



WA WTSC Traffic Survey 2023
Sampling Plan
02/03/2023

Research Objectives

The Washington (WA) Traffic Safety Commission (WTSC) has contracted Market Decisions Research (MDR) to conduct a comprehensive survey of Washington adults ages 18 and over to be used to collect data statewide and annually for planning and evaluating traffic safety programs. This document provides information for the sampling methodology of this project.

Recruitment

The sample for the 2023 WA Traffic Survey 2023 will consist of an address-based sampling frame (ABS), stratified into WA TSC's 17 Target Zero Regions. A total of 33,335 addresses will be randomly selected across the state. This will be push-to-web with pen-and-paper (PAPI) follow-up. The ABS sample will be combined with a complementary online panel.

ABS Sample

MDR will use a statewide address-based sample (ABS) frame. This frame will provide near complete population coverage of residences in Washington state. Sample will be purchased from our partner, Marketing Systems Group (MSG), a leading global provider of sampling solutions for the survey research industry. Their ABS sample is based on the United States Postal Service (USPS) Computerized Delivery Sequency file and includes over 158 million U.S. addresses for all known physical residential and business addresses and PO boxes. MSG can also append key demographic information about the household to the sample including age, income, education, and race/ethnicity of the household to allow for targeted oversampling and survey stratification.

MDR will use a stratified sampling plan to provide results that are representative of the state's 17 Target Zero Regions. This stratified sampling plan will draw samples within each of the state's Target Zero Regions to complete a minimum number of surveys within each region. Less populous areas will be oversampled, with a disproportionately large sample assigned to them relative to their population. This will allow us to make statements with a high level of precision at the region level, even in the most rural areas of the state, which would not be allowed under simple proportional random sampling.

In addition, MDR proposes to oversample less populated counties within regions with the goal of completing at least 50 surveys in every county (combining ABS and panel samples). This

would provide a large enough sample size by county to provide reliable data at the county level by Year 2 of data collection (see page 4).

MDR will use additional information on household demographic characteristics provided by our sampling partner to perform an oversample of households with residents more likely to be racial and ethnic minorities. This oversample serves two purposes. First, racial/ethnic minority populations tend to be less likely to respond to surveys in general. This results in lower response rates among these populations and survey results that often underrepresent these groups. Oversampling helps to offset the expected lower response rates among racial/ethnic minority populations. Second, oversampling will help complete sufficient surveys to provide results that are representative of key populations with a known level of precision. This will provide WA TSC with more accurate insights and allow WA TSC to more closely monitor the traffic habits and concerns of individuals who are racial/ethnic minorities.

It is important to note that our goal is to achieve an overall margin of error (MoE) of $\pm 5\%$ for the survey overall and within each region; and a $\pm 10\%$ MoE within each county after Year 2 of data collection.

Panel Sample

MDR will use the online panel sample to target groups who tend to have lower response rates to surveys in general; this includes males under the age of 35, as well as individuals who are racial and ethnic minorities. This will be in addition to a general/oversample sample of adults in Washington State as part of the ABS. The panel sample will be provided by Voxco Audience.

What is an online panel sample?

An online panel includes individuals who have consented to complete surveys on the internet for compensation. Respondents receive credits, for completing surveys, that vary in range and are often redeemable for cash, gift cards, etc.

It is important to note that most online panels are called non-probability panels, as they obtain their members using online ads, snowball sampling, river sampling, and direct enrollments; they also don't sample proportionally from the general population. Given their non-probabilistic nature, advanced weighting techniques should be used to reduce the bias associated with online panels and readers should keep these limitations in mind when making inferences on survey results derived from online panels. While there are several limitations associated with online panel surveys, their speed, cost-effectiveness and improved targeting when compared to telephone or paper surveys are the primary reasons for their increasing popularity.

Voxco Audience

MDR will obtain panel responses through Voxco Audience, which partners with the top online panel providers and has access to a vast network of registered participants. Voxco Audience's quality and security features, as well its coverage in Washington for the target population of the study, are crucial to an overall successful panel response. Voxco Audience has access to a robust set of samples: 60% of proprietary sample in North America. Voxco Audience allows

MDR to tap into an integrated marketplace of top panel providers like Dynata, Disqo, CINT, Lucid, Prodege, TapResearch, etc.

