 2022 NHTS collected data on several additional topics, including:

- The use of alternative means of transportation such as biking, walking, public transportation, taxi, ride share, e-scooters, etc.
- The short- and long-term impacts of the pandemic on working from home, schooling from home, and use of public transit.
- Whether respondents worked from home (and at what frequency) and/or usual mode of transportation to work.
- The use of outsourcing household tasks through delivery service or online purchasing as a form of trip substitution.
- The use of household vehicles for commercial purposes.
- Characteristics of the most recent long-distance trip.

The survey questionnaire underwent a careful development process to ensure the survey content was accurate, reliable, and could be used to create validated estimation metrics to measure respondent data and trend against prior NHTS results. A thorough examination was conducted on concepts relevant to the NHTS areas of interest to develop adequate measurements for each topic. Survey items were constructed to minimize the possibility for biases such as social desirability bias, item context effects, response order bias, and positivity bias. This included cognitive interviews among a small, but diverse, sample of nine individuals in early 2020 to evaluate the effectiveness of the instrument as a tool for

⁷ Godavarthy, R., & Hough, J. (2022, December). Interest of Shared Mobility and Emerging Vehicle Technologies in Rural America. SURCOM - interest of shared mobility and emerging vehicle technologies in rural America. <https://www.ugpti.org/resources/reports/details.php?id=1109&program=surcom>



measuring critical concepts and travel behaviors. The presentation of all questions and responses was examined to ensure a positive experience for all interview modes prior to final development.

An initial pilot study was conducted from September 2021 to December 2021. The main survey was conducted from January 2022 to January 2023. Travel days were assigned for all seven days of the week over the 12-month period.

2. The NHTS Program Evolution

2.1 2022 NHTS Survey Changes

The 2022 iteration of the NHTS implemented significant changes to survey methodology and procedure. The most notable change is the data collection approach shifted from a two-stage process (recruitment and travel day record) to a one-stage process. Previously, respondents were recruited via mail and asked to return the mail recruitment survey prior to being mailed travel diaries to be completed on a date scheduled by the survey contractor after the recruitment survey was returned.

2.2 Evolution of NHTS

Table 1 below highlights the past three NHTS surveys, in 2001, 2009, 2017, and the most recent survey in 2022. It includes an overview of the sample size, selection, unique attributes, add-ons, and other metrics that demonstrate how the survey has evolved over time.

Table 1: NHTS Evolution 2001-2022

NHTS Data Series				
	2022 NHTS	2017 NHTS	2009 NHTS	2001 NHTS
<i>Number of Households</i>	27,290 (7,893 national and 19,397 add-ons)	129,696 (26,099 national and 103,597 add-ons)	150,147 (25,510 national and 124,637 add-ons)	26,038 national and approx. 40,000 add-ons)
<i>Sample Selection</i>	Random sample of residential addresses selected from the U.S. Postal Service formed an ABS	Random sample of residential addresses selected from the U.S. Postal Service formed an ABS	List-assisted sample of telephone numbers formed a Random Digit Dialing (RDD) sample	List-assisted sample of telephone numbers formed a Random Digit Dialing (RDD) sample



NHTS Data Series				
	2022 NHTS	2017 NHTS	2009 NHTS	2001 NHTS
<i>Interview Methods</i>	Mail-push-to-web recruitment letter with link and login information for a single-stage web instrument. Option to request paper survey.	Mail-back (recruit survey), Web and telephone for both recruit and retrieval surveys carried out in two stages, recruitment and log	Telephone interviews in two-stages, recruitment and log	Telephone interviews in two-stages, recruitment and log
<i>Number of Contacts</i>	Two – one survey at the household level, one for each person in the household 5 years of ages and older	Two – one survey at the household level, one for each person in the household 5 years of ages and older	Two – one interview at the household level, one for each person in the household 5 years of ages and older	Two – one interview at the household level, one for each person in the household 5 years of ages and older
<i>Contractor</i>	Ipsos Public Affairs, Washington, D.C.	Westat, Rockville, MD	Westat, Rockville, MD	Westat, Rockville, MD
<i>Travel Day Data</i>	Retrospective recorded travel day (1 day prior)	Travel log used for one day	Travel log used for one day	Travel log used for one day
<i>Long-Distance Travel Definition</i>	A count of all trips of 50 miles or more and details on the most recent trip of 50 miles or more with a return home during the 30 days preceding and including travel day	No data collection to specifically record long-distance trips	No data collection to specifically record long-distance trips	All trips of 50 miles or more with a return home during the 28 days preceding and including travel day
<i>Unique Attributes</i>	<ul style="list-style-type: none"> ▪ Trip distance calculated using Google API based upon traffic patterns and likely route during time and date of trip ▪ Shared trip information auto-filled for accompanying household members 	<ul style="list-style-type: none"> ▪ Online geocoding in real time with Google API by participant or interviewer ▪ Trip distance calculated by shortest distance along path - trips starting and ending at the same location collected as loop trips 	<ul style="list-style-type: none"> ▪ Geocoding of addresses was conducted online during the CATI retrieval interview ▪ An experiment cellphone only HH sample of 1,254 (discussed separate from the main survey) 	<ul style="list-style-type: none"> ▪ First time the long trip survey (American Travel Survey) was combined with the daily trip survey (NPTS)



NHTS Data Series				
Response Rate	11.3%	15.6%	19.8%	41.0%
Add-ons	<p>Non-Statewide Add-Ons:</p> <ul style="list-style-type: none"> Oahu MPO <p>Statewide Add-Ons:</p> <ul style="list-style-type: none"> Tennessee Virginia 	<p>Non-Statewide Add-Ons:</p> <ul style="list-style-type: none"> Fort Worth, TX Des Moines, IA Tulsa, OK Waterloo, IA <p>Statewide Add-Ons:</p> <ul style="list-style-type: none"> Arizona California Georgia Maryland North Carolina New York South Carolina Texas Wisconsin 	<p>Non-Statewide Add-Ons:</p> <ul style="list-style-type: none"> Cedar Rapids, IA Omaha, NE Phoenix, AZ Piedmont, NC Tucson, AZ Chittenden County, VT <p>Statewide Add-Ons:</p> <ul style="list-style-type: none"> California Florida Georgia Iowa Indiana North Carolina New York South Carolina South Dakota Tennessee Texas Virginia Vermont Wisconsin 	<p>Non-Statewide Add-Ons:</p> <ul style="list-style-type: none"> Baltimore MPO Des Moines MPO Kentucky (4 counties) Lancaster, PA MPO Oahu MPO <p>Statewide Add-Ons:</p> <ul style="list-style-type: none"> Hawaii New York Texas Wisconsin



3. Survey Procedures and Methodology

Proper interpretation of data files for the 2022 NHTS requires an understanding of the survey fielding and data collection processes. Systematic procedures and techniques were used to collect, analyze, and interpret data collected for the 2022 NHTS. The survey design process, sampling methods, and questionnaire design were all customized for the purposes of eliciting the maximum usable data for the research goals of the Department of Transportation.

