

2025 Youth Tobacco Survey (YTS): AAPOR Transparency Initiative Immediate Disclosure Items

1. Data Collection Strategy. Describe the data collection strategies employed (e.g. surveys, focus groups, content analyses).

The 2025 YTS data were collected using a paper and web-based survey.

2. Who Sponsored the Research and Who Conducted It. Name the sponsor of the research and the party(ies) who conducted it. If the original source of funding is different than the sponsor, this source will also be disclosed.

The YTS is part of the Student Health and Risk Prevention (SHARP) Surveillance System that is sponsored by the Nebraska Department of Health and Human Services and the Nebraska Department of Education and administered by the Bureau of Sociological Research (BOSR) at the University of Nebraska-Lincoln.

3. Measurement Tools/Instruments. Measurement tools include questionnaires with survey questions and response options, show cards, vignettes, or scripts used to guide discussions or interviews. The exact wording and presentation of any measurement tool from which results are reported as well as any preceding contextual information that might reasonably be expected to influence responses to the reported results and instructions to respondents or interviewers should be included. Also included are scripts used to guide discussions and semi-structured interviews and any instructions to researchers, interviewers, moderators, and participants in the research. Content analyses and ethnographic research will provide the scheme or guide used to categorize the data; researchers will also disclose if no formal scheme was used.

<https://bosr.unl.edu/projects/sharp/yts/questionnaires/>

4. Population Under Study. Survey and public opinion research can be conducted with many different populations including, but not limited to, the general public, voters, people working in particular sectors, blog postings, news broadcasts, an elected official's social media feed. Researchers will be specific about the decision rules used to define the population when describing the study population, including location, age, other social or demographic characteristics (e.g., persons who access the internet), time (e.g., immigrants entering the US between 2015 and 2019). Content analyses will also include the unit of analysis (e.g., news article, social media post) and the source of the data (e.g., Twitter, Lexis-Nexis).

The YTS targeted public high school students in grades 9-12 during fall 2025 in Nebraska.

5. Method Used to Generate and Recruit the Sample. The description of the methods of sampling includes the sample design and methods used to contact or recruit research participants or collect units of analysis (content analysis).

- a. Explicitly state whether the sample comes from a frame selected using a probability-based methodology (meaning selecting potential participants with a known non-zero probability from a known frame) or if the sample was selected using non-probability methods (potential participants from opt-in, volunteer, or other sources).

The YTS uses a two-stage, cluster sampling design. The data file used to create the sampling frame is obtained through the Nebraska Department of Education (NDE) website. All school districts must register with NDE, so the list

has complete coverage of all public high schools in Nebraska. First, the Centers for Disease Control (CDC) selects a random sample of public high schools using a probability proportionate to size (i.e., school enrollment) method. Schools are then recruited to participate. Each participating school provides BOSR with a list of classrooms for a required period or subject (e.g., second period or English) that includes all students in grades 9-12 at their school. Classrooms are then randomly sampled using this list and all students in these sampled classrooms are targeted for participation. The CDC provides BOSR with this random classroom selection procedure.

- b. Probability-based sample specification should include a description of the sampling frame(s), list(s), or method(s).
 - i. If a frame, list, or panel is used, the description should include the name of the supplier of the sample or list and nature of the list (e.g., registered voters in the state of Texas in 2018, pre-recruited panel or pool).
 - ii. If a frame, list, or panel is used, the description should include the coverage of the population, including describing any segment of the target population that is not covered by the design.

The YTS uses a two-stage, cluster sampling design. The data file used to create the sampling frame is obtained through the Nebraska Department of Education (NDE) website. All school districts must register with NDE, so the list has complete coverage of all public high schools in Nebraska. First, the Centers for Disease Control (CDC) selects a random sample of public high schools using a probability proportionate to size (i.e., school enrollment) method. Schools are then recruited to participate. Each participating school provides BOSR with a list of classrooms for a required period or subject (e.g., second period or English) that includes all students in grades 9-12 at their school. Classrooms are then randomly sampled using this list and all students in these sampled classrooms are targeted for participation. The CDC provides BOSR with this random classroom selection procedure.

- c. For surveys, focus groups, or other forms of interviews, provide a clear indication of the method(s) by which participants were contacted, selected, recruited, intercepted, or otherwise contacted or encountered, along with any eligibility requirements and/or oversampling.