Given the sampling biases associated with individual panels (i.e., some tend to skew older respondents while others are skewed younger), tapping respondents from multiple panels helps to alleviate these issues. Voxco Audience provides high quality sample/panelists at a discount rate while reducing research time from weeks to days or hours. It has built-in quality controls including CAPTCHA, duplicate IP detection, AI validated open ends and third-party integration with Research Defender for bot, VPN and proxy detection. It also provides the option to create custom screening questions for enhanced logic and red herrings.

Sampling Distribution

We will collect two-thirds of completes (or 6,667) via the stratified random ABS sample and one-third of completes (3,333) via our online panel sample. MDR's suggestions for sample distributions by region are presented in the below.

Region	Region Name	Region Population	ABS Mailing	Total Completes	ABS Completes*	Online Panel Completes
Region 1	Clallam, Jefferson Counties	92,391	1,085	400	348	52
Region 2	Grays Harbor, Mason, Pacific Counties	130,993	1,165	400	298	102
Region 3	Cowlitz, Lewis, Wahkiakum Counties	151,584	1,165	400	304	96
Region 4	Thurston County	228,211	2,000	400	253	147
Region 5	Pierce County	696,507	3,000	1000	551	449
Region 6	Clark, Skamania Counties	388,204	2,165	600	363	237
Region 7	King County (North)	1,788,386	3,335	1000	520	480
Region 8	King County (South)	1,788,386	3,335	1000	520	480
Region 9	Kitsap County	217,388	1,500	400	280	120
Region 10	Snohomish County	634,352	3,000	1000	675	325
Region 11	Island, San Juan, Skagit, Whatcom Counties	367,108	2,000	600	428	172
Region 12	Chelan, Douglas, Kittitas, Okanogan Counties	160,515	1,165	400	326	74
Region 13	Klickitat, Yakima Counties	197,198	1,665	400	293	107
Region 14	Benton, Franklin Counties	214,467	1,665	400	285	115
Region 15	Adams, Ferry, Grant, Lincoln Counties	96,409	1,165	400	358	42
Region 16	Pend Oreille, Spokane, Stevens, Whitman Counties	501,489	2,835	800	496	304
Region 17	Asotin, Columbia, Garfield, Walla Wall Counties.	71,702	1,085	400	369	31
Statewide		5,936,904	33,335	10,000	6,667	3,333

** Estimated number of completed assuming 20% response rates.*

MDR’s suggestions for sample distributions by county are presented in the table below.

County	Region	Population 18+	Completes by Region	Targeted Completes by County	Panel Completes	ABS Completes	ABS Mailing
Clallam County	Region 1	63,702	400	276	40	236	1180
Jefferson County	Region 1	28,689	400	124	12	112	560
Grays Harbor County	Region 2	59,434	400	181	55	126	630
Mason County	Region 2	52,189	400	159	27	132	660
Pacific County	Region 2	19,370	400	59	20	39	195
Cowlitz County	Region 3	84,404	400	200	60	140	700
Lewis County	Region 3	63,549	400	150	35	115	575
Wahkiakum County	Region 3	3,631	400	50	1	49	245
Thurston County	Region 4	228,211	400	400	147	253	1265
Pierce County	Region 5	696,507	1000	1000	449	551	2755
Clark County	Region 6	378,392	600	550	233	317	1585
Skamania County	Region 6	9,812	600	50	4	46	230
King County	Region 7-8	1,788,386	2000	2000	960	1040	5200
Kitsap County	Region 9	217,388	400	400	120	280	1400
Snohomish County	Region 10	634,352	1000	1000	325	675	3375
Island County	Region 11	70,346	600	110	31	79	395
San Juan County	Region 11	15,338	600	50	5	45	225
Skagit County	Region 11	100,325	600	157	50	107	535
Whatcom County	Region 11	181,099	600	284	86	198	990
Chelan County	Region 12	60,073	400	150	25	125	625
Douglas County	Region 12	31,614	400	79	13	66	330
Kittitas County	Region 12	36,609	400	91	21	70	350
Okanogan County	Region 12	32,219	400	80	15	65	325
Klickitat County	Region 13	18,127	400	50	12	38	190
Yakima County	Region 13	179,071	400	350	95	255	1275
Benton County	Region 14	149,977	400	280	87	193	965
Franklin County	Region 14	64,490	400	120	28	92	460
Adams County	Region 15	13,114	400	54	1	53	265
Ferry County	Region 15	5,939	400	50	5	45	225
Grant County	Region 15	68,943	400	246	32	214	1070
Lincoln County	Region 15	8,413	400	50	4	46	230
Pend Oreille County	Region 16	10,710	800	50	4	46	230