3.1 Sampling Design and Selection

The sample design and selection for the 2022 NHTS ensured that a representative sample of the population was selected so that results can be generalized to the larger population. A few key points were considered when designing and selecting the sample:

1. The population was defined as the non-institutionalized households of the United States.
2. The sampling frame was identified and obtained on a monthly basis.
3. The sampling methodology was developed, and a stratified random sampling method was chosen.
4. The sample sizes were determined based upon the level of precision required and the allotted allocation of completes for the NHTS.
5. The sample was selected with careful consideration regarding minimizing any potential bias.
6. The selected sample was examined and assessed to ensure all research criteria were met. Checks were initiated to minimize the instance of outliers or biases affecting the sample selection.
7. The selected sample was fielded after all checks were completed.

The 2022 NHTS collected national data using an ABS sample created and maintained by Marketing Systems Group (MSG) derived from the United States Postal Service (USPS) Delivery Sequence File (DSF).

The ABS sample frame and selection approach for the 2022 NHTS is comparable to what was used for the 2017 NHTS administration, in which a nationally representative sample of addresses was chosen to provide Statistically Valid Representation (SVR) for all USDOT NHTS reporting subdomains. A total of 7,893 households completed the interview. The final survey estimates have error margins of no more than 5 percent with a 95 percent confidence level. This allows the ABS study's findings to be used to estimate the Federal Highway Administration's (FHWA) mandated categorization and reporting of personal travel behaviors based on SVR samples for key subdomains such as the Nation, census division, urban, and rural geographies.

3.1.1 Address Based Sample (ABS)

ABS frames are used to produce many household probability samples due to their ease of accessibility and high coverage. In addition to the benefits of high coverage, ABS designs are highly efficient because of their unclustered nature.



Ipsos obtained the ABS sampling frame MSG. Each month MSG reconstructs its database using the latest USPS Delivery Sequence File of all residential addresses. A sample of addresses was selected on a monthly basis and assigned to various strata beginning in December 2021 and extending to December 2022. This design allowed for sample size adjustments before each mail outgo in areas where response rates differed from original assumptions regarding the yield of completed interviews.

The sample selection process was initiated after addresses were assigned to sampling strata. Once the strata assignment was complete, addresses were randomly selected from each stratum. Figure 1 summarizes the monthly sample selection for the ABS sampling frame.

Figure 1: ABS Monthly Sample Selection

Monthly ABS Study Selection



3.1.2 Spanish Surveys

Data collection for the 2022 NHTS was conducted in English and Spanish. The survey material was printed in English and Spanish, and included an option for respondents to obtain a Spanish version of the mailed paper survey.

In addition, all respondents entering into the web-based instrument were greeted with an initial language selection screen, prompting them to choose whether they preferred to complete the survey in English or Spanish.

There are several benefits to having a Spanish survey option:

- **Reaching a wider audience:** Having a Spanish survey option allowed Ipsos to reach a wider audience, especially in areas with a high Spanish-speaking population. This helped in the collection of more diverse and representative data.



- **Improving accuracy:** Having a Spanish survey option allowed Spanish speakers to provide clearer answers to survey questions and helped improve the accuracy of their responses. Respondents were likely to feel more comfortable answering questions in their native language, which ultimately can lead to more thoughtful and accurate responses.
- **Enhancing engagement:** Conducting surveys in Spanish can also enhance engagement with the target audience. The option of a Spanish survey showed a willingness to communicate with Spanish-speakers in their native language. This allowed Ipsos to establish a connection with respondents and increase their willingness to participate.
- **Gaining insights into a specific demographic:** Conducting surveys in Spanish provided insights into the preferences and behaviors of Spanish-speaking populations, which is valuable for the Department of Transportation’s research efforts.

In the 2022 NHTS, respondents had the option of contacting a help desk with Spanish speaking assistance available. Conducting survey interviews in Spanish was an important factor in gaining the trust and cooperation of Spanish-speaking respondents throughout the country.

3.2 The Survey Process

The 2022 NHTS survey process involved gathering information about the travel habits of the U.S. population. This data will be used for a variety of purposes, including Department of Transportation initiatives, infrastructure planning, academic studies, and market research.

The following steps were employed to conduct the 2022 NHTS:

1. The research objectives of the Department of Transportation were defined, and a survey questionnaire was developed to collect data to fulfill the objectives.
2. The target population was identified, and a sampling methodology was developed to glean the relevant data from the U.S population.
3. The survey questions were developed and confirmed with Department of Transportation representatives. Ipsos’ methodologists refined the questionnaire to avoid leading questions, biased language, or confusing wording.
4. The survey methods were determined whereby the 2022 NHTS was fielded as a mailed invitation with an online survey questionnaire. An option for a paper version of the survey questionnaire was provided.
5. The 2022 NHTS was initially tested in a pilot survey with a smaller sample size to identify any issues with the questions, the survey method, the format, or the survey operations.
6. The full 2022 NHTS was administered to the U.S. population via a mailed invitation over the course of a year.
7. Survey data were collected and analyzed throughout the course of the NHTS field period and in post-field. Survey responses were analyzed to identify patterns or trends and report to Department of Transportation representatives.
8. Final deliverables included a presentation of the survey findings and discussion of the limitations of the survey methodology or potential sources of bias. A non-response bias analysis was conducted to ensure a representative sample of the U.S. population was surveyed.



3.3 Reducing Respondent Burden

Ipsos undertook extensive efforts to ensure a reduced respondent burden for the 2022 NHTS. All aspects of the NHTS underwent multiple review stages and thorough testing to ensure the web survey and the paper survey would not cause undue burden on respondents.

This effort included:

- **A concise questionnaire:** Long surveys can be overwhelming for respondents and lead to respondent fatigue. Ipsos kept the 2022 NHTS questionnaire as short as possible and only included questions deemed essential by Department of Transportation representatives.
- **The use of extensive skip logic:** Skip logic helps to reduce the number of questions that each respondent must answer. A reduced survey length results in greater response rates.
- **Use of plain language:** Ipsos ensured clear and concise language that is easy for respondents to understand. Technical jargon and complex language were limited to ensure respondents were not confused by the questionnaire.
- **Survey pretesting:** The 2022 NHTS was pretested during a 3-month pilot phase to a small sample of respondents. The pre-testing allowed Ipsos to identify areas where respondents experienced difficulty and help to refine the survey to elicit a higher response rate and useful data collection from all respondents.
- **Incentives:** Ipsos provided incentives to encourage survey completion and to help increase response rates. Respondent burden was reduced by compensating respondents for their time and effort.
- **Review materials:** Ipsos' leading survey methodologists reviewed survey materials.
- **Timed reviews:** Ipsos conducted timed reviews of the online questionnaire to ensure a modest survey length.
- **Thorough inspection:** Ipsos thoroughly inspected the paper questionnaire to confirm the validity and mapping of all question options.

