End Commercial Tobacco Campaign:
Retail Observation Survey
2022 Data Analysis Guidance

Issued: April 2022

Overall Guidelines

1) The Local Lead Agency (LLA) is responsible for the confidentiality and security of the dataset. LLAs must keep a record of anyone requesting the data and with whom the data are shared. (See TCEC website or contact TCEC for assistance).
2) Local partner programs or others interested in obtaining the End Commercial Tobacco Campaign Retail Observation survey dataset need to contact the LLA to request the data in writing and sign the form agreeing to these guidelines.
    a. A sample form is available on TCEC website.
3) After the signed form is received, the following resources will be provide along with the dataset:
    a. Codebook ("Codebook" tab in the Excel file of each dataset) and the online survey
    b. Data cleaning documentation ("Data Cleaning Summary" tab in the Excel file of each dataset)
    c. Training manual for question wording and explanations for each question
    d. This data analysis guidance document
4) Contact TCEC at tobaccoeval@ucdavis.edu with questions about the End Commercial Tobacco Campaign data analysis and reporting for any and all the waves of data.

Sampling and Weights

1) **Sampling method**:
    a. The California Tobacco Control Program (CTCP) drew the minimum required stores to be surveyed for each LLA.
    b. The ECTC retail observational survey attempted to survey all publicly accessible stores that sell tobacco. Data collectors were instructed not to survey retailers that had active smoking (e.g., hookah bars, cigar bars), retailers with substantial cost-consideration and membership fees (e.g., airport retailers, wholesale clubs), or retailers located in areas with restricted access (e.g., military base). More details about the sampling plan can be found on TCEC website.
2) **Weights:**

a. CTCP will calculate and provide survey weights to account for non-response 1-2 months after the completion of data collection by all LLAs.
b. Weighted analyses are recommended, as they allow for results to more accurately represent the LLA community surveyed.

If a census of stores were surveyed and the response rate is 100%, unweighted analyses can be performed. Completing a census means the LLA was able to observe and collect data at all stores.

i. Unweighted Analyses in **Excel**, **SPSS**, or **SAS**: Use a higher 99% confidence level and specify in report or footnote.
ii. Weighted Analyses: Use the default 95% confidence levels.
1. **In SPSS**, account for the calculated sampling weight using the following point and click instructions before conducting descriptive analyses: Data → Weight Cases → Weight cases by.

Example:

```
WEIGHT BY <var name>.
ONEWAY var1 BY var2
  /STATISTICS DESCRIPTIVES
  /MISSING ANALYSIS
  /CRITERIA=CILEVEL(0.95).
```

2. **In SAS**, use "**PROC SURVEYFREQ**" for categorical variables or "**PROC SURVEYMEANS**" for continuous variables and specify the weighting variable.

Example:

```
PROC SURVEYFREQ data = <dataset>;
table var1*var2 / row cl;
weight <var name>;
RUN;
```

If the LLA was able to complete a census of retail stores, they may use **Excel** to calculate a 95% confidence interval.

Data Analysis and Reporting

1) LLAs should consider their target audience and how they want to use 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.
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) Use the codebook, training manual, and this data analysis guidance document to understand which questions to use for key results.
4) Use the example reporting language provided in this document to frame the presentation of findings.

5) All data must be reported in aggregate. Do NOT report individual store results, store names, or store addresses. E.g., do NOT display individual store results on a map. Do NOT report results for areas with fewer than 5 retail stores.
6) Jurisdiction-level analyses are recommended to assess the availability of each major type of tobacco product, menthol cigarettes, and other flavored products; the average price for the cheapest pack of cigarettes; proportion of the stores where the cheapest pack of little cigars/cigarillos (LCCs) are sold for under $2.; proportion of stores with LCCs sold in packs less than 20; Proportion of stores with price promotions for any tobacco products.

Coding for variables with "Select all that apply"

Survey questions with response options indicating "Select all that apply" need to be recoded. Using vaping or other tobacco product availability as an example, below is the survey question and corresponding variables:

Which products are sold here? (Select all that apply)

1. Vaping products (e.g., e-cigarettes, vapor devices or e-liquids) if the store sells any products used to vape nicotine. Do not include non-nicotine devices like CBD or aromatherapy pens.
2. Any other tobacco (e.g., cigarettes, cigar products, chew) if the store sells at least one tobacco product other than e-cigarettes, vapor devices or e-liquids. This category includes cigarettes, chewing tobacco, snus, cigarillos/little cigars, large cigars, blunt wraps, and hookah.
3. None of the above if no products from these categories are sold. In this case, you will be taken to the end of the survey and will not be asked to answer any further questions.