Schools were first informed of their eligibility via email in March 2025. This email informed them of the opportunity to participate in one or more of the three surveys (i.e., YRBS, Youth Tobacco Survey (YTS), and/or Nebraska Risk and Protective Factor Student Survey (NRPFS)) under the Student Health and Risk Prevention (SHARP) surveillance system during fall 2025. BOSR began calling eligible schools in April 2025. A postcard reminder was mailed to eligible schools in May 2025. BOSR continued recruiting eligible non-responding schools by phone and email throughout the remainder of 2025 until an answer was received from the school. Participating schools were

contacted in fall 2025 to gather any remaining information BOSR needed (e.g., classroom lists, student enrollment numbers, administration dates, etc.) and to provide all materials schools needed for survey administration. Students eligible for participation and their parents were notified by each school using their district's approved processes. BOSR provided schools with parental consent form templates upon request. These templates can be found at <https://bosr.unl.edu/projects/sharp/yts/parent-consent-forms/>. For school districts with formal research request procedures, all communication and information related to the SHARP surveys was sent to the appropriate person at that district using the procedures provided by the district.

d. Describe any use of quotas.

Not applicable

e. Include the geographic location of data collection activities for any in-person research.

This is a school-based survey that is administered with students in classrooms.

f. For content analysis, detail the criteria or decision rules used to include or exclude elements of content and any approaches used to sample content. If a census of the target population of content was used, that will be explicitly stated.

Not applicable

g. Provide details of any strategies used to help gain cooperation (e.g., advance contact, letters and scripts, compensation or incentives, refusal conversion contacts) whether for participation in a survey, group, panel, or for participation in a particular research project. Describe any compensation/incentives provided to research subjects and the method of delivery (debit card, gift card, cash).

Schools were first informed of their eligibility via email in March 2025. This email informed them of the opportunity to participate in one or more of the three surveys (i.e., YRBS, Youth Tobacco Survey (YTS), and/or Nebraska Risk and Protective Factor Student Survey (NRPFS)) under the Student Health and Risk Prevention (SHARP) surveillance system during fall 2025. BOSR began calling eligible schools in April 2025. A postcard reminder was mailed to eligible schools in May 2025. BOSR continued recruiting eligible non-responding schools by phone and email throughout the remainder of 2025 until an answer was received from the school. Participating schools were contacted in fall 2025 to gather any remaining information BOSR needed (e.g., classroom lists, student enrollment numbers, administration dates, etc.) and to provide all materials schools needed for survey administration. Students eligible for participation and their parents were notified by each school using their district's approved processes. BOSR provided schools with parental consent form templates upon request. These templates can be found at <https://bosr.unl.edu/projects/sharp/yts/parent-consent-forms/>. For school districts with formal research request procedures, all communication and information related to the SHARP surveys was sent to the appropriate person at that district using the procedures provided by the district.

6. Method(s) and Mode(s) of Data Collection. Include a description of all mode(s) used to contact participants or collect data or information (e.g., CATI, CAPI, ACASI, IVR, mail, Web for survey; paper and pencil, audio or video recording for qualitative research, etc.) and the language(s) offered or included. For qualitative research such as in-depth interviews and focus groups, also include length of interviews or the focus group session.

The YTS is a self-administered survey that was available as a paper and web-based Qualtrics survey. This is a school-based survey that is administered with students in classrooms. It was administered in English and Spanish.

7. Dates of Data Collection. Disclose the dates of data collection (e.g., data collection from January 15 through March 10 of 2019). If this is a content analysis, include the dates of the content analyzed (e.g., social media posts between January 1 and 10, 2019).

The YTS was administered from August 2025 to December 2025.

8. Sample Sizes (by sampling frame if more than one frame was used) and (if applicable) Discussion of the Precision of the Results.

- Provide sample sizes for each mode of data collection (for surveys include sample sizes for each frame, list, or panel used).
- For probability sample surveys, report estimates of sampling error (often described as “the margin of error”) and discuss whether or not the reported sampling error or statistical analyses have been adjusted for the design effect due to weighting, clustering, or other factors.
- Reports of non-probability sample surveys will only provide measures of precision if they are defined and accompanied by a detailed description of how the underlying model was specified, its assumptions validated, and the measure(s) calculated.
- If content was analyzed using human coders, report the number of coders, whether inter-coder reliability estimates were calculated for any variables, and the resulting estimates.

The design effect due to weighting adjustments is 2.10, which represents the loss in statistical efficiency that results from unequal weights¹. Because the sample was drawn using probability proportionate to size, no sample weighting was necessary resulting in no design effect due to sampling.

