



# Tutorial for working with SEER, SEER\*STAT, and Joinpoint Software

# Tutorial Information

- **This tutorial will cover many topics that relate to Surveillance, Epidemiology, and End Results Program (SEER) data and SEER\*related software.**
- **SEER is a program of the National Cancer Institute which primarily provides information on cancer statistics, in an effort to reduce cancer burden on the U.S. population.**
- **SEER collects and maintains U.S. mortality data for all causes of death, not just cancer deaths.**



# Tutorial Information

- **SEER data is free to use but those who want to use it must get permission from SEER.**
- **This is accomplished through submitting a request on the SEER website and signing a data-use agreement. There is a detailed section in this tutorial that will walk you through this process step-by-step.**



# Tutorial Information (cont.)

- **There are many sections to this tutorial, and not all may be relevant to your needs. If you haven't used SEER data or software before, you will want to start with these two:**
  - **Getting Access to SEER**
  - **Downloading SEER\*Stat Software (the software used to query SEER data)**
- **Note: gaining initial access to SEER data may take at least a business day or two, so plan accordingly. Also, you may need to contact your IT department to be able to download SEER-related software (if you don't have administrative rights on your computer), so keep that in mind as well.**



# Tutorial Information (cont.)

- **There are different ways to access SEER data. Most people will use their internet connection and password from SEER to access data through SEER\*Stat.**
- **You can download SEER data for off-line use if you may not have internet access and want to query SEER data, or you can download SEER data to use in SAS. If these situations are not relevant to you, feel free to skip those sections.**
- **There are many different examples regarding how to use SEER data (e.g. frequencies, incidence rates), and each will walk you through an example. Again, jump to the section that is of interest to you.**



# Tutorial Contents

SEER Access/Download Topic	Slide Numbers
Getting Access to SEER	7 - 17
Downloading SEER*Stat Software	18 - 28
Downloading SEER Binary Data to use in SEER*Stat while off-line (optional)	29 - 39
Downloading SEER ASCII Data to use in SAS (optional)	40 - 55
Downloading Joinpoint Software	179 - 184

SEER Query/Analysis Topic	Slide Numbers
Incidence Count Data	56 - 76
Mortality Counts and Rates	
1. Comparisons between states	77 - 97
2. Comparisons between races within a state	98 - 115
Age-Adjusted Incidence Rates	
1. Example for table display	116 - 160
2. Modifying table to be able to use in Joinpoint	161 - 178
Joinpoint Example	185 - 191

# Getting Access to SEER



University of Nebraska  
Medical Center

# Getting Access to SEER Data

**This section will instruct you on how to get access to SEER data, using the following steps:**

- **Submitting a data request**
- **Receiving a verification email**
- **Sending a signed data use agreement**
- **Receiving a username and password**

**Once you receive your username and password, you will be able to download SEER data and SEER\*Stat software.**



Datasets	
<b>[ - ] SEER Data 1973-2014</b>	→
• Accessing the Data	
• Access Options	
• Submit Request for the Data	
• SEER*Stat Databases	
• Documentation for the Data Files	
• Suggested Citations	
<b>[ + ] Standard Population Data</b>	
U.S. Mortality Data	
<b>[ + ] U.S. Population Data</b>	
County Attributes	
<b>[ + ] Expected Survival Life Tables</b>	
SEER Linked Databases	
<b>[ + ] Specialized SEER*Stat Datasets</b>	
<b>Statistical Software</b>	
<b>[ + ] SEER*Stat</b>	
<b>[ + ] SEER*Prep</b>	
<b>[ + ] HD*Calc</b>	
Analytic Software	
<b>Documentation &amp; Recodes</b>	
Behavior Recode for Analysis	

## SEER Data, 1973-2014

*(Including July-December 2005 Hurricane Katrina Impacted Louisiana Cases)*

The SEER research data include SEER incidence and population data associated by age, sex, race, year of diagnosis, and geographic areas (including [SEER registry](#) and county). A signed Research Data Agreement is required to [access these data](#). Use of these data for publication purposes should contain a citation which includes submission and release dates. See [Citations for SEER Data and SEER\\*Stat Software](#) for specific suggestions regarding the citation format.

### Number of Records in the SEER Research Data

The numbers provided in the table below are for the most recent SEER research data release and the previous release. Malignant and In Situ cases are defined using the [SEER Behavior Recode for Analysis](#).

	All Cases*	Malignant Cases	Malignant + In Situ Cases
<b>1973-2014 Data (Nov 2016 Submission)</b>	9,675,661	8,662,369	9,429,379
<b>1973-2013 Data (Nov 2015 Submission)</b>			
<b>Increase from 2015 to 2016 Submission</b>			
<b>Number of 2014 cases</b>			

\* All Cases includes benign and both malignant and benign only malignant in ICD-O-3 or 2010

### SEER Database Details

- [PSA Values and SEER Data](#) -
- [Radiation/Chemotherapy Data](#) -

available through a custom data request.



Go to the SEER Data website here: <https://seer.cancer.gov/data>

Select the underlined text "access these data".

Cancer Statistics

Statistical Summaries | Interactive Tools | Publications

For Researchers  
Datasets and Software

For Cancer Registrars  
Coding Rules, Training and Support

About SEER  
Our Registries and Research

Home > For Researchers > SEER Data 1973-2014 > Accessing the Data

Datasets

- [-] SEER Data 1973-2014
  - **Accessing the Data**
  - Access Options
  - Submit Request for the Data
  - SEER\*Stat Databases
  - Documentation for the Data Files
  - Suggested Citations

[-] Standard Population Data

U.S. Mortality Data

[-] U.S. Population Data

County Attributes

[-] Expected Survival Life Tables

SEER Linked Databases

[-] Specialized SEER\*Stat Datasets

Statistical Software

[-] SEER\*Stat

[-] SEER\*Prep

[-] HD\*Calc

Analytic Software

Documentation & Recodes

Behavior Recode for Analysis

[-] Cancer Stage Variables

Cause of Death Recode

Cause-specific Death Classification

[-] Incidence Site Recode Variables

Insurance Recode

Months Survived Based on Complete Dates

Policy for Calculating Hispanic Mortality

Race Recode

SEER Coding and Staging Manuals

Site Specific Surgery (1983-1997)

Tools

Email

Print Page

### Accessing the 1973-2014 SEER Data

The 1973-2014 SEER Research Data are available in the binary format required by the SEER\*Stat software and in an ASCII text format that can be analyzed with your own statistical software. The data and SEER\*Stat software are available using three methods:

1. Use SEER\*Stat to access the data through your Internet connection (SEER\*Stat's client-server mode). This is the most common method for accessing the SEER Data.
2. Download compressed files from the Internet containing the same contents as the DVDs. SEER\*Stat, the binary data, and the text data are available for download.
3. Have a DVD containing the data and SEER\*Stat software shipped to you. This will contain SEER\*Stat, and the binary and text versions of the data in compressed format only.

These are described in more detail in [Options for Accessing the Data and SEER\\*Stat Software](#)

#### Datasets Included by Accessing the Data

##### All Access Options (Client-server)

- [SEER Research Data, 1973-2014](#)
- [County Attributes data](#)

##### Client-Server Only:

- [US Mortality](#)
- [Previous submissions of SEER\\*Stat](#)
- [Incidence with SEER Delay Factors](#)
- [US Populations](#)
- [Specialized SEER\\*Stat Datasets](#)
- [Multiple primaries-standardized mortality ratios \(MP-SMR\)](#)
- [Automatic SEER\\*Stat software updates](#)

##### Compressed Data Files Only:

- ASCII text version of the SEER 1973-2014 data

#### SEER Research Data Agreement

A signed SEER Research Data Agreement form is required to access the SEER data. Once you have determined if you want to access the data via the Internet or receive a DVD, please [submit a request](#) for access to the data and SEER\*Stat software. A personalized SEER Research Data Agreement will be created for you. The SEER program will process your request within 2 business days of receiving your signed agreement.

You may review the language of the SEER Research Data Agreement in the [sample agreement form](#) (this cannot be used to request data).

Please send questions or comments to: [seertrack@imsweb.com](mailto:seertrack@imsweb.com).

[Return to top](#)

On the next page, click "submit a request" under SEER Research Data Agreement.



<b>Datasets</b>
[+] SEER Data 1973–2014 <ul style="list-style-type: none"><li>Accessing the Data</li><li>Access Options</li><li>Submit Request for the Data</li><li>SEER*Stat Databases</li><li>Documentation for the Data Files</li><li>Suggested Citations</li></ul>
[+] Standard Population Data <ul style="list-style-type: none"><li>U.S. Mortality Data</li></ul>
[+] U.S. Population Data <ul style="list-style-type: none"><li>County Attributes</li></ul>
[+] Expected Survival Life Tables <ul style="list-style-type: none"><li>SEER Linked Databases</li></ul>
[+] Specialized SEER*Stat Datasets
<b>Statistical Software</b>
[+] SEER*Stat
[+] SEER*Prep
[+] HD*Calc
Analytic Software
<b>Documentation &amp; Recodes</b>
Behavior Recode for Analysis
[+] Cancer Stage Variables
Cause of Death Recode

### Request Access to the SEER 1973–2014 Research Data

There are two ways to obtain the data: through an Internet connection (requires username and password), or a DVD via US mail. [View detailed descriptions of the options.](#)

How would you like to access the SEER Data?

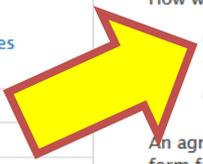
- Through my Internet connection (SEER\*Stat's client-server mode OR Download data files from the web site)
- Have a DVD delivered to my location via US mail

An agreement form is required for each data submission. If you submitted an agreement form for the current data, then you are not required to do so again. Forms submitted for previous data submissions are not valid for access to the current data.

Have you submitted an agreement form for the **current version** of the data, the SEER 1973–2014 Research Data?

- No (check this if you have not submitted a form since April 14, 2017)
- Yes, I previously signed an agreement for the SEER 1973–2014 Research Data.

Please send your comments to: [seertrack@seer.cancer.gov](#)



- Choose “Through my Internet Connection” as your access option.
- Choose the appropriate option for the next question, and then click “Submit”.

Datasets
(+) SEER Data 1973–2014 <ul style="list-style-type: none"><li>• Accessing the Data</li><li>• Access Options</li><li>• Submit Request for the Data</li><li>• SEER*Stat Databases</li><li>• Documentation for the Data Files</li><li>• Suggested Citations</li></ul>
(+) Standard Population Data <ul style="list-style-type: none"><li>• U.S. Mortality Data</li></ul>
(+) U.S. Population Data <ul style="list-style-type: none"><li>• County Attributes</li></ul>
(+) Expected Survival Life Tables <ul style="list-style-type: none"><li>• SEER Linked Databases</li></ul>
(+) Specialized SEER*Stat Datasets
Statistical Software
(+) SEER*Stat
(+) SEER*Prep
(+) HD*Calc <ul style="list-style-type: none"><li>• Analytic Software</li></ul>
Documentation & Records
Behavior Recode for Analysis
(+) Cancer Stage Variables <ul style="list-style-type: none"><li>• Cause of Death Recode</li><li>• Cause-specific Death Classification</li></ul>
(+) Incidence Site Recode Variables <ul style="list-style-type: none"><li>• Insurance Recode</li><li>• Months Survived Based on Complete Dates</li><li>• Policy for Calculating Hispanic Mortality</li><li>• Race Recode</li><li>• SEER Coding and Staging Manuals</li><li>• Site Specific Surgery (1983–1997)</li></ul>

Tools
Email
Print Page

### SEER 1973–2014 Research Data Access Request

First name:

Middle initial:

Last name:

Organization:

Address:

City:

State:

Zip:

Country:

Phone:

Fax:

Email:

Verify email:

To help us guide future development of the SEER database, we would appreciate your input. Responses to the following questions will be compiled separately without identifying information.