In the study, an online system was used to collect self-reported data from respondents, with built-in edit checks and subroutines to reduce respondent burden associated with completing the survey.

3.4 Data Collection Method

The 2022 NHTS was carried out beginning 1/20/2022 and closing on 1/19/2023. Prenotifications were first sent to sampled respondents on 1/17/2022 to inform them of the survey opening, and initial responses came in on the day of survey launch. The survey closed at midnight on 1/19/2023 in each given respondent's local time zone. If a respondent entered the travel day portion of the survey on 1/19/2023, they were able to complete their survey up until two weeks following that date.

2022 NHTS respondents were asked to use an online system to report travel data but were also given the opportunity to request paper mail-in surveys. The paper mail-in surveys were then entered into the same online system upon receipt. The automatic data gathering system comprised of the following:

- A dedicated website and custom software to enable the main household respondent to complete a household roster of persons and vehicles and then allow each household member

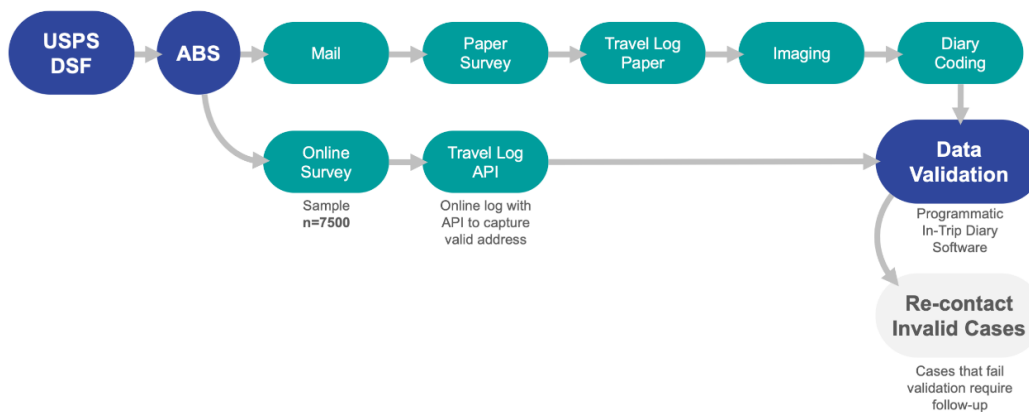
ages 16 and older to self-report travel activity (with automatic proxy reporting for those under age 16).

- Appropriate prompts and sophisticated question branching and skip patterns to facilitate accurate reporting and only present respondents with questions that were necessary and appropriate for that respondent.
- Comprehensive look-up tables to facilitate the reporting of vehicle make and model information.
- Household rostering (i.e., the listing of all vehicles and persons in the household) and a shared-trip reporting feature to allow a trip report from one household member that includes another household member (for example, spouses who travel to dinner together) to be inserted into the second person’s record, reducing reporting burden by not requiring respondents to re-report trips already reported by another household member.
- Real-time data range, consistency, and edit checks programmed to reduce reporting error, survey length, and maintenance of the flow of information processing.
- Integrated data consistency checks to ensure that the details entered passed basic data completeness and consistency requirements.
- A Google Maps application interface-enabled search engine to assist in identifying specific place names and locations.

Respondents were directed to an automated web-based data retrieval software system to complete the survey. Figure 2 outlines the general sample source and data capture approach for this study.

Figure 2: ABS Sample Data Flow

ABS Sample Data Flow



Mail surveys were returned using a return envelope issued to the respondent with their paper mail surveys. These surveys were then hand entered by a team of data entry specialists into an adapted version of the online survey program and stored initially in a separate dataset. This data was exposed to the same



validation checks and post-processing as the online data. This data was then appended to online survey data with a flag to clarify its origin as a mail response.

3.5 Household Reporting

In the 2022 NHTS, selected households received a letter inviting them to participate in the survey. Either on the online system or via a requested mail survey, the main household respondent completed a short roster to obtain general household information such as the number of household members and household vehicles. Respondents were then asked to record their travel for the previous day using the online travel day record or an equivalent paper form. Responding households were asked to report all trip information for a given day (typically the day before they entered the survey) for every household member 5 years of age and older.

If a household chose to do the survey online, the main respondent completed the household roster, filled out their travel day record, and acted as a proxy respondent for any children aged 5 to 15 in the home. Members of the household aged 16 and up were encouraged to create their own online diaries. Households that wanted to participate in the survey by mail received paper forms with comparable proxy-reporting instructions.

A completed household was one that met the following criteria:

- All household members ages 5 and older reported travel for households with 1-3 members. For households of 4 or more members, a minimum of 75% of members ages 5 and older reported travel.
- All respondents answered required questions.
- No more than 20% of non-required questions were missing for each household member.

The proxy reporting protocol was the following:

- Participants were presented with the proxy guidelines before person level details were collected for each household member:
 - Participants reported for themselves.
 - Household members younger than 16 required a proxy.
 - Proxy reporting was allowed for household members ages 16 and older who were unavailable or unable to respond for themselves.
- For all proxy reports, participants were required to identify who completed the survey on their behalf. Online participants could choose from a pre-populated list of all adults named in the survey or insert the name of a different adult.
- Participants were required to identify all household members aged 5 or older for whom proxy reporting was conducted.

Non-responders and partial respondents received reminders to finish the survey by mail and email. Any household was considered partly complete if the household had completed only a portion of the household profile or travel day record and/or if some of the household's travel day records were still incomplete. Households with incomplete forms received up to three reminder mailings and a reminder email (if an email address was provided) to encourage them to respond in a timely manner.



4. Data Validation & Edits

In addition to checks built into the 2022 NHTS survey platform (e.g., range restrictions, address confirmation), data collected as part of this study underwent daily data validation. Most daily data validation procedures focused on the recorded travel day and trips data. The daily data validation checks centered on identifying issues in the data collected via the travel day portion of the survey and addressing data concerns with a combination of checks, clerical review and, when needed, respondent follow-up.

In addition, there was at least one contact attempt made for any record requiring follow-up within one week of that record being “checked in” and exported from the survey platform. Records were exported from the survey platform daily in the evenings, and daily data validation checks began the following morning.

Table 2 below lays out daily data validation and quality control processes and how they feed into follow-up procedures and post-collection validation. Daily processes were based on a few assumptions:

1. Records that were output daily were for households where the recruitment/profile instrument was considered complete, and:
 - a. All eligible members ages 5 and older completed the travel record.
 - b. The required number of household members (100% of size 1-3, 75% of sizes 4 or more) completed the travel record.
 - c. The 2-week follow-up reminder window has passed.
2. Daily data exports ONLY contained the records that were newly completed between 8pm on the previous day and 8pm on the export day (i.e., new completes only).

