



G R E A T P L A I N S  
IDeA | Clinical and  
Translational Research

**Tutorial for working with  
BRFSS online data tools and downloading data to run  
analysis in statistical software  
April 2020**

# INTRODUCTION

- **This tutorial will cover topics that relate to use of data collected through Behavioral Risk Factor Surveillance system.**
- **BRFFS is the system of health-related telephone surveys that collect state data about non-institutionalized, adult U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive services.**
- **The states use a standardized core questionnaire, [optional modules](#), and state-added questions administering the surveys continuously throughout the year.**

# INTRODUCTION

- **The survey is conducted using Random Digit Dialing (RDD) techniques on both landlines and cell phones using two samples.**
- **Disproportionate stratified sampling (DSS) is used for landline sample i.e. dividing telephone numbers into two density-based strata. The landline sampling ratio of high to medium density is 1:1.5.**
- **The cellular telephone sample is randomly generated from a sampling frame of confirmed cellular area code and prefix combinations**

# INTRODUCTION

- For landline interviews, individual respondents are randomly selected from all adults, living in a household.
- Cellular interviews are treated as one-person households.
- BRFSS [data](#) and [comprehensive documentation](#) are freely available to download. There is a detailed section in this tutorial that will walk you through this process step-by-step.
- Detailed methodology of the survey has been explained in the [data user guide](#).



# TUTORIAL CONTENTS

Online Analysis Tool	Slide Numbers
Prevalence and Trends Data Tool	7
Chronic Disease Indicators (CDI) Tool	20
Web Enabled Analysis Tool (WEAT) for Cross Tabulation and Logistic Regression	27

Using data with software	Slide Numbers
Downloading BRFSS Data	55
Getting started with SAS	59
Getting started with SPSS	68

# Prevalence And Trends Data Tool

# PREVALENCE AND TRENDS DATA TOOL

- Quick tool to explore prevalence using a wide range of variables
- Produces charts and maps for individual states or the nation by health topics.
- **Particularly useful to produce trends over multiple years.**
- Trends are produced as line graphs and single year estimates are produced as maps and bar charts.

CDC - BRFSS Prevalence Data & Tools | x +

cdc.gov/brfss/data\_tools.htm

CDC Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

Behavioral Risk Factor Surveillance System

CDC > BRFSS


BRFSS

- About BRFSS +
- Archived
- Prevalence Data and Data Analysis Tools**
- Survey Data and Documentation +
- GIS Maps Data +
- SMART: City and County Survey Data +
- Statistical Briefs
- Questionnaires
- Publications and Resources +
- State Information +
- Fact Sheets

Get Email Updates

## Prevalence Data & Data Analysis Tools

Find city and county data collected through the Selected Metropolitan/Micropolitan Area Risk Trends (SMART) project, the Web Enabled Analysis Tool (WEAT), interactive maps, and other resources provided through BRFSS.



### Prevalence and Trends Data

Using the Prevalence and Trends Data Tools, users may produce charts for individual states or the nation by health topic. Users may select for request multiple year data. The Prevalence and Trends Data Tools will produce line graphs for multiple years and bar charts for single years for each indicator.

### Web Enabled Analysis Tool (WEAT)

The Web Enabled Analysis Tool (WEAT) permits users to create custom crosstabulation of indicators within selected states. Variables may be included to within each category of control. Data may be used to create logistic data. Users are prompted to

### SMART: City and County Data

Selected Metropolitan/Micropolitan Area Risk Trends (SMART) is an ongoing project that uses BRFSS data to produce some local area estimates. Counties and Metropolitan/Micropolitan Areas (MMSAs) were selected for SMART if there were 500 or more respondents BRFSS combined landline and cell phone data for any year.

**Go to the BRFSS data website here:**  
[https://www.cdc.gov/brfss/data\\_tools.htm](https://www.cdc.gov/brfss/data_tools.htm)

**Click "Prevalence and Trends Data"**

BRFSS Prevalence & Trends Data: x +

cdc.gov/brfss/brfssprevalence/index.html

CDC Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives. Protecting People™

BRFSS Prevalence & Trends Data

Search


**Starts with selection of geographical location and produces trends or single year prevalence**

**Starts with selection of a health topic and produces only single year prevalence**

Explore BRFSS Data By Location

Explore BRFSS Prevalence & Trends data for all questions for one location.

States & Territories



[Metropolitan/Micropolitan Statistical Areas \(MMSAs\)](#)

Explore BRFSS Data By Topic

Explore BRFSS Prevalence & Trends data for one question for all available states and MMSAs.

Class

Topic

**Work with BRFSS Data Directly**

Go to the Behavioral Risk Factors Data Portal to create your own filtered BRFSS Prevalence dataset, customize visualizations, download data, and more.

[Behavioral Risk Factors Data Portal](#)

**Behavioral Risk Factor Surveillance System**

**BRFSS** The [Behavioral Risk Factor Surveillance System](#) (BRFSS) is the nation's premier system of health-related telephone surveys that collect state data about U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive services.



# PREVALENCE OF BINGE DRINKING

- **This section will instruct you on how to get prevalence for a specific topic**
- **For this example, let us see how to get:**
  - **Age-adjusted prevalence of binge drinking in Nebraska in 2018**
  - **Trend of age-adjusted prevalence of binge drinking in Nebraska for all available years (2011-2018)**
  - **Prevalence of binge drinking in Nebraska by gender in 2018 and trend over years (only gives crude prevalence)**


- Drop the menu down
- Select Nebraska and click “GO”

CDC BRFSS Prevalence & Trends Data: [cdc.gov/brfss/brfssprevalence/index.html](#)

CDC Centers for Disease Control and Prevention  
 CDC 24/7: Saving Lives. Protecting People™

BRFSS Prevalence & Trends Data

Welcome! Here you will find the enhanced version of the BRFSS Prevalence and Trends Tool. You have the option of exploring the prevalence data by location or topic. Both routes enable you to view and download prevalence estimates through charts, graphs, and maps based on the path you direct. This version provides access to prevalence estimates from the BRFSS core data at the state level as well as data from Selected Metropolitan/Micropolitan Area Risk Trends (SMART). The prevalence estimates have been updated to include both crude prevalence and age-adjusted prevalence. The tool will continue to be updated as new functions and data become available.



Explore BRFSS Data By Location  
 Explore BRFSS Prevalence & Trends data for all questions for a location.

States & Territories

Iowa  
 Kansas  
 Kentucky  
 Louisiana  
 Maine  
 Maryland  
 Massachusetts  
 Michigan  
 Minnesota  
 Mississippi  
 Missouri  
 Montana  
 Nebraska  
 Nevada  
 New Hampshire  
 New Jersey

Explore BRFSS Data By Topic  
 Explore BRFSS Prevalence & Trends data for one question for all available states and MMSAs.

Class

Topic

Work with BRFSS Data Directly  
 Go to the Behavioral Risk Factors Data Portal to create your own filtered BRFSS Prevalence dataset, customize visualizations, download data, and more.

[Behavioral Risk Factors Data Portal](#)

Behavioral Risk Factor Surveillance System

The screenshot shows the CDC BRFSS Prevalence & Trends Data website. The browser address bar shows the URL: [nccd.cdc.gov/BRFSSPrevalence/rdPage.aspx?rdReport=DPH\\_BRFSS.ExploreByLocati...](https://nccd.cdc.gov/BRFSSPrevalence/rdPage.aspx?rdReport=DPH_BRFSS.ExploreByLocati...). The page header includes the CDC logo and the text "Centers for Disease Control and Prevention" and "CDC 24/7: Saving Lives, Protecting People™". The main title is "BRFSS Prevalence & Trends Data". Below the title is a navigation bar with links: "Home", "Explore by Topic", "Explore by Location", "Data Portal", and "Help". The breadcrumb trail reads: "Behavioral Risk Factor Surveillance System > Prevalence Data and Analysis Tools > BRFSS Prevalence & Trends Data". The section "Explore by Location" is active, showing social media icons for Facebook, Twitter, and a plus sign. Below these are radio buttons for "States" (selected) and "MMSAs". The search form includes four dropdown menus: "Location" (set to "Nebraska"), "Class" (set to "Alcohol Consumption"), "Topic" (set to "Binge Drinking"), and "Year" (set to "2018"). A green "GO" button is to the right of the "Year" dropdown. Four yellow arrows with black numbers point to the search fields: arrow 1 points to the "Class" dropdown, arrow 2 points to the "Topic" dropdown, arrow 3 points to the "Year" dropdown, and arrow 4 points to the "GO" button.

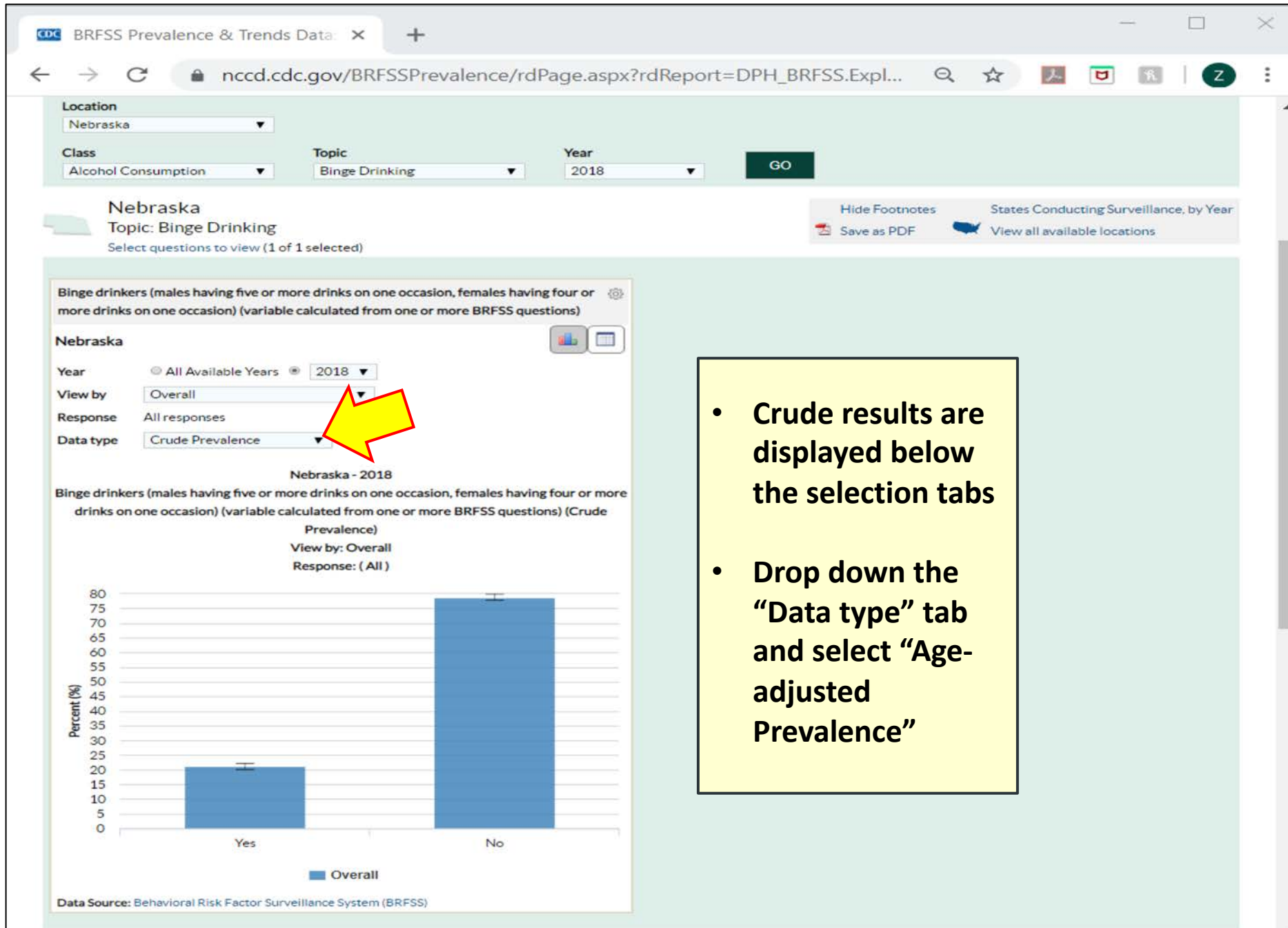
1

2

3

4

- Under “class” select “Alcohol Consumption”
- Under “topic” select “Binge Drinking”
- Under “Year” select “2018”
- Click “go”



- Crude results are displayed below the selection tabs
- Drop down the “Data type” tab and select “Age-adjusted Prevalence”



Binge drinkers (males having five or more drinks on one occasion, females having four or more drinks on one occasion) (variable calculated from one or more BRFSS questions)

Nebraska

Year ☐ All Available Years ☒ 2018

View by Overall

Response All responses

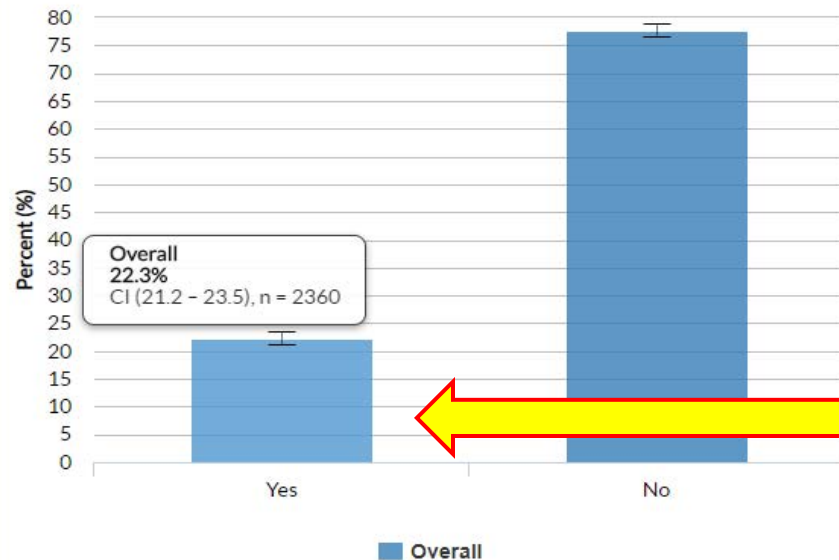
Data type Age-adjusted Prevalence

Nebraska - 2018

Binge drinkers (males having five or more drinks on one occasion, females having four or more drinks on one occasion) (variable calculated from one or more BRFSS questions) (Age-adjusted Prevalence)

View by: Overall

Response: ( All )



Data Source: Behavioral Risk Factor Surveillance System (BRFSS)

- Click this button to get the results in tabular form below

Binge drinkers (males having five or more drinks on one occasion, females having four or more drinks on one occasion) (variable calculated from one or more BRFSS questions)

Nebraska

Year ☐ All Available Years ☒ 2018

View by Overall

Response All responses

Data type Age-adjusted Prevalence

Nebraska - 2018

Binge drinkers (males having five or more drinks on one occasion, females having four or more drinks on one occasion) (variable calculated from one or more BRFSS questions) (Age-adjusted Prevalence)

View by: Overall

Response: ( All )

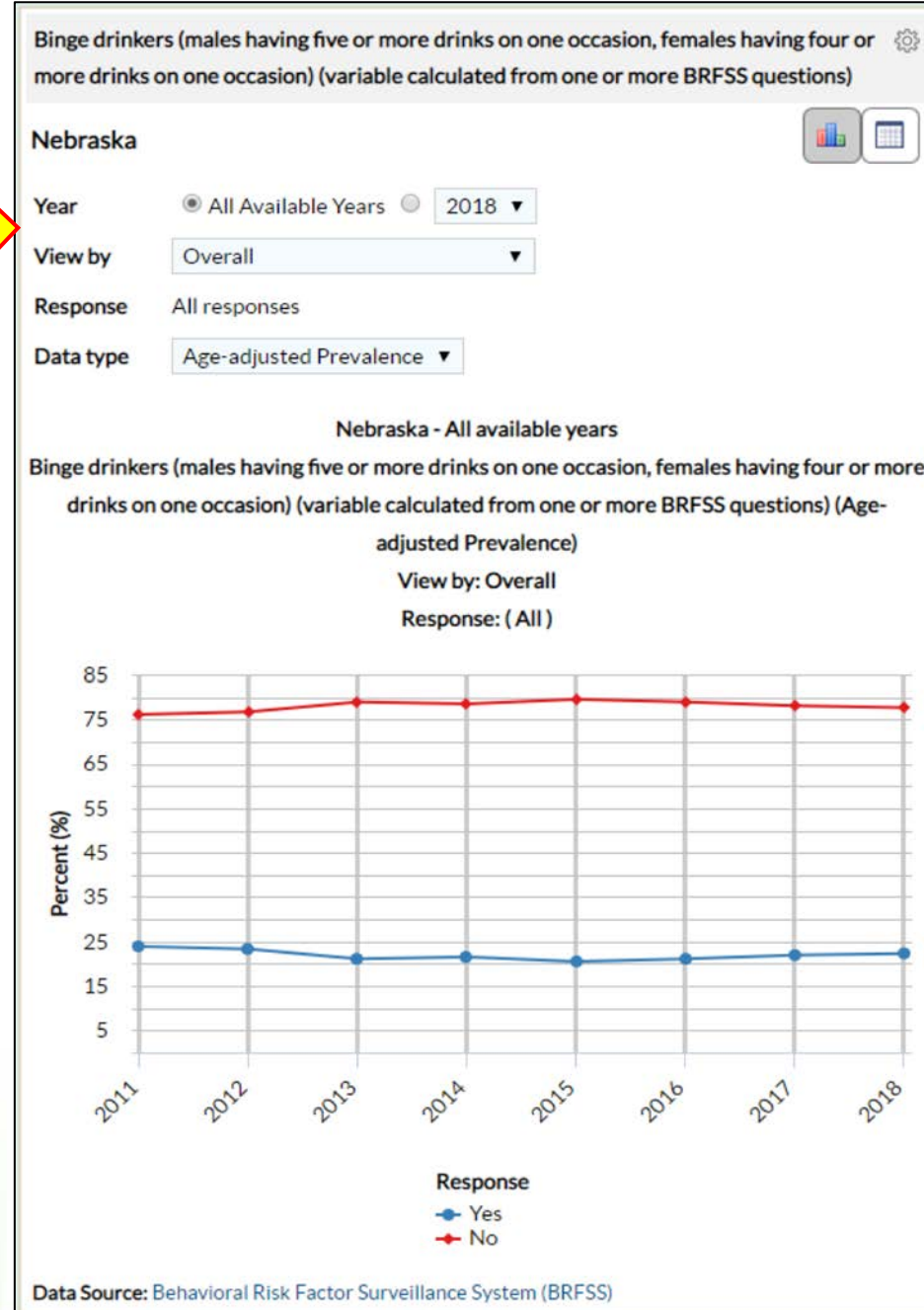
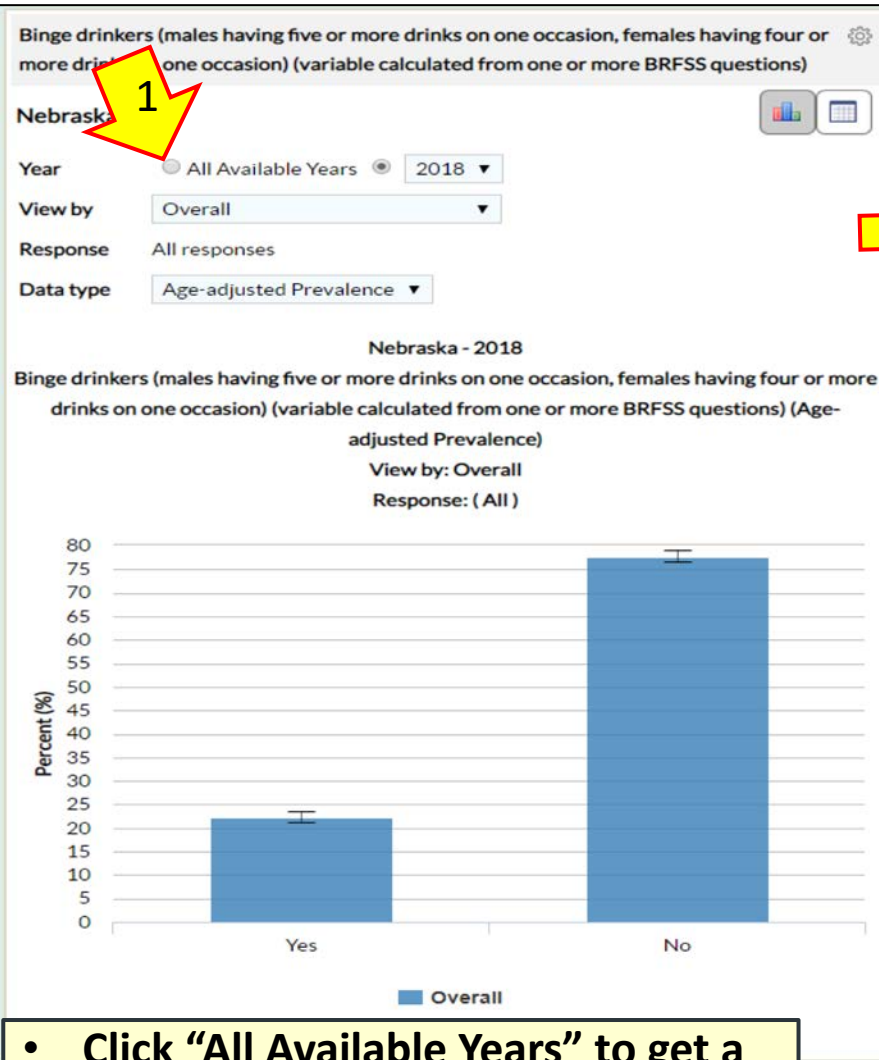
	Yes	No
Percent (%)	22.3	77.7
95% CI	21.2 - 23.5	76.5 - 78.8
n	2360	11580

Data Source: Behavioral Risk Factor Surveillance System (BRFSS)

