

# End Commercial Tobacco Campaign: Sidewalks Outdoor Secondhand Smoke Wave 1 Data Analysis Guidance

Issued: April 2022

## Overall Guidelines

- 1) The Local Lead Agency (LLA) is responsible for the dataset and must keep a record of anyone requesting the data and with whom the data are shared (see the sample [Data Request Tracking Form](#) or contact the Tobacco Control Evaluation Center [TCEC] for assistance).
- 2) Local partner programs or others interested in obtaining the End Commercial Tobacco Campaign (ECTC) Sidewalks Outdoor Secondhand Smoke dataset need to contact the LLA to request the data in writing and sign a form agreeing to data sharing and use guidelines.
  - a. See sample [Data Request Form](#) or contact TCEC for assistance.
- 3) The following resources will be provided by TCEC along with the dataset to assist with LLA-level analysis:
  - a. Codebook ("Sidewalks Codebook" tab in the Excel file of each dataset) and [the online survey](#)
  - b. This data analysis guidance document
  - c. Data cleaning documentation ("Data Cleaning Summary" tab in the Excel file of each dataset)
  - d. Training manual for question wording, explanations for each question, and online survey instructions saved as a PDF
- 4) Contact TCEC at [tobaccoeval@ucdavis.edu](mailto:tobaccoeval@ucdavis.edu) with questions about the End Commercial Tobacco Campaign data analysis and reporting.

## Sampling and Weights

- 1) **Sampling method:**
  - a. The California Tobacco Control Program (CTCP) instructed LLAs to purposively sample sidewalks anywhere within each jurisdiction. At a minimum, one sidewalk was observed per park sampled. LLAs that had 20 or more parks were required to select a minimum total of 20 sidewalks. See the [TCEC Website](#) for the complete sampling plan.
- 2) **Weights:**
  - a. Weights are neither needed nor included for LLA level analysis of Sidewalks data due to the purposive sampling design. LLAs should be aware that relying on unweighted data changes the interpretation and representativeness of results.
  - b. Confidence intervals should be generated due to the sample plan (i.e., purposive sampling) and potential measurement error (e.g., data collector accidentally recording the wrong data). These can increase

variability of the estimates and introduce bias. In addition, the longitudinal nature of ECTC requires confidence intervals.

- i. In **Excel**, **SPSS**, or **SAS**: Use a higher 99% confidence level and specify in report or footnote.

**SPSS Example:**

```
ONEWAY var1 BY var2
  /STATISTICS DESCRIPTIVES
  /MISSING ANALYSIS
  /CRITERIA=CILEVEL(0.99) .
```

**SAS Example:**

```
PROC SURVEYFREQ data = <dataset>;
table var1*var2 / row cl alpha=0.01;
RUN;
```

## Data Analysis and Reporting

- 1) LLAs should consider their target audience and how they want to use the results of their data when developing the data analysis plan and framing any reports of results. Different variables, different sub-analyses, and different language may be appropriate for different audiences.
  - a. Sub-groups of Sidewalks may be combined and analyzed to assess Sidewalks with similar demographic or geographic characteristics (e.g., neighborhood socioeconomic status, urban/suburban/rural location) as deemed useful by the LLA.
  - b. Use the example reporting language provided in this document to frame the presentation of findings.
- 2) TCEC will provide guidance on analyzing data using Excel. Example SAS code for recoding variables and answering some evaluation questions are presented in Appendices 1 and 2. Assistance with other statistical programs is available upon request.
- 3) LLA-level analyses will typically aim to assess Tobacco Product Waste (TPW) and Observed Active Smoking (OAS) for various jurisdictions and building use types (e.g., Housing - Multi-family, Restaurant WITH outdoor dining).
  - a. Use the codebook, training manual, and this data analysis guidance document to understand which questions to use for results that may be of interest. Examples of key questions for analysis are included in this document, but LLAs may decide to analyze other variables as well.
- 4) Due to the potential for sensitivity of results for areas near individual businesses, LLAs should use caution when reporting information on individual sidewalks, particularly sidewalks adjacent to businesses.