Which of these best describes you?:

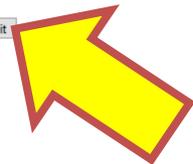
If other, please specify:

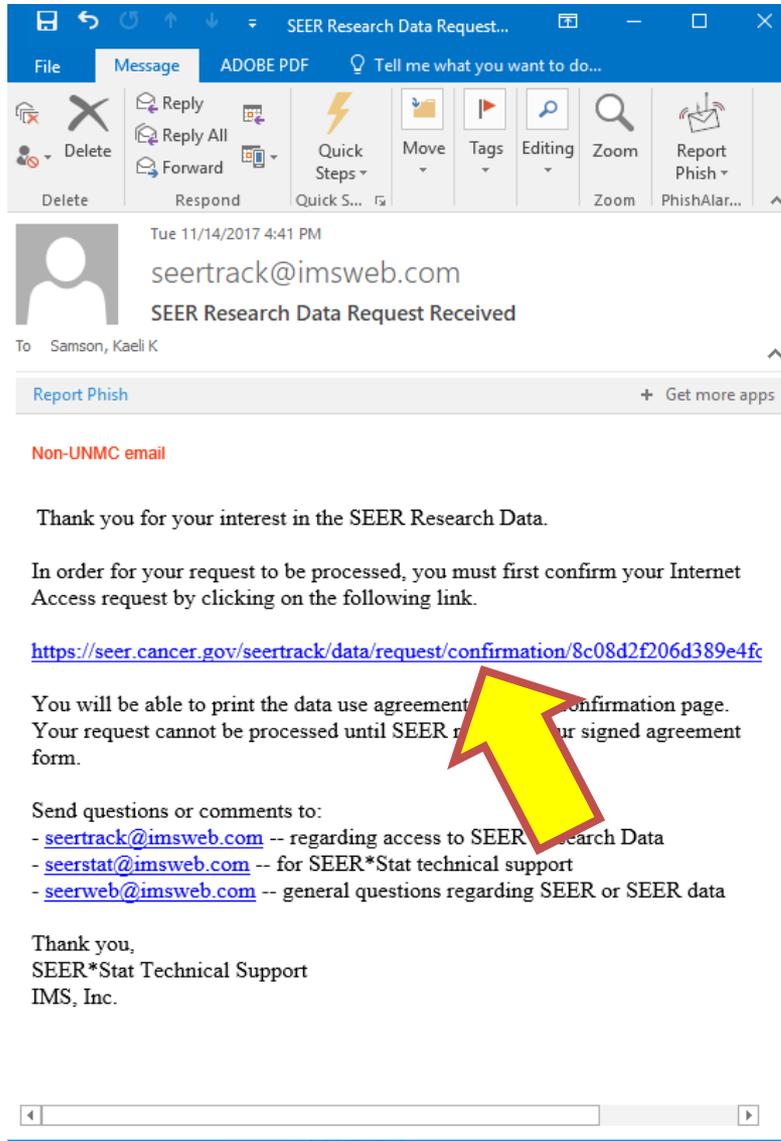
What is your purpose for using the data?:

\* required item

Submit

- Fill out and submit your info on this page.
- Click “Submit”.
- Check the email account you listed for a verification link!





- Check your email for a link to verify your access request.
- Click the link to take you to a website with the agreement form.

<b>Datasets</b>
[+] SEER Data 1973–2014 <ul style="list-style-type: none"><li>• Accessing the Data</li><li>• Access Options</li><li>• Submit Request for the Data</li><li>• SEER*Stat Databases</li><li>• Documentation for the Data Files</li><li>• Suggested Citations</li></ul>
[+] Standard Population Data <ul style="list-style-type: none"><li>• U.S. Mortality Data</li></ul>
[+] U.S. Population Data <ul style="list-style-type: none"><li>• County Attributes</li></ul>
[+] Expected Survival Life Tables <ul style="list-style-type: none"><li>• SEER Linked Databases</li></ul>
[+] Specialized SEER*Stat Datasets
<b>Statistical Software</b>
[+] SEER*Stat
[+] SEER*Prep
[+] HD*Calc <ul style="list-style-type: none"><li>• Analytic Software</li></ul>
<b>Documentation &amp; Recodes</b>

Your data request has been confirmed.

Please sign and return the required [agreement form](#). Your request cannot be processed until SEER receives your **signed agreement form**.

Please send any questions or comments to [seerhelp@imsweb.com](mailto:seerhelp@imsweb.com)



Click "agreement form".

Data Request Confirma X Data Request Confirma X SEER Research Data Agri X +

https://seer.cancer.gov/seertra 110% SEER

**Request Type: Internet Access**

**SURVEILLANCE, EPIDEMIOLOGY, AND END RESULTS PROGRAM  
Data-Use Agreement for the SEER 1973-2014 Research Data File**

It is of utmost importance to protect the identities of cancer patients. Every effort has been made to exclude identifying information on individual patients from the computer files. Certain demographic information - such as sex, race, etc. - has been included for research purposes. All research results must be presented or published in a manner that ensures that no individual can be identified. In addition, there must be no attempt either to identify individuals from any computer file or to link with a computer file containing patient identifiers.

**In order for the Surveillance, Epidemiology, and End Results Program to provide access to its Research Data File to you, it is necessary that you agree to the following provisions.**

1. I will not use - or permit others to use - the data in any way other than for statistical reporting and analysis for research purposes. I must notify the SEER Program if I discover that there has been any other use of the data.
2. I will not present or publish data in which an individual patient can be identified. I will not publish any information on an individual patient, including any information generated on an individual case by the case listing session of SEER\*Stat. In addition, I will avoid publication of statistics for very small groups.
3. I will not attempt either to link - or permit others to link - the data with individual level records in another database.
4. I will not attempt to learn the identity of any patient whose cancer data is contained in the supplied file(s).
5. If I inadvertently discover the identity of any patient, then (a) I will make no use of this knowledge, (b) I will notify the SEER Program of the incident, and (c) I will inform no one else of the discovered identity.
6. I will not either release - or permit others to release - the data - in full or in part - to any person except with the written approval of the SEER Program. In particular, all members of a research team who have access to the data must sign this data-use agreement.
7. I will use appropriate safeguards to prevent use or disclosure of the information other than as provided for by this data-use agreement. If accessing the data from a centralized location on a time sharing computer system or LAN with SEER\*Stat or another statistical package, I will not share my logon name or password with any other individuals. I will also not allow any other individuals to use my computer account after I have logged on with my logon name and password.
8. For all software provided by the SEER Program, I will not copy it, distribute it, reverse engineer it, profit from its sale or use, or incorporate it in any other software system.
9. I will cite the source of information in all publications. The appropriate citation is associated with the data file used. (Please see either Suggested Citations on the SEER\*Stat Help menu or the Readme.txt associated with the ASCII text version of the SEER data.)

My signature indicates that I agree to comply with the above stated provisions.

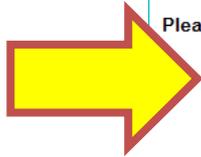
\_\_\_\_\_  
Signature

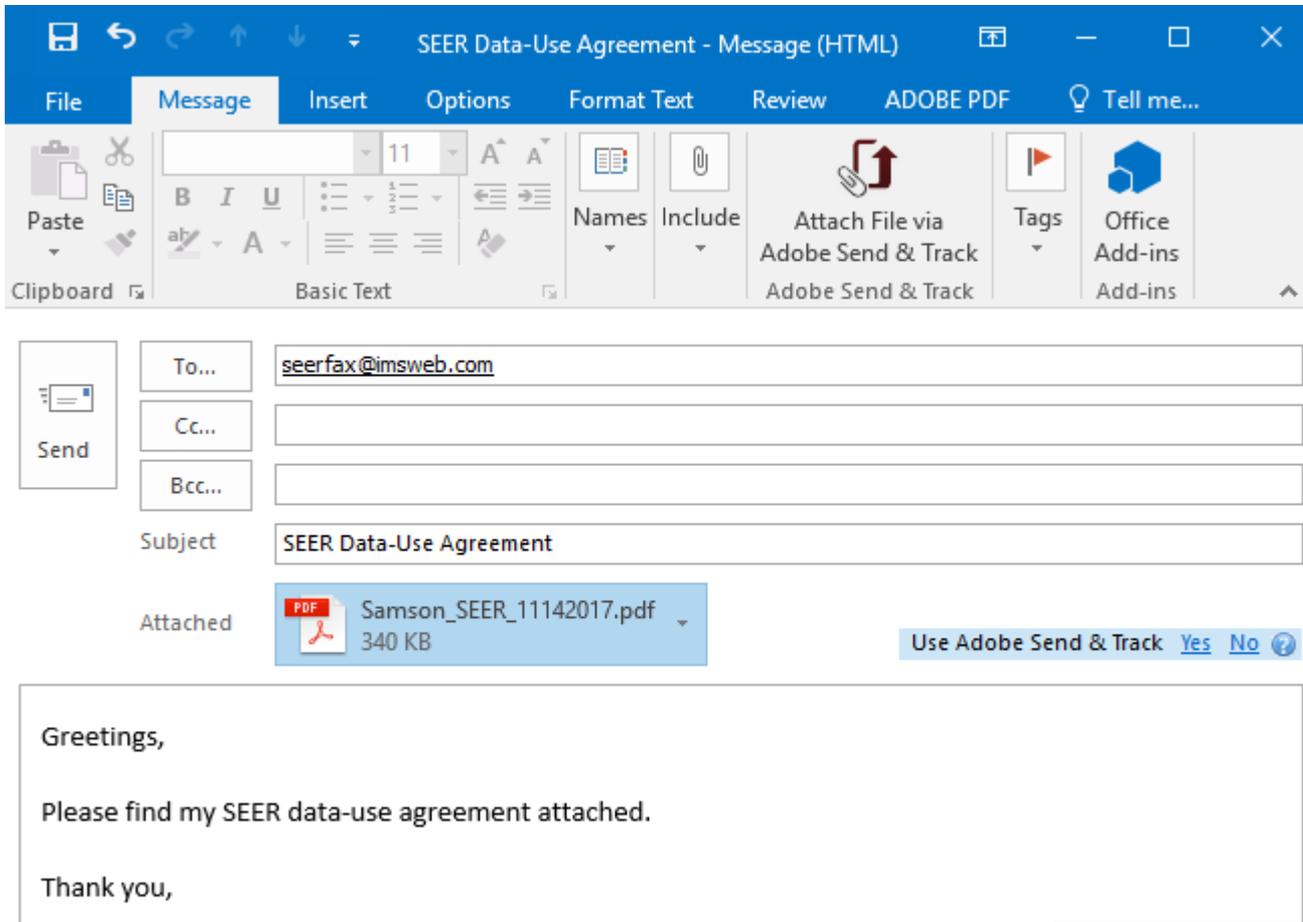
\_\_\_\_\_  
Date

**Please print, sign, and date the agreement. Send the form to The SEER Program:**

- By fax to 301-680-9571
- Or, e-mail a scanned form to [seerfax@imsweb.com](mailto:seerfax@imsweb.com)

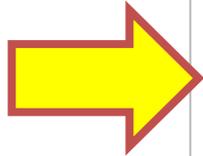
- Print off this page, then sign and date it.
- Scan and email it to the email address listed at the bottom of the page, or fax it to the number shown at the bottom of the page





Example of scanning and emailing the signed and dated agreement.