In the Retail Observation Survey dataset these options appear as variables:

RSPSA1 for Vaping products
RSPSA2 for Any other tobacco
RSPSA3 for none of these

If the data collector selected "Yes" for an option, then "1" appears in the dataset. If the data collector selected "No" or if the question did not appear to the data collector because of survey logic, the cell in the data set would be blank.

Note: Please create a new variable to analyze the proportion of retail stores selling ANY Tobacco products: If data collector selected "Yes "for either RSPSA1 or RSPSA2, then create a new variable (e.g., Any_productsold) = "Yes".

For example SAS code, please refer to appendix 1.

Example Evaluation Questions with Data Analysis Suggestions

1. What proportion of tobacco retailers sold each major type of tobacco product per jurisdiction?				
Survey Questions	**Outcome Variable(s)**	**Numerator & Denominator**	**Reported Estimates**	**Notes**
1.1 Individual products Sold: For vaping products, use RSPS: Which products are sold here? (Select all that apply) For other tobacco products, use RSOT: Which other tobacco products are sold here? (Select all that apply)	Type of tobacco products RSPSA1=1 Menthol cigarettes RSOTA1=1 Non-Menthol cigarettes RSOTA2=1 Little cigars/cigarillos/cigar wrap RSOTA3=1 Premium/Large cigars RSOTA4=1 Smokeless tobacco RSOTA5=1 Hookah/Shisha RSOTA6=1 Roll-your-own tobacco RSOTA7=1	**Numerator:** Number of stores that sold each product **Denominator**: total number of retailers that sold any of tobacco products (RSPSA1= 1 or RSPSA2= 1)	Percentage (%) and confidence interval estimate of tobacco retailers selling each major type of tobacco products	Analysts may run separate analysis for each type of product as needed. LLAs should do data checks on the Jurisdiction (Community) to ensure that the variable information is correct

2. What proportion of retailers sold any kind of flavored tobacco (menthol cigarettes and/or other flavored products) per jurisdiction?				
Survey Questions	**Outcome Variable(s)**	**Numerator & Denominator**	**Reported Estimates**	**Notes**
2.1 Menthol Cigarettes Use RSOT: Which other tobacco products are sold here? (Select all that apply)	Menthol cigarettes sold RSOTA1 (If RSOTA1=1)	**Numerator:** Number of stores that sold menthol cigarettes **Denominator**: total number of retailers that sold any of tobacco products	Percentage (%) and confidence interval estimate of tobacco retailers selling menthol cigarettes	
2.2 Other Favored products Use RSFS: Which flavors of non-cigarette tobacco products are sold here? (Select all that apply)	Flavored other tobacco products sold RSFSA1=1 or RSFSA2=1 or RSFSA3=1 or RSFSA4=1	**Numerator:** Number of stores that sold flavored non-cigarette tobacco products **Denominator**: total number of retailers that sold any tobacco products	Percentage (%) and confidence interval estimate of tobacco retailers selling flavored non-cigarette tobacco products	IF stores ONLY sell tobacco products with another concept flavor (RSFSA1=0 and RSFSA2=0 and RSFSA3=0 and RSFSA4=1) Please check write-in responses for Q27 to determine if there is a concept flavor sold. Please consult with TCEC with more questions

2. What proportion of retailers sold any kind of flavored tobacco (menthol cigarettes and/or other flavored products) per jurisdiction?				
Survey Questions	**Outcome Variable(s)**	**Numerator & Denominator**	**Reported Estimates**	**Notes**
2.3 Menthol cigarette or other Favored products Use RSOT: Which other tobacco products are sold here? (Select all that apply) AND Use RSFS: Which tobacco products are available in explicit flavors (e.g., menthol, mint, fruit, sweet, or liquor)?	Flavored Tobacco products sold RSOTA1=1 or RSFSA1=1 or RSFSA2=1 or RSFSA3=1 or RSFSA4=1	**Numerator:** Number of stores that sold ANY flavored tobacco product (menthol cigarettes or other flavored tobacco products) **Denominator**: total number of retailers that sold any tobacco products	Percentage (%) and confidence interval estimate of tobacco retailers selling any flavored tobacco product	IF stores ONLY sell tobacco products with another concept flavor (RSFSA1=0 and RSFSA2=0 and RSFSA3=0 and RSFSA4=1) Please check write-in responses for Q27 to determine if there is a concept flavor sold. Please consult with TCEC with more questions