Table 2: Daily Data Validation Steps

Applicable Files	Data Validation Checks
ALL	Confirm all survey variables present on output data
	Create programs/scripts to read files in for processing
	Create variable guide that lists variables, formats, & possible values/ranges.
HOUSEHOLD PERSON TRIP VEHICLE	Confirm all Household IDs are present in the PERSON, VEHICLES, and TRIPS files.
	Using the household file, create household size variable (HHSIZE). Add HHSIZE to the PERSON, VEHICLES, and TRIPS files.
HOUSEHOLD TRIP PERSON	Check records on all files for: <ul style="list-style-type: none"> • Comments • Missing location data (home, school, work, origin, destination) • Improper AM/PM use for departure or arrival time.

Applicable Files	Data Validation Checks
	Reported trip duration looks unreasonable (50% higher or lower than reported)
	Flag records that require review.
	Clerical review of Daily Follow-up Report.
	Append repaired/updated records to files.
TRIP PERSON	Create long-distance travel flag.
TRIP PERSON TRIP	Set long-distance travel flag to 1 if a destination is 50 miles or more from a respondent's home address (regardless of origin point). If no long-distance trips were entered in the recorded travel day, set long-distance travel flag to 0.
	Copy long-distance trip information from Trips file to long-distance section of Person file, where applicable.
	Restructure Trips file.
TRIP PERSON	Using the restructured Trips file, count trips per person and add to Person file.
	Using the restructured Trips file, calculate trips per household and add to Household file.
	Confirm that the number of Person records per household in the Person files matches household size in Household file.
VEHICLE	Confirm that the number of household vehicles in the Household file matches the number of Vehicle records in the Vehicle file
LOCATION	Create an executable program for pulling location data from other files and adding to its own file.

*Note: As corrections take-place, derived/created/composite variables were recalculated. For example, if a start or end time was corrected, that could impact trip duration. The addition of trips based on recontacts, etc. may require a recalculation of trips per person and trips per household.

4.1 Clerical Review

Automated data reports for clerical review were generated daily that flagged extreme or illogical values. These reports allowed for the identification of issues that required manual review or respondent follow-up.

The Google API estimate of trip duration was used to identify extreme values by mode of travel (car, bike, transit, walking). This included trips where the duration of the trip was unusually long (more than



twice the time estimated by the Google API given the mode of the trip), unusually short (less than half the estimated duration), or very unlikely to have happened at all (i.e., where the destination was Alaska or Hawaii and the mode was land travel, etc.).

In addition, open-ended responses were flagged for review. These open-ended comments included information from respondents that was not captured in their original household roster or trip log. For example, a respondent may indicate in the open-end response that they incorrectly entered the time the trip started/stopped or the destination/purpose of the trips.

The flags listed below triggered Ipsos’ clerical review process. During this stage of review, the designated Clerical Review Team manually reviewed each flagged record to identify required data edits and/or recontact prompts:

- Long duration: if the trip duration variable is more than double the Google API estimated trip duration. Trip duration is the self-reported time it took for the respondent to get from point A to point B. Trips with duration shorter than 30 minutes were excluded from this flag and automatically moved onto the next stage of data validation.
- Short duration: if the trip duration variable is less than half the Google API estimated trip duration. Trips with duration shorter than 30 minutes were excluded from this flag and automatically moved onto the next stage of data validation.
- Short duration and plane mode: Travel mode is airplane and actual trip duration is less than 30 minutes.
- Respondent provided a comment indicating a correction was needed in the data.
- Home address is PO Box: The confirmed home address is listed as a PO Box.

Table 3 below shows the most common error flags encountered during the clerical review process, sorted from most frequent to least frequent:

Table 3: Common Errors, in Order of Frequency

Common Errors
The respondent's comment indicates a correction was needed in the data
Long duration issues, where the trip duration variable is more than double the Google API estimated trip duration
Short duration issues, where the trip duration variable is less than half the Google API estimated trip duration
Short duration and plane mode, where the travel mode is airplane and actual trip duration is less than 30 minutes
The respondent's comment indicates one or more missing trips
A loop trip has one or multiple specific destinations
A question with a required response has missing data
The contact information for incentives is incorrect or invalid
The confirmed home address is listed as a PO Box
The respondent provides Spanish language comments that need translation



During this process, these files may have been subject to logic edits by the Clerical Review Team in order to ensure the information was entered accurately and could be included in the final data files. Any edits applied to the data at this stage are listed below:

- Missing values for identifying vehicle driver were filled in according to trip or household attributes.
- Travel mode values were populated based upon information in the vehicle file for records where a HH vehicle was used on the trip.
- Travel mode values in the other specify category were upcoded based on preliminary keyword analysis.
- Trip purpose values in the other specify category were upcoded based on preliminary keyword analysis.
- Vehicle make and models were upcoded using rules derived from keyword analysis of the Make and Model text in the other specify fields.
- Other corrections stemming from clerical review were applied to variables in the trip file as needed.

4.2 Recontact Process

All clerical review files were read into the cumulative csv files and households with trips that required a recontact were flagged. The updated cumulative csv files were loaded to the database and recontact sample files were exported. Household files with logical errors not resolved during the initial clerical review process were then loaded to an online survey tool to initiate the recontact process. Respondents were sent a recontact form via email and invited to correct the entries themselves. The respondents who provided an email or phone number in their initial survey response consented to recontact via that option.

When follow-up was necessary and the respondent provided an email address, the initial recontact attempt was conducted online via email using an online form. If the email went unanswered after 2 reminders and 1-2 weeks, the households who did not respond were sent to CATI (Computer Assisted Telephone Interviewing) for phone recontact.

Additionally, any complex cases that were not resolved online or via phone recontact (CATI), were identified as custom phone call cases. In these cases, a custom call was made to help mitigate problems while resolving complicated record issues such as multiple issues across trips, confusing comments, etc.

After responses were received from any of these review methods, the Clerical Review team generated the recontact response forms to enter the necessary edits or resolve any outstanding issues. If, after going through all phases in the validation and recontact process, there was still insufficient information to make the necessary corrections, the records were excluded from cumulative CSV files.

Once the recontact codes were assigned, the Clerical Review Team updated the corrected files using the Python “updates.py” script and the corresponding edits were applied to the cumulative CSV files.



5. Description of Data Files

5.1 Basic Data File Structure

The files associated with the 2022 NHTS Data User Guide represent the households that completed the 2022 NHTS survey between January 20, 2022, and January 19, 2023. They include data on the people, trips, vehicles, and locations associated with each household.

As in previous iterations of the NHTS, there are four data files: HOUSEHOLD, PERSON, VEHICLE, and TRIP. Each data file is output in CSV, SAS, and SPSS formats. These files are hierarchically structured (see Figure 3) and can be merged using shared ID variables as outlined in Table 4:

Table 4: Data Files

File	Record Level Description	ID Variable
HOUSEHOLD	One record per HH unit	HOUSEID
VEHICLE	One record per HH Vehicle (If present)	HOUSEID VEHID
PERSON	One record per HH member	HOUSEID PERSONID
TRIP	One record per HH member's travel day trip (If at least one trip is made)	HOUSEID PERSONID TRIPID

5.2 Data File Relationships

HOUSEHOLD

The HOUSEHOLD file contains primary household profile characteristics (e.g., household size and other relevant demographic information). This file holds the primary key for merging all files, the HOUSEID.