- Wait for a response from SEER.
- This will typically take 1-3 business days.
- Their response will include a username and password.



-----Original Message-----  
From: seertrack@imsweb.com [mailto:seertrack@imsweb.com]  
Subject: SEER Data Request Approved

Thank you for your interest in the SEER Research Data. Your signed Research Data Agreement is on file at SEER. Your username and password have been generated for Internet access and they are shown below. Please note that both the username and password are case sensitive.

Username: XXXXXXXX|  
Password: XXXXXXXX

These will allow you to utilize the SEER\*Stat client-server system and/or download the files which make up the SEER Research Data DVD. These options are described at the following URL:

<http://seer.cancer.gov/data/options.html>

You can change your password once you log into SEER\*Stat from the "Client Server User Information" option located under the Profile menu.

The review of PSA data has been completed for 2004+ cases, therefore PSA values are available for 2004-2014 prostate cases. For more information, see <https://seer.cancer.gov/data/psa-values.html>.

Radiation treatment variables have been removed from the public research database starting with the November 2016 data submission. These variables are available through a custom data request process after signing an additional data use agreement that describes the completeness of the radiation treatment variable and the potential biases associated with use of the radiation data. To request access to

# Downloading SEER\*Stat software to query SEER data



University of Nebraska  
Medical Center

# Downloading SEER\*Stat software to query SEER data

- **SEER\*Stat is a versatile tool for making queries of SEER data.**
- **It is particularly useful for easily calculating age-adjusted rates.**
- **Note: SEER\*Stat is a Windows application and will not run on Mac OSX!**



# Downloading SEER\*Stat software to query SEER data (cont.)

- **After downloading SEER\*Stat you can instantly access SEER data by simply entering your username and password, given you have an internet connection. If you may not always have regular access to the internet, you may want to consider downloading the binary SEER data for off-line use in SEER\*Stat (there is a section for doing this later in the tutorial, if needed).**



← → ↻ 🏠 <https://seer.cancer.gov/seerstat/> 110% 🔍 seer stat →

**NIH** NATIONAL CANCER INSTITUTE  
Surveillance, Epidemiology, and End Results Program

Search SEER

**Cancer Statistics** | Statistical Summaries | Interactive Tools | Publications

**For Researchers** | Datasets and Software

**For Cancer Registrars** | Coding Rules, Training and Support

**About SEER** | Our Registries and Res

Home ▶ For Researchers ▶ SEER\*Stat Software

**Datasets**

- [+] SEER Data 1973–2014
- [+] Standard Population Data
  - U.S. Mortality Data
- [+] U.S. Population Data
  - County Attributes
- [+] Expected Survival Life Tables
- SEER Linked Databases
- [+] Specialized SEER\*Stat Datasets

**Statistical Software**

- [–] SEER\*Stat →
  - Getting Started
  - Tutorials
  - FAQs
  - Analysis Data
  - Installation Files
- [+] SEER\*Prep
- [+] HD\*Calc
  - Analytic Software

**Documentation & Recodes**

- Behavior Recode for Analysis

### SEER\*Stat Software

Latest Release: Version 8.3.4 – March 23, 2017

The SEER\*Stat statistical software mechanism for the analysis of cancer data is a powerful PC tool to view and analyze cancer statistics for studying the impact of cancer. The software is distributed with access to the data before use.

Visit the following sections:

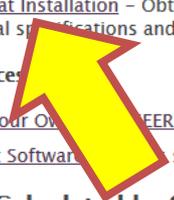
- [Getting Started](#) – Learn the answers to frequent questions and get support.
- [Tutorials](#) – Follow step-by-step instructions on how to use SEER\*Stat and to calculate the many statistics available.
- [Analysis Data](#) – Find out what databases can be analyzed in SEER\*Stat.
- [SEER\\*Stat Installation](#) – Obtain software updates, learn how to use the auto-update feature, and access a technical specifications and revision history.

**Other Resources**

- [Using Your Own Data \(SEER\\*Prep\)](#) – Software to convert your data to the file format required by SEER\*Stat.
- [Analytic Software](#) – Several other software tools developed for the analysis of cancer statistics.

### Statistics Calculated by SEER\*Stat

- To download SEER\*Stat software, go the website:  
<https://seer.cancer.gov/seerstat/software/>
- Once there, click on the “SEER\*Stat Installation” link.



Browser address bar: <https://seer.cancer.gov/seerstat/software/>

NIH NATIONAL CANCER INSTITUTE  
Surveillance, Epidemiology, and End Results Program

Search SEER

**Cancer Statistics**  
Statistical Summaries | Interactive Tools | Publications

**For Researchers**  
Datasets and Software

**For Cancer Registrars**  
Coding Rules, Training and Support

**About SEER**  
Our Registries and Res

Home > For Researchers > SEER\*Stat > Installation

**Datasets**

- [+] SEER Data 1973–2014
- [+] Standard Population Data
  - U.S. Mortality Data
- [+] U.S. Population Data
  - County Attributes
- [+] Expected Survival Life Tables
  - SEER Linked Databases
- [+] Specialized SEER\*Stat Datasets

**Statistical Software**

- [+] SEER\*Stat
  - Getting Started
  - Tutorials
  - FAQs
  - Analysis Data
  - Installation Files →
- [+] SEER\*Prep
- [+] HD\*Calc
  - Analytic Software

**Documentation & Recodes**

- Behavior Recode for Analysis

## SEER\*Stat Installation

The installation program(s) for SEER\*Stat are provided below. All new software releases will be posted here and include the version number in the filename. SEER\*Stat's title bar includes the version number so that you can see you are using the latest version.

The SEER\*Stat software has an auto-update feature. The software can be configured to automatically upgrade with each SEER\*Stat release or to delay the update for a couple weeks after the release. See [Auto-updating in SEER\\*Stat](#) for more information.

Current SEER\*Stat users may [sign up for e-mail notifications](#).

### Installation Files

- [ss8\\_3\\_4.exe](#) – modified 23-Mar-2017
- [Readme](#) – modified 23-Mar-2017

In most cases, to use SEER\*Stat, you must have access to the SEER Research Data. The installation program provided does not include access to these data. If you have not done so already, please follow the steps to [Access SEER Research Data](#).

### SEER\*Stat Readme File

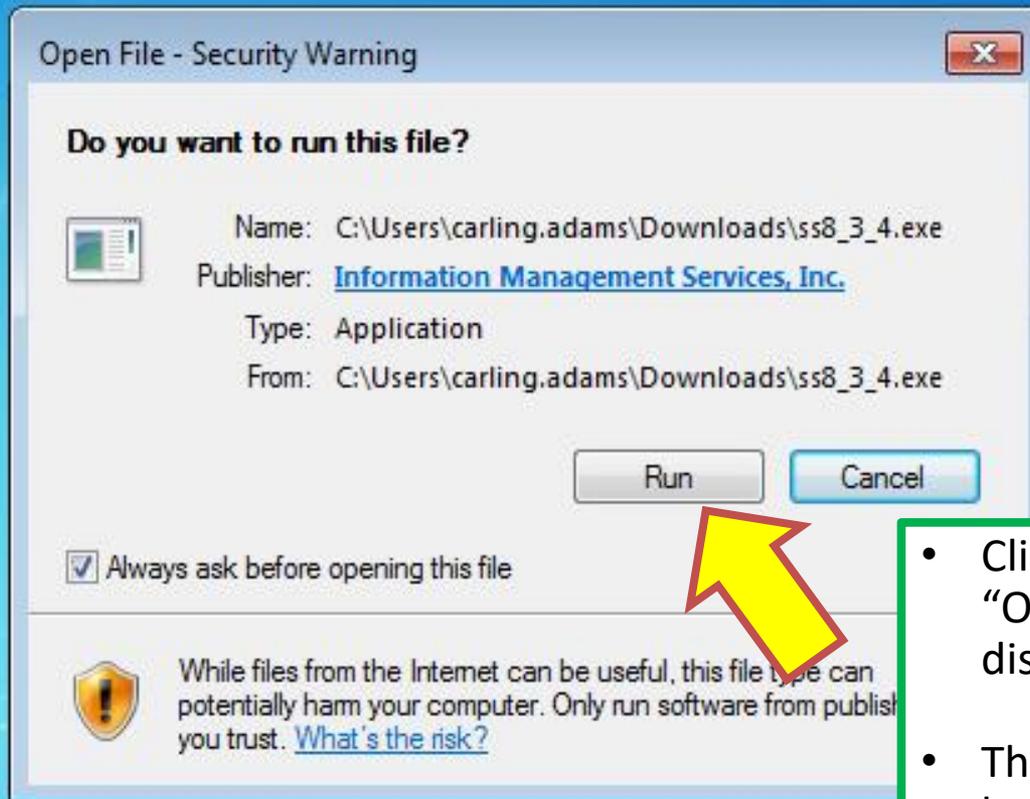
```
-----  
SEER*Stat 8.3.4  
-----
```

The SEER\*Stat statistical software provides a convenient, intuitive mechanism for the analysis of SEER and other cancer-related databases. It is a powerful PC tool to view individual cancer

**Click on the installation file labeled "ss8\_3\_4.exe".**

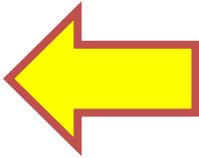
The screenshot shows the SEER\*Stat installation page on the National Cancer Institute website. A yellow arrow labeled '1' points to the download icon in the browser's top right corner. A second yellow arrow labeled '2' points to the download dialog box that has appeared, showing the file 'ss8\_3\_4(1).exe'. A third yellow arrow labeled '3' points to the 'Save File' button in the file opening dialog box. The background page includes a navigation menu with 'Cancer Statistics', 'For Researchers', 'For Cancer Registrars', and 'About SEER'. The main content area is titled 'SEER\*Stat Installation' and contains text about the installation program and a 'SEER\*Stat Readme File' section.

- After clicking on the executable file, follow the prompts to download the software.