County	Region	Population 18+	Completes by Region	Targeted Completes by County	Panel Completes	ABS Completes	ABS Mailing
Spokane County	Region 16	413,775	800	632	267	365	1825
Stevens County	Region 16	36,132	800	55	16	39	195
Whitman County	Region 16	40,872	800	63	17	46	230
Asotin County	Region 17	17,676	400	80	9	71	355
Columbia County	Region 17	3,227	400	50	1	49	245
Garfield County	Region 17	1,837	400	50	1	49	245
Walla Walla County	Region 17	48,962	400	220	20	200	1000
Total				10,000	3,333	6,667	33,335

Survey Instrument

The final survey includes 104 questions and was designed by WTSC and Montana State University. The WTSC contracted with the Montana State University (MSU) Center for Health and Safety Culture (CHSC) to develop and validate the traffic safety survey instrument. MDR will work with our partner, Hispanic Marketing Solutions, to create a translated and culturally appropriate version of the English language questionnaire in Spanish. The final survey will be programmed into our Voxco® online survey software in both English and Spanish. To legitimize the look and feel of the questionnaire, MDR will include WTSC's branding, logos, and color schemes into the survey. MDR's research team will test the survey program to ensure it is functioning as intended. This includes running simulations of surveys in Voxco to test for all possible skip logic and survey patterning issues. WTSC staff will be provided a link to test the survey and provided feedback that will be implemented before final approval to launch the survey.

All survey participants will be screened for eligibility at the start of the survey. Screening criteria included age, county, and state of residence. Throughout the survey, participants will be reminded that participation is completely voluntary and confidential, and responses will not be linked to specific users. Users will be able to take the survey in either English or Spanish and can choose to utilize the desktop or mobile-friendly version, based on their device.

Weighting

Panel samples can be carefully selected based on demographic and geographic characteristics to accurately represent the target population or sample frame, even though they are not considered probability sample. Considering the non-probability component of the sampling plan, MDR will use a propensity weighting or quasi-randomization approach based on recommendations from Elliot and Valliant (2017) to combine the probability and non-probability samples and produce one final, weighted dataset. This approach uses known benchmarks from the probability sample (or other trusted sources such as the Census) to estimate a probability of selection for each record in the non-probability sample. Propensity weighting balances the covariate distributions to match those of a probability sample, which is used as a reference. Under this approach, we fit a multivariate statistical model to estimate the inclusion probability of the nonprobability units, and then use the predicted probabilities to derive the nonprobability sample weights or pseudo weights.

Data Quality Checks

To ensure the integrity of the data collected, MDR will develop a comprehensive set of data quality checks to identify potentially suspicious, fraudulent, and invalid surveys that could be completed by surveys bots, professional survey takers, or in bad-faith by legitimate respondents seeking multiple monetary incentives.

Surveys will be considered complete and retained in the dataset for weighting if at least 50% of the questions have been answered. Any partial surveys, invalid surveys and surveys that do not pass the quality checks listed below will be removed from the final dataset for weighting.

Cases will be flagged as invalid if they met one or more of the following criteria:

- **Incomplete surveys.**
 - **Survey “speeders”:** The industry standard for speeder checks would flag anything less than one-third the median length of a survey.
 - **Surveys from duplicate and suspicious IP Addresses.**
 - **Surveys that fail the logic check questions.**
 - *It is important for us to ensure that the answers that we collect are accurate. For these reasons, please select “disagree a lot” from the options below to continue taking the survey.*
 - *Agree a lot*
 - *Agree a little*
 - *Neither Agree nor Disagree*
 - *Disagree a little*
 - *Disagree a lot*
 - *I prefer not to answer.*
- (Note, these logic check questions will be randomly added at 3 different points of the survey.)
- **Surveys that fail the age check question in which we will ask respondent to enter their age both at the beginning and end of the survey.**
 - **Surveys in which respondents provide incoherent open-ended responses.**

Limitations

While extensive precautions will be taken to ensure that the results and findings of this survey are sound and based upon statistically valid methods and analyses, there are some limitations to note. Due to the use of a non-probability sample in the sampling design, it is important to use caution when assessing significant differences between subgroups in the survey. Margin of errors (which are based on the assumptions of a random probability sample and normality) may not be approximated appropriately and variances statistics may be underestimated.