Table 1 presents margins of sampling error for some of the most likely sample sizes not taking the design effect from the weighting into account. Exact margins of error for alternative specifications of sample size and reported percentages can be easily computed by using the following formula for the 95% confidence level:

Margin of error = 1.96 * square root (p(1-p)/n)

p = the expected proportion selecting the answer

n = number of responses

¹ The formula used is: $1 + cv^2(w) = \frac{n(\sum_1^n w_i^2)}{(\sum_1^n w_i)^2}$

Table 1. Approximate Margins of Error of Percentages by Selected Sample Size NOT Accounting for Design Effect

	Full Sample*	75% Sample	50% Sample	33.3% Sample	25% Sample	10% Sample
Reported Percentage	n=585	n=438	n=292	n=195	n=146	n=58
50	4.05%	4.68%	5.74%	7.02%	8.11%	12.87%
40 or 60	3.97%	4.59%	5.62%	6.88%	7.95%	12.61%
30 or 70	3.71%	4.29%	5.26%	6.43%	7.43%	11.79%
20 or 80	3.24%	3.75%	4.59%	5.61%	6.49%	10.29%
10 or 90	2.43%	2.81%	3.44%	4.21%	4.87%	7.72%
5 or 95	1.77%	2.04%	2.50%	3.06%	3.54%	5.61%

* 95% confidence interval states that in 95 out of 100 samples drawn using the same sample size and design, the interval will contain the population value

When accounting for design effects due to weighting, the adjusted sampling error will be increased as is shown when comparing Table 1 to Table 2 where the sampling design effect is incorporated:

Margin of error = square root (deff) * 1.96 * square root (p(1-p)/n)

deff = design effects

p = the expected proportion selecting the answer

n = number of responses

Table 2. Approximate Margins of Error of Percentages by Selected Sample Size Accounting for the Design Effect of Weighting

	Full Sample*	75% Sample	50% Sample	33.3% Sample	25% Sample	10% Sample
Reported Percentage	n=585	n=438	n=292	n=195	n=146	n=58
50	5.87%	6.79%	8.31%	10.17%	11.76%	18.66%
40 or 60	5.76%	6.65%	8.15%	9.97%	11.52%	18.28%
30 or 70	5.38%	6.22%	7.62%	9.32%	10.78%	17.10%
20 or 80	4.70%	5.43%	6.65%	8.14%	9.41%	14.92%
10 or 90	3.52%	4.07%	4.99%	6.10%	7.05%	11.19%
5 or 95	2.56%	2.96%	3.62%	4.43%	5.13%	8.13%

* 95% confidence interval states that in 95 out of 100 samples drawn using the same sample size and design, the interval will contain the population value

9. How the Data Were Weighted. Describe how the weights were calculated, including the variables used and the sources of the weighting parameters.

The data were weighted in two ways to account for nonresponse and population characteristics. Sample weights were not needed given that the sample was drawn using the probability proportionate to size method. First, data were weighted to account for school nonresponse. Schools were grouped into two categories based on their enrollment size. A weight was then calculated for each category. Next, the data were weighted to account for student nonresponse by calculating weights for each classroom. As much as possible, classrooms with less than 10 completes were merged with other classrooms with a similar dominate grade level. Lastly, poststratification weights were applied based on gender, race, and grade level in order for the data to more closely resemble the population. The final weight in the dataset is called Pwate.

10. How the Data Were Processed and Procedures to Ensure Data Quality. Describe validity checks, where applicable, including but not limited to whether the researcher added attention checks, logic checks, or excluded respondents who straight-lined or completed the survey under a certain time constraint, any screening of content for evidence that it originated from bots or fabricated profiles, re-contacts to confirm that the interview occurred or to verify respondent's identity or both, and measures to prevent respondents from completing the survey more than once. Any data imputation or other data exclusions or replacement will also be discussed. Researchers will provide information about whether any coding was done by software or human coders (or both); if automated coding was done, name the software and specify the parameters or decision rules that were used.

Any duplicate or blank responses were removed from the dataset and one response was removed for early break-off and for falling outside of the target high school age range. Validity checks were done by looking for students who provided inconsistent, unlikely, or illogical responses.

11. A General Statement Acknowledging Limitations of the Design and Data Collection. All research has limitations and researchers will include a general statement acknowledging the unmeasured error associated with all forms of public opinion research.

All surveys contain errors that cannot be directly measured. These errors may arise when response rates are less than 100%, when weights do not fully account for potential differences in the representation of the respondents to the target population, or when respondents do not understand or cannot answer all of the questions that are asked in the survey. This survey was only offered in English and Spanish and excludes those who do not understand these languages. The YTS is only administered in public high schools, and therefore data collection is limited only to students in public high schools. Students that are in private schools or that are homeschooled are not included in the design.