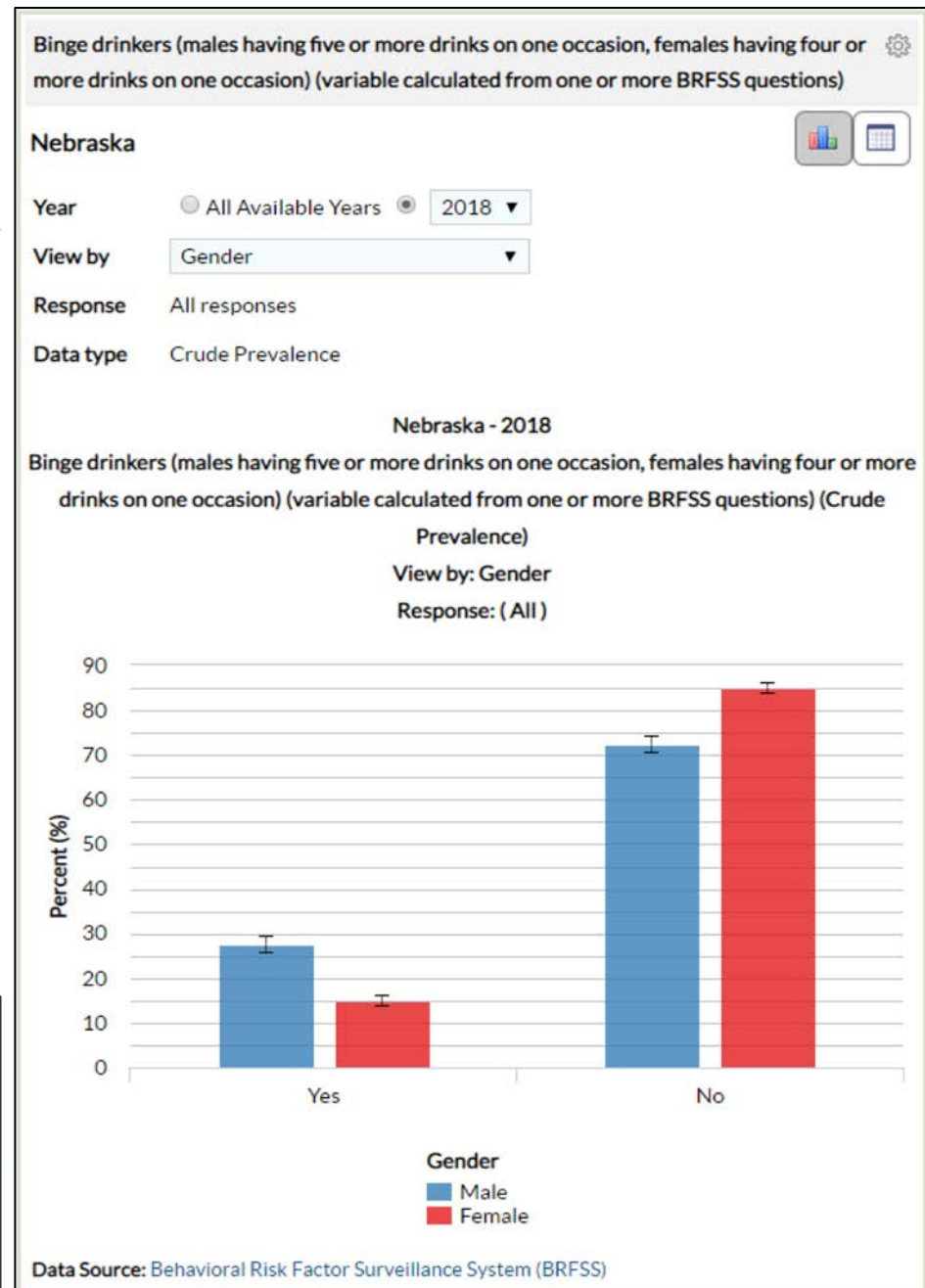
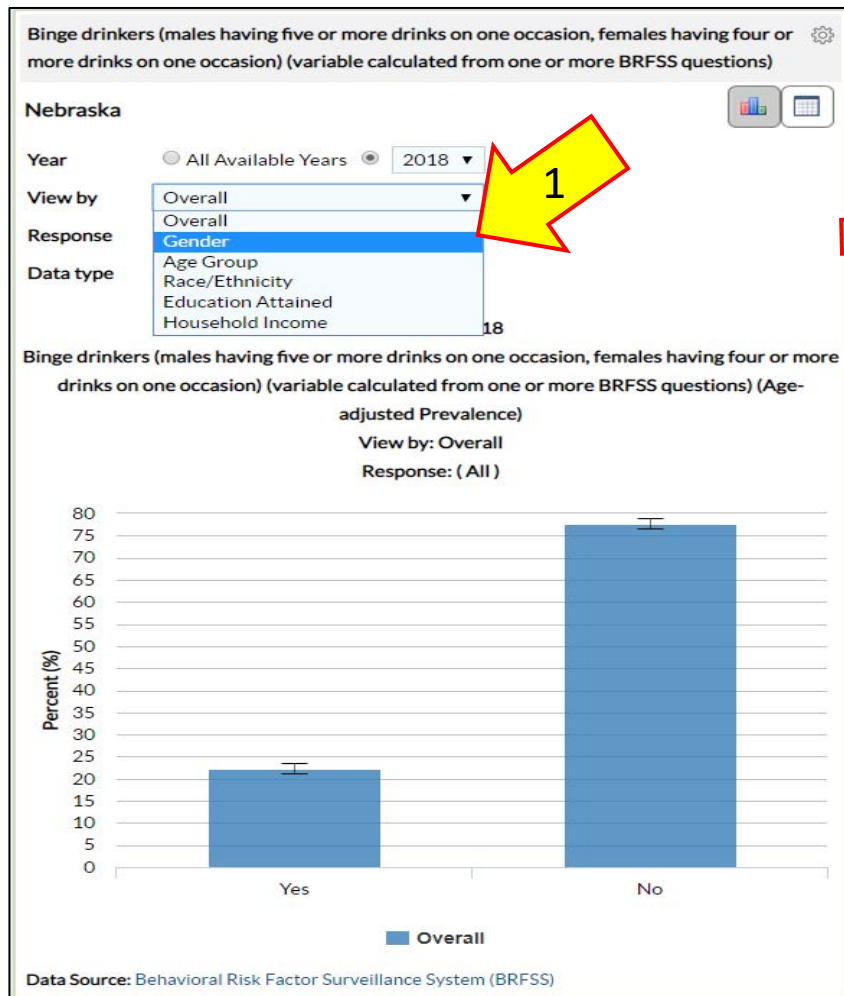
- Hover the Mouse over the bars to see the numbers
- The results show that in 2018, the age-adjusted prevalence of binge drinking in Nebraska was 22.3%





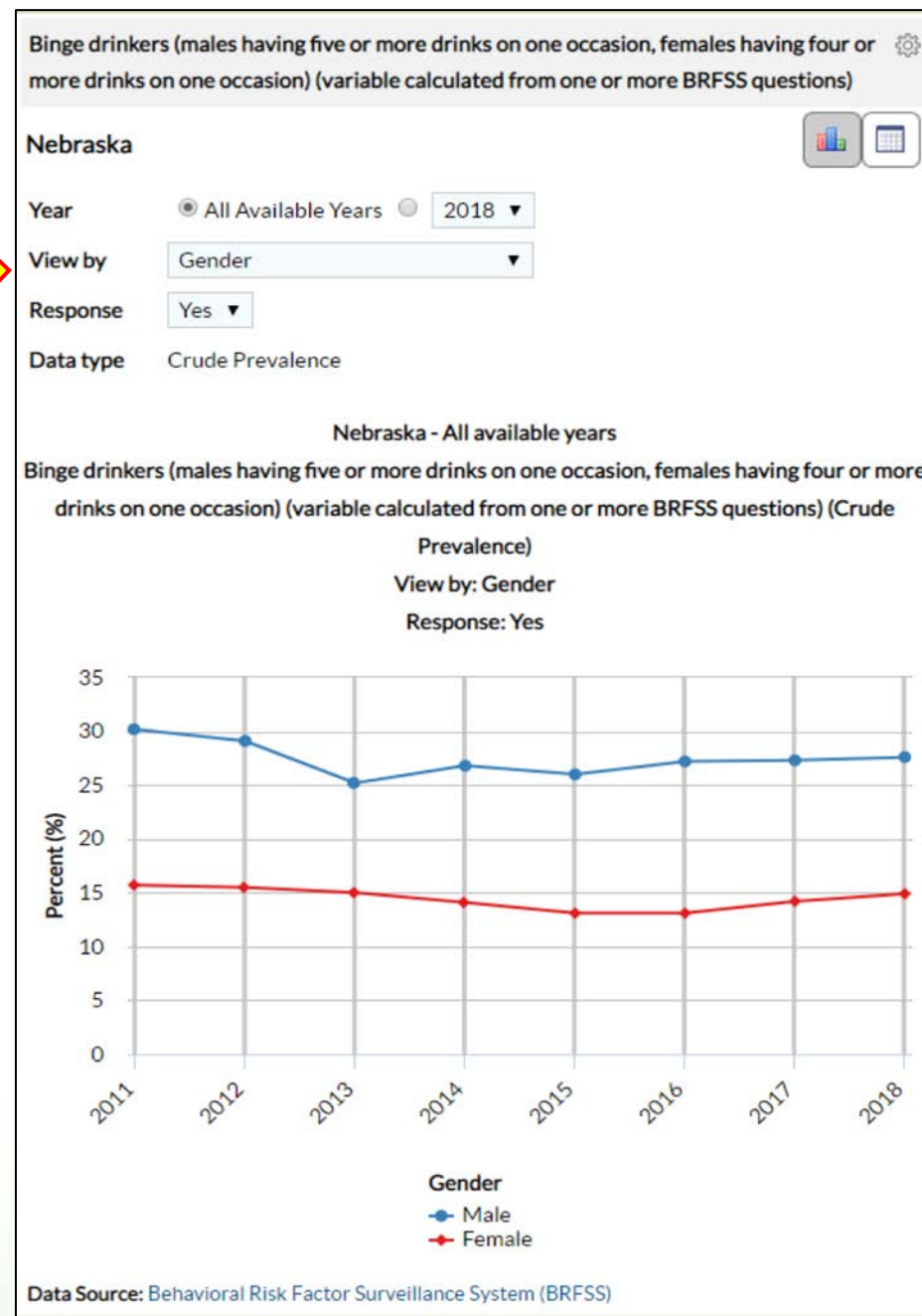
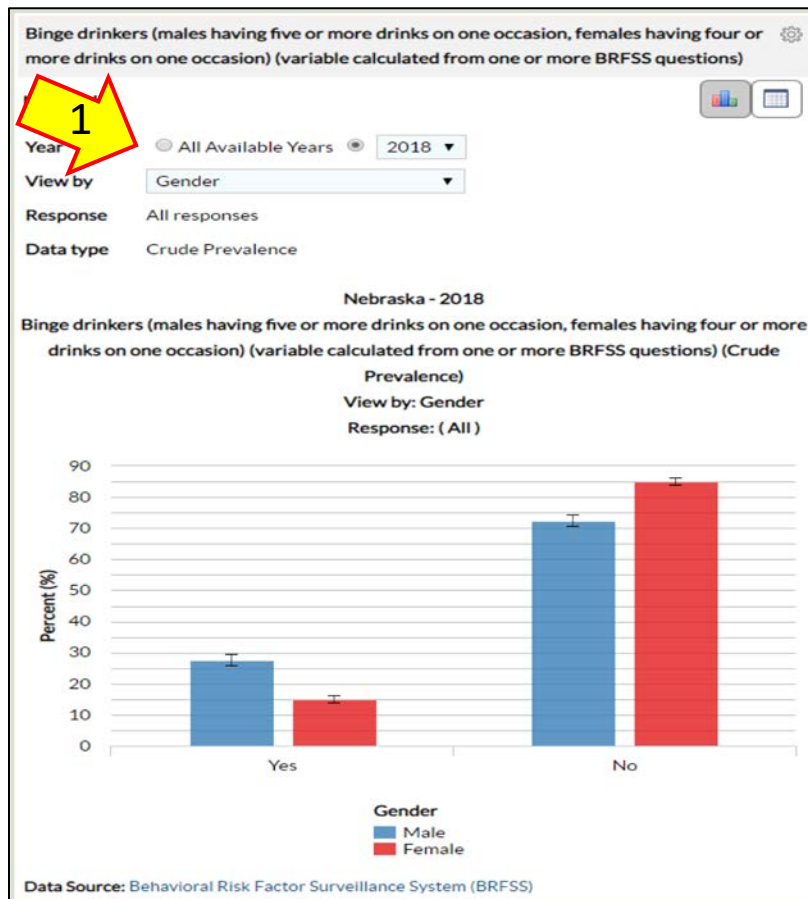


- Click "All Available Years" to get a trend of binge drinking prevalence as shown on the right
- The age-adjusted prevalence of binge drinking in Nebraska has followed nearly a constant trend from 2011 to 2018



- Drop down “view by” and select gender to get the results on the right. Only crude prevalence is displayed
- In 2018, males had a higher prevalence(27.6%) of binge drinking compared to females (14.9%) in Nebraska (Hover over the bars or convert into tabular form to see the numbers).





- Click “All Available Years” to get trends of binge drinking by gender (crude prevalence)
- Among males, the prevalence of binge drinking in Nebraska shows a decrease in 2013 followed by a steady trend
- The prevalence of binge drinking among females has followed a steady trend from 2011 to 2018

# CHRONIC DISEASE INDICATORS TOOL (CDI)



# PREVALENCE AND TRENDS DATA TOOL

- Quick tool to explore prevalence using a wide range of variables
- Produces charts and maps for individual states or the nation by health topics.
- **Particularly useful to produce side by side tables of user selected indicators.**
- Has a wider variety of variables compared to the 'prevalence and trends data' tool.

CDC - BRFSS Prevalence Data & Tools

cdc.gov/brfss/data\_tools.htm

CDC Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives. Protecting People™

Behavioral Risk Factor Surveillance System


CDC > BRFSS

BRFSS

- About BRFSS
- Archived
- Prevalence Data and Data Analysis Tools**
- Survey Data and Documentation
- GIS Maps Data
- SMART: City and County Survey Data
- Statistical Briefs
- Questionnaires
- Publications and Resources
- State Information
- Fact Sheets

## Prevalence Data & Data Analysis Tools

Find city and county data collected through the Selected Metropolitan/Micropolitan Area Risk Trends (SMART) project, the Web Enabled Analysis Tool (WEAT), interactive maps, and other resources provided through BRFSS.



**Prevalence and Trends Data**  
Using the Prevalence and Trends Data Tools, users may produce charts for individual states or the nation by health topic. Users may select specific years or request multiple year data. The Prevalence and Trend Data Tools will produce line graphs for multiple years and bar charts for single years for each selected indicator.

**Web Enabled Analysis Tool (WEAT)**  
The Web Enabled Analysis Tool (WEAT) produces custom crosstabulation tables for health indicators within selected states. Up to two control variables may be included to create crosstab tables within each category of control variables. WEAT also produces logistic equations using BRFSS data to make selections of year, state, and health topic.

**SMART: City and County Data**  
Selected Metropolitan/Micropolitan Area Risk Trends (SMART) is an ongoing project that uses BRFSS data to produce some local area estimates. Counties and Metropolitan/Micropolitan Areas (MMSAs) were selected for SMART if there were 500 or more respondents BRFSS combined landline and cell phone data for any year.

**Chronic Disease Indicators (CDI)**  
The Chronic Disease Indicators Tool allows users to select two or more geographic areas such as states, Metropolitan/Micropolitan Areas (MMSAs), or regions within states. The tool then creates a table illustrating differences among selected health topics.

Go to the BRFSS data website here:  
[https://www.cdc.gov/brfss/data\\_tools.htm](https://www.cdc.gov/brfss/data_tools.htm)

Click “Web Enabled Analysis Tool”


Chronic Disease Indicators (CDI) | x +

cdc.gov/cdi/

## Chronic Disease Indicators

The chronic disease indicators (CDI) are a set of surveillance indicators developed by consensus among CDC, the Council of State and Territorial Epidemiologists (CSTE), and the National Association of Chronic Disease Directors (NACDD). CDI enables public health professionals and policymakers to retrieve uniformly defined state and selected metropolitan-level data for chronic diseases and risk factors that have a substantial impact on public health. These indicators are essential for surveillance, prioritization, and evaluation of public health interventions. [More](#)


[Help](#)



### Explore CDI Data By Location

Explore Chronic Disease Indicators (CDI) data for a specific location.

Location:



Guam Puerto Rico Virgin Islands

### Explore CDI Data by Indicator

Explore Chronic Disease Indicators (CDI) data for one indicator or all available locations.

Category:

#### View Comparison Report

[View Comparison Report](#)  
View a report that contains all indicators for the locations of your choice.

#### View Indicator Definition

View an indicator definition by selecting a category.

#### Use the Data Portal

Need to work with CDI data directly?

Go to the Chronic Disease Indicators Data Portal to create your own filtered CDI dataset, customize visualizations, download CDI data, and more.

[Chronic Disease Indicators Data Portal >](#)

### CDI MMWR

MMWR Recommendations & Reports

[Indicators for Chronic Disease Surveillance - United States, 2013](#) [1.76 MB]

1. To start with a geographic location, drop down 'Location', select the state and click 'Go'.

2. To start with a topic drop down 'Category', select topic and click 'Go'.

3. Use 'View Indicator Definition' to see how an indicator has been defined

We will start with location. Select Nebraska and click 'Go'



Chronic Disease Indicators: Explore

nccd.cdc.gov/cdi/rdPage.aspx?rdReport=DPH\_CDI.E...

**CDC** Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

## Chronic Disease Indicators

Home Explore by Indicator Explore by Location Comparison Report Data Portal Help

CDC > Division of Population Health > Chronic Disease Indicators

### Explore by Location

f t +

**Location**  
Nebraska

**Category**  
Select one  
Select one  
**Alcohol**  
Arthritis  
Asthma  
Cancer  
Cardiovascular Disease  
Chronic Kidney Disease  
Chronic Obstructive Pulmonary Disease  
Diabetes  
Disability  
Immunization  
Mental Health  
Nutrition, Physical Activity, and Weight Status  
Older Adults  
Oral Health  
Overarching Conditions  
Reproductive Health  
School Health  
Tobacco

CDC Media

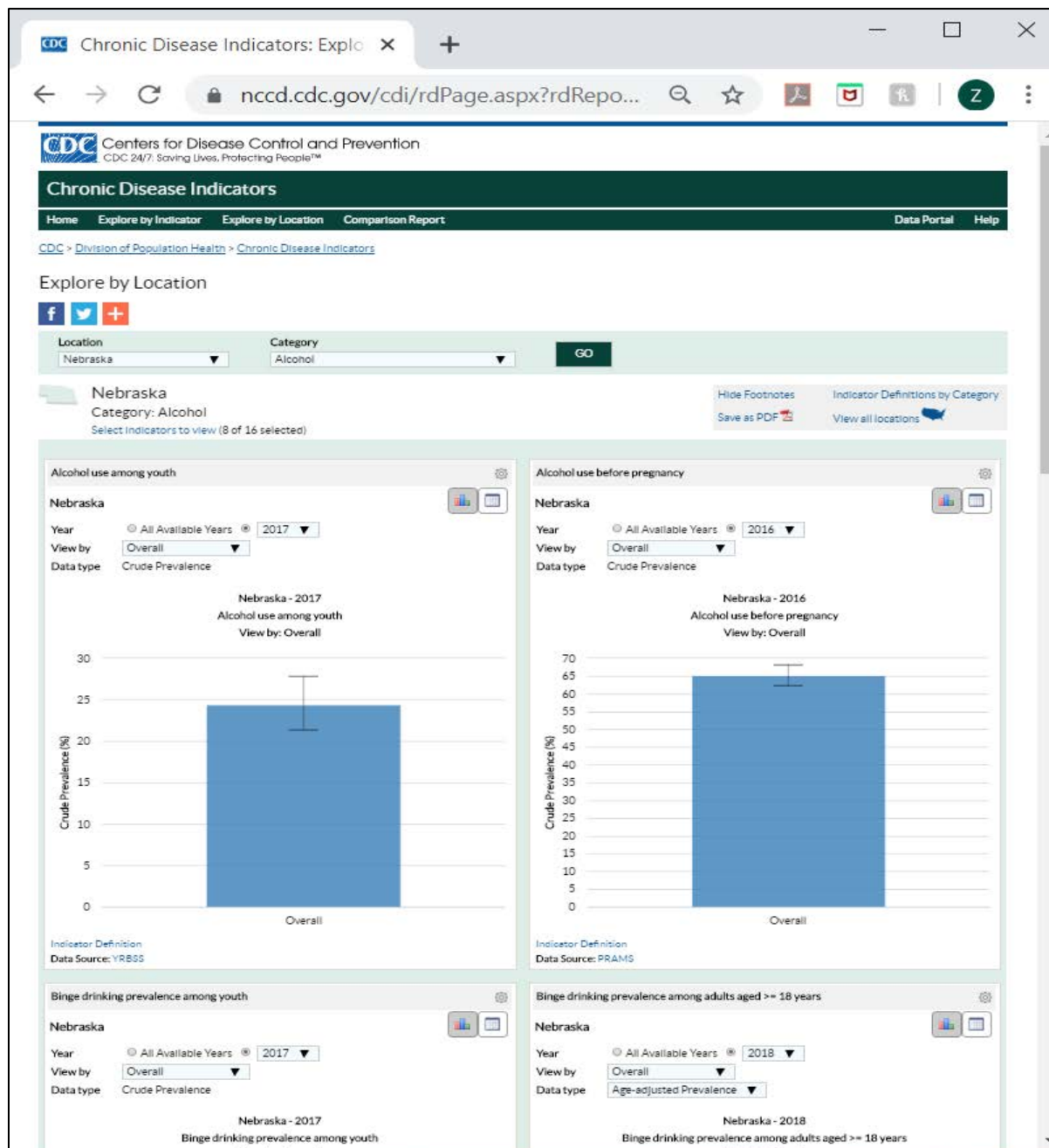
CDC  
for Disease Control and  
n  
n Road, Atlanta, GA 30329-4027  
NFO (800-232-4636)  
DC-INFO

Go to the BRFSS data website here:

[https://www.cdc.gov/brfss/data\\_tools.htm](https://www.cdc.gov/brfss/data_tools.htm)

Click "Chronic Disease Indicators"

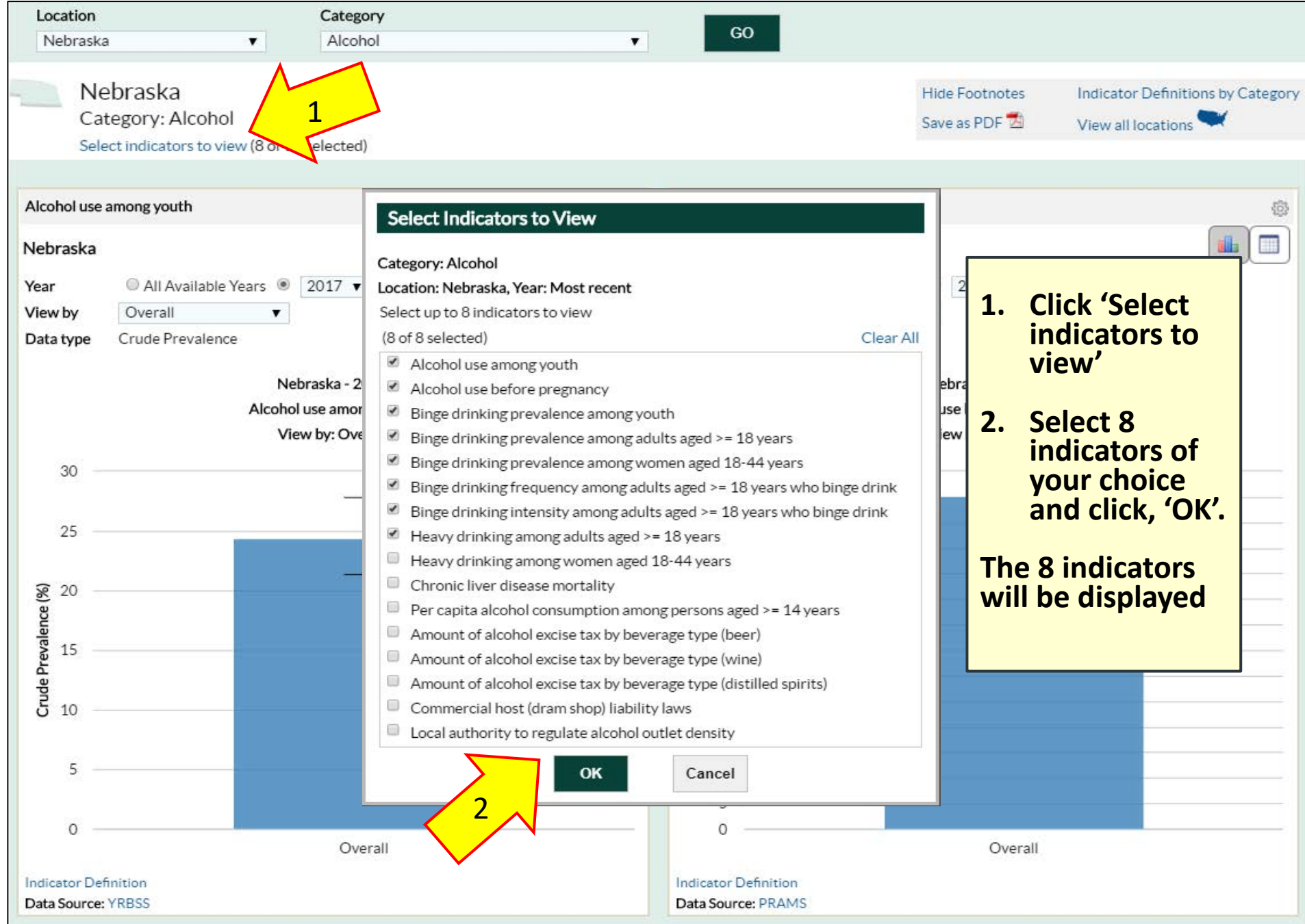
Select alcohol and click 'Go'