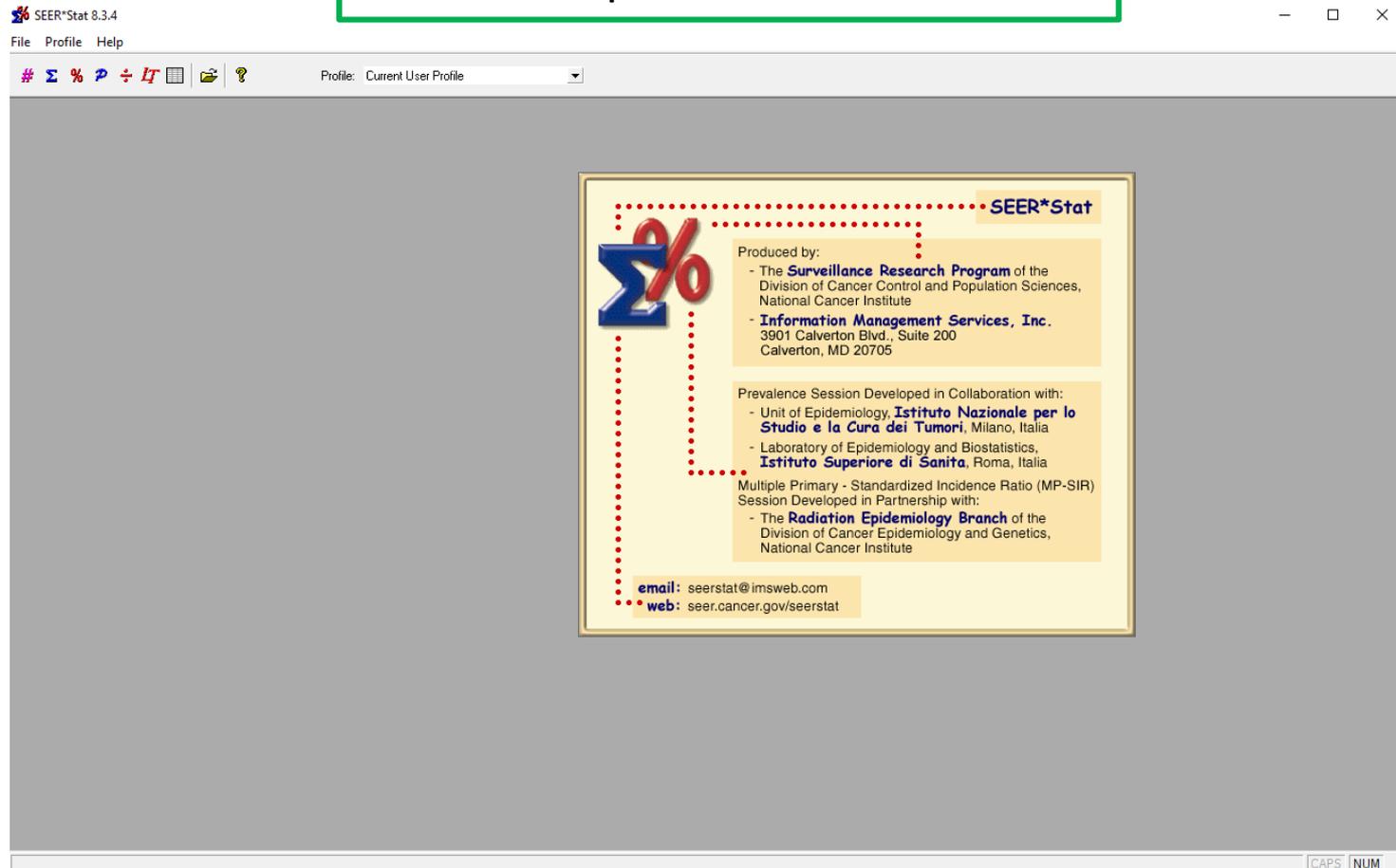


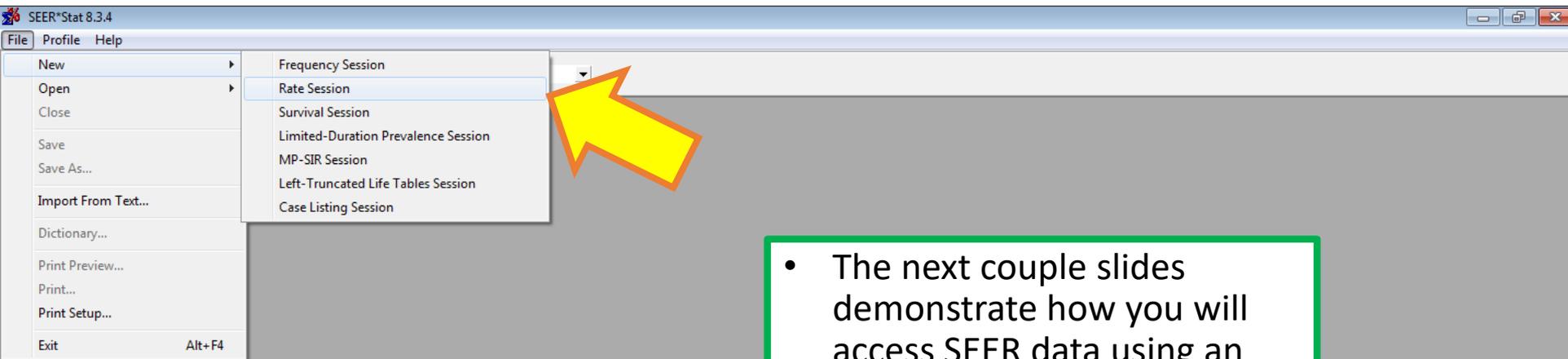
- Click on "Run" when the "Open File" window displays.
- The installation window will be displayed. Click through the steps in this window to complete installation.

At some institutions, you may need to contact IT or a system administrator to allow the download on your computer.



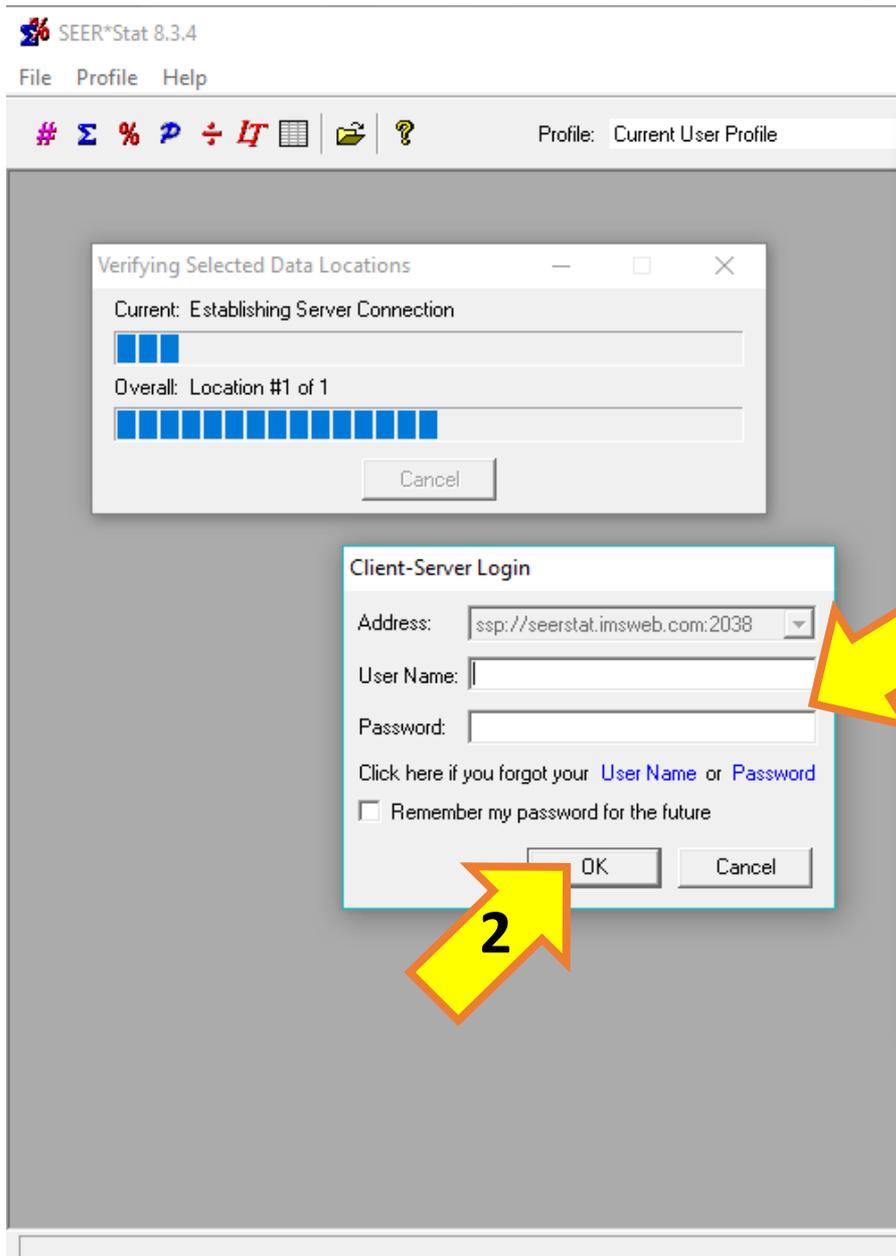
- After following the prompts to download the software, it should create a shortcut on your desktop.
- Double-click on the desktop icon to open the software.





- The next couple slides demonstrate how you will access SEER data using an internet connection.
- When you start a new session in SEER\*Stat, you'll go to "File", "New", and then choose a session.
- As a demonstration, we'll choose "Rate Session"

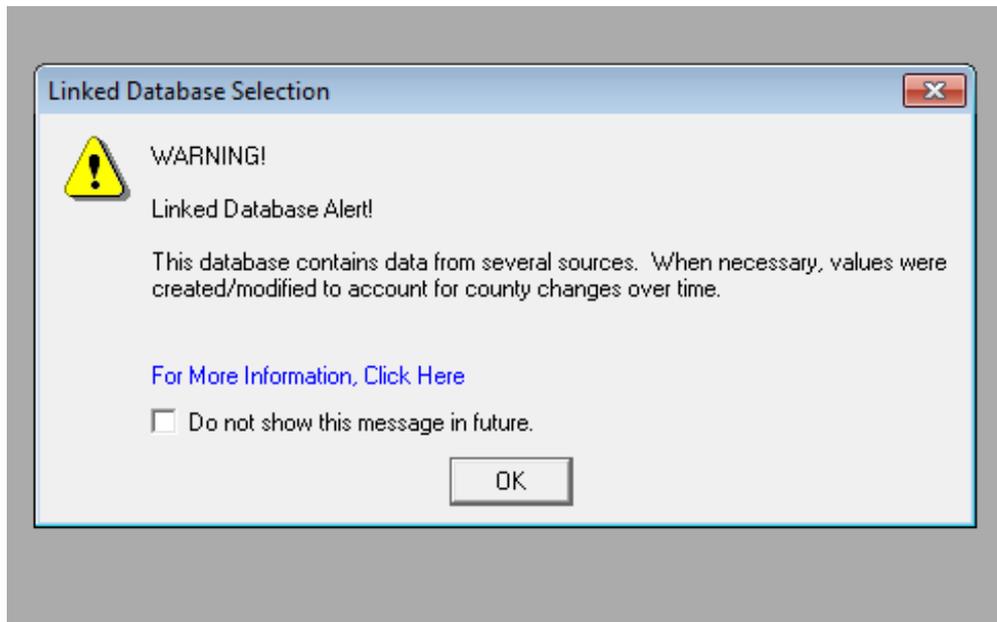
Creates a new rate session



- A login window may appear to connect you to the server.
- Log in using the username and password given to you through an email from SEER.
- Click "OK".



- Messages will pop up.
- You may click on the link in the “Linked Database Selection” window to read more about that alert.
- Click “OK” for both the welcome and warning messages to continue.
- Now you may proceed with your query.



# Downloading binary SEER Data to use in SEER\*Software while off- line



University of Nebraska  
Medical Center

# Downloading SEER Data to use in SEER\*Software while off-line

- **This section will instruct you on how to download binary SEER data to be able to query the data in SEER\*Stat while not connected to the internet.**
- **If you have an internet connection, you can instantly access SEER data by simply entering your username and password when you open SEER\*Stat. If you have regular access to the internet, you may want to skip this section.**



Datasets
(-) SEER Data 1973-2014
• Accessing the Data

## Options for Accessing the Data and SEER\*Stat Software

There are three ways to obtain the research data: two require access through an Internet connection, and the third is a DVD that is best suited for users who do not have an Internet connection.