VEHICLE

The VEHICLE file contains characteristics regarding any household vehicles (e.g., make and model).

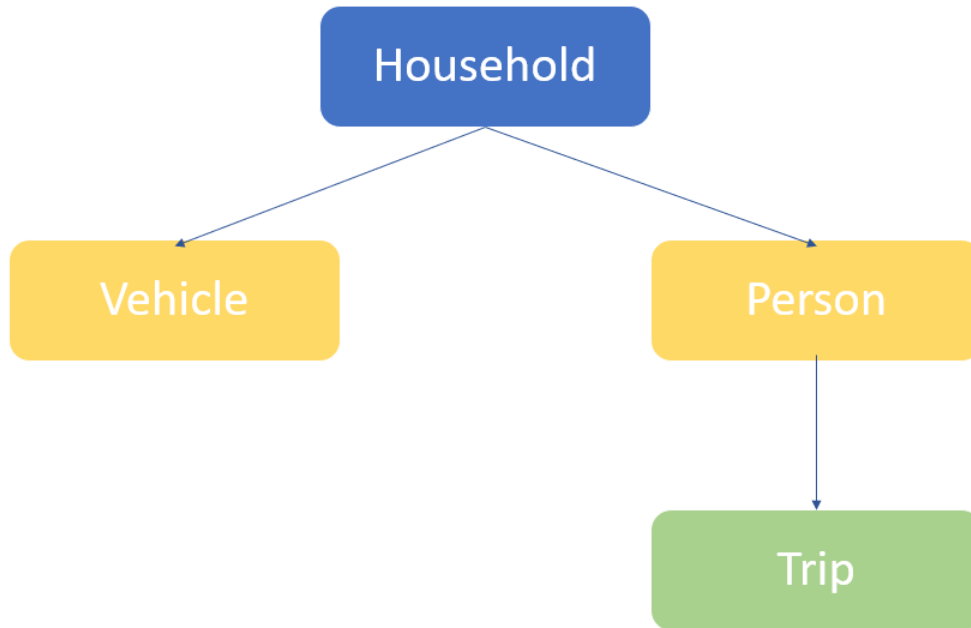
PERSON

The PERSON file contains characteristics of individual Household members (e.g., age and other demographic attributes).

TRIP

The TRIP file contains characteristics of trip(s) for each member of the household who took a trip (e.g., start times, end times and destinations).

Figure 3: 2022 NHTS Data File Relationships



5.3 Records Included in Data Files

A record is included in the files when it is complete. The requirements for a completed record depend on the size of the household. For a household with one to three members, a travel day record must be completed for all members of that household. A household with four or more members required a travel day record for a minimum of 75% of eligible household members to be considered complete.

HOUSEHOLD: A record is created when:

- A travel day record is completed for all eligible household members in households with 1-3 members over age 5.
- A travel day record is completed for at least 75% of eligible household members in households with four or more members over age 5.

VEHICLE: A record is created when a household vehicle is reported by the primary respondent in the survey. All vehicles leased, owned, or available for use are requested for declaration. If a household does not have any vehicles, that household ID is not present in this file.

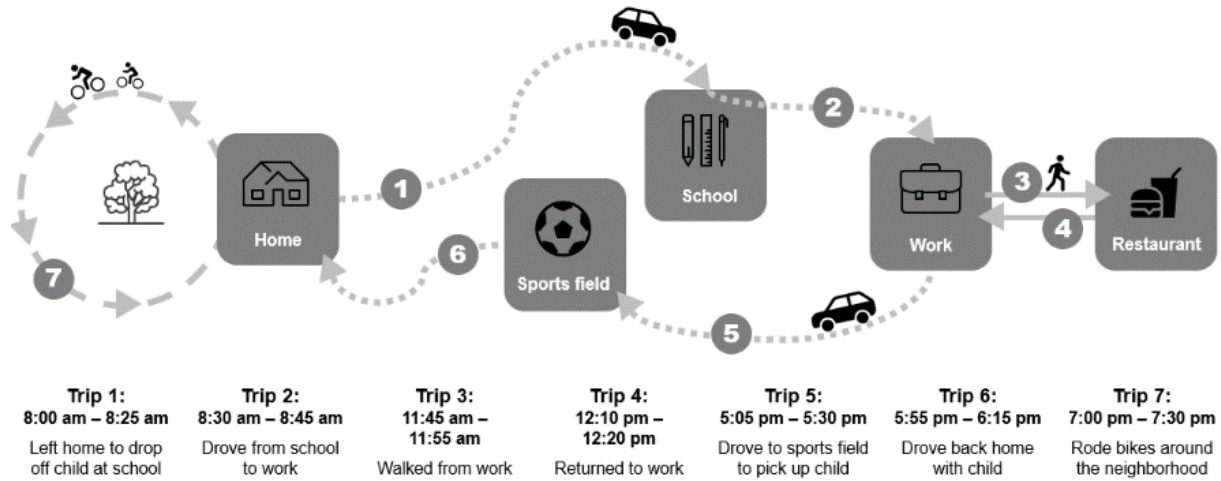
PERSON: Each record represents a reported individual member of a household. Each person is assigned a number that corresponds to the order they were entered into the roster by the primary respondent.

TRIP: Each record represents a trip as reported for a household member aged 5 and older on the travel day date. Respondents were instructed to report all locations traveled from 4:00 A.M. on the assigned travel day to 4:00 A.M. on the following day. They were also asked to include their mode of travel for each trip. The system is designed to capture 'loop trips' that start and end in the same location with no

other stops on the route. Figure 4 provides an illustration of a household’s full travel day. It includes the necessary information for a complete record, as well as examples of potential trips that would be recorded during the travel day.

Figure 4: Example of a Full Travel Day

Example Travel Day:



5.4 Data Documentation

Along with this User’s Guide, the NHTS codebook and data dictionary provide those who use this data with important context, metadata, and information about the 2022 NHTS files, variables, and variable values. This information includes, but is not limited to, sample size, weighted totals, as well as the complete question and answer text.

5.5 Data File Conventions and Special Codes

Several special codes are in place to capture specific cases present in the data files. Table 5 below contains a list of special codes used throughout the data files and their meanings:

Table 5: Special Codes

VALUE	LABEL	DESCRIPTION
-1	Valid Skip	Input during data processing when question is appropriately skipped, meaning by default no value exists
-9	Not Ascertained	Input during data processing when question is inappropriately skipped, and no value exists
-7	I Prefer Not to Answer	Input by the respondent as a valid answer choice
-8	I Don’t Know	Input by the respondent as a valid answer choice



5.5 Trip Purpose Variables

A key objective of the 2022 NHTS was to capture the reasons for travel during a respondents' travel day. When reporting travel, respondents were asked to categorize the reason for each trip, with the options varying depending on whether the trip was one-way trip (where the origin and destination of the trip were distinct) or if it was a loop trip (where the respondent started and ended their trip in the same place).