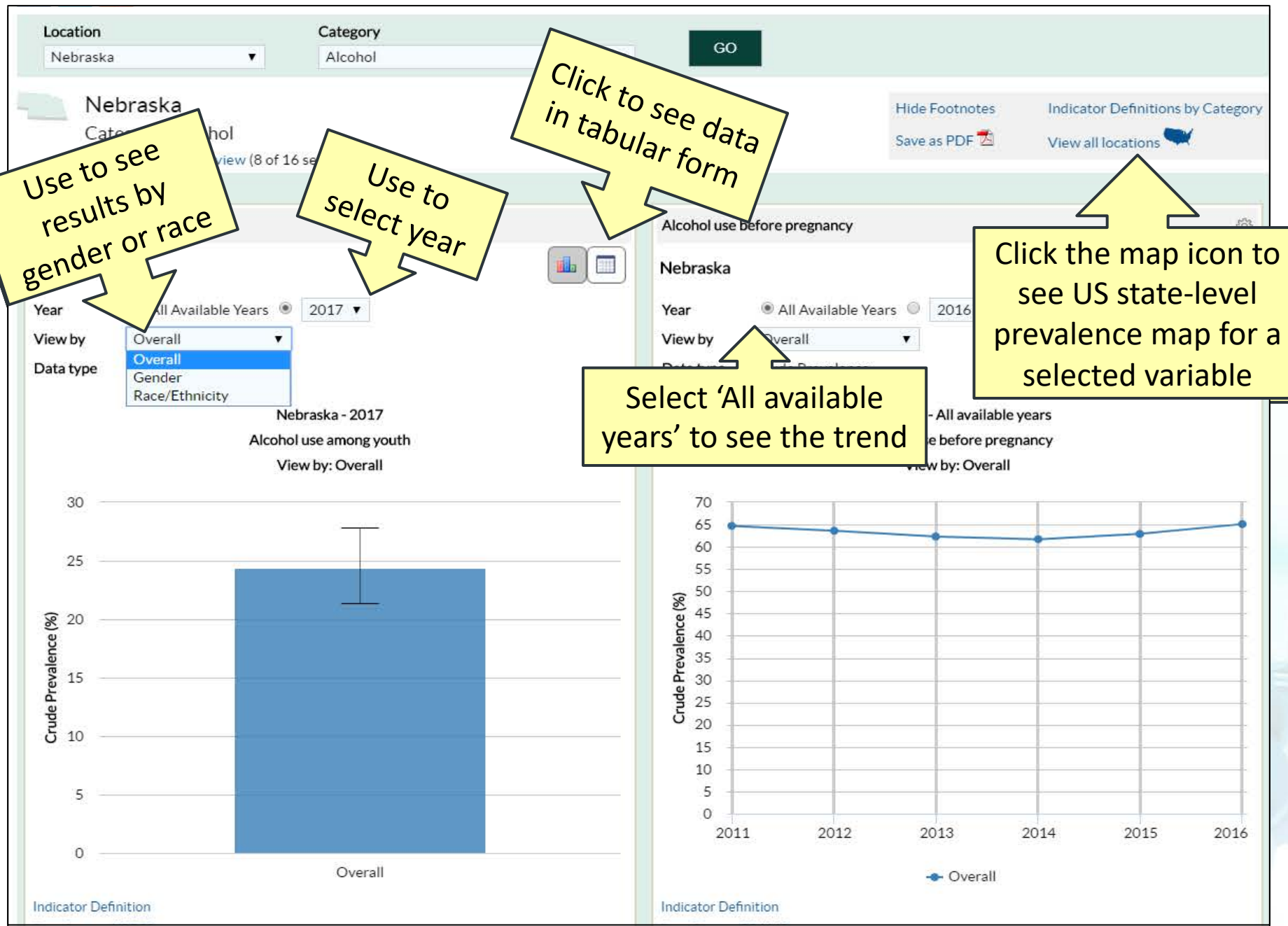


Shows multiple alcohol indicators side by side.

Scroll down your browser to see a total of eight indicator at a time.







# WEB ENABLED ANALYSIS TOOL (WEAT)

# WEB ENABLED ANALYSIS TOOL (WEAT)

- **Enables users to create custom crosstabulation tables for variables within selected states**
- **Can be used to run logistic regression**



CDC - BRFSS Prevalence Data & Tools | x +

cdc.gov/brfss/data\_tools.htm

CDC Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

Behavioral Risk Factor Surveillance System

CDC > BRFSS


BRFSS

- About BRFSS +
- Archived
- Prevalence Data and Data Analysis Tools**
- Survey Data and Documentation +
- GIS Maps Data +
- SMART: City and County Survey Data +
- Statistical Briefs
- Questionnaires
- Publications and Resources
- State Information +
- Fact Sheets

Get Email Updates

## Prevalence Data & Data Analysis Tools

Find city and county data collected through the Selected Metropolitan/Micropolitan Area Risk Trends (SMART) project, the Web Enabled Analysis Tool (WEAT), interactive maps, and other resources provided through BRFSS.



### Prevalence and Trends Data

Using the Prevalence and Trends Data Tools, users can produce charts for individual states or the national health topic. Users may select specific years or multiple year data. The Prevalence and Trends Data Tools will produce line graphs for multiple years and bar charts for single years for each selected indicator.

### Web Enabled Analysis Tool (WEAT)

The Web Enabled Analysis Tool (WEAT) permits users to create custom crosstabulation tables for health indicators within selected states. Up to two control variables may be included to create crosstab tables within each category of control variables. WEAT also may be used to create logistic equations using BRFSS data. Users are prompted to make selections of year, state, and variables to be included in the analyses.

more respondents BRFSS combined landline and cell phone data for any year.

### Chronic Disease Indicators (CDI)

The Chronic Disease Indicators Tool allows users to select two or more geographic areas such as states, Metropolitan/Micropolitan Areas (MMSAs), or regions within states. The tool then creates a table illustrating differences among selected health indicators.

**Go to the BRFSS data website here:**

**[https://www.cdc.gov/brfss/data\\_tools.htm](https://www.cdc.gov/brfss/data_tools.htm)**

**Click “Web Enabled Analysis Tool”**

# ORAL HEALTH AND TOBACCO USE

- **This section will instruct you on how to use WEAT tool to explore if an association exists between oral health and tobacco use**
- **For this example, let us see how to get:**
  - **A cross-tabulation between current tobacco use and permanent teeth extraction stratified by gender among Nebraska residents in 2018**
  - **Gender-adjusted measure of association (odds ratio in this case) between current tobacco use and permanent teeth extraction among Nebraska residents in 2018 (as the outcome variable is binary, we will use logistic regression)**

Behavioral Risk Factor Surveillanc

+

← → ↻

nccd.cdc.gov/weat/#/analysis

🔍


☆

📄

📱

👤

⋮



Centers for Disease Control and Prevention

CDC 24/7: Saving Lives, Protecting People™

SEARCH

🔍

BRFSS Web Enabled Analysis Tool

[CDC](#) > [BRFSS](#) > [WEAT](#)

BRFSS WEAT

Analysis -

Cross Tabulation

Logistic Regression

Help +

Related Information +

The BRFSS is the world's largest telephone health survey system, tracking health risks in the United States. Information from the survey is used to improve the health of US residents. The BRFSS Web Enabled Analysis Tool (WEAT) allows users to conduct real-time state-level data analysis.

BRFSS Analysis

Cross Tabulation

Logistic Regression

A cross tabulation, or "crosstab," produces frequencies or percentages for one or more variables, in one or more tables. For example, one can use the cross-tabulation procedure in the BRFSS to generate a table showing numbers and percentages of respondents with diabetes by age group. A general formula for cross tabulation can be depicted as  $A \times B$ , where A is the dependent variable or outcome (e.g., diabetes) and B is the independent or exposure variable (e.g., age). For our purposes, "crosstab" includes frequencies or percentages for a single variable.

Logistic regression is a calculation of the contribution of one or more predictors on a particular outcome, such as "Risk factor:At risk for binge drinking." The results provide a predictive model and can be converted to log odds. The basic logistic formula using one predictor is  $Y = \exp(a + B1X1)/1 + \exp(a + B1X1)$ .

About CDC

Employment

Newsroom

Training/Education

Funding

CDC's Organization

Mission and Vision

Using this Site

Legal

Link to Us

Policies

FOIA

Accessibility

Privacy

No FEAR Act

Inspector General

USA.gov

Contact CDC


Centers for Disease Control and Prevention

1600 Clifton Road

Atlanta, GA 30329-4027 USA

800-CDC-INFO (800-232-4636)

[Contact CDC-INFO](#)





Behavioral Risk Factor Surveillance System | nccd.cdc.gov/weat/#/crossTabulation/selectYear

**CDC** Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

## BRFSS Web Enabled Analysis Tool

CDC > BRFSS > WEAT > Cross Tabulation Years

**BRFSS WEAT**

- Analysis -
- Cross Tabulation
- Logistic Regression
- Help +
- Related Information +

### Cross Tabulation

Select Year

☒ Single Year
 ☐ Combine Years
 ☐ Compare Years

Select Year ▼ **1**  
 Select Year  
 2018  
 2017  
 2016  
 2015  
 2014  
 2013  
 2012  
 2011  
 2010  
 2009  
 2008  
 2007  
 2006  
 2005  
 2004

Select Methodology ▼  
 Select Methodology

Select Methodology ▼  
 Select Methodology

Back Next **2**

**Help with this step...**

Start building your cross tabulation table by selecting one or more years.

You may choose to analyze a **Single Year** of data, **Combine** two or more years of data, or **Compare** data across two or more years.

Start the cross tabulation analysis over at any time by clicking the "Reset" button to the left.

**My Analysis**

Year(s)	
Location(s)	
Row Variable	
Column Variable	
Control Variable(s)	
Number Of Source	0

Using these options, years can be combined or compared. We will not use these in this example

1. Drop down "Select Year" and click "2018"

2. Click next



Behavioral Risk Factor Surveillance System

nccd.cdc.gov/weat/#/crossTabulation/selectLocation

**CDC** Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

**BRFSS Web Enabled Analysis Tool**

[CDC](#) > [BRFSS](#) > [WEAT](#) > [Cross Tabulation Years](#) > [Location](#)

**BRFSS WEAT**

- Analysis -
- Cross Tabulation
- Logistic Regression
- Help +

**Related Information** +

**Cross Tabulation**

**Select Location**

☒ Single Location ☐ Combine Locations ☐ Compare Locations

Select Location

- Iowa
- Kansas
- Kentucky
- Louisiana
- Maine
- Maryland
- Massachusetts
- Michigan
- Minnesota
- Mississippi
- Missouri
- Montana
- Nebraska**
- Nevada
- New Hampshire
- New Jersey
- New Mexico
- New York
- North Carolina
- North Dakota
- Ohio

Single Location of data, to Combine data from two or more states, or to Compare data across two or more states.

Back Next

**My Analysis** [Reset](#)

Year(s)	2018 (single)
Location(s)	
Row Variable	
Column Variable	
Control Variable(s)	
Number Of Source Records	437436

Using these options, locations can be combined or compared. We will not use these in this example

1. Drop down "Select Location" and click "Nebraska"

2. Click next

Behavioral Risk Factor Surveillance System

nccd.cdc.gov/weat/#/crossTabulation/selectRowVariable

**CDC** Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

**BRFSS Web Enabled Analysis Tool**

CDC > BRFSS > WEAT > Cross Tabulation Years > Location > Row

**BRFSS WEAT**

Analysis

Cross Tabulation

Logistic Regression

**Cross Tabulation**

Select Row Variable (Required)

Variable

Topic

Search All

Search All

Alcohol Consumption

Chronic Health Conditions

Colorectal Cancer Screening

Demographic Information

Disability

Drinking and Driving

Exercise

Falls

Health Care Access

Health Status

Healthy Days

HIV/AIDS

Immunization

Inadequate Sleep

Oral Health

Prostate Cancer Screening

Seatbelt Use

**Tobacco Use**

Veteran Status

Back

Next

My Analysis

Reset

Year(s)	2018 (single)
Location(s)	Nebraska (single)
Row Variable	
Column Variable	
Control Variable(s)	

We will use our exposure (current tobacco use) as the row variable

Drop down "Topic" and select "Tobacco use"

Different tobacco related variables will appear

Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

SEARCH

BRFSS WEAT

Analysis

Cross Tabulation

Logistic Regression

Help

Related Information

Cross Tabulation

Select Row Variable (Required)

☒ Variable

Topic

Tobacco Use

Calculated variable for 4 level smoker status: everyday/someday/former/non-smoker (S...)

Calculated variable for adults who are current smokers (RFSMOK3)

Quit smoking for 1 day or more in past 12 months (STOPSMK2)

Back

Next

Help with this step...

You have selected a variable for the row. This means that each level of this variable will be displayed as a separate row in the output cross tabulation.

My Analysis

Year(s)

2018 (single)

Location(s)

Nebraska (single)

Reset

Select “Calculated variable for adults who are current smokers”

Click “Next”



Behavioral Risk Factor Surveillance

nccd.cdc.gov/weat/#/crossTabulation/selectColumnVariable

**CDC** Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

**BRFSS Web Enabled Analysis Tool**

CDC > BRFSS > WEAT > Cross Tabulation Years > Location > Row > Column

**BRFSS WEAT**

Analysis

Cross Tabulation

Logistic Regression

### Cross Tabulation

Select a Column Variable (Optional)

☒ Include a Column Variable

☐ Variable

Topic

Annual Alcohol Consumption  
 Are you Chronic Health Conditions  
 Blind or when wearing glasses (BLIND) (Disability)  
 Calculated per-day (DROCDY3\_ categorical) (Alcohol Consumption)  
 Calculated Disability  
 Calculated Drinking and Driving  
 Calculated Exercise  
 Calculated Falls  
 Health Care Access  
 Health Status  
 Healthy Days  
 HIV/AIDS  
 Immunization  
 Inadequate Sleep  
 Oral Health  
 Prostate Cancer Screening  
 Seatbelt Use  
 Tobacco Use  
 Veteran Status

Demographic Information  
 (status)  
 on when wearing glasses (BLIND) (Disability)  
 per-day (DROCDY3\_ categorical) (Alcohol Consumption)  
 ory (\_AGEG5YR) (Demographic Information)  
 ry (\_AGE65YR, 18-64, 65+) (Demographic Information)  
 mental health status (\_MENT14D) (Healthy Days)  
 physical health status (\_PHYS14D) (Healthy Days)  
 al/mental health status from doing usual activities (POOP14D) (Healthy Days)

Back Next

### My Analysis

[Reset](#)

Year(s)	2018 (single)
Location(s)	Nebraska (single)
Row Variable	Tobacco Use: Calculated variable for adults who are current smokers (_RFSMOK3)
Column Variable	
Control Variable(s)	
Number Of Source	14634

We will use our outcome (permanent tooth extraction) as the column variable

Drop down "Topic" and select "Oral Health"

Different oral health related variables will appear



Behavioral Risk Factor Surveillance × +

nccd.cdc.gov/weat/#/crossTabulation/selectColumnVariable

**CDC** Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

SEARCH Q

**BRFSS Web Enabled Analysis Tool**

[CDC](#) > [BRFSS](#) > [WEAT](#) > [Cross Tabulation Years](#) > [Location](#) > [Row](#) > [Column](#)

**BRFSS WEAT**

- Analysis −
- Cross Tabulation
- Logistic Regression
- Help +
- Related Information +

### Cross Tabulation

Select a Column Variable (Optional)

☒ Include a Column Variable

☒ Variable

Topic: Oral Health ▾

Calculated variable for adults aged 18+ who have had permanent teeth extracted (\_EXTETH3)

Calculated variable for adults aged 65+ who have had all their natural teeth extracted (\_ALTETH3)

Calculated variable for adults who have visited a dentist, dental hygienist or dental clinic within the past year (\_DENVST3)

Number of permanent teeth removed (RMVTETH4)

Time since last visited dentist or dental clinic (LASTDEN4)

Back Next

Help with this step...

You have selected a variable for the column. Each level of your selected variable will be displayed as a separate column in the output cross tabulation.

Click "Next" to select your optional control variables or click "Back" to change your row variable selection.

My Analysis [Reset](#)

Year(s)	2018 (single)
Location(s)	Nebraska (single)
Row Variable	Tobacco Use: Calculated variable for adults who are current smokers (_RFSMOK3)
Column Variable	Oral Health: Calculated variable for adults aged 18+ who have had permanent teeth extracted (_EXTETH3)

Select “calculated variable for adults aged 18+ who have had permanent teeth extracted”

Click “Next”

Behavioral Risk Factor Surveillance System

nccd.cdc.gov/weat/#/crossTabulation/selectControlVariables

**CDC** Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

BRFSS Web Enabled Analysis Tool

SEARCH

Location Years > Location > Row > Column > Controls

1 Tabulation

Control Variables (Optional)

☒ Include a Control Variable

☐ Variable

Topic

Search All

Search All

Alcohol Consumption

Chronic Health Conditions

Colorectal Cancer Screening

Demographic Information

Disability

Drinking and Driving

Exercise

Falls

Health Care Access

Health Status

Healthy Days

HIV/AIDS

Immunization

Inadequate Sleep

Oral Health

Prostate Cancer Screening

Seatbelt Use

Tobacco Use

Veteran Status

Wearing glasses (BLIND) (Disability)

PROCDY3\_categorical (Alcohol Consumption)

ory (AGE65YR) (Demographic Information)

ry (AGE65YR, 18-64, 65+) (Demographic Information)

ental health status (\_MENT14D) (Healthy Days)

ysical health status (\_PHYS14D) (Healthy Days)

al/mental health status from doing usual activities (DOGB14D) (Healthy Days)

Back

Next

Help with this step

My Analysis

Reset

As we want to see results stratified by gender, we will select gender as our control variable

Click “Include a Control Variable”

Drop down “topic” and select “Demographic Information”

We can select two controls per analysis but for this one we will select only one

Behavioral Risk Factor Surveillance

nccd.cdc.gov/weat/#/crossTabulation/selectControlVariables

CDC Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

BRFSS Web Enabled Analysis Tool

CDC > BRFSS > WEAT > Cross Tabulation Years > Location > Row > Column > Controls

**BRFSS WEAT**

- Analysis -
- Cross Tabulation
- Logistic Regression
- Help +
- Related Information +

### Cross Tabulation

#### Select Control Variables (Optional)

☒ Include a Control Variable

☐ Variable

Topic: Demographic Information

- Calculated variable for reported height in inches (HTIN4, categorical)
- Calculated variable for reported height in meters (HTM4, categorical)
- Calculated variable for reported weight in kilograms (WTKG3, categorical)
- Calculated variable for white non-Hispanic (W\_NH, categorical)
- Education level (EDUCA)
- Employment Status (EMPLOY1)
- Gender (SEX1)
- Marital status (MARITAL)
- Pregnancy status (PREGNANT)

☐ Include an Additional Control Variable

☐ Variable

Topic:

Back Next

Select "Gender (SEX1)"  
Click "Next"

Behavioral Risk Factor Surveillance System

nccd.cdc.gov/weat/#/crossTabulation/selectStatistics

CDC Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

BRFSS Web Enabled Analysis Tool

CDC > BRFSS > WEAT > Cross Tabulation Years > Location > Row > Column > Controls > Statistics

**BRFSS WEAT**

- Analysis -
- Cross Tabulation
- Logistic Regression
- Help +
- Related Information +

**Cross Tabulation**

**Calculated Statistics**

- Sample Size
- Chi Square / Degrees of Freedom /p-value
- Row Percentage
- Column Percentage
- Total Percentage

**Select Additional Statistics**

- ☐ Standard Error
- ☒ 95% Confidence Interval

**Other**

- ☐ Include Weighted N
- ☐ Include Non Response Categories

**Help with this step...**

Certain calculated statistics have already been selected for you. These statistics are standard when running a cross tabulation.

"Check" the **Standard Error** and/or **95% Confidence Interval** boxes under the "Select Additional Statistics" section if you would like to see additional row and column statistics in your analysis.

**Click "Run Report"**

**Back** **Run Report**

**My Analysis** [Reset](#)

Year(s)	2018 (single)
Location(s)	Nebraska (single)
Row Variable	Tobacco Use: Calculated variable for adults who are current smokers (RFSMOK3)
Column Variable	Oral Health: Calculated variable for adults aged 18 and over

These will be automatically calculated

You can choose the statistic you want



Behavioral Risk Factor Surveillance System  
Cross Tabulation, Nebraska, 2018  
of Tobacco Use by Oral Health

Sample Size	
Row %	(95% Confidence Interval)
Column %	(95% Confidence Interval)
Total % (Weighted)	(95% Confidence Interval)

Calculated variable for adults aged 18+ who have had permanent teeth extracted (\_EXTETH3)

No Yes Total

Calculated variable for adults who are current smokers (\_RFSMOK3)

Former smoker or never smoked	n	6,665	5,396	12,061
	Row%	64.8% (63.6 - 66.1)	35.2% (33.9 - 36.4)	100.0% (100.0 - 100.0)
	Col%	87.6% (86.5 - 88.8)	78.2% (76.6 - 79.8)	84.1% (83.1 - 85.0)
	%	54.5% (53.3 - 55.8)	29.6% (28.5 - 30.6)	84.1% (83.1 - 85.0)
Current smoker	n	729	1,155	1,884
	Row%	48.3% (45.0 - 51.6)	51.7% (48.4 - 55.0)	100.0% (100.0 - 100.0)
	Col%	12.4% (11.2 - 13.5)	21.8% (20.2 - 23.4)	15.9% (15.0 - 16.9)
	%	7.7% (6.9 - 8.4)	8.2% (7.6 - 8.9)	15.9% (15.0 - 16.9)
Total	n	7,394	6,551	13,945
	Row%	62.2% (61.0 - 63.4)	37.8% (36.6 - 39.0)	
	Col%	100.0% (100.0 - 100.0)	100.0% (100.0 - 100.0)	
	%	62.2% (61.0 - 63.4)	37.8% (36.6 - 39.0)	

Wald Chi-Square Value	Degrees of Freedom	p-value
85.39	1	<0.0001

Number of records on the BRFSS dataset for the year and location selected	14,634
Number of records excluded from the analysis	689
Sample Size (Number of records used in the analysis)	13,945
Total Weighted N (Population)	1,395,494

Multiple result tables are produced.