3. What is the average price for the cheapest pack of cigarettes in retail settings per jurisdiction ?				
Survey Questions	**Outcome Variable(s)**	**Numerator & Denominator**	**Reported Estimates**	**Notes**
RSCP. Cheapest single pack of cigarettes: $_. _ _ (Enter X.XX) RSTI. Is sales tax included in the cheapest cigarette price?	Cheapest single pack of cigarettes IF RSTI= 1, then create a new variable RSTI_1 to get the non-taxed price of cheapest pack of cigarettes	**Numerator:** total sum of the cheapest price for a single pack of cigarettes across all stores that sold cigarettes **Denominator**: total number of retailers that sold cigarettes (RSOTA1=1 or RSOTA2=1)	Average/mean (confidence interval) of the price of the cheapest pack of cigarettes sold in stores	If sales tax was included into the price, be sure to exclude the tax from the price first before calculating the average price Please refer this website to determine local sales tax

4. What proportion of stores sold packs of little cigars/cigarillos (LCCs) for under $2.00 per jurisdiction?				
Survey Questions	**Outcome Variable(s)**	**Numerator & Denominator**	**Reported Estimates**	**Notes**
RSPC. What is the price of the CHEAPEST pack of little cigars/cigarillos?	Cheapest pack of LCCs Survey question has several categories, for the price under $2.00, create a new variable (RSPC=1 or RSPC=2)	**Numerator:** Number of stores that sold LCCs for under $2. **Denominator:** total number of stores that sold LCCs (RSOTA3=1)	Percentage (%) and confidence interval estimate of the stores where cheapest pack of little cigars/cigarillos	LLAs can select another cutoff of interest, instead of under $2, based on the available response options but need to create a new variable to make it dichotomous.

5. What proportion of stores with LCCs sold LCCs in packs less than 20?				
Survey Questions	**Outcome Variable(s)**	**Numerator & Denominator**	**Reported Estimates**	**Notes**
RSCC. If little cigars/cigarillos are sold here, what is the SMALLEST pack of little cigars/cigarillos?	LCCs sold in packs less than 20 RSCC=1 or RSCC=2 or RSCC=3 or RSCC=4)	**Numerator:** Number of stores that sold LCCs in packs less than 20. **Denominator:** total number of stores that sold LCCs	Percentage (%) and confidence interval estimate of the stores with LCCs sold in packs less than 20	LLAs can select another cutoff of interest based on the available response options

6. What proportion of stores displayed price promotions for any tobacco products (exterior or interior)?				
Survey Questions	**Outcome Variable(s)**	**Numerator & Denominator**	**Reported Estimates**	**Notes**
RAPP. OUTSIDE the store, are there any price promotions for the following? (Select all that apply) RSPP. INSIDE the store, are there any price promotions for the following? (Select all that apply)	Retail Ads Price Retail Sales Price Promotion (RAPP=1 OR RSPP=1)	**Numerator:** Number of stores with price promotions at stores (inside or outside) **Denominator:** total number of tobacco retailers (RQAC=1 or RQAC=6)	Percentage (%) and confidence interval estimate of the stores with any exterior or interior price promotion	

Interpretation and Write-Up Guidance

1) The inclusion criteria should be clearly reported.
   Example:
   i. **Weighted Analysis:** [Final sample size] publicly accessible retailers licensed to sell tobacco in X jurisdiction were surveyed.
   ii. **Unweighted:** [Final sample size] publicly and safely accessible retailers licensed to sell tobacco in X jurisdiction were surveyed.
2) The Confidence intervals should be generated regardless of whether a census was attempted. Confidence intervals are needed due to potential non-response (e.g., data collector unable to survey a store), sampling frame errors (e.g., missing retailer that sells tobacco products not in the sampling frame), and measurement error (e.g., data collector accidentally recording the wrong data). These can increase the variability of the estimates and introduce bias. In addition, the longitudinal nature of ECTC requires confidence intervals E.g., data could be reported as: In 2022 in X LLA, the estimated average price for the cheapest pack of cigarettes was $5.45, 95% CI (5.37, 5.53).

Appendix 1
Example SAS Code to create a new variable:

Create a new variable ANY Tobacco products:

Which products are sold here? (Select all that apply)

1. Vaping products (e.g., e-cigarettes, vapor devices or e-liquids) if the store sells any products used to vape nicotine. Do not include non-nicotine devices like CBD or aromatherapy pens.
2. Any other tobacco (e.g., cigarettes, cigar products, chew) if the store sells at least one tobacco product other than e-cigarettes, vapor devices or e-liquids. This category includes cigarettes, chewing tobacco, snus, cigarillos/little cigars, large cigars, blunt wraps, and hookah.
3. None of the above if no products from these categories are sold. In this case, you will be taken to the end of the survey and will not be asked to answer any further questions.