## Variables

- 1) Variable names are listed in the Codebook ("Sidewalks Codebook" tab in the Excel file of each dataset).

## 2) **Predictor Variables:**

- a. Sidewalk Street Name (SQSN) and Jurisdiction (SQCJ)
  - i. Categorical Variables
- b. Building Use Type (SUSS; e.g., Housing - Multi-family [SUSSA2], Restaurant WITH outdoor dining [SUSSA6], etc.)
  - i. "Select all that apply" Categorical Variable: Survey questions with response options indicating "Select all that apply" have been formatted so that each response option represents a separate variable in the codebook.

## 3) **Outcome Variables:**

- a. The main outcome variables to include in the descriptive analysis are Tobacco Product Waste (TPW) and Observed Active Smoking (OAS). LLAs may also be interested in analyzing No Smoking or Vaping Signage.
  - i. Numeric outcome measures (e.g., TPW) might be skewed due to outliers. LLAs should conduct data checks of skew and/or kurtosis for the presence of outliers. If skew/kurtosis is high ( $\pm 2$ ), it is suggested that median values (instead of means) are reported.
  - ii. LLAs may wish to recode these variables into a categorical variable prior to analysis (see Appendix 1).
- b. Tobacco Product Waste (TPW):
  - i. *Survey Question*: How many pieces of TPW are in the observation area?
  - ii. *Numeric Measure*: Open-ended response scale assessing count (ranging from 0 to 99)
  - iii. *Categorical Recode*: Recode the numeric measure of TPW to binary/categorical variable where:
    - a. 0 = 0: no TPW
    - b. 1 thru 99 = 1: yes, TPW observed
- c. Observed Active Smoking (OAS):
  - i. *Survey Question*: During your observation, how many times did you see or smell tobacco or marijuana smoke or vapor?
  - ii. *Numeric Measure*: Response options ranging from 0 to 5 or more
  - iii. *Categorical Recode*: Recode the numeric measure of OAS to binary/categorical variable where:
    - a. 0 = 0: no Observed Active Smoking
    - b. 1 thru 5 = 1: yes, Observed Active Smoking
- d. "No Smoking/No Vaping" Signage:
  - i. *Survey Question*: Are there "Designated smoking/vaping area" signs?
  - ii. *Categorical Measure*: Response options were yes/no.

## Interpretation and Write-Up Guidance

- 1) For each analysis, present an estimate (estimated average or estimated percentage) and corresponding confidence interval.
- 2) Results of analyses that assess TPW should be reported per 100 square feet.
- 3) If the LLA is assessing TPW by jurisdiction, then 99% confidence intervals should be calculated, and data can be reported as:
  - a. **In 2022 across sidewalks surveyed in X Jurisdiction, approximately 3.0 (99% CI: 2.9, 4.1) pieces of tobacco product waste per 100 square feet were observed.**
- 4) If the LLA is assessing the percentage of sidewalks that had any TPW (> 0) across all jurisdictions, 99% confidence intervals should be calculated, and data can be reported as:
  - a. **In 2022 across sidewalks surveyed in X Jurisdiction, tobacco product waste was observed on approximately 20.0% (99% CI: 19.1%, 20.9%) of sidewalks.**

## Example Evaluation Questions with Data Analysis Suggestions

Predictor Variable(s)	Outcome Variable(s)	Outcome Variable Type	Reported Estimates
<b>What is the average amount of TPW per jurisdiction?</b>			
Jurisdiction (SQCJ)	TPW (STPW)	Numeric	Average/mean <sup>b</sup> and confidence interval of TPW for each jurisdiction
<b>What percentage of sidewalks had any TPW?</b>			
None	TPW (STPW) Recoded as No/Yes TPW (STPW_cat) <sup>a</sup>	Categorical	Percentage (%) and confidence interval estimate of sidewalks with any TPW (1 or Yes TPW) across all jurisdictions
<b>What percentage of sidewalks had any TPW per jurisdiction?</b>			
Jurisdiction (SQCJ)	TPW (STPW) Recoded as No/Yes TPW (STPW_cat) <sup>a</sup>	Categorical	Percentage (%) and confidence interval estimate of sidewalks with TPW (1 or Yes TPW) for each separate jurisdiction
<b>Which building use types adjacent to sidewalks (e.g., offices, bars) had the most TPW on average per jurisdiction?</b>			
Each Building Use Type (SUSSA1, SUSSA2, etc.) <sup>a</sup> and Jurisdiction (SQCJ)	TPW (STPW)	Numeric	Average/mean <sup>b</sup> and confidence interval of TPW for each building use type variable for each separate jurisdiction