The first table, as shown here, shows overall sample without stratification

In 2018, among former/never smokers living in Nebraska, the prevalence of permanent tooth extraction was 32.5%

In 2018, among current smokers living in Nebraska, the prevalence of permanent tooth extraction was 51.7%

The Chi-Square p-value (<0.0001) suggests a significant association between smoking status and permanent tooth extraction

Controlling  
for

Gender (SEX1) = Male

Sample Size  
Row % (95% Confidence Interval)  
Column % (95% Confidence Interval)  
Total % (Weighted) (95% Confidence Interval)

Calculated variable for adults aged 18+ who have had permanent teeth  
extracted (\_EXTETH3)

No Yes Total

Calculated variable for adults who are current smokers (\_RFSMOK3)

Former smoker or never smoked	n	2,978	2,338	5,316
	Row%	64.7% (62.9 - 66.6)	35.3% (33.4 - 37.1)	100.0% (100.0 - 100.0)
	Col%	86.1% (84.4 - 87.9)	76.9% (74.5 - 79.3)	82.6% (81.2 - 84.1)
	%	53.5% (51.6 - 55.3)	29.2% (27.6 - 30.7)	82.6% (81.2 - 84.1)
Current smoker	n	394	548	942
	Row%	49.5% (45.0 - 54.1)	50.5% (45.9 - 55.0)	100.0% (100.0 - 100.0)
	Col%	13.9% (12.1 - 15.6)	23.1% (20.7 - 25.5)	17.4% (15.9 - 18.8)
	%	8.6% (7.5 - 9.7)	8.8% (7.8 - 9.8)	17.4% (15.9 - 18.8)
Total	n	3,372	2,886	6,258
	Row%	62.1% (60.3 - 63.8)	37.9% (36.2 - 39.7)	
	Col%	100.0% (100.0 - 100.0)	100.0% (100.0 - 100.0)	
	%	62.1% (60.3 - 63.8)	37.9% (36.2 - 39.7)	

Wald Chi-Square Value	Degrees of Freedom	p-value
36.83	1	<0.0001

The second table, as shown here, shows results only for males

In 2018, among male former/never smokers living in Nebraska, the prevalence of permanent tooth extraction was 35.3%

In 2018, among male current smokers living in Nebraska, the prevalence of permanent tooth extraction was 50.5%

The Chi-Square p-value (<0.0001) suggests a significant association between smoking status and permanent tooth extraction in males

Controlling for Gender (SEX1) = Female

Sample Size  
 Row % (95% Confidence Interval)  
 Column % (95% Confidence Interval)  
 Total % (Weighted) (95% Confidence Interval)

Calculated variable for adults aged 18+ who have had permanent teeth extracted (\_EXTETH3)

No Yes Total

Calculated variable for adults who are current smokers (\_RFSMOK3)

Former smoker or never smoked	n	3,681	3,055	6,736
	Row%	65.0% (63.4 - 66.7)	35.0% (33.3 - 36.6)	100.0% (100.0 - 100.0)
	Col%	89.1% (87.5 - 90.6)	79.5% (77.3 - 81.6)	85.5% (84.2 - 86.7)
	%	55.6% (53.9 - 57.3)	29.9% (28.5 - 31.4)	85.5% (84.2 - 86.7)
Current smoker	n	335	606	941
	Row%	46.9% (42.0 - 51.7)	53.1% (48.3 - 58.0)	100.0% (100.0 - 100.0)
	Col%	10.9% (9.4 - 12.5)	20.5% (18.4 - 22.7)	14.5% (13.3 - 15.8)
	%	6.8% (5.8 - 7.8)	7.7% (6.8 - 8.6)	14.5% (13.3 - 15.8)
Total	n	4,016	3,661	7,677
	Row%	62.4% (60.8 - 64.0)	37.6% (36.0 - 39.2)	
	Col%	100.0% (100.0 - 100.0)	100.0% (100.0 - 100.0)	
	%	62.4% (60.8 - 64.0)	37.6% (36.0 - 39.2)	

Wald Chi-Square Value	Degrees of Freedom	p-value
48.98	1	<0.0001

The third table, as shown here, shows results only for females

In 2018, among female former/never smokers living in Nebraska, the prevalence of permanent tooth extraction was 35.0%

In 2018, among female current smokers living in Nebraska, the prevalence of permanent tooth extraction was 53.1%

The Chi-Square p-value (<0.0001) suggests a significant association between smoking status and permanent tooth extraction in females



Controlling for		Gender (SEX1) = Don't know/Not Sure		
Sample Size				
Row %	(95% Confidence Interval)			
Column %	(95% Confidence Interval)			
Total % (Weighted)	(95% Confidence Interval)			
		Calculated variable for adults aged 18+ who have had permanent teeth extracted (_EXTETH3)		
		No	Yes	Total
		Calculated variable for adults who are current smokers (_RFSMOK3)		
Former smoker or never smoked	n	2	2	4
	Row%	*	*	100.0% (100.0 - 100.0)
	Col%	*	*	*
	%	*	*	*
Current smoker	n	0	1	1
	Row%	*	*	100.0% (100.0 - 100.0)
	Col%	*	*	*
	%	*	*	*
Total	n	2	3	5
	Row%	*	*	
	Col%	100.0% (100.0 - 100.0)	100.0% (100.0 - 100.0)	
	%	*	*	
		Wald Chi-Square Value	Degrees of Freedom	p-value
		0.69	1	0.4056

\* Estimate not available if the unweighted sample size for the denominator was < 50 or the Relative Standard Error (RSE) is > 0.3.

Controlling for		Gender (SEX1) = Refused		
Sample Size				
Row %	(95% Confidence Interval)			
Column %	(95% Confidence Interval)			
Total % (Weighted)	(95% Confidence Interval)			
		Calculated variable for adults aged 18+ who have had permanent teeth extracted (_EXTETH3)		
		No	Yes	Total
		Calculated variable for adults who are current smokers (_RFSMOK3)		
Former smoker or never smoked	n	4	1	5
	Row%	*	*	100.0% (100.0 - 100.0)
	Col%	*	*	*
	%	*	*	*
Current smoker	n	0	0	0
	Row%	*	*	*
	Col%	*	*	*
	%	*	*	*
Total	n	4	1	5
	Row%	*	*	
	Col%	100.0% (100.0 - 100.0)	100.0% (100.0 - 100.0)	
	%	*	*	
		Wald Chi-Square Value	Degrees of Freedom	p-value
		-	-	-

- Estimate could not be computed. Some estimates cannot be computed when the denominator sample size is 0, though other causes may exist as well.

As the variable “Gender” had 4 categories including male, female, don’t know/not sure and refused, we get the cross tabulations for the last two categories.

We can ignore these tables as the sample size is too small for any estimate computation. These categories may be converted to missing data in a proper statistical software. This online analysis tool does not provide an option for recoding the variables



Behavioral Risk Factor Surveillance

nccd.cdc.gov/weat/#/analysis

SEARCH

CDC

Centers for Disease Control and Prevention

CDC 24/7: Saving Lives, Protecting People™

BRFSS Web Enabled Analysis Tool

CDC > BRFSS > WEAT

BRFSS WEAT

Analysis

Cross Tabulation

Logistic Regression

Help

Related Information

The BRFSS is the world's largest telephone health survey in the United States. Information from the survey is used to monitor and prevent chronic diseases and injuries. The BRFSS Web Enabled Analysis Tool (WEAT) allows users to analyze the data.

BRFSS Analysis

Cross Tabulation

Logistic Regression

A cross tabulation, or "crosstab," produces frequency tables. For example, one can use the cross-tabulation procedure in the BRFSS to generate a table showing numbers and percentages of respondents with diabetes by age group. A general formula for cross tabulation can be depicted as  $A \times B$ , where A is the dependent variable or outcome (e.g., diabetes) and B is the independent or exposure variable (e.g., age). For our purposes, the cross tabulation includes frequencies or percentages for a single variable.

Logistic Regression

Logistic regression is a calculation of the contribution of one or more predictors on a particular outcome, such as "Risk factor:At risk for binge drinking." The results provide a predictive model and can be converted to log odds. The basic logistic formula using one predictor is  $Y = \exp(a + B1X1) / 1 + \exp(a + B1X1)$ .

About CDC

Employment

Newsroom

Training/Education

Legal

Link to Us

Policies

FOIA

Contact CDC

Centers for Disease Control and Prevention


1600 Clifton Road

Atlanta, GA 30329-4027 USA

800-CDC-INFO (800-232-4636)

DEPARTMENT OF HEALTH & HUMAN SERVICES (HHS)

**To run logistic regression, go to the starting webpage at <https://nccd.cdc.gov/weat/#/analysis> and click "Logistic Regression"**


  
 GREAT PLAINS
   
 IDeA | Clinical and
   
 Translational Research

45

Behavioral Risk Factor Surveillance System | nccd.cdc.gov/weat/#/logisticRegression/selectYear

**CDC** Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

SEARCH

## BRFSS Web Enabled Analysis Tool

[CDC](#) > [BRFSS](#) > [WEAT](#) > [Logistic Regression Years](#)

**BRFSS WEAT**

Analysis -

Cross Tabulation

Logistic Regression

Help +

**Related Information** +

### Logistic Regression

#### Select Year

☒ Single Year

Select Year ▼

- Select Year
- 2018
- 2017
- 2016
- 2015
- 2014
- 2013
- 2012
- 2011
- 2010
- 2009
- 2008
- 2007
- 2006
- 2005
- 2004

Drop down "Select year" and click "2018"

Click "Next"

1

2

Back Next

#### Help with this step...

Start building your logistic regression model by selecting one or more years.

You may choose to analyze a **Single Year** of data, **Combine** two or more years of data, or **Compare** data across two or more years.

#### My Analysis

Year(s)	
Location(s)	
Equation	=
Number Of Source	0

Behavioral Risk Factor Surveillance System

nccd.cdc.gov/weat/#/logisticRegression/selectLocation

**CDC** Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

SEARCH

## BRFSS Web Enabled Analysis Tool

[CDC](#) > [BRFSS](#) > [WEAT](#) > [Logistic Regression Years](#) > [Location](#)

**BRFSS WEAT**

- Analysis -
- Cross Tabulation
- Logistic Regression**
- Help +

**Related Information** +

### Logistic Regression

#### Select Location

☒ Single Location ☐ Compare

Select Location

- Iowa
- Kansas
- Kentucky
- Louisiana
- Maine
- Maryland
- Massachusetts
- Michigan
- Minnesota
- Mississippi
- Missouri
- Montana
- Nebraska**
- Nevada
- New Hampshire
- New Jersey
- New Mexico
- New York
- North Carolina
- North Dakota
- Ohio

Help

Select

You

Single Location of data, to **Combine** data from two or more states, or to **Compare** data across two or more states.

Back

**Next**

My Analysis

[Reset](#)

Year(s)	2018 (single)
Location(s)	
Equation	=
Number Of Source Records	437436

Drop down "Select Location" and click "Nebraska"

Click "Next"

1

2



Behavioral Risk Factor Surveillance System

nccd.cdc.gov/weat/#/logisticRegression/selectDependentVariable

**CDC** Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

SEARCH

## BRFSS Web Enabled Analysis Tool

[CDC](#) > [BRFSS](#) > [WEAT](#) > [Logistic Regression Years](#) > [Location](#) > [Dependent](#)

**BRFSS WEAT**

Analysis -

Cross Tabulation

Logistic Regression

Help +

Related Information +

### Logistic Regression

#### Select Dependent Variable (Required)

☒ Variable

Topic

Search All

Search All

Blind or visually impaired when wearing glasses (BLIND) (Disability)

Mental health status (\_MENT14D) (Healthy Days)

Physical health status (\_PHYS14D) (Healthy Days)

Physical/mental health kept you from doing usual activities (POOR14D)

Who have had permanent teeth extracted (\_EXTETH3) (Oral Health)

Who have any form of health care coverage (\_HCVU651) (Health Care Access)

Who have fully met the USPSTF recommendation (\_CRCREC)

Alcohol Consumption

Chronic Health Conditions

Colorectal Cancer Screening

Demographic Information

Disability

Drinking and Driving

Exercise

Health Care Access

Health Status

Healthy Days

HIV/AIDS

Immunization

Inadequate Sleep

Oral Health

Prostate Cancer Screening

Seatbelt Use

Tobacco Use

Women's Health

Back

Next

Help with Analysis [Reset](#)

Year(s)	2018 (single)
Location(s)	Nebraska (single)
Equation	=
Number Of Source Records	14634
Estimated Execution Time	2 Seconds

**Our dependent variable is permanent teeth extraction**

**Drop down "Topic" and click "Oral health"**



Behavioral Risk Factor Surveillance System | nccd.cdc.gov/weat/#/logisticRegression/selectDependentVariable

**CDC** Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

**BRFSS Web Enabled Analysis Tool**

[CDC](#) > [BRFSS](#) > [WEAT](#) > [Logistic Regression Years](#) > [Location](#) > [Dependent](#)

**BRFSS WEAT**

- Analysis -
- Cross Tabulation
- Logistic Regression
- Help +

**Related Information +**

### Logistic Regression

#### Select Dependent Variable (Required)

☒ Variable

Topic:

Calculated variable for adults aged 18+ who have had permanent teeth extracted (**\_EXTETH3**)

Calculated variable for adults aged 65+ who have had all their natural teeth extracted (**\_ALTETH3**)

Calculated variable for adults who have visited a dentist, dental hygienist or dental clinic within the past year (**\_DENVST3**)

**1**

**2**

**Click "Calculated variable for adults aged 18+ who have had permanent teeth extracted"**

**Click "Next"**

Back Next

#### My Analysis

[Reset](#)

Year(s)	2018 (single)
Location(s)	Nebraska (single)
Equation	Oral Health: Calculated variable for adults aged 18+ who have had permanent teeth extracted ( <b>_EXTETH3</b> ) =
Number Of Source	14634

**BRFSS WEAT**

Analysis

Cross Tabulation

Logistic Regression

Help

Related Information

### Logistic Regression

Select Independent Variable(s)

Variable

Topic: Search All

Search All

Alcohol Consumption

Chronic Health Conditions

Colorectal Cancer Screening

Demographic Information

Disability

Drinking and Driving

Exercise

Falls

Health Care Access

Health Status

Healthy Days

HIV/AIDS

Immunization

Inadequate Sleep

Oral Health

Prostate Cancer Screening

Seatbelt Use

**Tobacco Use**

Women's Health

1

**Our primary independent variable (exposure) is current tobacco use and we have age as another independent variable**

**Drop down “Topic” and click “Tobacco Use”**

**Click “Calculated variable for adults who are current smokers”**

**Click “Add”**

**BRFSS WEAT**

Analysis

Cross Tabulation

Logistic Regression

Help

Related Information

### Logistic Regression

Select Independent Variable(s)

Variable

Topic: Tobacco Use

Calculated variable for 4 level smoker status: everyday/someday/former/non-smoker (SMOKE)

**Calculated variable for adults who are current smokers (RFSMOK3)**

Quit smoking for 1 day or more in past 12 months (STOPSMK2)

3

Add

Selected Variables

**Logistic Regression**

Select Independent Variable(s)

☐ Variable

Topic: Tobacco Use

Search All

Alcohol Consumption

Chronic Health Conditions

Colorectal Cancer Screening

**Demographic Information**

Disability

Drinking and Driving

Exercise

Falls

Health Care Access

Health Status

Healthy Days

HIV/AIDS

Immunization

Inadequate Sleep

Oral Health

Prostate Cancer Screening

Seatbelt Use

Tobacco Use

Women's Health

Add

Selected Variables

Calculated variable for adults who are current smokers (RFSMOK3)

Drop down "Topic" and click  
"Demographic Information"

Click "Gender"

Click "Add"

Click "Next"

**Logistic Regression**

Select Independent Variable(s)

☐ Variable

Topic: Demographic Information

Calculated variable for reported weight in kilograms (WTG3 continuous)

Calculated variable for reported weight in kilograms (WTG3, categorical)

Calculated variable for white non-hispanic race group (RACEG21)

Education level (EDUCA)

Employment Status (EMPLOY1)

**Gender (SEX1)**

Marital status (MARITAL)

Pregnancy status (PREGNANT)

Add

Selected Variables

Calculated variable for adults who are current smokers (RFSMOK3)

Remove Remove All

Back Next

Help with this step... My Analysis [Reset](#)

Behavioral Risk Factor Surveillance

nccd.cdc.gov/weat/#/logisticRegression/setReference

CDC Centers for Disease Control and Prevention  
CDC 24/7: Saving Lives, Protecting People™

BRFSS Web Enabled Analysis Tool

CDC > BRFSS > WEAT > Logistic Regression Years > Location > Dependent > Independent > Reference

BRFSS WEAT

Analysis

Cross Tabulation

### Logistic Regression

#### Set Reference Category

This procedure allows you to calculate the effect of one or more predictors, or independent variables, on a dependent variable that has two levels (for example, "at risk" and "not at risk").

Select the Category for the Dependent Variable

Dependent Variable (Calculated variable for adults aged 18+ who have had permanent teeth extracted (\_EXTETH3))

☐ No

☒ Yes

Select the Category for the Independent Variables

1st independent Variable (Calculated variable for adults who are current smokers (\_RFSMOK3))

☒ Former smoker or never smoked

☐ Current smoker

2nd independent Variable (Gender (SEX1))

☒ Male

☐ Female

☐ Don't know/Not Sure

☐ Refused

Back

Run Report

Change the category for dependent variable to "Yes"

1

Any of the categories for the independent variables can be selected as references. We will select former smoker and male as reference

Click "Run Report"

2



Behavioral Risk Factor Surveillance System  
Logistic Regression, Nebraska, 2018

Dependent Variable: Oral Health: Calculated variable for adults aged 18+ who have had permanent teeth extracted (\_EXTETH3)

Independent Variables: Tobacco Use: Calculated variable for adults who are current smokers (\_RFSMOK3)

Demographic Information: Gender (SEX1)

	Sample Size	Beta	Standard Error	t-statistic	p-value for t-statistic	Odds Ratio	95% CI for Odds Ratio
Calculated variable for adults aged 18+ who have had permanent teeth extracted (_EXTETH3), Modeling category: Yes							
Intercept	13,945	-0.62	0.04	-15.52	<0.0001	0.54	(0.50 - 0.58)
Calculated variable for adults who are current smokers (_RFSMOK3)							
Former smoker or never smoked (REF)	12,061	0.00	0.00	.	.	1.00	(1.00 - 1.00)
Current smoker	1,884	0.68	0.07	9.30	<0.0001	1.98	(1.71 - 2.28)
Gender (SEX1)							
Male (REF)	6,258	0.00	0.00	.	.	1.00	(1.00 - 1.00)
Female	7,677	0.01	0.05	0.14	0.8921	1.01	(0.91 - 1.11)
Don't know/Not Sure	5	2.15	1.08	1.99	0.0465	8.62	(1.03 - 71.92)
Refused	5	0.06	1.25	0.05	0.9608	1.06	(0.09 - 12.25)

Model test statistics

Contrast	Degrees of Freedom	Wald chi-square	p-value for Wald chi-square
Overall Model	5	499.18	<0.0001
Model minus intercept	4	90.59	<0.0001
Intercept	.	.	.
Calculated variable for adults who are current smokers (_RFSMOK3)	1	86.48	<0.0001
Gender (SEX1)	3	3.97	<0.0001

The value of Odds Ratio against “Current smoker” is 1.98 and the p-value is <0.0001.

The odds of permanent teeth extraction is 1.98 times higher for current smokers compared to former/never smokers, after adjusting for gender.

As p-value is less than 0.05, the association is statistically significant.

The odds ratio against “Female” is 1.01 and the p-value is also insignificant (0.8921). So there is no significant difference in permanent tooth extraction for females compared to males after adjusting for smoking status.

As the sample sizes for gender categories of “Don’t know/Not sure” and “Refused” are very small, the results associated with them are not precise. This is evident from the very broad confidence intervals

# DOWNLOADING BRFSS DATA

CDC - 2018 BRFSS Survey Data

cdc.gov/brfss/annual\_data/annual\_2018.html

2018 Weighting Formula CDC [PDF - 473 KB]

this page, enter your email address:

Email Address

What's this? Submit


### Related Links

[Health Indicator Sortable Stats](#)

[HealthData.gov](#)

[Data Catalog](#)

[CDC's Major Disease Surveillance Systems](#)



### Data Files

There are 437,436 records for 2018. More information on participation is available in the [states conducting surveillance, by year table](#). The data files are provided in ASCII and SAS Transport formats. The November update includes the addition of E-Cigarettes optional module data from California and a correction for the Lung Cancer Screening optional module variable LCSLAST in two states (MD, TX).

[2018 BRFSS Data \(ASCII\)](#) [ZIP - 66.2 MB]  
November, 2019  
This file for the combined landline and cell phone data set is in ASCII format. It has a fixed record length of 2033 positions.

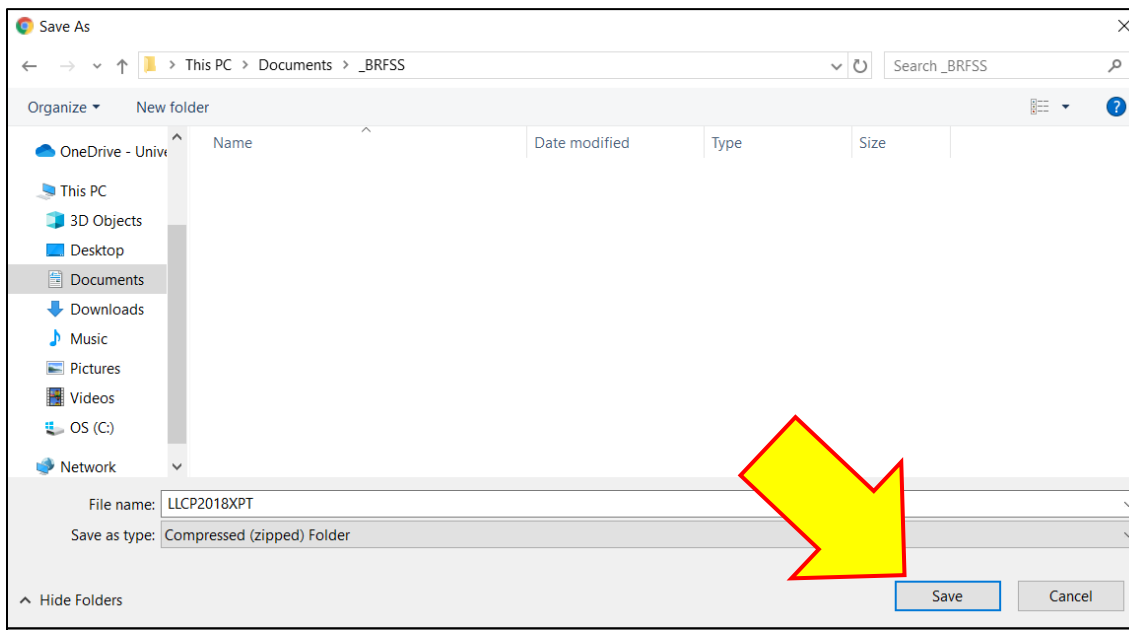
[2018 BRFSS Data \(SAS Transport Format\)](#) [ZIP - 101 MB]  
November, 2019  
This file for the combined landline and cell phone data set was exported from SAS V9.3 in the XPT transport format. This file contains 275 variables. This format can be imported into SPSS or STATA. Please note: some of the variable labels get truncated in the process of converting to the XPT format so they may be slightly different from what is on the SASOUT18.SAS program.

[Variable Layout](#)  
Format information on variable name by column position.

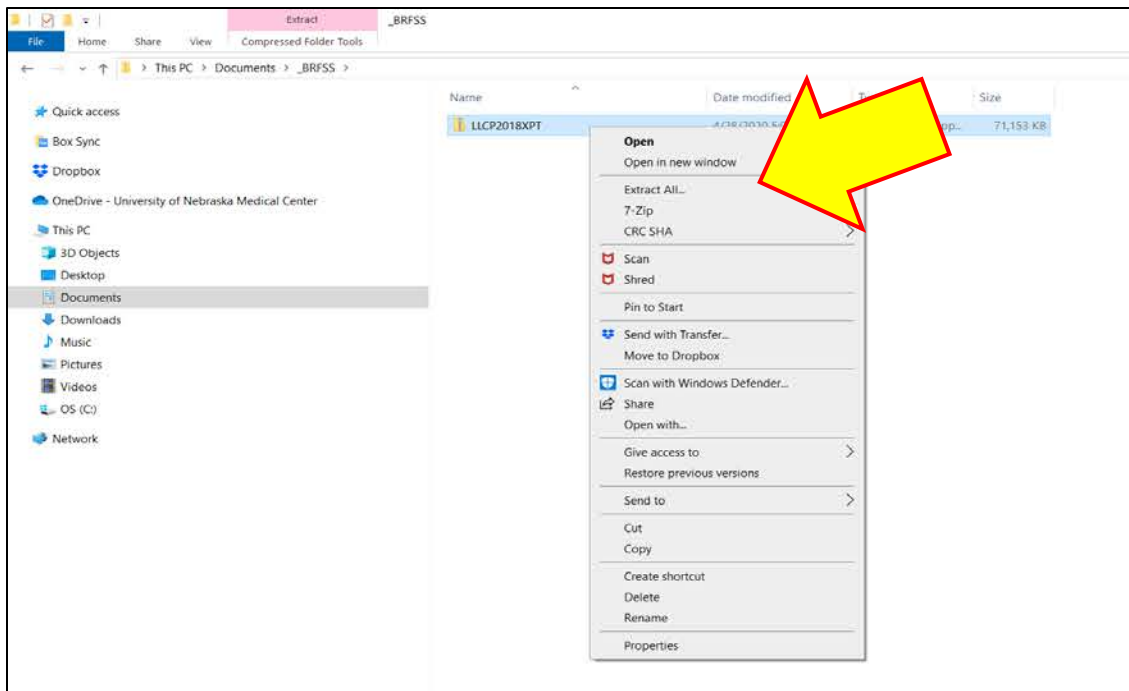
[The Combined Landline and Cellular Telephone Survey: Multiple Questionnaire Version Data-includes Optional Modules](#)  
The combined landline and cellular telephone multiple version questionnaire data sets contain the data from the states which conducted more than one version of their landline and cellular telephone survey questionnaires and used optional modules in 2018.