Trip purpose could be categorized as one of twenty activities, largely grouped by location. For example, a respondent could select if the trip's purpose was for at home activities; work; school, religion, or daycare; medical or dental; shopping or errands; social or recreational; or other activities, with some overlap allowed between these categories (i.e., "Buying meals" could be considered both a shopping/errand task and a social/recreational task). Loop trips could be categorized as one of three activities: went for a walk, jog, or run; went for a bike ride; or walked the dog. Both trip types could also be classified as "Something else" that the respondent could input. Additional trip purpose variables were derived based on the one-way WHYTO variable and are outlined in Table 6 below:

Table 6: Trip Purpose Variables

Variable Name	Derived Status	Description
LOOPTRIP1	No	Asked for all loop trips. "Why did you leave and return to...?"
WHYTO	No	Asked for all trips. "Why did you go to this location?"
WHYFROM	Yes	Derived from WHYTO for all but the first trip recorded as equal to WHYTO for the preceding trip
WHYTRP90	Yes	Derived from WHYTO to aggregate response options into categories consistent with 1990 NHTS design
WHYTRP1S	Yes	Derived from WHYTO and WHYFROM to aggregate response options into general purpose categories
TRIPPURP	Yes	Derived from WHYTO to aggregate response options into general purpose categories

6. Survey Response Rates

As described in Section 3.1, the 2022 NHTS included some significant changes to survey methodology. It is imperative to understand how these updates impacted overall response rates.

Response rates were calculated based on all households that were sent survey invitations between January 18 – December 29, 2022, with fielding between January 18, 2022 – January 19, 2023.

Two response rates were calculated. The first is an Overall qualified response rate based on households that completed the survey screener and reported travel for the minimum number of household



members (based on size and age). This rate is calculated using AAPOR response rates 3 (RR3)⁸. The second calculated rate is a Profile response rate that, in addition to completes, includes partial completes, defined as anyone who completed at least the initial age screener on the survey, to ensure at least one household member was age 18 or older. This is calculated using AAPOR response rate 4 (RR4)⁹. For each of these rates, guidelines from the 9th edition of the American Association for Public Opinion Research's (AAPOR) Standard Definitions were utilized. These response rates are further outlined in the following sections.

6.1 Classification of all Sampled Addresses

Given the ABS sampling approach, calculations were based on the AAPOR guidelines for mail surveys of unnamed persons.

Definitions for the dispositions are as follows:

- **Complete:** Households with one to three members aged 5 and up completed all diaries, and households with four or more members aged 5 and up completed at least 75% of all diaries.
- **Partial:** Began Household survey, and, at minimum, completed one question. The first survey question was a required age eligibility screener to confirm the main household respondent was at least 18 years old.
- **Refusal:** Household contacted Ipsos indicating they would not be participating in the survey. This includes both households that refused without starting the survey and those who refused and broke off.
- **Housing unit, unknown if eligible respondent:** Household was sampled, but no response was received. Eligibility was not confirmed.

⁸ The Qualified Response rate was calculated following AAPOR guidelines and using the RR3 formula $I/((I+P)+(R+NC+O)+e(UH+UO))$ or the number of Qualified Completed Households divided by all sample sent out (since there were no ineligible households) where:

I = Qualified Survey Complete

P = Partial/Profile Complete (Household did not have qualified complete)

R = Refusal

NC = Non-Contact

O = Other

UH = Unknown Household

UO = Unknown Other

e = estimated proportion of cases of unknown eligibility that are eligible.

Due to the lack of ineligible cases in the known sample, e=1 for these calculations.

⁹ Profile Response Rate was calculated following AAPOR guidelines and using the RR4 formula

$(I+P)/((I+P)+(R+NC+O)+e(UH+UO))$ where:

I = Qualified Survey Complete

P = Partial/Profile Complete (Household did not have qualified complete)

R = Refusal

NC = Non-Contact

O = Other

UH = Unknown Household

UO = Unknown Other

e = estimated proportion of cases of unknown eligibility that are eligible.

Due to the lack of ineligible cases in the known sample, e=1 for these calculations.



- **No screener completed:** Household entered survey program, but did not pass initial survey screener or confirm age eligibility.

Table 7 below provides the counts, where applicable, for each of these dispositions.

Table 7: Household Dispositions

Disposition Type	Count
Qualified Complete	7,893
Partial/Profile Complete	6,940
Refusal	34
Deceased respondent	1
Housing unit, unknown if eligible respondent	57,895
No screener complete	59
Total sample used	72,822
e (estimated proportion of cases with unknown eligibility treated as eligible)	1.000

6.2 Profile and Overall Survey Response Rate

Primary survey respondents were required to complete an initial age screener to confirm in order to participate in the survey; this was asked of the primary respondent as the first survey question, and the survey would not progress if the respondent refused or confirmed they were under age 18. Using AAPOR RR4, a profile response rate was calculated to capture the rate at which this screener was completed.

An overall response rate was calculated based on reaching the 75% completion threshold. This was reached when households of three or fewer had all eligible household members record their travel day and households of four or more had at least 75% of eligible members respond. This was calculated using AAPOR RR3. Table 8 below shows the profile and overall response rates.

Table 8: Profile and Overall Response Rates

Response Rate Type	Rate
Profile Response Rate	20.4%
Overall Response Rate	10.8%

7. Weighting

7.1 Weighting the Data

The weights reflect selection probabilities for each household, adjustments to account for eligibility, nonresponse, and undercoverage, and are smoothed across day and month of response. The 2022 NHTS samples were designed to select households in proportion to population within the study's strata (Census division by urban/rural status). The weighting process provided correctly balanced estimates by



the geographic strata, day and month of response, and adjustments for demographic factors, such as household size, education, race, and ethnicity. The weights, when used, provide estimates that are minimally biased. Weighted tabulations may be significantly different than unweighted estimates, indicating potentially large biases for the unweighted estimates. Thus, estimates with minimal biases are obtained by multiplying each data value by its appropriate weight and summing the results.

7.2 Which Weight to Use?

There are several different weights, and it is important that the appropriate weight be used for a particular estimate. There are household weights, person weights and trip weights, and each of these are further differentiated by Seven-day, Five-day, and Two-day versions. Seven-day weights should be used when estimating travel for the entire year for 2022. Five-day weights should be used for weekday travel estimates, whereas two-day weights are designed for weekend travel estimates.

- **Household weights (WTHHFIN, WTHHFIN5D, WTHHFIN2D):** Use these when tabulating an estimate at the household level (e.g., number of households by household vehicle ownership and distribution of households by number of household drivers).
- **Vehicle weights (WTHHFIN):** These are the same as the household weight since the vehicle is considered a household attribute. Use the household weight for items such as vehicles by vehicle type or by vehicle age.
- **Trip weights (WTRDFIN, WTRDFIN5D, WTRDFIN2D):** Use these for estimates involving numbers of trips or miles of travel, for example, number of vehicle trips by trip purpose. Only trips in privately operated vehicles (POV) that are reported by the driver should be counted in estimating vehicle trips. For example, if a person reports being a passenger in a vehicle driven by another household member, that trip would not be counted.
- **Person weights (WTPERFIN, WTPERFIN5D, WTPERFIN2D):** Use these for person-level estimates of non-household and non-travel day items of interest, for example workers by gender, drivers by annual miles estimated, etc.