Predictor Variable(s)	Outcome Variable(s)	Outcome Variable Type	Reported Estimates
<b>What was the average amount of TPW on sidewalks adjacent too restaurants with outdoor dining per jurisdiction?</b>			
Restaurant WITH outdoor dining Building Use Type (SUSSA6) and Jurisdiction (SQCJ)	TPW (STPW)	Numeric	Average/mean <sup>b</sup> and confidence interval of TPW on sidewalks adjacent too restaurants with outdoor dining for each separate jurisdiction
<b>What percentage of sidewalks had any Observed Active Smoking per jurisdiction?</b>			
Jurisdiction (SQCJ)	Observed Active Smoking (SAM) Recoded as No/Yes OAS (SAM_cat) <sup>a</sup>	Categorical	Percentage (%) and confidence interval estimate of sidewalks with any Observed Active Smoking (1 or Yes OAS) for each separate jurisdiction
<b>Which building use types had the most Observed Active Smoking per jurisdiction?</b>			
Each Building Use Type (SUSSA1, SUSSA2, etc.) <sup>a</sup> and Jurisdiction (SQCJ)	Observed Active Smoking (SAM)	Numeric	Average/mean <sup>b</sup> and confidence interval of Observed Active Smoking for each building use type variable for each separate jurisdiction
<b>What percentage of sidewalks had any No Smoking or Vaping Signage per jurisdiction?</b>			
Jurisdiction (SQCJ)	No Smoking or Vaping Signage (SADS) Recoded as No/Yes Signage (SADS_cat) <sup>a</sup>	Categorical	Percentage (%) and confidence interval estimate of sidewalks with any No Smoking or Vaping Signage (1 or Yes Signage) for each separate jurisdiction
<b>Which building use types were most likely to have No Smoking or Vaping Signage per jurisdiction?</b>			
Each Building Use Type (SUSSA1, SUSSA2, etc.) <sup>a</sup> and Jurisdiction (SQCJ)	No Smoking or Vaping Signage (SADS) Recoded as No/Yes Signage (SADS_cat) <sup>a</sup>	Categorical	Percentage (%) and confidence interval estimate of No Smoking or Vaping Signage (1 or Yes Signage) for each building use type variable for each separate jurisdiction

Note. Example dataset variable names are listed in blue.

a. See Appendix 1.

b. If skew/kurtosis is high ( $\pm 2$ ), median values (instead of mean values) may be reported.

## Appendix 1

### Detailed Examples of SAS Variable Re-Coding

#### Tobacco Product Waste (TPW):

##### Categorical Recode:

###### Option 1:

```
if STPW=0 then STPW_cat='No TPW';  
else if STPW in (1:99) then STPW_cat='Yes TPW';
```

###### Option 2:

```
if STPW=0 then STPW_cat=0;  
else if STPW in (1:99) then STPW_cat=1;  
proc format;  
value STPWformat  
0 = ' 0: No TPW '  
1 = ' 1: Yes TPW ';
```

#### Observed Active Smoking (OAS):

##### Categorical Recode:

###### Option 1:

```
if SAM=0 then SAM_cat='No OAS';  
else if SAM in (1:5) then SAM_cat='Yes OAS';
```

###### Option 2:

```
if SAM=0 then SAM_cat=0;  
else if SAM in (1:5) then SAM_cat=1;  
proc format;  
value SAMformat  
0 = ' 0: No OAS '  
1 = ' 1: Yes OAS ';
```