SAS Resources

Go to the webpage,  
“[https://www.cdc.gov/brfss/annual\\_data/annual\\_2018.html](https://www.cdc.gov/brfss/annual_data/annual_2018.html)”  
and select “2018 BRFSS Data (SAS Transport Format)”




**Select destination folder and save the zip file**



**Right click the zip folder and Extract All**



18XPT

Name	Date modified
 LLCP2018	4/28/2020 5:55 PM

**Within the extracted folder you will have this SAS export transport file**

# GETTING STARTED WITH SAS

CDC - 2018 BRFSS Survey Data | X +

← → ↻ cdc.gov/brfss/annual\_data/annual\_2018.html 🔍 ☆ 📄 📅 📱 👤 ⋮

### SAS Resources

The following files are included for use with SAS software developed by SAS Institute, Inc. The SAS files are programs that can be read using any text editor. Save SAS files as plain text files.

[Format18.sas CDC](#) [SAS – 685 KB]  
SAS for Windows program used to generate the 2018 format library.

[Formats18.\[SAS7BCAT Direct Download – 493 KB\]](#)  
Format library used with the 2018 BRFSS SAS Data File in 32 bit SAS.

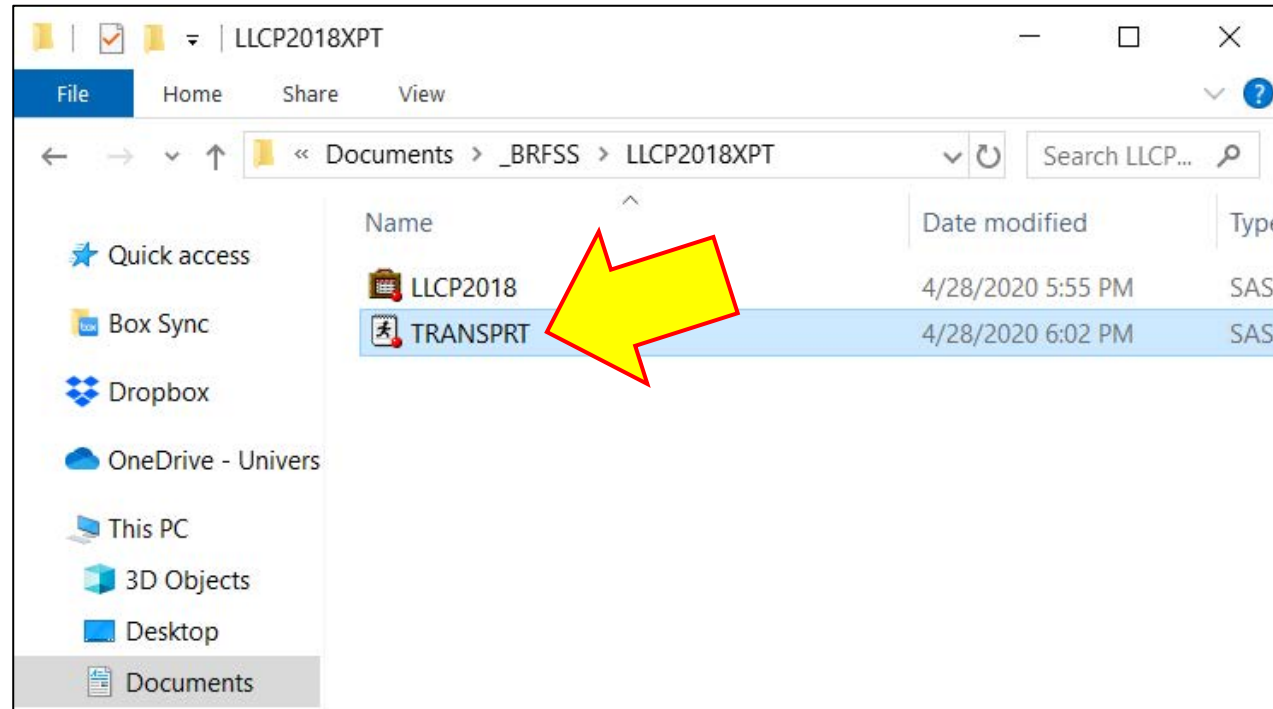
[Formats18.\[SAS7BCAT Direct Download – 493 KB\]](#)  
Format library used with the 2018 BRFSS SAS Data File in 64 bit SAS.

[SASOUT18\\_LLCP.SAS \[SAS – 34 KB\]](#)  
SAS for Windows program to convert the SAS file into a SAS dataset.

[TRANSPRT.SAS \[SAS – 3 KB\]](#)  
SAS for Windows program to convert the XPT (SAS Transport) file into a SAS dataset.

[Formas18.sas CDC](#) [SAS – 64 KB]  
SAS for Windows format assignment statements.

Go again to the webpage,  
“[https://www.cdc.gov/brfss/annual\\_data/annual\\_2018.html](https://www.cdc.gov/brfss/annual_data/annual_2018.html)”  
scroll down and click,  
“TRANSPORT.SAS [SAS – 3 KB]”.  
Save this file in the same folder  
as your data



**First, we will convert the SAS Xport transport file to permanent SAS data set**

**Double click on the SAS system program, “TRANSPRT” that we downloaded from the BRFSS website. This will open the SAS program**



**We will have to make changes to two SAS commands in the already provided SAS codes**

```

* DEFINE SAS ENVIRONMENT OPTIONS *
*****;
OPTIONS PAGESIZE=1 NOFMERR;
*****;
* CLEAR EXISTING TITLES AND FOOTNOTES *
*****;
TITLE ;
FOOTNOTE ;
RUN ;

* INSTRUCTIONS: *
* SPECIFY PATH OF WHERE TRANSPORT DATAFILE IS STORED *
* SPECIFY PATH OF WHERE SAS DATASET IS TO BE STORED *
*****;

*LIBNAME TRANSPRT XPORT '<FOLDER WHERE SAS TRANSPORT FILE IS STORED>';
*****;

* EXAMPLE: *
* LIBNAME TRANSPRT XPORT 'C:\BRFSS\2013\LLCP2013.XPT' *
* NOTE: MAKE SURE THE ASTERISK IS REMOVED BEFORE THE WORD LIBNAME AND *
* MAKE SURE THAT THE GREATER THAN AND LESS THAN SIGNS < > ARE REMOVED *
* FROM THE FILENAME STATEMENT *
* THE QUOTATION MARKS MUST BE THERE FOR THE PROGRAM TO WORK *
*****;

*LIBNAME DATAOUT V7 '<PATH OF PERMANENTLY STORED SAS DATASET VERSION OF DATAIN FILEREF>' ;
*****;

* EXAMPLE: *
* LIBNAME DATAOUT 'C:\' *
* A SAS DATABASE WILL BE STORED AT C:\SASDATA.SAS7BDAT *
* NOTE: MAKE SURE THE ASTERISK IS REMOVED BEFORE THE WORD LIBNAME AND *
* MAKE SURE THAT THE GREATER THAN AND LESS THAN SIGNS < > ARE REMOVED *
* FROM THE LIBNAME STATEMENT *
* THE QUOTATION MARKS MUST BE THERE FOR THE PROGRAM TO WORK *
*****;

PROC COPY IN=TRANSPRT OUT=DATAOUT;
RUN;

```

**We will specify the path where our SAS Xport transport file is saved along with the name of the file. See the example below the command**

**We will specify the path of the folder where we want to save the permanent SAS data set. See the example below the command**

```

*****
OPTIONS PAGENO=1 NOFMterr;
*****
* CLEAR EXISTING TITLES AND FOOTNOTES *
*****
TITLE ;
FOOTNOTE ;
RUN ;
*****
* INSTRUCTIONS:
* SPECIFY PATH OF WHERE TRANSPORT DATAFILE IS STORED
* SPECIFY PATH OF WHERE SAS DATASET IS TO BE STORED
*****;

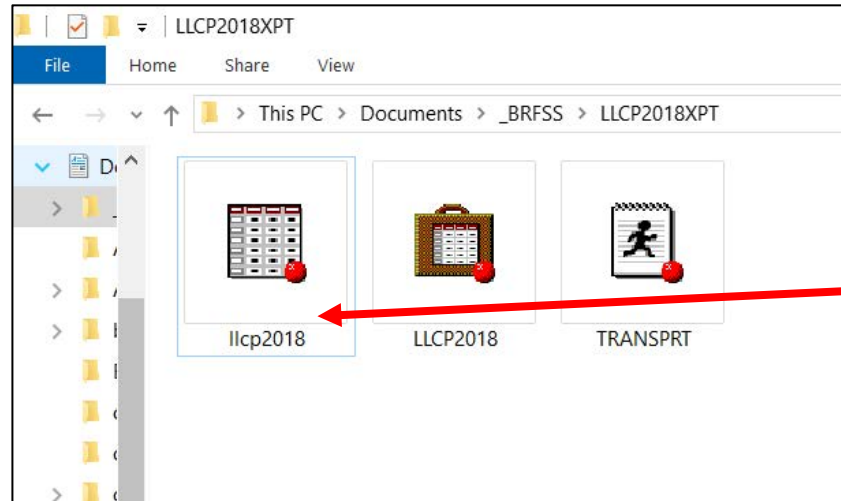
LIBNAME TRANSPRT XPORT 'C:\Users\zaeem\Documents\_BRFSS\LLCP2018XPT\LLCP2018.XPT';
*****
* EXAMPLE:
* LIBNAME TRANSPRT XPORT 'C:\BRFSS\2013\LLCP2013.XPT'
* NOTE: MAKE SURE THE ASTERISK IS REMOVED BEFORE THE WORD LIBNAME AND
* MAKE SURE THAT THE GREATER THAN AND LESS THAN SIGNS < > ARE REMOVED
* FROM THE FILENAME STATEMENT
* THE QUOTATION MARKS MUST BE THERE FOR THE PROGRAM TO WORK
*****;

LIBNAME DATAOUT V7 'C:\Users\zaeem\Documents\_BRFSS\LLCP2018XPT' ;
*****
* EXAMPLE:
* LIBNAME DATAOUT 'C:\'
* A SAS DATABASE WILL BE STORED AT C:\SASDATA.SAS7BDAT
* NOTE: MAKE SURE THE ASTERISK IS REMOVED BEFORE THE WORD LIBNAME AND
* MAKE SURE THAT THE GREATER THAN AND LESS THAN SIGNS < > ARE REMOVED
* FROM THE LIBNAME STATEMENT
* THE QUOTATION MARKS MUST BE THERE FOR THE PROGRAM TO WORK
*****;

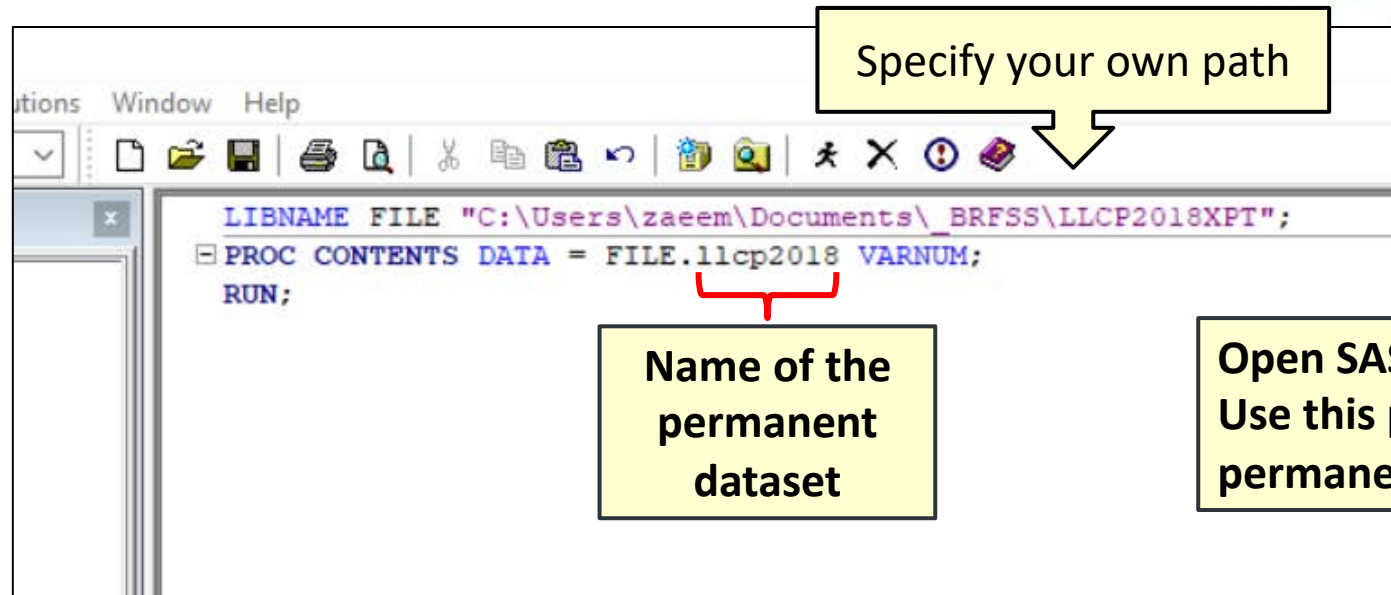
PROC COPY IN=TRANSPRT OUT=DATAOUT;
RUN;
  
```

- See the changes that have been made
- Specify your own folder pathways
- Make sure to remove the Asterix
- Make sure to remove <>

Select the code and click the running man.  
Close the SAS program.



**Permanent SAS dataset has been saved to the folder**



**Specify your own path**

**Name of the permanent dataset**

**Open SAS**  
Use this program to open the permanent data set



# SAS CODE FOR CROSS-TABULATION

```
*READ THE PERMANENT SAS DATASET INTO SAS;
LIBNAME FILE "C:\Users\zaeem\Documents\_BRFSS\LLCP2018XPT";
*PROC CONTENTS SHOWS ALL THE VARIABLES IN THE DATASET;
PROC CONTENTS DATA = FILE.llcp2018 VARNUM;
RUN;

*CREATING A NEW TEMPORARY DATASET FROM PERMANENT DATA SET AND RECODING TO ASSIGN MISSING
VALUES TO MATCH THE RESULTS FROM WEAT TOOL;

DATA BRFSS;
SET FILE.llcp2018;
IF SEX1=9 THEN SEX1=.;
IF _RFSMOK3=9 THEN _RFSMOK3=.;
IF _EXTETH3 = 9 THEN _EXTETH3=.;RUN;

*CROSS TABULATION BETWEEN CURRENT SMOKING AND PERMANENT TEETH EXTRACTION BY GENDER;
PROC SURVEYFREQ DATA = BRFSS;
WHERE _STATE = 31; *SPECIFIES STATE = NEBRASKA (31 IS THE FIPS CODE);
CLUSTER _PSU; *SPECIFIES CLUSTER VARIABLE;
STRATA _ststr; *SPECIFIES THE STRATUM VARIABLE;
WEIGHT _llcpwt; *SPECIFIES THE WEIGHT VARIABLE;

TABLE SEX1* RFSMOK3* EXTETH3/CHISQ COL ROW;*PRODUCES THE CROSS-TABULATION BY SEX WITH
CHI SQUARE TEST. SEX1 IS SEX, _RFSMOK3 IS CURRENT SMOKING AND _EXTETH3 IS EXTRACTION OF
PERMANENT TEETH;

run;
```



Table of _RFSMOK3 by _EXTETH3										
Controlling for SEX1=1										
_RFSMOK3	_EXTETH3	Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent	Row Percent	Std Err of Row Percent	Column Percent	Std Err of Col Percent
1	1	2978	368277	8818	53.4669	0.9332	64.7066	0.9432	86.1449	0.8941
	2	2338	200871	5977	29.1628	0.8070	35.2934	0.9432	76.8782	1.2229
	Total	5316	569148	9757	82.6297	0.7288	100.000			
2	1	394	59232	4071	8.5993	0.5706	49.5060	2.3216	13.8551	0.8941
	2	548	60414	3627	8.7710	0.5144	50.4940	2.3216	23.1218	1.2229
	Total	942	119645	5353	17.3703	0.7288	100.000			
Total	1	3372	427508	9350	62.0663	0.8835			100.000	
	2	2886	261285	6778	37.9337	0.8835			100.000	
	Total	6258	688794	10230	100.000					

Rao-Scott Chi-Square Test	
Pearson Chi-Square	88.1499
Design Correction	2.2828
Rao-Scott Chi-Square	38.6144
DF	1
Pr > ChiSq	<.0001
F Value	38.6144
Num DF	1
Den DF	13878
Pr > F	<.0001
Sample Size = 13940	

Here is one of the resulting tables. The estimates are for men (SEX=1)

We can compare these results to the results of the table produced by WEAT Tool as they are similar

We use proc surveyfreq instead of the traditional proc freq because it takes strata and clusters into account

The variable names and details can be looked up in the [code book](#) for BRFSS 2018 data

# SAS CODE FOR LOGISTIC REGRESSION

```
*READ THE PERMANENT SAS DATASET INTO SAS;

LIBNAME FILE "C:\Users\zaeem\Documents\_BRFSS\LLCP2018XPT";

*PROC CONTENTS SHOWS ALL THE VARIABLES IN THE DATASET;

PROC CONTENTS DATA = FILE.llcp2018 VARNUM;

RUN;

*CREATING A NEW TEMPORARY DATASET FROM PERMANENT DATA SET AND RECODING TO ASSIGN MISSING
VALUES TO MATCH THE RESULTS FROM WEAT TOOL;

DATA BRFSS;

SET FILE.llcp2018;

IF SEX1=9 THEN SEX1=.;

IF _RFSMOK3=9 THEN _RFSMOK3=.;

IF _EXTETH3 = 9 THEN _EXTETH3=.;RUN;

*LOGISTIC REGRESSION PERMANENT TEETH EXTRACTION AS DEPENDENT VARIABLE AND CURRENT
SMOKING AND GENDER AS INDEPENDENT VARIABLES;

PROC SURVEYLOGISTIC DATA= BRFSS;

WHERE _STATE = 31;

CLUSTER _PSU;

STRATA _STSTR;

WEIGHT _llcpwt;

CLASS _RFSMOK3 (REF="1") SEX1 (REF='1'); *SPECIFYING REFERENCE CATEGORIES. NO CURRENT
SMOKING AND MALES GENDER ARE REFERENCE CATEGORIES;

MODEL _EXTETH3 (EVENT='2')= _RFSMOK3 SEX1; *MODEL WITH TEETH EXTRACTION AS DEPENDENT
VARIABLE AND CURRENT SMOKING AND GENDER AS INDEPENDENT VARIABLES;

RUN;
```

Analysis of Maximum Likelihood Estimates					
Parameter		Estimate	Standard Error	t Value	Pr >  t
Intercept		0.4448	0.3618	1.23	0.2190
_RFSMOK3	2	0.3410	0.0367	9.30	<.0001
SEX1	2	-0.7134	0.3615	-1.97	0.0484
SEX1	7	1.4339	0.7212	1.99	0.0468
NOTE: The degrees of freedom for the t tests is 13878.					

Odds Ratio Estimates			
Effect	Point Estimate	95% Confidence Limits	
_RFSMOK3 2 vs 1	1.978	1.713	2.284
SEX1 2 vs 1	1.007	0.910	1.114
SEX1 7 vs 1	8.622	1.034	71.928
NOTE: The degrees of freedom in computing the confidence limits is 13878.			

The results are shown to the left.

We can compare these results to the results of the logistic regression produced by WEAT Tool as they are similar.

P value for gender is different because we removed one category of gender that was included by the WEAT tool (SEX1=9)


We use proc surveylogistic instead of the traditional proc logistic because it takes strata and clusters into account

Similarly, we use proc surveyreg for linear regression

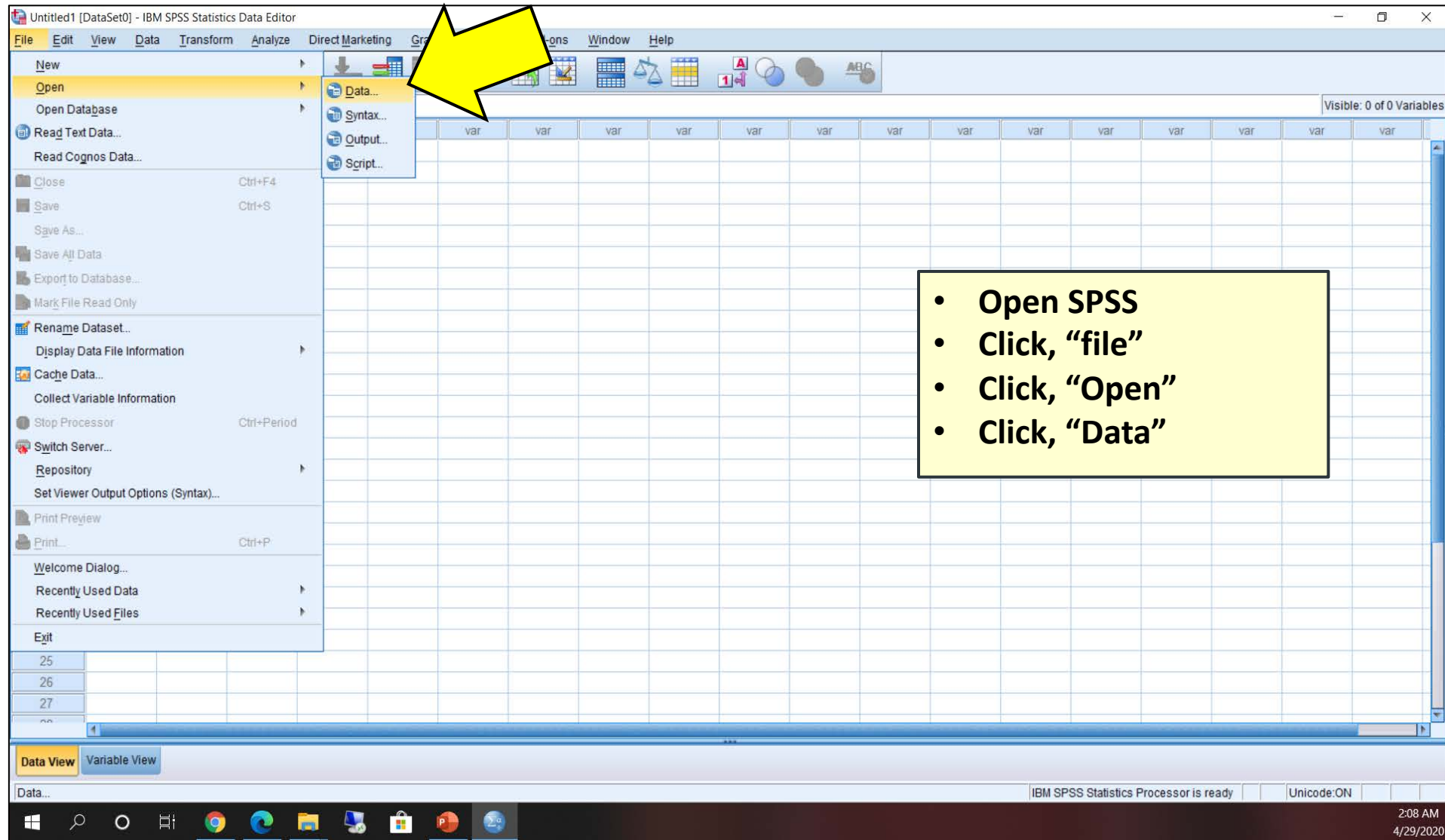
# GETTING STARTED WITH SPSS



18XPT

Name	Date modified
 LLCP2018	4/28/2020 5:55 PM

**Please review the section, “Downloading BRFFS DATA” starting sliding number 55 as we will start with opening this data set into SPSS**



Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Visible: 0 of 0 Variables

- Browse to the file where your data is saved
- Drop down, “Files of types” and select, “SAS”

Open Data

Look in: LLCP2018XPT

File name:

Files of type: SPSS Statistics (\*.sav)

Encoding:

Minimize

Excel (\*.xls, \*.xlsx, \*.xlsm)  
Lotus (\*.w\*)  
Sylk (\*.slk)  
dBase (\*.dbf)  
SAS (\*.sas7bdat, \*.sd7, \*.sd2, \*.ssd01, \*.ssd04, \*.xpt)  
Stata (\*.dta)  
Text (\*.txt, \*.dat, \*.csv, \*.tab)  
All Files (\*.\*)

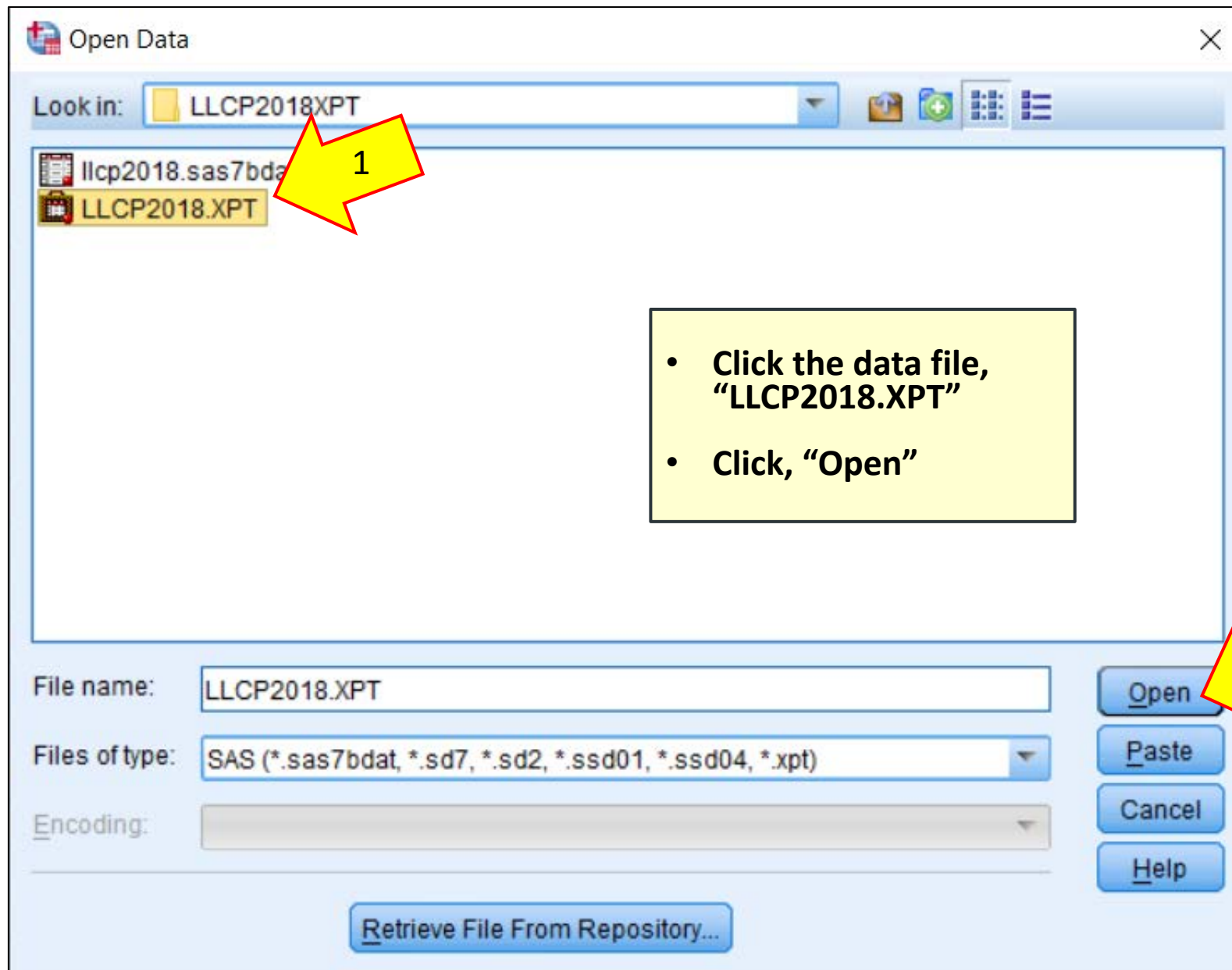
Open  
Paste  
Cancel

Data View Variable View

Data...

IBM SPSS Statistics Processor is ready Unicode:ON

2:11 AM  
4/29/2020





\*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Define Variable Properties...  
Set Measurement Level for Unknown...  
Copy Data Properties...  
New Custom Attribute...  
Define Dates...  
Define Multiple Response Sets...  
Validation  
Identify Duplicate Cases...  
Identify Unusual Cases...  
Compare Datasets...  
Sort Cases...  
Sort Variables...  
Transpose...  
Merge Files  
Restructure...  
Rake Weights...  
Propensity Score Matching...  
Case Control Matching...  
Aggregate...  
Split into Files  
Orthogonal Design  
Copy Dataset  
Split File...  
**Select Cases...**  
Weight Cases...

Visible: 275 of 275 Variables

	IMONTH	IDAY	IYEAR	DISPCODE	SEQNO	@_PSU	CTELEM1	PVTRES1	COLG1
1	01	05	2018	1100	2018000001	2018000001	1	1	
2	01	12	2018	1100	2018000002	2018000002	1	1	
3	01	08	2018	1100	2018000003	2018000003	1	1	
4	01	03	2018	1100	2018000004	2018000004	1	1	
5	01	12	2018	1100	2018000005	2018000005	1	1	
6	01	11	2018	1100	2018000006	2018000006	1	1	
7	01	10	2018				1	1	
8	01	13	2018				1	1	
9	01	09	2018				1	1	
10	01	10	2018				1	1	
11	01	10	2018				1	1	
12	01	03	2018				1	1	
13	01	11	2018				1	1	
14	01	12	2018				1	1	
15	01	25	2018				1	1	
16	01	19	2018				1	1	
17	01	10	2018				1	1	
18	01	18	2018				1	1	
19	01	11	2018				1	1	
20	01	18	2018				1	1	
21	05	05	2018	1100	2018000021	2018000021	1	1	
22	03	03	2018	1100	2018000022	2018000022	1	1	
23	10	10	2018	1100	2018000023	2018000023	1	1	
24	11	11	2018	1100	2018000024	2018000024	1	1	
25	11	11	2018	1100	2018000025	2018000025	1	1	
26	10	10	2018	1100	2018000026	2018000026	1	1	
27	03	03	2018	1100	2018000027	2018000027	1	1	

1

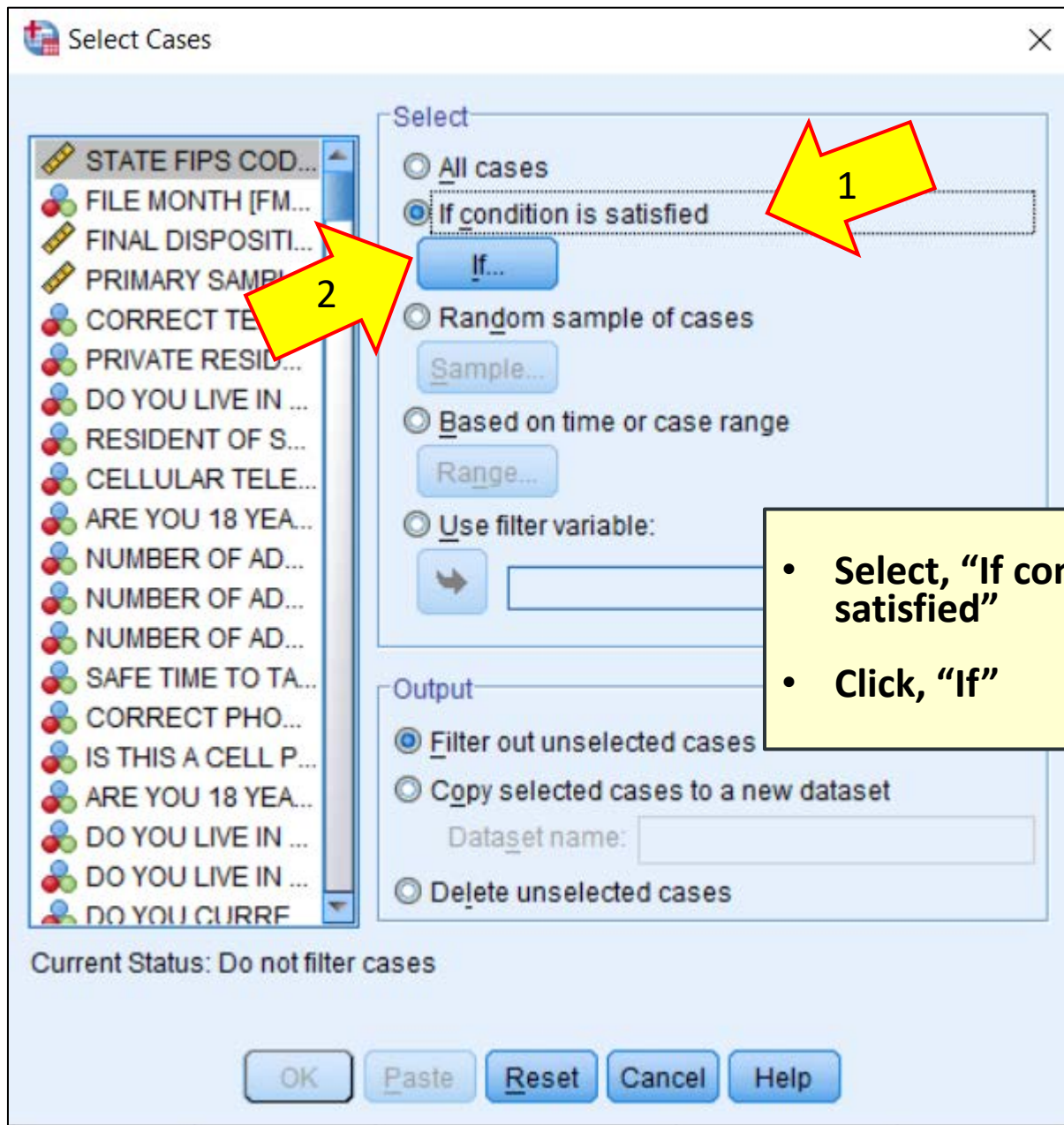
- First, we will restrict data to Nebraska
- Click, "Data"
- Click, "Select Cases"

Data View Variable View

Select Cases...

IBM SPSS Statistics Processor is ready Unicode:ON

2:19 AM  
4/29/2020



- Select, "If condition is satisfied"
- Click, "If"

Select Case

1

2

3

4

STATE FIPS COD...

FILE MONTH [FM...

INTERVIEW DAT...

INTERVIEW MON...

INTERVIEW DAY ...

INTERVIEW YEA...

FINAL DISPOSITI...

ANNUAL SEQUE...

PRIMARY SAMPL...

CORRECT TELE...

PRIVATE RESID...

DO YOU LIVE IN ...

RESIDENT OF S...

CELLULAR TELE...

ARE YOU 18 YEA...

NUMBER OF AD...

NUMBER OF AD...

NUMBER OF AD...

SAFE TIME TO TA...

CORRECT PHO...

IS THIS A CELL P...

ARE YOU 18 YEA...

DO YOU LIVE IN

@\_STATE=31

Function group:

All

al Variables:

Continue Cancel Help

- Select, "STATE FIPS CODE"
- Click the arrow to bring it over
- Put, "=31". This is FIPS code for Nebraska
- Click, "Continue"



Select Cases

Select

☐ All cases

☒ If condition is satisfied

If... @\_STATE = 31

☐ Random sample of cases

Sample...

☐ Based on time or case range

Range...

☐ Use filter variable:

Output

☒ Filter out unselected cases

☐ Copy selected cases to a new dataset

Dataset name:

☐ Delete unselected cases

Current State: Not filter cases

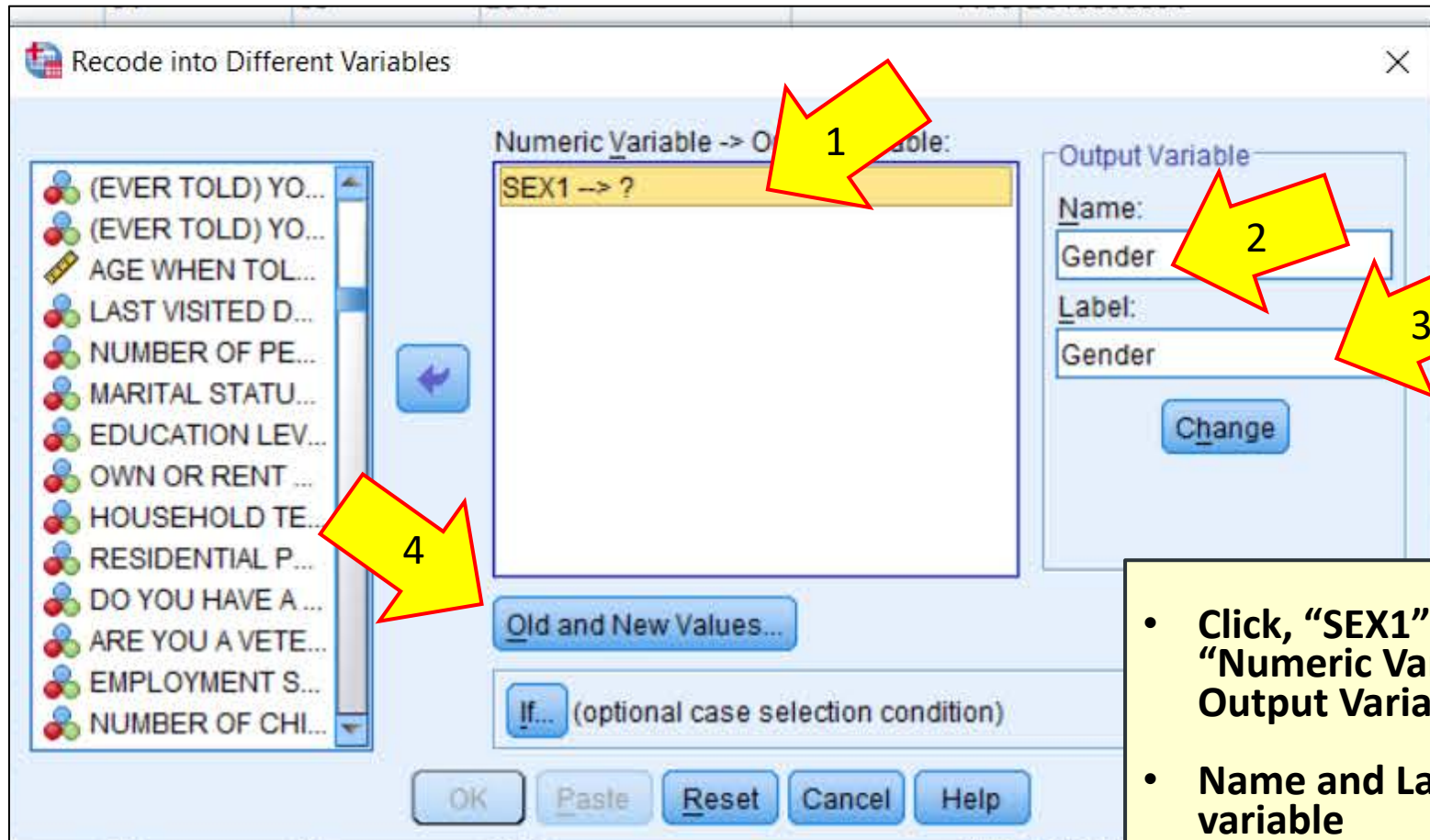
4

OK Paste Reset Cancel Help

- Click, "OK"







- Click, “SEX1” and bring it “Numeric Variable→ Output Variable”
- Name and Label the variable
- Click, “Old and New Values”

SPSS Recode into Different Variables: Old and New Values dialog box.

**Old Value:**

- ☒ Value: 1
- ☐ System-missing
- ☐ System- or user-missing

**New Value:**

- ☒ Value: 1
- ☐ System-missing
- ☐ Copy old value(s)

**Old --> New:**

Buttons: Add, Change, Remove

☐ Output variables are strings Width: 8

☐ Convert numeric strings to numbers ('5' -> 5)

Buttons: Continue, Cancel, Help

- The variable, “SEX1” has four categories; 1, 2, 7 and 9.
- We want to set 9 to missing and keep the rest as is.
- Write 1 under Old value and 1 under new value. This tells SPSS that we want category 1 as 1 for the new variable Gender.
- Click “Add”



SPSS Recode into Different Variables: Old and New Values

Old Value

☒ Value:

☐ System-missing

☐ System- or user-missing

New Value

☐ Value:

☒ System-missing

☐ Copy old value(s)

Old → New:

1	→	1
2	→	2
7	→	7
9	→	SYSMIS

Add Change Remove

☐ Output variables are strings Width: 8

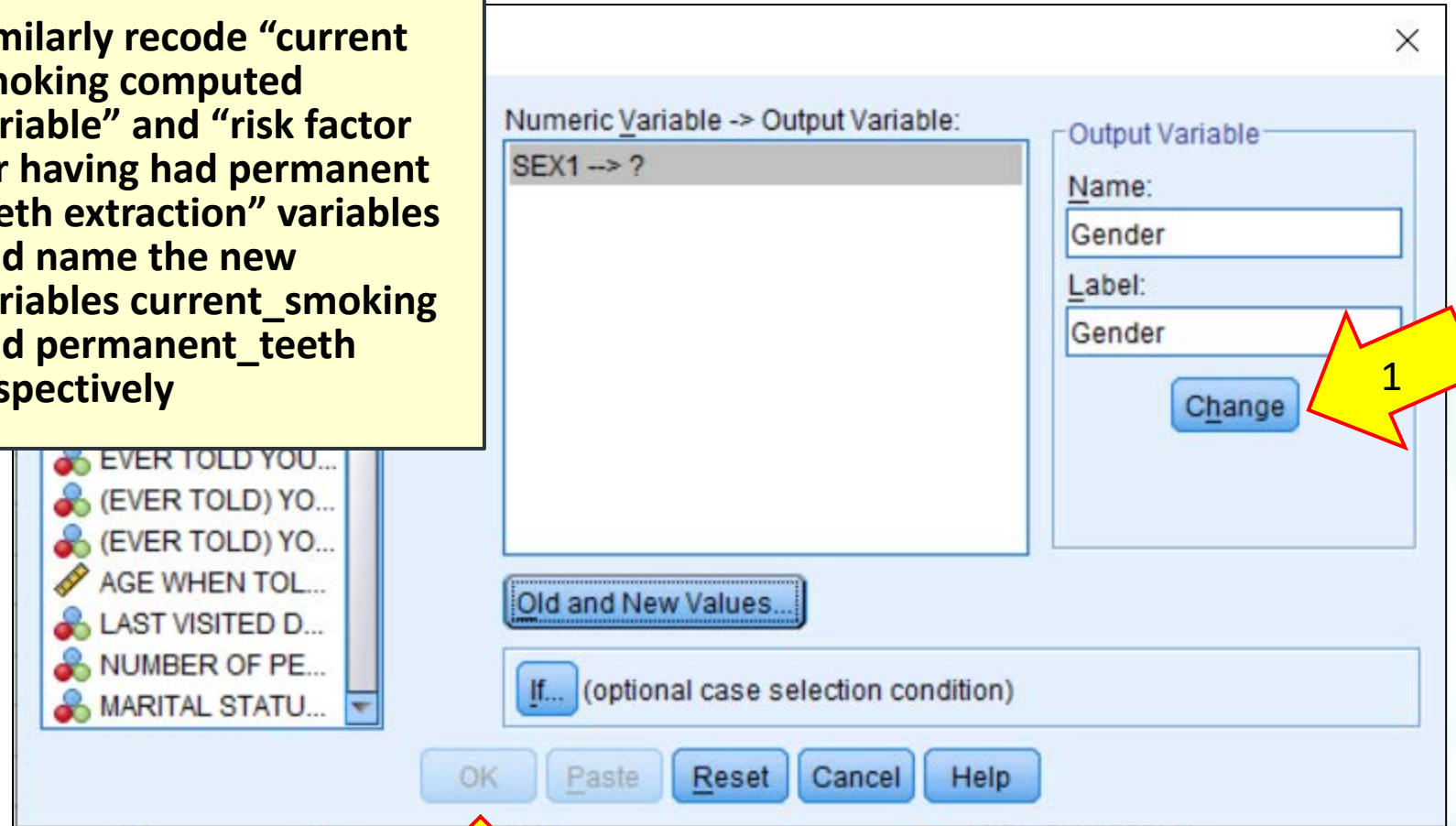
☐ Convert numeric strings to numbers (5'→5)

Continue Cancel Help

- Do the same for 2 and 7
- For “9”, instead of a value click system missing
- Click, “Continue”



- Click, “Change”
- Click, “OK”
- Similarly recode “current smoking computed variable” and “risk factor for having had permanent teeth extraction” variables and name the new variables current\_smoking and permanent\_teeth respectively



The image shows the 'Numeric Variable -> Output Variable' dialog box in SPSS. On the left, a list of variables is partially visible, including 'EVER TOLD YOU...', '(EVER TOLD) YO...', 'AGE WHEN TOL...', 'LAST VISITED D...', 'NUMBER OF PE...', and 'MARITAL STATU...'. The main area of the dialog has 'SEX1 --> ?' in the 'Numeric Variable -> Output Variable:' field. To the right, the 'Output Variable' section contains 'Name:' with 'Gender' entered, and 'Label:' with 'Gender' entered. A yellow arrow labeled '1' points to the 'Change' button in this section. At the bottom of the dialog, there are buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help'. A yellow arrow labeled '2' points to the 'OK' button.

The screenshot shows the SPSS Data Editor interface. The 'Analyze' menu is open, and the 'Complex Samples' option is selected. The 'Prepare for Analysis' option is highlighted. A yellow box contains the following instructions:

- Now, let us set up the data for complex analysis
- Click, “Analyze”
- Click, “Complex Samples”
- Click, “Prepare for Analysis”

The data table shows variables: @\_STATE, FMONTH, IMONTH, IDAY, IYEAR, DISPCODE, SEQNO, @\_PSU, CTELENM1, PVTRES1, and COLG. The status bar at the bottom indicates 'IBM SPSS Statistics Processor is ready' and the date '4/29/2020'.

Welcome to the Analysis Preparation Wizard

The Analysis Preparation Wizard helps you describe your complex sample and sample weights and other information needed for accurate estimation of standard errors.

Your selections will be saved to a plan file that you can use in any of the analysis

- Make sure “Create a plan” file is selected
- Click, “Browse”



What would you like to do?

☒ Create a plan file

Choose this option if you have sample data but have not created a plan file.

File:

Browse...

☐ Edit a plan file

Choose this option if you want to add, remove, or modify stages of an existing plan.

File:

Browse...



If you already have a plan file you can skip the Analysis Preparation Wizard and go directly to any of the analysis procedures in the Complex Samples Option to analyze your sample.