As an example, to estimate the number of daily trips per household by Census region, for each region, calculate:

- The weighted count of households = Sum of the household weights, and
- The weighted count of trips = Sum of the trip weights.

The estimate of daily trips per household for that region is then simply its weighted trip count divided by its weighted household count.

7.3 Weighting Targets and Benchmarks

Weighting targets and benchmarks are values that exist independent of the NHTS with the utility of adapting and adjusting weights for survey non-response and non-coverage. Targets and benchmarks were used for the NHTS to address the number of U.S. Households and the number of inhabitants in those households.



Appendix 1:

Resources on the NHTS Website

The NHTS website (<http://nhts.ornl.gov>) is the data user's first stop to determine which of the many NHTS resources can assist the user in finding or creating the data needed. Many users seek data that is already available in the Frequently Asked for Tables or ask questions documented in the Frequently Asked Questions tab. In addition, the Data Explorer is an online analysis tool that allows users to create tables without having to download the files. The website also provides for user support (see the "Contact US" section).

The 2022 NHTS data files contain a large number of variables of different types, originating from different points of the data collection process. To facilitate navigation of these, we have provided a codebook and a data dictionary. The codebook is a straightforward source of primary information about each data file. It contains info on each of the HOUSEHOLD, PERSON, TRIP, and VEHICLE files with the names of variables, basic descriptive information, data characteristics, associated question, response ranges, and frequencies. It is a good source for initial quality checking of fundamental frequencies associated with each variable to ensure sound analysis.

The Data Dictionary contains a roster of all variables contained in the 2022 NHTS data files, listed alphabetically, and a codex of which file or files each variable is housed within.

Travel Concepts

The travel concepts described below are central to using the NHTS Trip data and are provided primarily for data users who are not familiar with NHTS data. All data users may find these concepts helpful, as the use of certain terms and concepts often vary by individual survey.

Trip

A trip is movement from one point to another on a respondent's travel day. It does not matter which member of the household, how far they went, where they began or where they went – movement from one point to another is a trip. This includes something like a jog or walking a pet, where the origin and destination are the same. A trip is starting and ending movement from one location to another using any mode of transportation over any period of time.

Person Trip

A person trip involves just one person using any mode of transportation. It is the simplest and most fundamental trip in the data. Each piece of trip data in the TRIP file is one person trip. If one person takes any trip to any place using any mode, that is one person trip. If a family of four travels together using any mode of transportation, that is four person trips.

Person Miles of Travel (PMT)

Person Miles of Travel (PMT) are the total number of miles traveled by each member of a trip utilizing any mode of transportation. It accounts for ALL miles traveled by ALL persons. If one person travels one mile, that is one PMT. If four people travel one mile, that is four PMT.



Vehicle Trips

A vehicle trip involves a single privately operated vehicle (POV) by a household regardless of the number of persons in the vehicle. It further does not consider distance, start, or end points. Vehicle trips are calculated by filtering for privately owned vehicle trips where a household member was the driver (DRVR_FLG = 01 and TRPTRANS = (01, 02, 03, 04, 06, 07)). If an entire household takes a trip in a car they own and a member of that household drives, that is a vehicle trip. A person living alone taking their personal car on a trip is also a vehicle trip.

Vehicle Miles of Travel (VMT)

Vehicle Miles of Travel (VMT) is the total number of vehicles that a single vehicle privately owned by a household travels. It does not consider starting point, end point, or occupancy. It is subject to the same conditions as Vehicle Trips, above. This concept is illustrated as follows: if a person drives their POV ten miles, then ten VMT is generated. If two people travel five miles in a POV, then five VMT is generated.

Vehicle Occupancy

Vehicle Occupancy is defined and output as the number of people in a single vehicle on a single trip, so a family of four on a ten-mile trip in a car would be a Vehicle Occupancy of four.

Merging Data from Multiple Files

Despite the effort to include key variables on multiple files (see Chapter 6), an analyst may need to use information from separate files. For example, to study the daily trip patterns of different types of household-owned vehicles, one would need to append the VEHICLE file characteristics into the TRIP file. In these types of circumstances, one needs to merge together two or more of the four files.

File merging can be complicated and confusing, and a mistake can lead to invalid results. However, understanding the structure and relationship of the four files can significantly clarify the process.

ID Numbers – Each unit (i.e., households, persons) in the survey has a unique 10-digit identification number (HOUSEID). Within each household, household members are identified by a two-digit person number (PERSONID) and, similarly, household vehicles are identified by a two digit vehicle number (VEHID). Finally, trips made by an individual are numbered by a trip number (TRIPID) for each travel day trip.

Table 9 shows the most common data linking of any two data files. The linking ID must be common to both the “from” and “to” files. For example, in linking Person file data with Trip file data, the common IDs are HOUSEID and PERSONID.

Table 9: Examples of Link Variables between the 2022 NHTS Data Files

File 1 (From)	File 2 (To)	Linking ID Variables
Household File	Person file	HOUSEID + PERSONID
Household file	Vehicle file	HOUSEID
Household file	Trip file	HOUSEID
Person file	Vehicle file	HOUSEID + PERSONID
Person file	Trip file	HOUSEID + PERSONID
Vehicle file	Trip file	HOUSEID + VEHID



Merging Data Files Example

Below is an example of a scenario in which one would need to merge multiple data files before performing analysis. In this example, the user wants to analyze the impact of occasional telecommuting on the number of daily trips. WKFMHM, the variable indicating occasional telecommuting, is located in the Person file and the trip information is in the Trip file.

Consequently, the two files need to be merged on common identifiers. The variables HOUSEID and PERSONID combined enable one to use the Person file to identify those who occasionally telecommute and those who do not. Using the combined identification number for HOUSEID and PERSONID, one can identify trips taken by that person in the Trip file. In this case, HOUSEID and PERSONID combined is the common identification needed to merge the Trip and Person files.

The illustration below shows how the two files are “linked” by common household and person identifiers. After the merge, each record in the resulting table should correspond with a unique trip, like the Trip file. This is because the Trip file contains information that is more “granular” than the Person file. Thus, the variable, WKRMHM, is repeated for each person trip, as shown below in Figure 5 below.