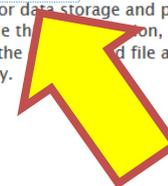
- There are two types of SEER data for download:
  - Binary (for use in off-line SEER\*Stat)
  - ASCII (for use in SAS)
- To download the binary dataset to use off-line in SEER\*Stat, go to <https://seer.cancer.gov/data/options.html>
- Click “Windows Executable” found in binary version option.

(-) SEER*Stat
(+) SEER*Prep
(+) HD*Calc
Analytic Software
Documentation & Recodes
Behavior Recode for Analysis
(+) Cancer Stage Variables

### 2. Download compressed files from the Internet

The combination of these compressed (Windows self-extracting ZIP format) files contain the exact contents of the DVD that is shipped. The username and password you receive from SEER is needed to download the files.

- **Binary version of the data and the SEER\*Stat software:** [Windows Executable](#) [EXE - 1.2 GB]  
Download this file if you would like to use your computer system for data storage and processing. These binary data files can only be analyzed using the SEER\*Stat software. To use the files, download the file, uncompress it, and install SEER\*Stat on your local system. Storing the compressed file and its uncompressed contents will consume approximately 4.5 gigabytes of disk capacity.



Access Options - SEER Datasets

https://seer.cancer.gov/data/options.html

U.S. Mortality Data

- U.S. Population Data
  - County Attributes
- Expected Survival Life Tables
  - SEER Linked Databases
- Specialized SEER\*Stat Datasets
- Statistical Software
  - SEER\*Stat
  - SEER\*Prep
  - HD\*Calc
- Analytic Software
- Documentation & Recodes
  - Behavior Recode for Analysis
- Cancer Stage Variables
  - Cause of Death Recode
  - Cause-specific Death Classification
- Incidence Site Recode Variables
  - Insurance Recode
  - Months Survived Based on Complete Dates
  - Policy for Calculating Hispanic Mortality
  - Race Recode
- SEER Coding and Staging Manuals
- Site Specific Surgery (1983-1997)

Tools

https://seerstat.imsweb.com/.cd\_images/SEER\_1973\_2014\_SEERSTAT.d04062017.exe

When you submit a request for the data and choose the access option, "Through my Internet connection", you will be given a username and password within two business days of SEER receiving your signed research data agreement. This username and password may be used to access the data through either of these two Internet-based options.

### 1. SEER\*Stat's Client-Server Mode

Use SEER\*Stat to access the data through your Internet connection. With this, you only need to download and install the current version of the [SEER\\*Stat Installation](#) program. You must be connected to the Internet while using SEER\*Stat. SEER\*Stat will access the Internet as needed. For example, each time you execute an analysis, the parameters of your request will be transferred from your computer to the SEER\*Stat server and the results will be transferred back to your computer. SEER\*Stat's local and client-server modes are compared in [SEER\\*Stat's help system](#).

### 2. Download compressed files from the Internet

Download this file if you want to use your own programs to analyze the data in text format. You do not need to download these files if you are using SEER\*Stat. SEER does not provide programming support for the analysis of these data. To use this data, you must download the file in the preferred format and uncompress it. File descriptions and documentation are included and are available online in [Documentation for the ASCII Text Data Files](#). Storing the downloaded and uncompressed contents will consume approximately 4 gigabytes of disk capacity.

Download this file if you want to use your own programs to analyze the data in text format. You do not need to download these files if you are using SEER\*Stat. SEER does not provide programming support for the analysis of these data. To use this data, you must download the file in the preferred format and uncompress it. File descriptions and documentation are included and are available online in [Documentation for the ASCII Text Data Files](#). Storing the downloaded and uncompressed contents will consume approximately 4 gigabytes of disk capacity.

Download this file if you want to use your own programs to analyze the data in text format. You do not need to download these files if you are using SEER\*Stat. SEER does not provide programming support for the analysis of these data. To use this data, you must download the file in the preferred format and uncompress it. File descriptions and documentation are included and are available online in [Documentation for the ASCII Text Data Files](#). Storing the downloaded and uncompressed contents will consume approximately 4 gigabytes of disk capacity.

### DVD via US Mail

### 3. SEER Research Data and SEER\*Stat

The SEER research data and the SEER\*Stat software are available for download. If you are using SEER\*Stat and do not have high speed internet access, you may want to download the data via internet connection for the initial download of the data, and the ASCII data files that are available for download in compressed format only.

Authentication Required

https://seerstat.imsweb.com is requesting your username and password. The site says: "SEER Limited-Use CDs"

User Name: XXXXXXXX

Password: ●●●●●●

OK Cancel

Enter the exact contents of the DVD to download the files.

[EXE - 1.2 GB]

storage and processing. These binary configuration, download the file, downloaded file and its uncompressed

[ZIP - 358 MB]

you are using the data. The require the format of and ASCII

Return to top

Entering the username and password you received from SEER via email and then click "OK".

Interactive Tools | Publications | **For Researchers** Datasets and Software | **For Cancer Registrars** Coding Rules, Training and Support | **About** Our P...

SEER Data 1973–2014

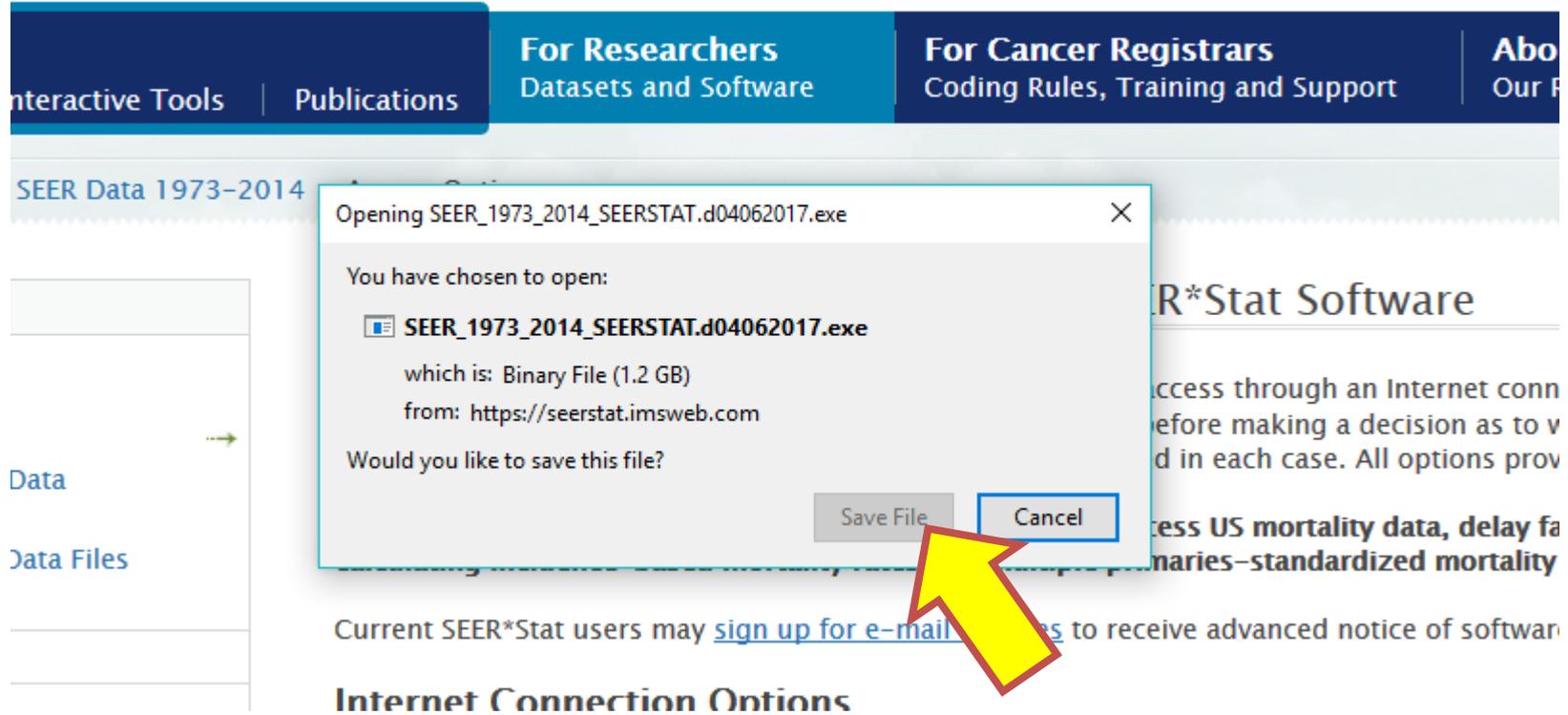
**SEER\*Stat Software**

access through an Internet connection before making a decision as to whether to proceed in each case. All options provided...

Access US mortality data, delay factors, primary-standardized mortality rates...

Current SEER\*Stat users may [sign up for e-mail](#) to receive advanced notice of software updates.

**Internet Connection Options**



Click "Save File".

Access Options - SEER Datasets x +

https://seer.cancer.gov/data/options.html

seer stat

**Cancer Statistics**  
 Statistical Summaries | Interactive Tools | Publications

**For Researchers**  
 Datasets and Software

**For Cancer Registrars**  
 Coding Rules, Training and Support

**About SEER**  
 Our Registries and P

Home ▶ For Researchers ▶ SEER Data 1973–2014 ▶ Access Options

**Datasets**

- [-] SEER Data 1973–2014
  - Accessing the Data
  - Access Options
  - Submit Request for the Data
  - SEER\*Stat Databases
  - Documentation for the Data Files
  - Suggested Citations
- [+] Standard Population Data
  - U.S. Mortality Data
- [+] U.S. Population Data
  - County Attributes
- [+] Expected Survival Life Tables
  - SEER Linked Databases
- [+] Specialized SEER\*Stat Datasets
- Statistical Software**
  - [+] SEER\*Stat
  - [+] SEER\*Prep
  - [+] HD\*Calc
    - Analytic Software
- Documentation & Recodes**
  - Behavior Recode for Analysis
- [+] Cancer Stage Variables
  - Cause of Death Recode
  - Cause-specific Death Classification
- [+] Incidence Site Recode Variables
  - Insurance Recode
- Months Survived Based on Complete

**Options for Accessing the Data and SEER\*Stat Software**

There are three ways to obtain the research data: two require a DVD sent via US mail. Please read through each option below to determine which option best meets your needs. A signed [SEER Research Data Agreement](#) is required for all options.

**You must use SEER\*Stat in client-server mode in order to calculate incidence-based mortality rates and multiple primary cancer rates.**

Current SEER\*Stat users may [sign up for e-mail updates](#) to receive the latest information.

**Internet Connection Options**

When you submit a request for the data and choose the access option, you will be given a username and password within two business days of your request. The username and password may be used to access the data through the Internet.

**1. SEER\*Stat's Client-Server Mode**

Use SEER\*Stat to access the data through your Internet connection. You must use the current version of the [SEER\\*Stat Installation](#) program. You must use the SEER\*Stat software to access the Internet as needed. For example, each request will be transferred from your computer to the SEER\*Stat computer. SEER\*Stat's local and client-server modes are compared below.

**2. Download compressed files from the Internet**

The combination of these compressed (Windows self-extracting) files and the software that is shipped. The username and password you receive from SEER\*Stat will be used to access the data through the Internet.