#### No Smoking or Vaping Signage:

##### Categorical Recode:

###### Option 1:

```
if SADS =0 then SADS_cat='No Signage';  
else if SADS in (1:7) then SADS_cat='Yes Signage';
```

###### Option 2:

```
if SADS =0 then SADS_cat=0;  
else if SADS in (1:7) then SADS_cat=1;  
proc format;  
value SADSformat  
0 = ' 0: No Signage '  
1 = ' 1: Yes Signage ';
```

## Appendix 2

### Detailed Examples of SAS Analysis Code for Obtaining Overall Estimate and Corresponding Confidence Interval

**Example 1:** Unweighted estimated average/mean value and corresponding confidence interval for numeric outcome measures

**Use:** PROC SURVEYMEANS

**SAS Code:** TPW across all Sidewalks:

```
DATA <dataset>; set <original dataset>;
RUN;
PROC SURVEYMEANS data = <dataset name> alpha=0.01
plots=none;
var STPW;
RUN;
```

**SAS Results Output:**

Statistics					
Variable	N	Mean	Std Error of Mean	99% CL for Mean	
STPW	45	3.466667	0.437624	2.28845990	4.64487344

**Interpretation:**

In 2022 across sidewalks surveyed in Jurisdiction X, an estimated average of 3.5 (99% CI: 2.3, 4.6) pieces of tobacco product waste per 100 square feet was observed.

**Example 2:** Unweighted estimated percentage and corresponding confidence interval for categorical outcome measures

**Use:** PROC SURVEYFREQ

**SAS Code:** Observed Active Smoking (OAS) by Jurisdiction:

```
DATA <dataset>; set <original dataset>;
if SAM=0 then SAM_cat='No OAS';
else if SAM in (1:5) then SAM_cat='Yes OAS';
RUN;
PROC SURVEYFREQ data = <dataset>;
table SQCJ*SAM_cat/ row cl alpha=0.01;
RUN;
```

Table of Jurisdiction by SAM_cat										
SQCJ	SAM_cat	Frequency	Percent	Std Err of Percent	99% Confidence Limits for Percent		Row Percent	Std Err of Row Percent	99% Confidence Limits for Row Percent	
1	No OAS	9	19.5652	5.9137	3.6599	35.4706	36.0000	9.7061	9.8947	62.1053
	Yes OA	16	34.7826	7.1000	15.6866	53.8786	64.0000	9.7061	37.8947	90.1053
	Total	25	54.3478	7.4253	34.3768	74.3189	100.0000			
2	No OAS	10	21.7391	6.1488	5.2015	38.2767	47.6190	11.0189	17.9827	77.2554

Table of Jurisdiction by SAM_cat											
				Std Err of Percent	99% Confidence Limits for Percent		Row Percent	Std Err of Row Percent	99% Confidence Limits for Row Percent		
SQCJ	SAM_cat	Frequency	Percent								
	Yes OA	11	23.9130	6.3587	6.8109	41.0152	52.3810	11.0189	22.7446	82.0173	
	Total	21	45.6522	7.4253	25.6811	65.6232	100.0000				

### Interpretation:

In 2022 across sidewalks surveyed, active smoking was observed on approximately 64.0% (99% CI: 37.9%, 90.1%) of sidewalks in Jurisdiction 1 and 52.4% (99% CI: 22.7%, 82.0%) of sidewalks in Jurisdiction 2.

**Example 3:** Unweighted estimated average/mean value and corresponding confidence interval for numeric outcome measures for multiple jurisdictions within the same county

**Use: PROC SURVEYMEANS**

**SAS Code: TPW for Multi-Family Housing Building Use Type:**

```
DATA <dataset>; set <original dataset>;
RUN;
PROC SURVEYMEANS data = <dataset> alpha=0.01 plots=none;
domain SUSSA2;
var STPW;
RUN;
```

### SAS Results Output:

Statistics for SUSSA2 Domains						
SUSSA2	Variable	N	Mean	Std Error of Mean	99% CL for Mean	
1	STPW	17	3.591176	0.737048	1.43842048	5.74393246

### Interpretation:

In 2022 across sidewalks surveyed in Jurisdiction X, an average of 3.6 (99% CI: 1.4, 5.7) pieces of tobacco product waste per 100 square feet was observed on sidewalks near Multi-Family Housing.