Figure 5: Merging Data Files

Trip File			Person File		
HOUSEID	PERSONID	SEQ_TRIPID	HOUSEID	PERSONID	WKFMHM
1234000000	01	01	1234000000	01	02
1234000000	01	02	1234000000	02	01
1234000000	02	01	3234000000	01	02
1234000000	02	02			
3234000000	01	01			
3234000000	01	02			
3234000000	01	03			

Combined Trip and Person file

HOUSEID	PERSONID	SEQ_TRIPID	WKFMHM
1234000000	01	01	02
1234000000	01	02	02
1234000000	02	01	01
1234000000	02	02	01
3234000000	01	01	02
3234000000	01	02	02
3234000000	01	03	02

Here is another example. Let us say your goal is to analyze the impact of occasional telecommuting on the number of daily trips. To do this you would need to merge the trip and person data files. Instructions on how to accomplish this follow. Step-by-step Instructions:



1. Read the person table into your software, selecting WKRMMH and the person table identifiers, HOUSEID and PERSONID.
2. Read the trip table into your software, selecting the trip table identifiers, HOUSEID, PERSONID, and SEQ_TRIPID.
3. Instruct the software to join the person table and trip table on common identifiers, HOUSEID and PERSONID, so that every trip record is matched to its corresponding person record.

ID Variables Not Always Sequential

The various ID variables within the file vary and do not always conform to a clear sequence. There are various reasons for this, including:

- When persons or vehicles were reported by a household but later determined not to belong to it, and therefore deleted,
- Trips segments reported as separate and later determined to be part of one trip, and therefore combined,
- Some trips were reported as whole and needed to be divided into multiple segments and were therefore separated.

Sum of Weights / Control Totals

The final weights for Household and Persons were devised to sum to each of their population totals for the nation and their sampling plans' strata, using the 2021 American Community Survey 1-year estimates. The sum of the trip weights represents an estimate of the total number of trips nationwide. These values are helpful to users verify the correctness of their data tabulations. Table 10 presents the unweighted counts and weighted totals for the 2022 NHTS national ABS study.

Table 10: Number of Cases and Sum of Weights for Households, Persons and Trips

	Number of Cases	Sum of Weights
Households	7,893	127,544,707
Persons	16,997	305,560,925
Personally Owned Vehicle	14,641	230,496,688
Trips (Annualized)	31,074	231,715,789,629

Table 11 presents unweighted counts and weighted totals for other travel indicators. These values serve as control totals when a user is tabulating the NHTS data.



Table 11: Number of Cases and Projected Numbers for Key Travel-Related Results

	Number of Cases	Sum of Weights
Workers	7,866	158,009,510
Drivers	13,828	240,744,989
Person Trips	31,074	231,715,789,629
Person Miles of Travel	4,770,028	3,031,433,636,129
Vehicle Trips	27,597	201,218,455,064
Vehicle Miles of Travel	2,124,313	1,743,820,794,277

Non-sampling Errors

In addition to sampling errors that may arise during study design, various non-sampling errors can also occur during the data collection and processing stages of a study. Examples of such errors at the data collection stage could include:

- misunderstanding or misinterpretation of a survey question and its response options,
- failure to accurately report a travel day,
- failure to answer a survey question (item non-response),
- and survey non-response.

Following data collection, errors could occur during processing; this could include errors in response coding or weighting. Additionally, error may arise as a result of coverage issues; undercoverage could occur if new households were not present in the sampling frame. Household-level undercoverage could occur if all members of a household were not accurately reported.

Regardless of how such errors are introduced, they can impact study results by introducing bias and variance. Efforts were taken to minimize introducing any non-sampling errors and the impact they may have had.

Sampling Errors

While the likelihood of sampling error was low for this study, some error may have been introduced when selecting sample for the ABS study. For part of the field period, incorrect response rates were assumed which led to selecting the wrong number of records and subsequently skewed the distribution of completes for the ABS sample across the months. However, this error has been corrected through weighting, though at the cost of reduced precision.

Standard errors for estimates should be calculated using the Taylor series method, which is available in most statistical packages with sample survey analysis techniques. Specification of the analysis weight, the sampling stratum, and the primary sampling unit (HOUSEID) is generally required in the analysis. The variable HHSTRATUMID_2010 should be used as the strata variable for analysis of the national ABS and PFS, while HHSTRATUMID_STATE_2010 should be used when analyzing any of the state add-on studies (described in Appendix 2). Replicate weights used in resampling techniques for estimating standard errors are not available for the Nextgen NHTS.



Appendix 2: Trip Purpose and Mode Changes, 2017 to 2022

Table 12: Trip Purpose Codes, 2022 and 2017 NHTS

2022 Code	2022 Category Name	2017 Code	2017 Category Name
-9	Not ascertained	-9	Not ascertained
		-8	Don't know
		-7	Refused
1	Regular activities at home	1	Reg. home activities
2	Work from home	2	Work from home
3	Work at a non-home location	3	Work
4	Work activity to drop-off/pickup someone/something	4	Work related/Trip
5	Other work-related activities		
6	Attend school as a student	8	Attend school as a student
7	Attend child or adult care	9	Attend child care
		10	Attend adult care
8	Volunteer activities	5	Volunteer activities
9	Change transportation type	7	Change type of transportation
10	Drop off/pick up someone	6	Drop off/pickup someone
11	Health care visit	18	Health care visit
12	Buy meals	13	Buy meals
13	Shop/buy/pick-up or return goods	12	Buy services
		11	Buy goods
14	Other family/personal errands	14	Other general errands
15	Recreational activities	15	Recreational activities
16	Exercise	16	Exercise
17	Visit friends or relatives	17	Visit friends and relatives
18	Religious or other community activities	19	Religious or other community activities
19	Rest or relaxation/vacation	97	Something else
97	Something else		



Table 13: Means of Transportation Codes, 2022 and 2017 NHTS

Summary Mode of Travel	Model of Travel 2022 NHTS	Mode of Travel 2017 NHTS
Private Vehicle	01 Car	03 Car 18 Rental Car (Inc. Zipcar and Car2Go)
	03 SUV/Crossover	04 SUV
	02 Van	05 Van (Minivan)
	04 Pick Up Truck	06 Pick Up Truck
	07 Motorcycle	08 Motorcycle/Moped
Other Vehicle	06 Recreational Vehicle	09 RV (motorhome, ATV, Snowmobile) 07 Golf cart/Segway
	15 Taxicab or limo service	17 Taxi/Limo (including Uber/Lyft)
	16 Other ride-sharing services	
Walk	20 Walk	01 Walk
Bike	18 Bicycle (including bikeshare, ebike, etc.)	02 Bicycle
School Bus	09 School bus	10 School bus
Public Transit	08 Public or commuter bus	11 Public or Commuter Bus
	12 Commuter Rail	15 Amtrak/Commuter Rail
	11 Subway or Elevated Rail	16 Subway/Elevated/Light Rail/Streetcar
	10 Streetcar or trolley car	
Other Transit	22 Ferryboat	20 Boat/Ferry/Water Taxi 13 Private/Charter/Tour/Shuttle Bus 14 City-to-City Bus (Greyhound, Megabus)
	13 Amtrak	15 Amtrak/Commuter Rail
	17 Paratransit/Dial-a-Ride	12 Paratransit/Dial-a-Ride
Other	14 Airplane	19 Airplane
	19 E-scooter	
	21 Other	97 Something else