- Binary version of the data and the SEER\*Stat software:** Download this file if you would like to use your computer to analyze the data. The data files can only be analyzed using the SEER\*Stat software. You must decompress it, and install SEER\*Stat on your local system. Storing the downloaded file and its uncompressed contents will consume approximately 4.5 gigabytes of disk capacity.
- ASCII text version of the data:** [Windows Executable](#) [EXE - 358 MB] | [ZIP](#) [ZIP - 358 MB] Download this file if you would like to use your own programs to analyze the data in text format. You do not need to download these files if you are using SEER\*Stat. SEER does not provide programming support for the analysis of these data. To use this configuration, download the file in the preferred format and uncompress it. File descriptions and documentation are included and are available online in [Documentation for the ASCII Text Data Files](#). Storing the downloaded file and its uncompressed contents will consume approximately 4 gigabytes of disk capacity.

- Wait for the file to download and open your downloads folder from your internet browser.
- Where you access your downloads will differ based on the internet browser you are using.

Options for Accessing

There are three ways to obtain the DVD sent via US mail. Please read the instructions for your needs. A signed [SEER Research](#)

You must use SEER\*Stat in client-mode for calculating incidence-based mortality rates.

Current SEER\*Stat users may [sign up for e-mail](#) to receive advanced notice of software releases.

Internet Connection Options

When you submit a request for the data and choose the access option, "Through my Internet connection," you will receive the data in a self-extractor file.

• ASCII text version of the data: [Windows Executable](#) [EXE - 358 MB] | [ZIP](#) [ZIP - 358 MB]

Download this file if you would like to use your own programs to analyze the data in text format. You do not need to download these files if you are using SEER\*Stat. To use this configuration, download the [SEER\\*Stat](#) descriptions and documentation are included in the [Files](#). Storing the downloaded file and its unzipped files may require significant hard drive capacity.

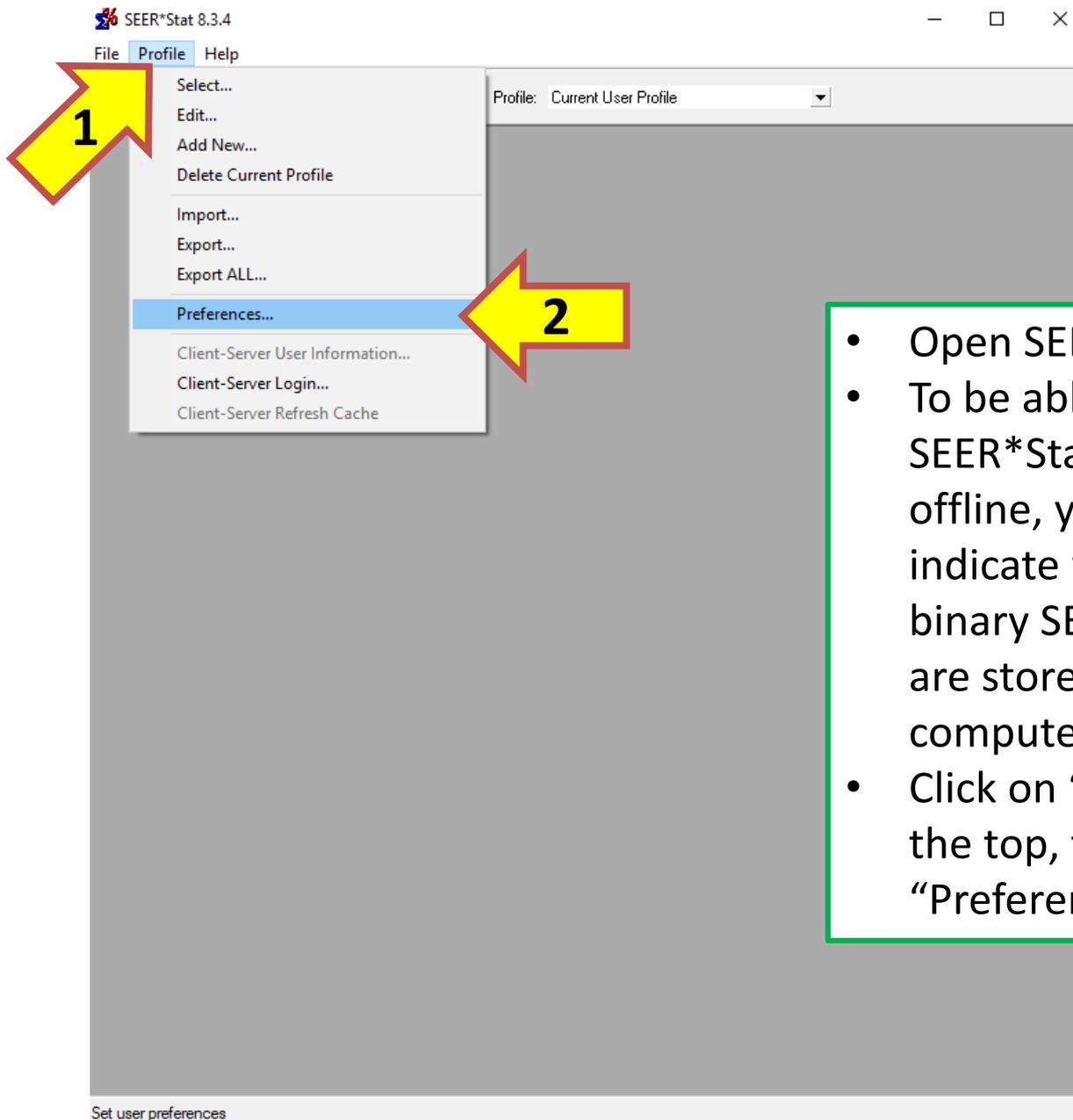
DVD via US Mail

- Click on the .exe file to download the binary dataset.
- It will ask you where you want to unzip the file to store the data.
- Specify the location and then click "Unzip".
- That is the location where your binary SEER data will be stored on your computer.

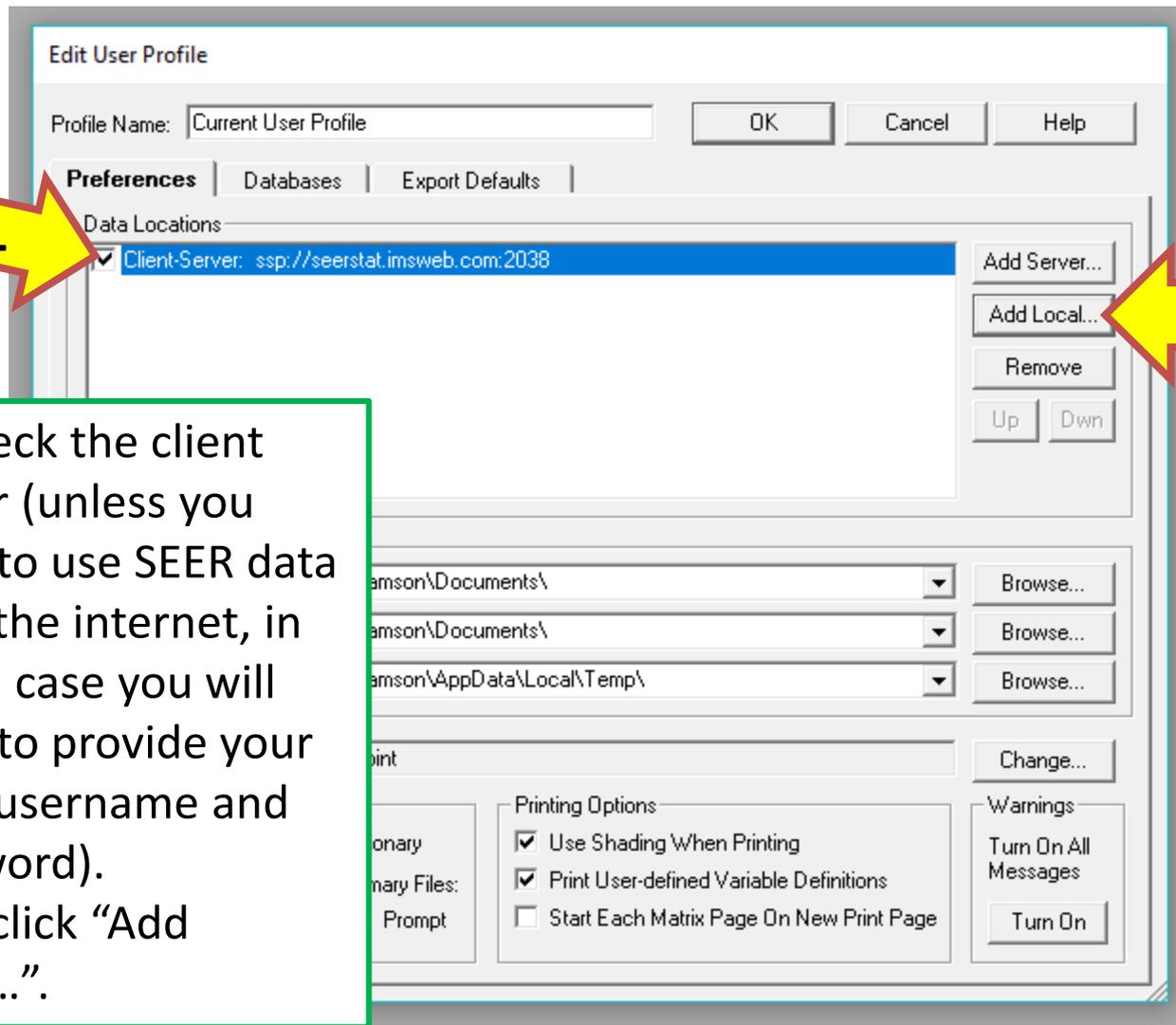
# Setting up SEER\*Stat to use downloaded binary data

- **Open up your SEER\*Stat software by double clicking on the SEER\*Stat icon on your desktop.**