< Back

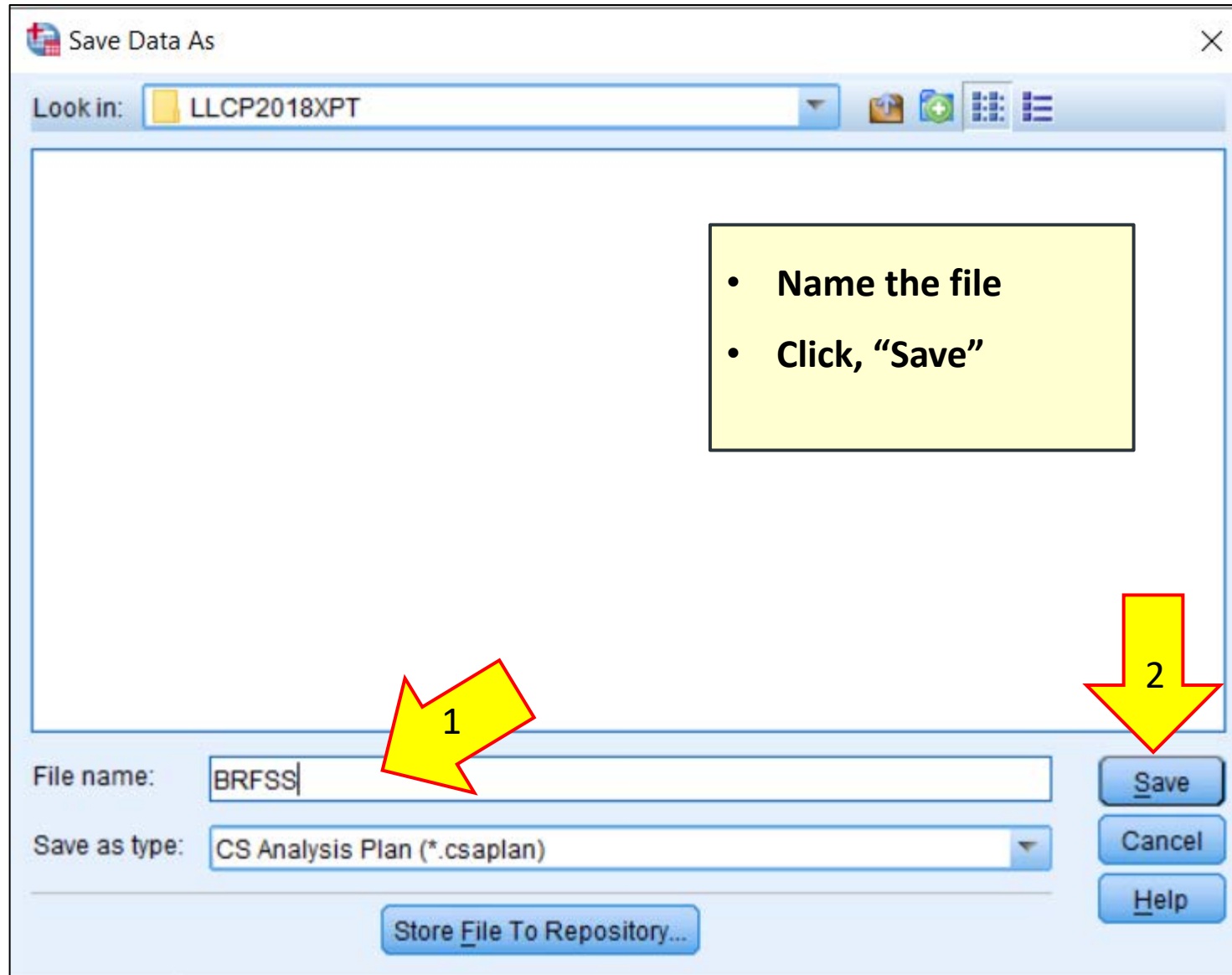
Next >

Finish

Cancel

Help







Analysis Preparation Wizard

Welcome to the Analysis Preparation Wizard

The Analysis Preparation Wizard helps you describe your complex sample and choose an estimation method. You will be asked to provide sample weights and other information needed for accurate estimation of standard errors.

Your selections will be saved to a plan file that you can use in any of the analysis procedures in the Complex Samples Option.

What would you like to do?

☒ Create a plan file

Choose this option if you have sample data but have not created a plan file.

File: C:\Users\zaeem\Documents\\_BRFSS Browse...

☐ Edit a plan file

Choose this option if you want to add, remove, or modify stages of an existing plan.

File: Browse...

**i** If you already have a plan file you can skip the Analysis Preparation Wizard and go directly to any of the analysis procedures in the Complex Samples Option to analyze your sample.

< Back Next > Finish Cancel Help

• Click, "Next"

- Click, “SAMPLE DESIGN STRATIFICATION VARIABLE” from the variables list and send it under Strata
- Click “Gender of participants” and send it under strata. This is the variable that we have just created setting 9 to missing
- Click, “PRIMARY SAMPLING UNIT” from the variable list and send it under Clusters
- Click “FINAL WIRGHT: LAND AND CELL-PHONE” send it under Sample Weight
- Click “Next”

The screenshot shows the 'Preparation Wizard' dialog box with the 'Design Variables' tab selected. The 'Variables' list on the left contains various demographic and survey variables. The 'Strata' section contains 'SAMPLE DESIGN STRATIFICATION VA...' and 'Gender of participants [Gender]'. The 'Clusters' section contains 'PRIMARY SAMPLING UNIT [@\_PSU]'. The 'Sample Weight' section contains 'FINAL WEIGHT: LAND-LINE AND CELL-...'. The 'Label' field is empty. The 'Next >' button is highlighted with a yellow arrow.

Numbered arrows indicate the following steps:

- Arrow 1 points to the 'SAMPLE DESIGN STRATIFICATION VA...' variable in the Strata list.
- Arrow 2 points to the 'Gender of participants [Gender]' variable in the Strata list.
- Arrow 3 points to the 'PRIMARY SAMPLING UNIT [@\_PSU]' variable in the Clusters list.
- Arrow 4 points to the 'FINAL WEIGHT: LAND-LINE AND CELL-...' variable in the Sample Weight list.
- Arrow 5 points to the 'Next >' button at the bottom of the dialog.

Analysis Preparation Wizard

Stage 1: Estimation Method

In this panel you select a method for estimating standard errors.

The estimation method depends on assumptions about how the sample was drawn.

Which of the following sample designs should be assumed for estimation?

☒ **WR** (with replacement)

If you select this option you will not be able to add additional stages. Any sample stages after the current stage will be ignored when the data are analyzed.

☒ Use finite population correction (FPC) when estimating variance under simple random sampling assumption

☐ **E**qual WOR (equal probability sampling without replacement)

The next panel will ask you to specify inclusion probabilities or population sizes.

☐ **U**nequal WOR (unequal probability sampling without replacement)

Joint probabilities will be required to analyze sample data. This option is available in stage 1 only.

< Back Next > Finish Cancel Help

- Uncheck 'Use the finite population correction (FPC) when estimating variance under simple random sampling assumption'
- Click "Next"



- Click, “Next”
- Click, “Finish”

Analysis Preparation Wizard

Stage 1: Plan Summary

This panel summarizes the plan so far. The next step is the Completion panel.

Summary:

Stage	Label	Strata	Clusters	Weights	Size	Method
1	(None)	@_STSTR Gender	@_PSU	@_LLCPWT	(n/a)	WR

File: C:\Users\zaeem\Documents\\_BRFSS\ILLCP2018\PT\BRFSS.csaplan

1

< Back Next > Finish Cancel Help

Analysis Preparation Wizard

Completing the Analysis Wizard

You have provided all of the information needed to create a plan.

You can use the plan file in any Complex Samples analysis procedure when you are ready to analyze the data.

What do you want to do?

☒ Save your specifications to a plan file

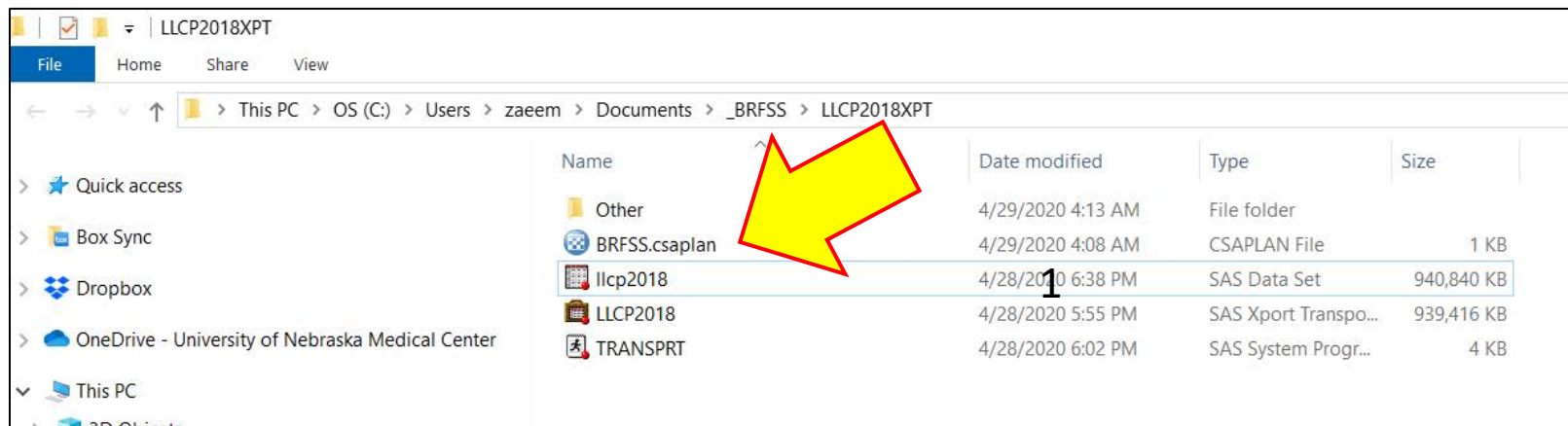
☐ Paste the syntax generated by the Wizard into a syntax window

To close this wizard, click

2

< Back Next > Finish Cancel Help





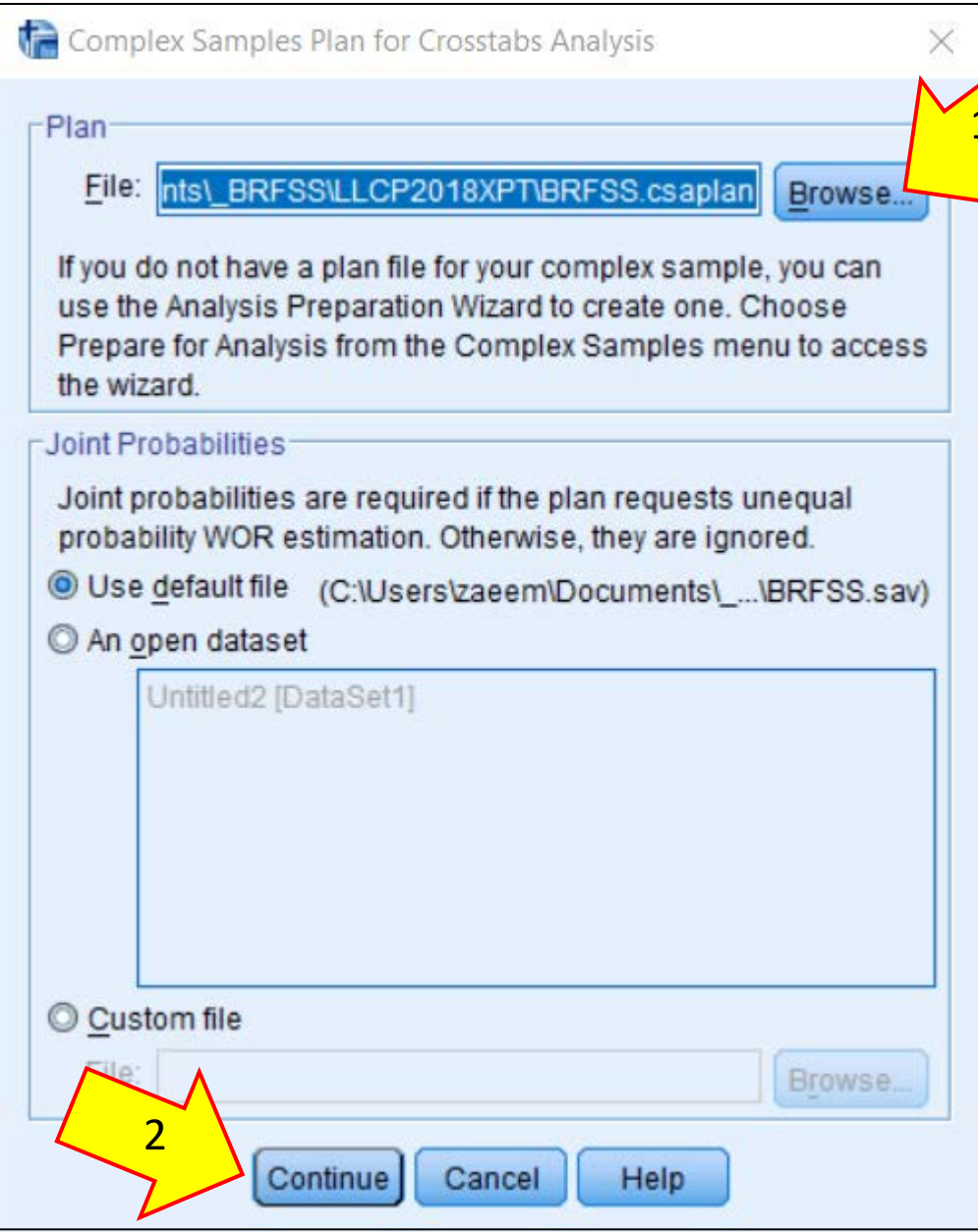
- The complex sample plan file is saved in your specified folder

The screenshot shows the IBM SPSS Statistics interface. The 'Analyze' menu is open, and the 'Complex Samples' option is selected. The 'Crosstabs' option is also visible in the submenu. Three yellow arrows with numbers 1, 2, and 3 indicate the steps to follow:

- Click, "Analyze"
- Click "Complex Samples"
- Click "Crosstabs"

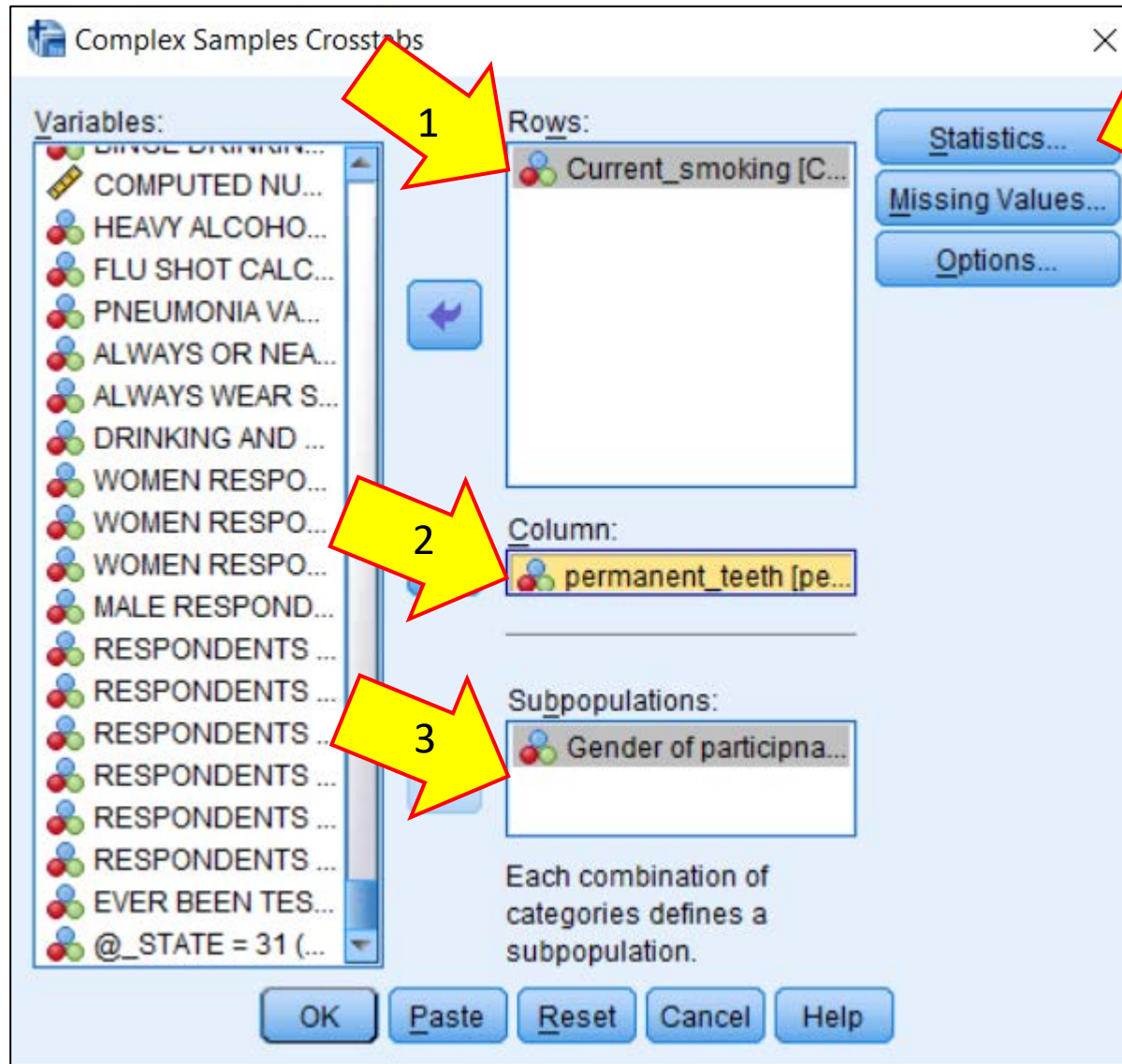
The 'Crosstabs' dialog box is also visible, showing the 'Design Variables' section with 'Stratification' and 'Cluster' options. The 'Analysis Information' section shows 'Estimator Assumption'.

- Now we will run the cross tabulation
- Click, "Analyze"
- Click "Complex Samples"
- Click "Crosstabs"



- **Whenever you open SPSS, browse and select the plan file that we just created. This plan file has the information about cluster, strata and weight**
- **Click, “Continue”**





- Bring “Current\_smoking” under Rows
- Bring Permanent\_teeth under column
- Bring “Gender of participants” under Subpopulations
- These variables are the ones that we recoded into new variables specifying the missing values
- The recoded variables always appear at the bottom of the “Variables”
- Click, “Statistics”



Complex Samples Crosstabs: Statistics

**Cells**

☒ Population size ☒ Column percent  
☒ Row percent ☒ Table percent

**Statistics**

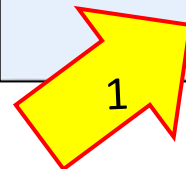
☐ Standard error ☒ Unweighted count  
☐ Confidence interval ☐ Design effect  
Level(%): 95 ☐ Square root of design effect  
☐ Coefficient of variation ☐ Residuals  
☐ Expected values ☐ Adjusted residuals

**Summaries for 2-by-2 Tables**

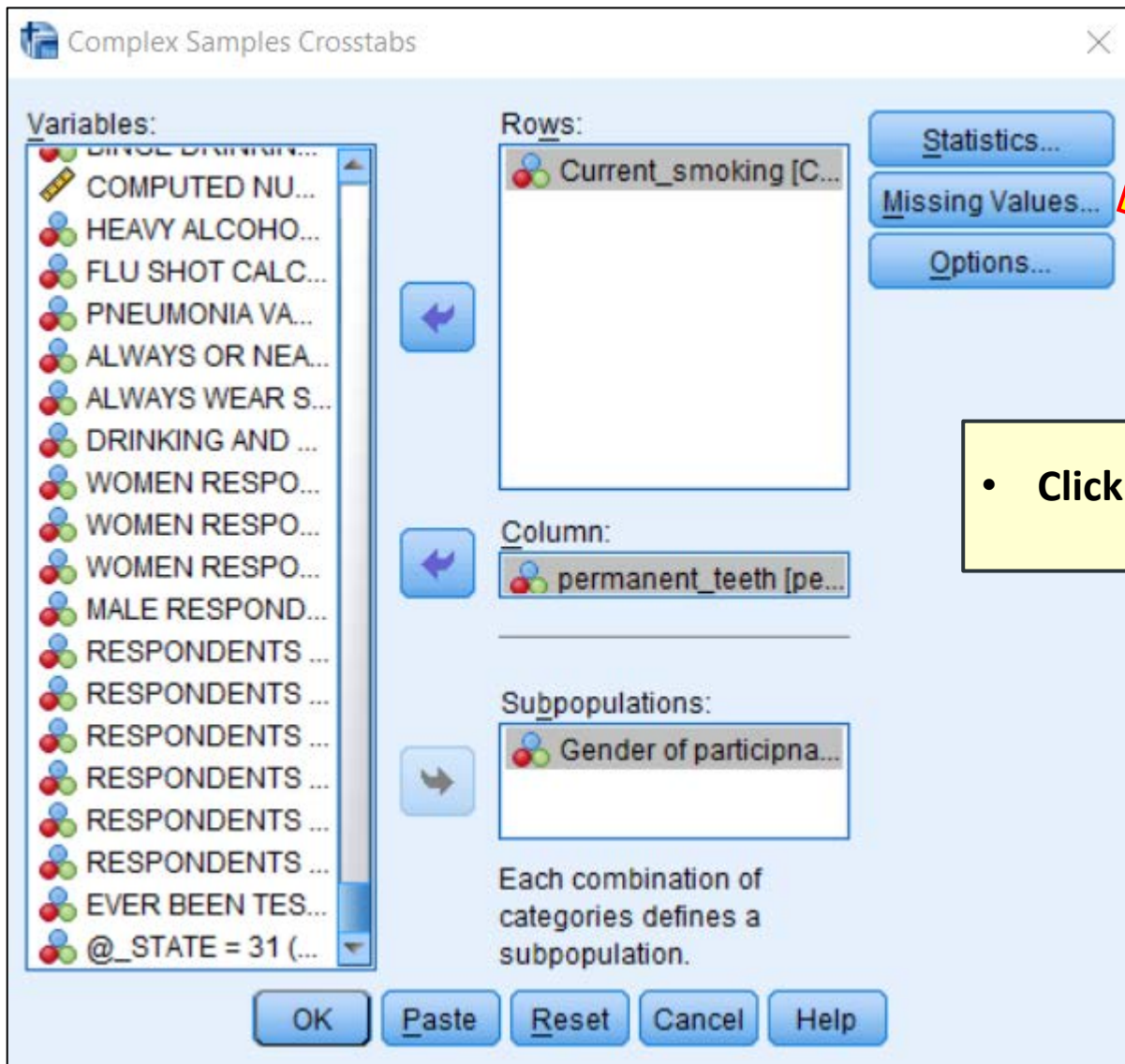
☐ Odds ratio ☐ Risk difference  
☐ Relative risk

☒ Test of independence of rows and columns

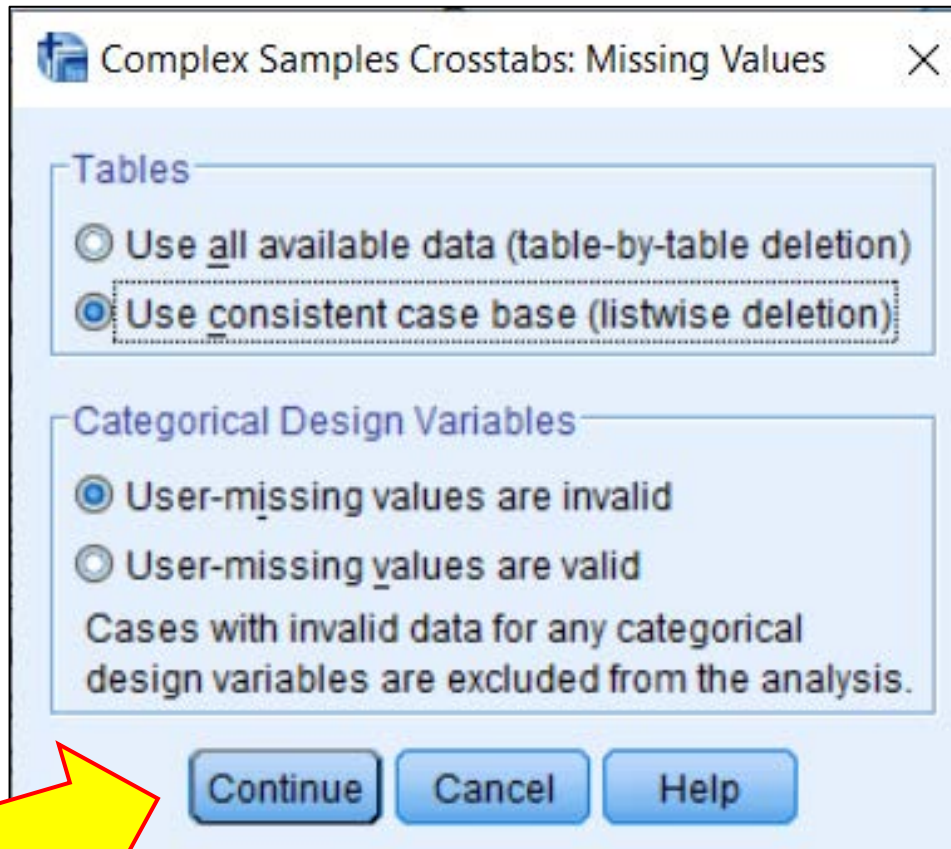
**Continue** **Cancel** **Help**



- You can select the options of your choice or follow what has been shown here
- Click, “Continue”