In the Retail Observation Survey dataset these options appear as variables:

RSPSA1 for Vaping products
RSPSA2 for Any other tobacco
RSPSA3 for none of these

SAS Code:

```
if RSPSA1= 1 or RSPSA2= 1 then Any_productsold=1;
else Any_productsold=0;

proc format;
value Any_productsoldformat
      0 = ' 0: Tobacco product Not Sold '
      1 = ' 1: Any Tobacco Product Sold';
```

Detailed Examples of SAS Cross-Tab Analysis Code

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

Use: PROC SURVEYMEANS

SAS Code: cheapest pack of cigarettes by jurisdiction:

```
DATA <dataset>; set <original dataset>;
RUN;

PROC SURVEYMEANS data = <dataset name> plots=none;
domain Jurisdiction;
var RCCP;
weight <var name>;
RUN;
```

SAS Results Output:

Statistics for Jurisdiction Domains						
Jurisdiction	Variable	N	Mean	Std Error of Mean	95% CL for Mean	
X	RCCP	15	5.601345	0.744871	3.06278934	6.14278435
Y	RCCP	18	6.504765	0.572567	4.31468912	7.69145927

Interpretation:

In 2022, in Jurisdiction X, the average price for the cheapest pack of cigarettes was estimated to be $5.60 (95% CI: 3.06, 6.14). In Jurisdiction Y, the average price was estimated to be $6.50 (95% CI: 4.31, 7.69).

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

Use: PROC SURVEYFREQ

SAS Code: Types of tobacco sold (In this example-Vaping products sold) by jurisdiction

```
DATA <dataset>; set <original dataset>;
RUN;

PROC FORMAT;
value Q18format
      0 = '0: No vaping products sold'
      1 = '1: Yes, vaping products sold';
RUN;

PROC SURVEYFREQ data = <dataset>;
weight <var name>;
table Jurisdiction*RSPSA1 / row cl;
format RSPS_1 Q18format.;
RUN;
```

SAS Results Output:

Table of Jurisdiction by RSPSA1												
Jurisdiction	RSPSA1	Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent	95% Confidence Limits for Percent		Row Percent	Std Err of Row Percent	95% Confidence Limits for Row Percent	
X	0: No Vaping products Sold	12	10.74000	3.75132	10.8201	7.4414	5.0000	26.2525	39.8800	4.7362	43.0000	89.4022
	1: Yes, vaping products Sold	14	14.60000	3.94534	14.7894	7.9391	6.0000	31.2541	60.1200	4.7362	48.3601	100.0000
	Total	26	25.40000	5.10123	25.6095	10.0624	6.7413	46.4777	100.0000			
Y	0: No Vaping products Sold	12	10.74000	3.76002	10.8201	7.4414	5.0000	26.2525	46.9406	3.6123	68.0000	97.9833
	1: Yes, Vaping products Sold	13	12.14000	3.82105	12.2305	7.6080	8.0000	28.0085	53.0594	3.6123	45.0167	100.0000
	Total	25	22.88000	5.07324	23.0506	9.8907	6.5385	30.5626	100.0000			

Interpretation:
In 2022, in Jurisdiction X, approximately 60.12% (95% CI: 48.36%, 100.00%) of tobacco retailers sold vaping products. In Jurisdiction Y, approximately 53.06% (95% CI: 45.02%, 100.00%) of tobacco retailers sold vaping products.

Example 3: Unweighted estimated average/mean value and corresponding confidence interval for numeric outcome measures
Use: PROC SURVEYMEANS
SAS Code: Cheapest pack of cigarettes by jurisdiction:

```
DATA <dataset>; set <original dataset>;
RUN;

PROC SURVEYMEANS data = <dataset name> alpha=0.01
plots=none;
domain Jurisdiction;
var RSPC;
RUN;
```

SAS Results Output:

Statistics for Jurisdiction Domains						
Jurisdiction	Variable	N	Mean	Std Error of Mean	99% CL for Mean	
X	RSPC	13	5.461538	0.492632	3.07292897	5.85014796
Y	RSPC	10	5.200000	1.070429	4.18272185	6.21727815

Interpretation:
In 2022, the average price for the cheapest pack of cigarette was estimated to be $5.46 (99% CI: 3.07, 5.85) in Jurisdiction X. In Jurisdiction Y, the average price was estimated to be $5.20 (99% CI: 4.18, 6.21).