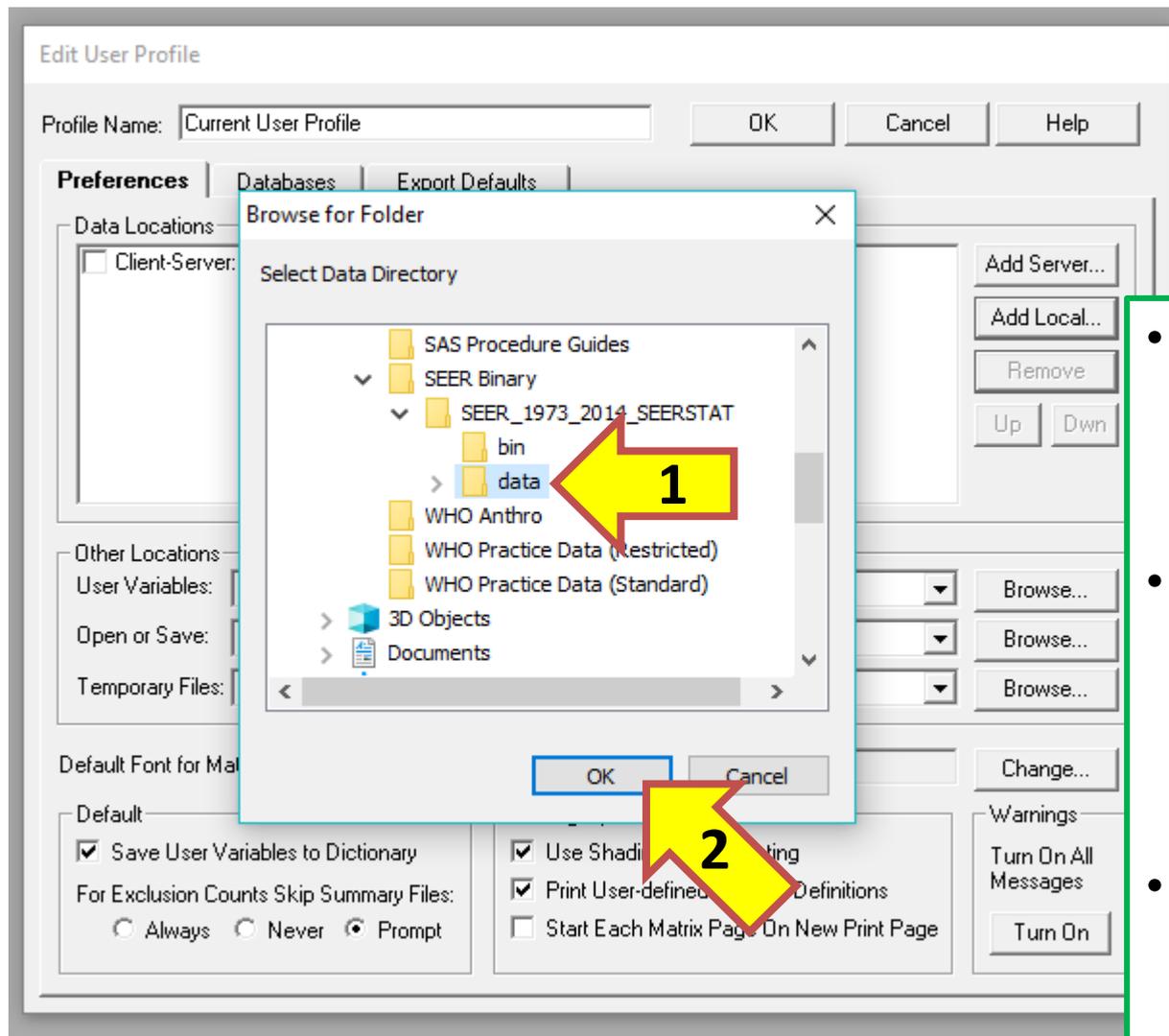




- Open SEER\*Stat.
- To be able to utilize SEER\*Stat software offline, you need to indicate where your binary SEER data files are stored on your computer.
- Click on “Profile” at the top, then “Preferences”.



- Uncheck the client server (unless you want to use SEER data from the internet, in which case you will need to provide your SEER username and password).
- Next click “Add Local...”.



- Locate the SEER Binary data folder you downloaded from the SEER website.
- Select the “data” sub-folder inside the “SEER\_1973\_2014\_SEERSTAT” folder and click “OK”.
- Then click “OK” to close the “Edit User Profile” window.
- Now you may run analyses offline.

# Downloading ASCII SEER Data to use in SAS (and reading the file into SAS)



# Downloading SEER Data to use in SAS

- **This section will instruct you on how to download SEER data to be able to use in SAS.**
- **Not many people will use this option, as SEER\*Stat is the most user-friendly way to access SEER data and calculate age-adjusted rates.**
- **Note: this ASCII data cannot be used in SEER\*Stat; for that, you need to download the binary data (which is covered in the previous section).**



Browser address bar: <https://seer.cancer.gov/data/options.html>

Navigation: Home > For Researchers > SEER Data 1973–2014 > Access Options

**Options for Accessing the Data and SEER\*Stat Software**

There are three ways to obtain the research data: two require access through an Internet connection, and the third is a DVD sent via US mail. Please read through each option below before making a decision as to which methods best suit your needs. A signed [SEER Research Data Agreement](#) is required in each case. All options provide access to the SEER Data.

**You must use SEER\*Stat in client-server mode in order to access US mortality data, delay factors, and databases for calculating incidence-based mortality rates and multiple primaries-standardized mortality ratios (MP-SMRs).**

Current SEER\*Stat users may [sign up for e-mail updates](#) to receive advanced notice of software releases.

**Internet Connection Options**

When you submit a request for the data and choose the access option, "Through my Internet connection", you will be given a username and password within two business days of SEER receiving your signed research data agreement. This username and password may be used to access the data through either of these two Internet-based options.

**1. SEER\*Stat's Client-Server Mode**

Use SEER\*Stat to access the data through your Internet connection. With this, you only need to download and install the current version of the [SEER\\*Stat Installation](#) program. You must be connected to the Internet while using SEER\*Stat. SEER\*Stat will access the Internet as needed. For example, each time you execute an analysis, the parameters of your request will be transferred from your computer to the SEER\*Stat server and the results will be transferred back to your computer. SEER\*Stat's local and client-server modes are compared in [SEER\\*Stat's help system](#).

**2. Download compressed files from the Internet**

The combination of these compressed (Windows self-extracting ZIP format) files contain the exact contents of the DVD that is shipped. The username and password you receive from SEER is needed to download the files.

- **Binary version of the data and the SEER\*Stat software:** [Windows Executable](#) [EXE - 358 MB] | [ZIP](#) [ZIP - 358 MB] [2 GB]  
Download this file if you would like to use your computer system for data analysis and processing. These binary data files can only be analyzed using the SEER\*Stat software. To use this configuration, download the file, uncompress it, and install SEER\*Stat on your local system. Storing the downloaded file and its uncompressed contents will consume approximately 4.5 gigabytes of disk capacity.
- **ASCII text version of the data:** [Windows Executable](#) [EXE - 358 MB] | [ZIP](#) [ZIP - 358 MB]  
Download this file if you would like to use your own programs to analyze the data in text format. You do not need to download these files if you are using SEER\*Stat. SEER does not provide programming support for the analysis of these data. To use this configuration, download the file in the preferred format and uncompress it. File descriptions and documentation are included and are available online in [Documentation for the ASCII Text Data Files](#). Storing the downloaded file and its uncompressed contents will consume approximately 4 gigabytes of disk capacity.

- Go to: <https://seer.cancer.gov/data/options.html>
- Click the zip link associated with the ASCII text version of the data.
- You will be prompted to enter the username and password SEER sent.



Datasets
<a href="#">[-] SEER Data 1973-2014</a>
<ul style="list-style-type: none"><li><a href="#">Accessing the Data</a></li><li><a href="#">Access Options</a></li><li><a href="#">Submit Request for the Data</a></li><li><a href="#">SEER*Stat Databases</a></li><li><a href="#">Documentation for the Data Files</a></li><li><a href="#">Suggested Citations</a></li></ul>
<a href="#">[+] Standard Population</a>
<a href="#">U.S. Mortality Data</a>
<a href="#">[+] U.S. Population Data</a>
<a href="#">County Attribute Data</a>
<a href="#">[+] Expected Survival</a>
<a href="#">SEER Linked Data</a>
<a href="#">[+] Specialized SEER Data</a>
Statistical Software
<a href="#">[+] SEER*Stat</a>
<a href="#">[+] SEER*Prep</a>
<a href="#">[+] HD*Calc</a>
Analytic Software
Documentation & Downloads
<a href="#">Behavior Recode for Analysis</a>
<a href="#">[+] Cancer Stage Variables</a>
<a href="#">Cause of Death Recode</a>
<a href="#">Cause-specific Death Classification</a>
<a href="#">[+] Incidence Site Recode Variables</a>
<a href="#">Insurance Recode</a>
<a href="#">Months Survived Based on Complete Dates</a>

## Options for Accessing the Data and SEER\*Stat Software

There are three ways to obtain the research data: two require access through an Internet connection, and the third is a DVD sent via US mail. Please read through each option below before making a decision as to which methods best suit your needs. A signed [SEER Research Data Agreement](#) is required in each case. All options provide access to the SEER Data

You must use SEER\*Stat in client-server mode in order to access the data and calculate incidence-based mortality rates and multiple primary cancer rates.

Opening SEER\_1973\_2014\_TEXTDATA.d04122017.zip

You have chosen to open:

- SEER\_1973\_2014\_TEXTDATA.d04122017.zip

which is: Compressed (zipped) Folder (359 MB)  
from: https://seerstat.imsweb.com

What should Firefox do with this file?

Open with: Windows Explorer (default)

Save File

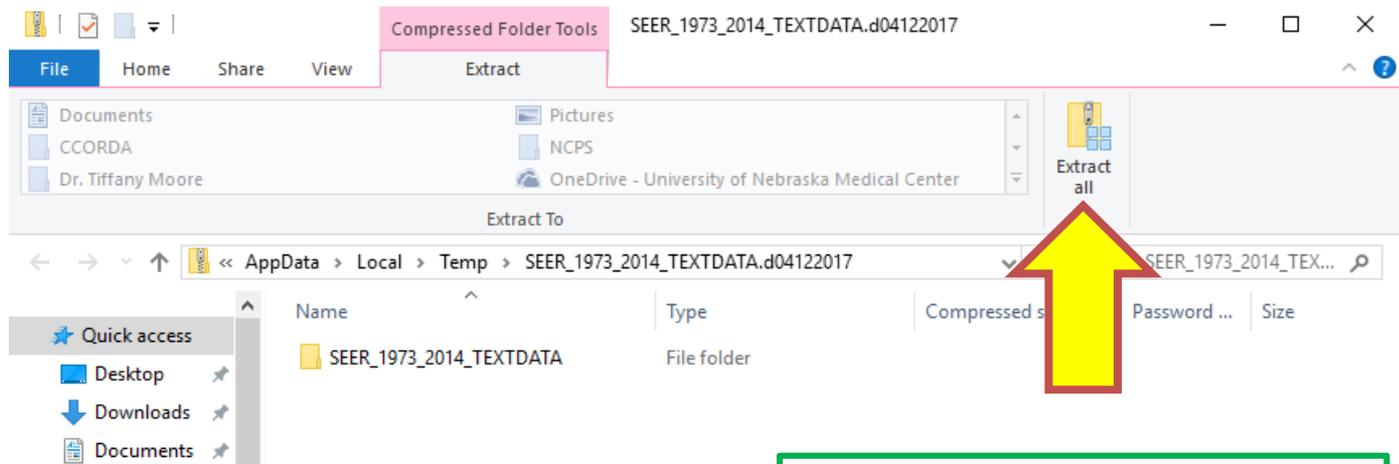
Do this automatically for files like this from now on.