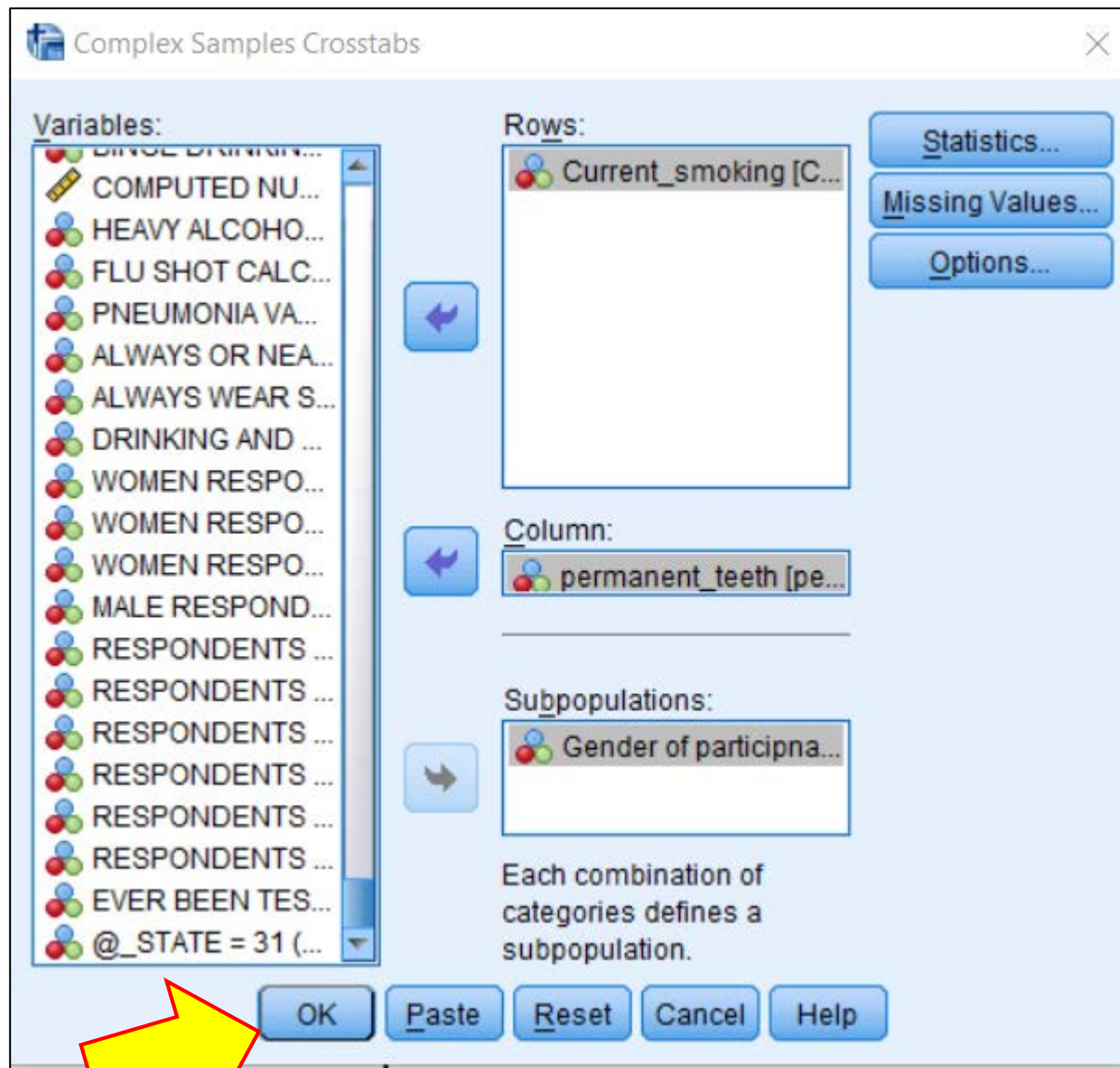


- Click “Missing Values”



- Select the options as shown
- Click, “Continue”





- Click, "OK"



Current_smoking * permanent_teeth						
				permanent_teeth		
Gender of participnats	Current_smoking			1.00	2.00	Total
1.00	1.00	Population Size	Estimate	368276.735	200871.480	569148.216
			Unweighted Count	2978	2338	5316
		% within Current_smoking	Estimate	64.7%	35.3%	100.0%
			Unweighted Count	2978	2338	5316
		% within permanent_teeth	Estimate	86.1%	76.9%	82.6%
			Unweighted Count	2978	2338	5316
	2.00	Population Size	Estimate	59231.644	60413.756	119645.401
			Unweighted Count	394	548	942
		% within Current_smoking	Estimate	49.5%	50.5%	100.0%
			Unweighted Count	394	548	942
		% within permanent_teeth	Estimate	13.9%	23.1%	17.4%
			Unweighted Count	394	548	942
	Total	Population Size	Estimate	427508.380	261285.236	688793.616
			Unweighted Count	3372	2886	6258
		% within Current_smoking	Estimate	62.1%	37.9%	100.0%
			Unweighted Count	3372	2886	6258
		% within permanent_teeth	Estimate	100.0%	100.0%	100.0%
			Unweighted Count	3372	2886	6258

Tests of Independence						
Gender of participnats			Chi-Square	Adjusted F	df1	df2
1.00	Current_smoking * permanent_teeth	Pearson	88.150	38.523	1	6196
		Likelihood Ratio	86.078	37.618	1	6196
2.00	Current_smoking * permanent_teeth	Pearson	133.683	50.685	1	7615
		Likelihood Ratio	129.795	49.211	1	7615
7.00	Current_smoking * permanent_teeth	Pearson	.203	.	.	.
		Likelihood Ratio	.341	.	.	.

The adjusted F is a variant of the second-order Rao-Scott adjusted chi-square statistic. Significance is based on the adjusted F and its degrees of freedom.

- These are two tables from the results.
- The upper table is the cross tabulation between current smoking and permanent teeth extraction for males.
- As you scroll down in you SPSS, you will see results for other gender categories
- The bottom table shows the results of chi-square test
- Note that we have received the same results as those of the WEAT tool and SASS

# LOGISTIC REGRESSION

Analysis Preparation Wizard

Stage 1: Design Variables

In this panel you can select variables that define strata or clusters. A sample weight variable must be selected in the first stage.

Variables used in the output:

Strata:

- SAMPLE DESIGN STRATIFICATION VA...
- Gender of participnats [Gender]

Clusters:

- PRIMARY SAMPLING UNIT [@\_PSU]

Sample Weight:

- FINAL WEIGHT: LAND-LINE AND CELL-...

Stage Label:

< Back Next > Finish Cancel Help

- To perform logistic regression and to use gender as a covariate, we cannot include gender under “Strata” as shown on slide number 86
- Create a complex sample plan file as shown in the previous slides and do not include gender under “Strata”

SPSS Statistics Data Editor

File Edit View Data Analyze Direct Marketing Graphs Utilities Add-ons Window Help

1

2

3

- Click "Analyze"
- Click "Complex Samples"
- Click "Logistic Regression"

Visible: 279 of 279 Variables

	@_STATE	IMONTH	IDAY	IYEAR	ENM1	PVTRES1	COLG1
1	1	01	05	2018	1	1	
2	1	01	12	2018	1	1	
3	1	01	08	2018	1	1	
4	1	01	03	2018	1	1	
5	1	01	12	2018	1	1	
6	1	01	11	2018	1	1	
7	1	01	10	2018	1	1	
8	1	01	13	2018	1	1	
9	1	01	09	2018	1	1	
10	1	01	10	2018	1	1	
11	1	01	10	2018	1	1	
12	1	01	03	2018	1	1	
13	1	01	11	2018	1100	2018000013	2018000013
14	1	01	12	2018	1100	2018000014	2018000014
15	1	01	25	2018	1100	2018000015	2018000015
16	1	01	19	2018	1100	2018000016	2018000016
17	1	01	10	2018	1100	2018000017	2018000017
18	1	01	10	2018	1100	2018000018	2018000018
19	1	01	10	2018	1100	2018000019	2018000019
20	1	01	10	2018	1200	2018000020	2018000020
21	1	01	10	2018	1100	2018000021	2018000021
22	1	01	10	2018	1100	2018000022	2018000022
23	1	01	10	2018	1100	2018000023	2018000023
24	1	01	10	2018	1100	2018000024	2018000024
25	1	01	10	2018	1100	2018000025	2018000025
26	1	01	10	2018	1100	2018000026	2018000026
27	1	01	10	2018	1100	2018000027	2018000027

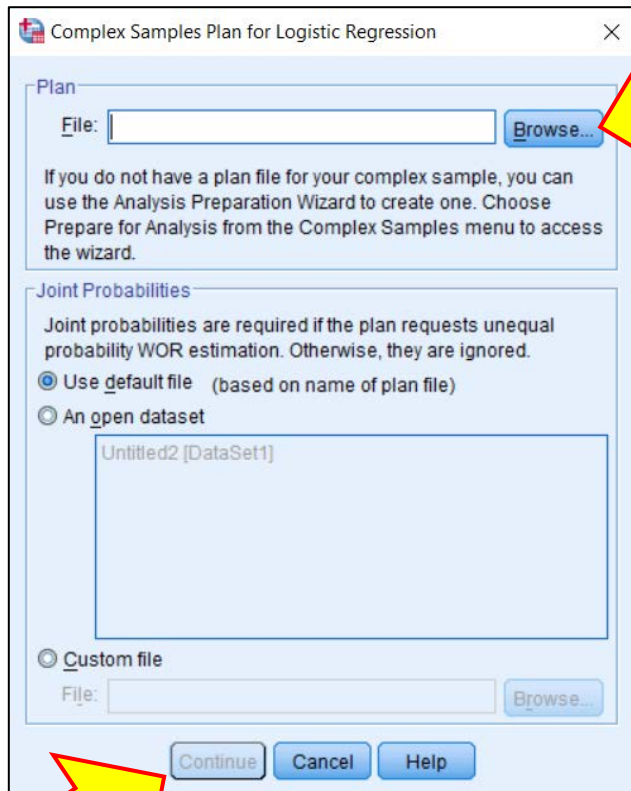
Data View Variable View

Logistic Regression...

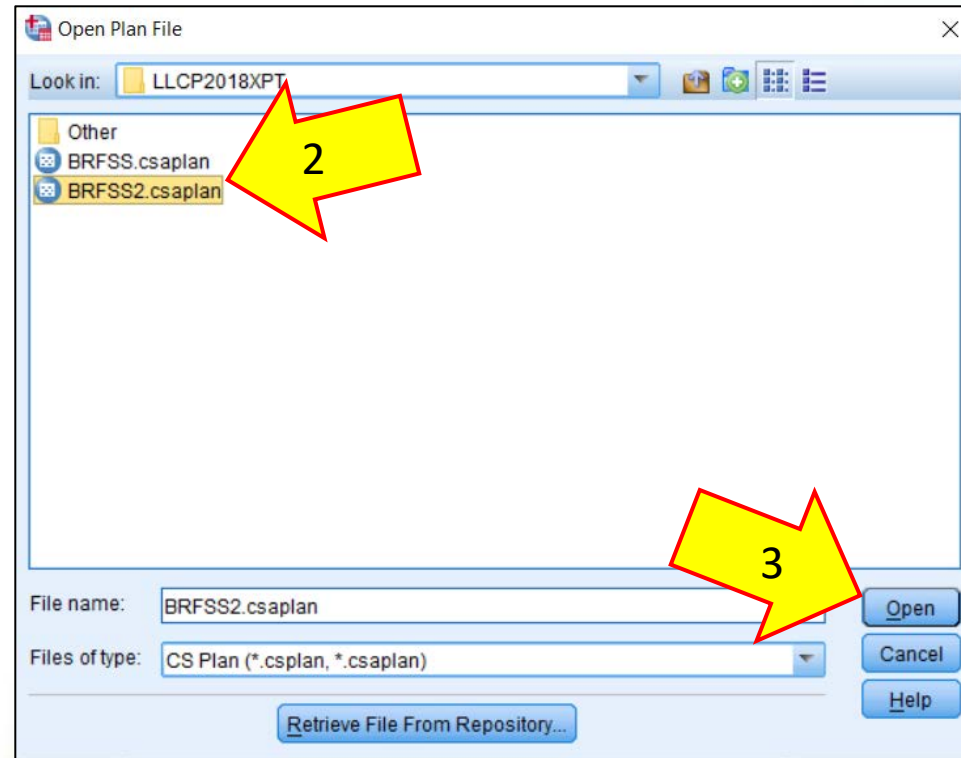
IBM SPSS Statistics Processor is ready Unicode: ON Filter On

6:31 PM 5/3/2020

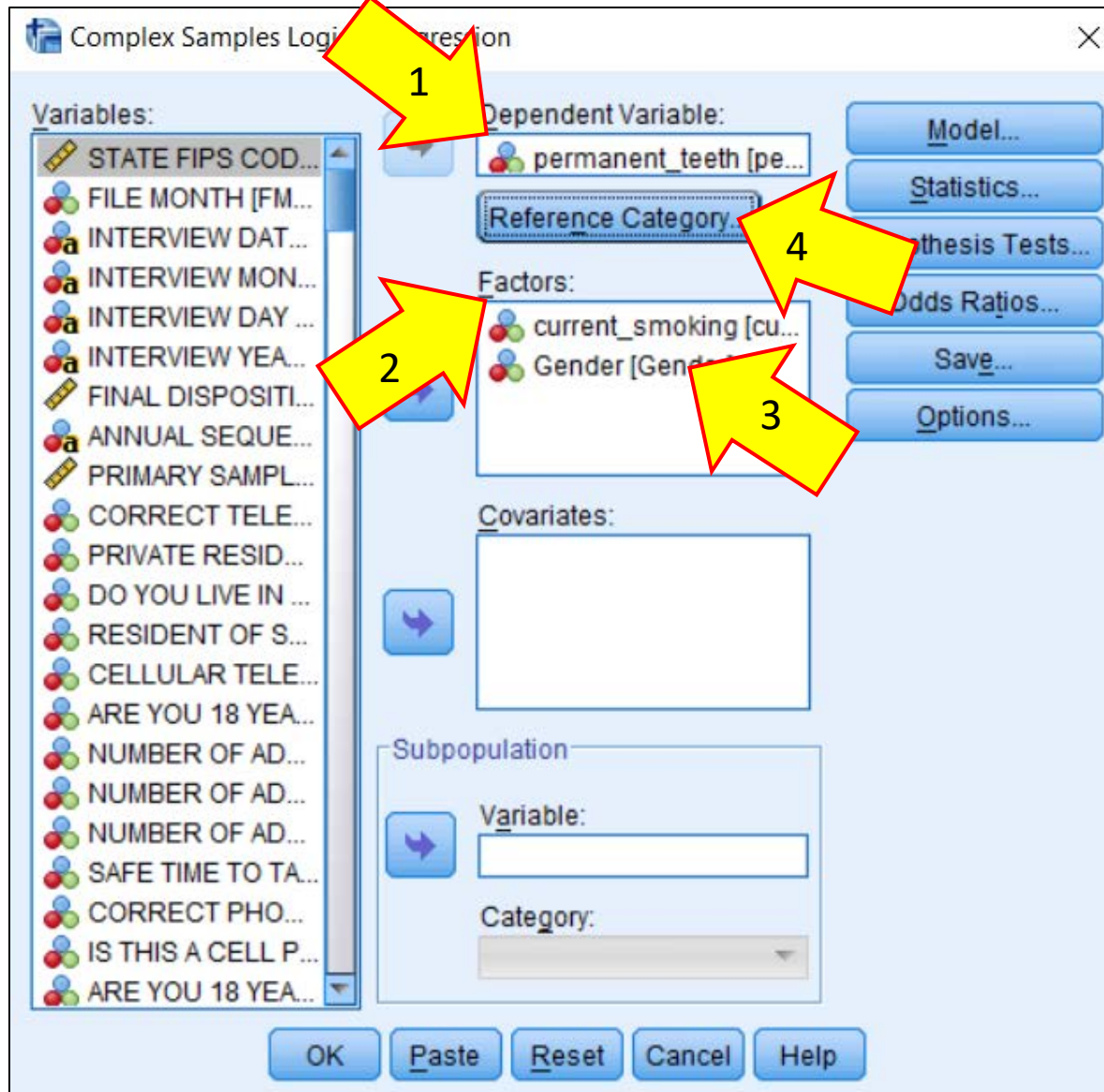




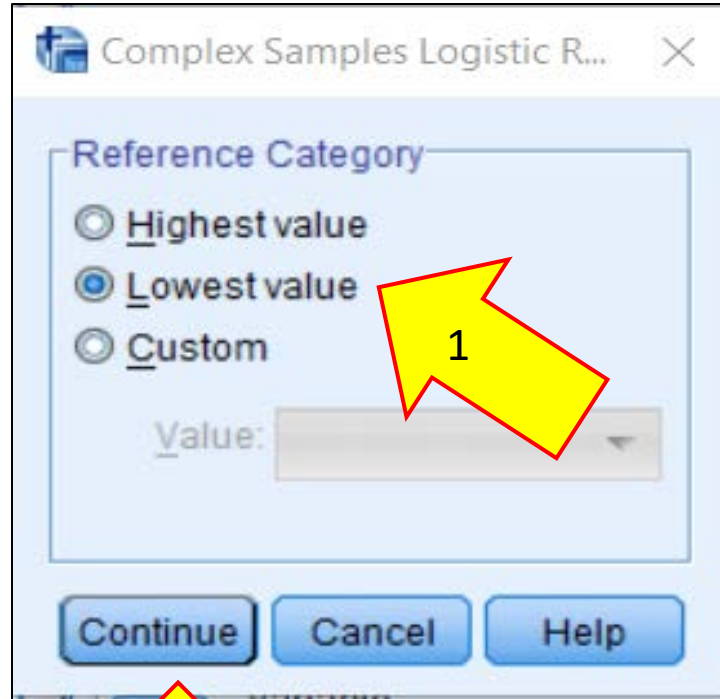
- Click “Browse”
- Select your saved complex sample plan file.
- Click open
- Click “Continue”







- Select “permanent\_teeth” from variables list and transfer it under “Dependent Variable”
- Select “current\_smoking” from variables list and transfer it under “Factors”
- Select “Gender” from variables list and transfer it under “Factors”
- Click “Reference Category”



- Set the reference category to “lowest value”
- The variable “permanent-teeth” has two categories i.e.
  - 1=not at risk
  - 2=at risk
- We want to see the results for those who are at risk, thus we will put 1 as reference (lowest value)
- Click Continue

Complex Samples Logistic Regression

Variables:

- STATE FIPS COD...
- FILE MONTH [FM...
- INTERVIEW DAT...
- INTERVIEW MON...
- INTERVIEW DAY ...
- INTERVIEW YEA...
- FINAL DISPOSITI...
- ANNUAL SEQUE...
- PRIMARY SAMPL...
- CORRECT TELE...
- PRIVATE RESID...
- DO YOU LIVE IN ...
- RESIDENT OF S...
- CELLULAR TELE...
- ARE YOU 18 YEA...
- NUMBER OF AD...
- NUMBER OF AD...
- NUMBER OF AD...
- SAFE TIME TO TA...
- CORRECT PHO...
- IS THIS A CELL P...
- ARE YOU 18 YEA...

Dependent Variable:

permanent\_teeth [pe...

Reference Category...

Factors:

- current\_smoking [cu...
- Gender [Gender]

Covariates:

Subpopulation

Variable:

Category:

OK Paste Reset Cancel Help



- Click “Odds Ratios”
- On the resulting screen, transfer both, “current\_smoking” and “gender” in the factors list to the cell on the right

Complex Samples Logistic Regression: Odds Ratios

Factors:

- current\_smoking [c...
- Gender [Gender]

Covariates:

Odds Ratios for Comparing Factor Levels:

Factor	Reference Category

Odds Ratios for Change in Covariate Values:

Covariate	Units of Change

One set of odds ratios is produced for each variable in the Odds Ratios grids. For each set, all other factors in the model are evaluated at their highest levels; all other covariates are evaluated at their means.

Continue Cancel Help





Complex Samples Logistic Regression: Odds Ratios

**Factors:**

current\_smoking [c...  
Gender [Gender]

**Odds Ratios for Comparing Factor Levels:**

Factor	Reference Category
current_smoking [c...	(Highest value)
Gender [Gender]	(Highest value)

**Covariates:**

**Odds Ratios for Change in Covariate Values:**

Covariate	Units of Change
-----------	-----------------

One set of odds ratios is produced for each variable in the Odds Ratios grids. For each set, all other factors in the model are evaluated at their highest levels; all other covariates are evaluated at their means.

Continue Cancel Help

- The the options under “Reference Category” one by one and change tem to lowest as shoen in the lower figure
- It is very important to set your references correctly to get the desired measurements
- Click Continue

Complex Samples Logistic Regression: Odds Ratios

**Factors:**

current\_smoking [c...  
Gender [Gender]

**Odds Ratios for Comparing Factor Levels:**

Factor	Reference Category
current_smoking [c...	(Lowest value)
Gender [Gender]	(Lowest value)

**Covariates:**

**Odds Ratios for Change in Covariate Values:**

Covariate	Units of Change
-----------	-----------------

One set of odds ratios is produced for each variable in the Odds Ratios grids. For each set, all other factors in the model are evaluated at their highest levels; all other covariates are evaluated at their means.

Continue Cancel Help



Complex Samples Logistic Regression

Variables:

- STATE FIPS COD...
- FILE MONTH [FM...
- INTERVIEW DAT...
- INTERVIEW MON...
- INTERVIEW DAY ...
- INTERVIEW YEA...
- FINAL DISPOSITI...
- ANNUAL SEQUE...
- PRIMARY SAMPL...
- CORRECT TELE...
- PRIVATE RESID...
- DO YOU LIVE IN ...
- RESIDENT OF S...
- CELLULAR TELE...
- ARE YOU 18 YEA...
- NUMBER OF AD...
- NUMBER OF AD...
- NUMBER OF AD...
- SAFE TIME TO TA...
- CORRECT PHO...
- IS THIS A CELL P...
- ARE YOU 18 YEA...

Dependent Variable:

permanent\_teeth [pe...

Reference Category...

Factors:

- current\_smoking [cu...
- Gender [Gender]

Covariates:

Subpopulation

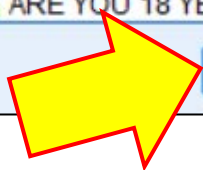
Variable:

Category:

Model...  
Statistics...  
Hypothesis Tests...  
Odds Ratios...  
Save...  
Options...

OK Paste Reset Cancel Help

• Click "OK"



### Odds Ratios 1<sup>a</sup>

permanent_teeth	Odds Ratio	95% Confidence Interval	
		Lower	Upper
current_smoking 2.00 vs. 1.00 2.00	1.978	1.713	2.284

Dependent Variable: permanent\_teeth (reference category = 1.00)

Model: (Intercept), current\_smoking, Gender

a. Factors and covariates used in the computation are fixed at the following values:  
current\_smoking=2.00; Gender=7.00

### Odds Ratios 2<sup>a</sup>

permanent_teeth	Odds Ratio	95% Confidence Interval	
		Lower	Upper
Gender 2.00 vs. 1.00 2.00	1.007	.910	1.114
7.00 vs. 1.00 2.00	8.623	1.034	71.918

Dependent Variable: permanent\_teeth (reference category = 1.00)

Model: (Intercept), current\_smoking, Gender

a. Factors and covariates used in the computation are fixed at the following values:  
current\_smoking=2.00; Gender=7.00

- See the SPSS results above
- They are similar to what we get from SAS and WEAT tool



# University of Nebraska<sup>SM</sup> Medical Center

**BREAKTHROUGHS** FOR LIFE.<sup>®</sup>