OK Cancel

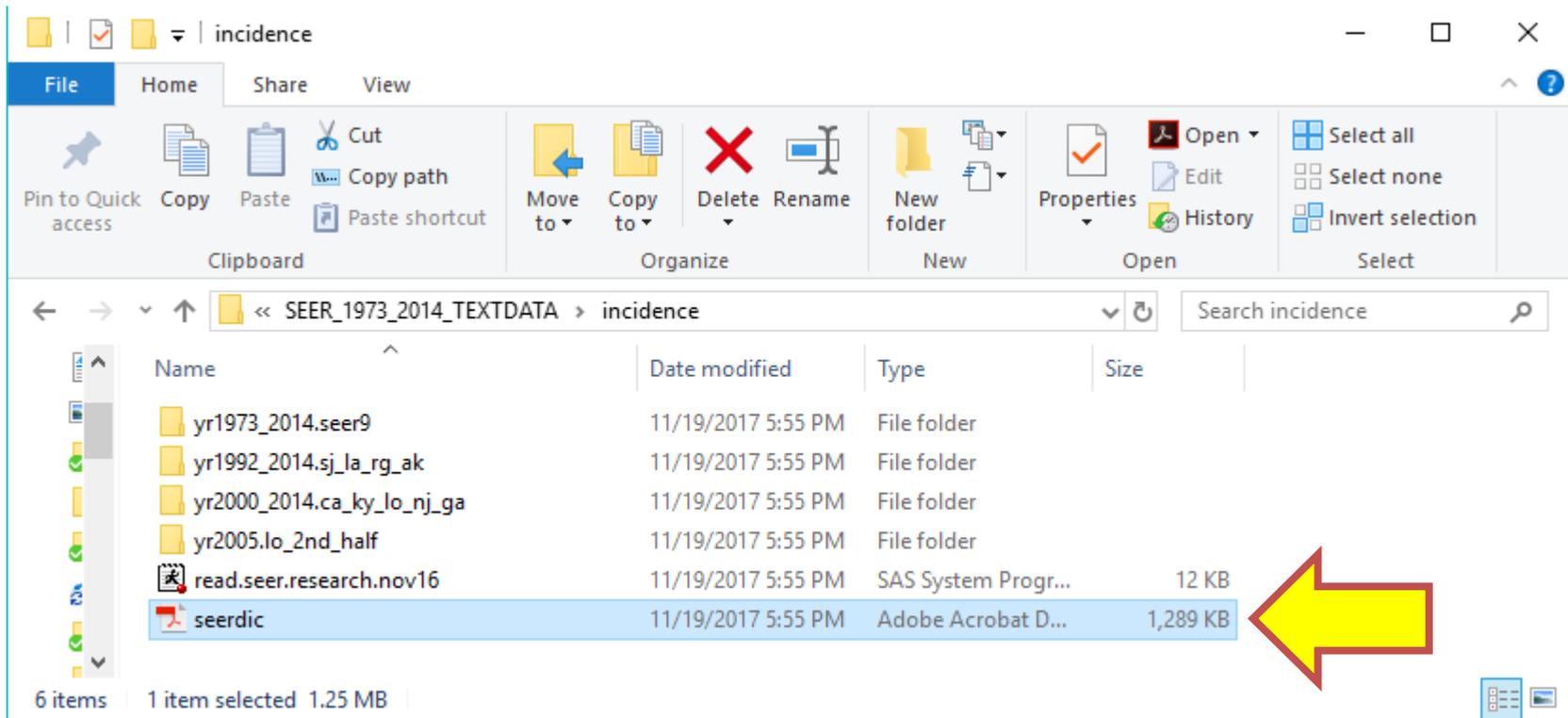
After entering your username and password, wait for the zip file to download and open your downloads folder from your internet browser, or click "OK" if you get a similar pop-up window.



- Download text version of the data:** [Windows Executable](#) [EXE - 1.2 GB] | [ZIP](#) [ZIP - 358 MB]
  - Download text version of the data:** [Windows Executable](#) [EXE - 358 MB] | [ZIP](#) [ZIP - 358 MB]
- Download this file if you would like to use your own programs to analyze the data in text format. You do not need to download the SEER\*Stat software. SEER does not provide programming support for the analysis of these data. To use this configuration, download the file in the preferred format and uncompress it. File descriptions and documentation are included and are available online in [Documentation for the ASCII Text Data Files](#). Storing the downloaded file and its uncompressed contents will consume approximately 4 gigabytes of disk capacity.



- Extract the contents of the zipped folder by clicking the “Extract all” button.
- This will prompt you for the location where you would like to save the contents.



- This is what the contents of the incidence extracted folder looks like.
- Double click on the “seerdic” PDF, which is the documentation for this data.

---

### MARITAL STATUS AT DX

**NAACCR Item #: 150**

**SEER\*Stat Name: Marital status at diagnosis**

**Item Length: 1**

*Field Description:* This data item identifies the patient's marital status at the time of diagnosis for the reportable tumor.

Code	Description
1	Single (never married)
2	Married (including common law)
3	Separated
4	Divorced
5	Widowed
6	Unmarried or domestic partner (same sex or opposite sex or unregistered)
9	Unknown

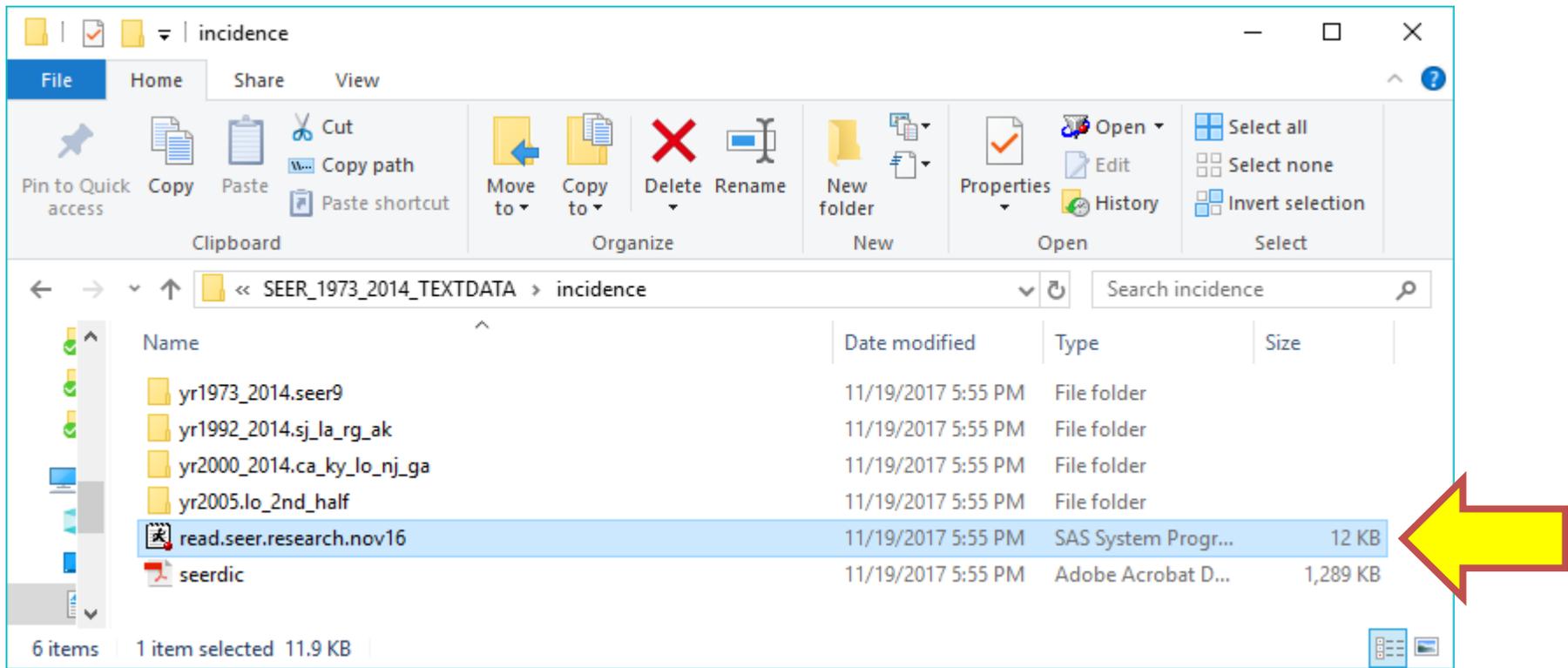
---

- This is an example of what is included in the documentation.
- Here, it tells you specific information about a specific variable.

# Reading in SEER data using SAS

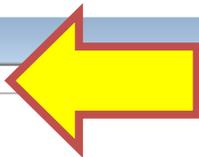
- **Example: Read in colorectal cancer incidence data from the “yr2000\_2014.ca\_ky\_lo\_nj\_ga” data folder downloaded from the SEER website.**



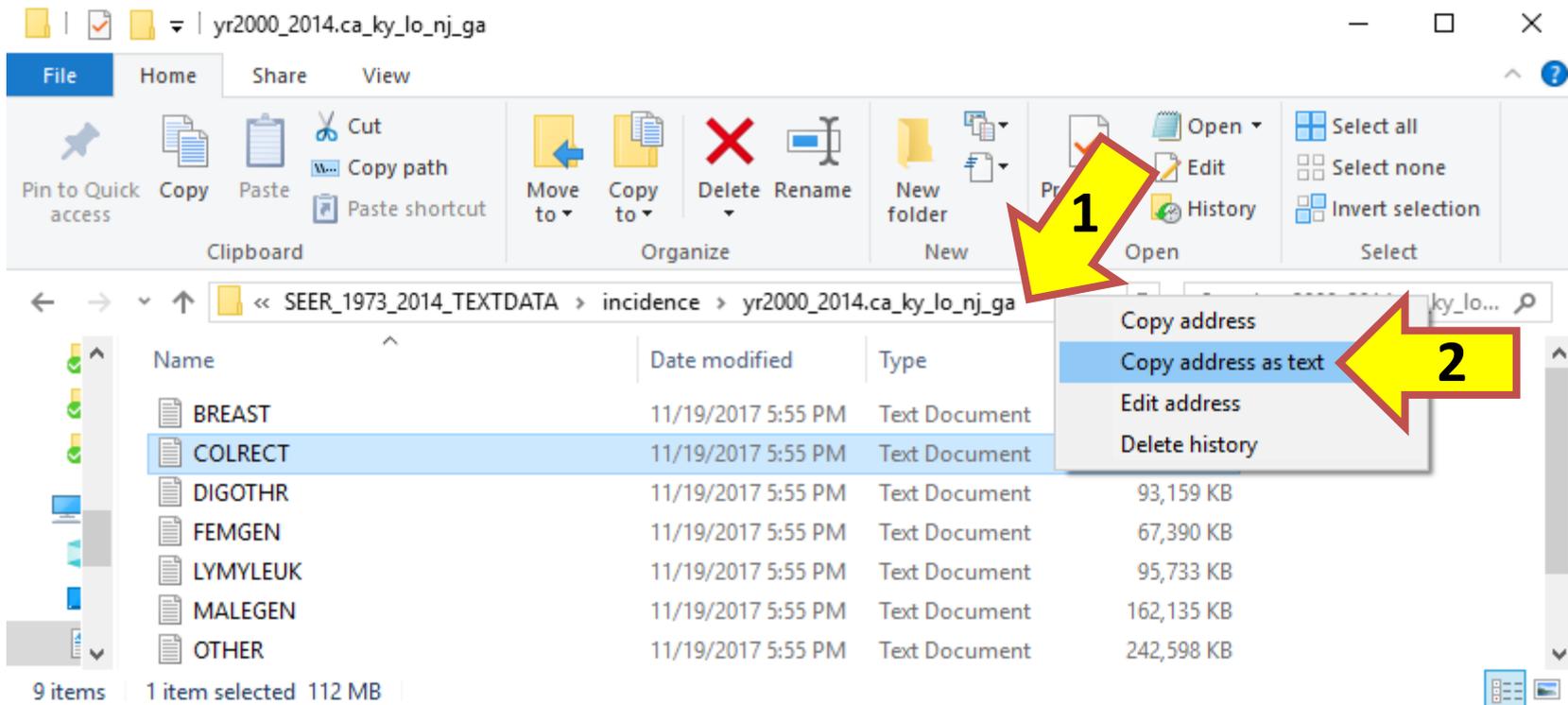


- The “read.seer.research.nov16” in the “incidence” folder is the SAS file you can use to read in the individual .txt data files.
- The .txt files are contained in the four folders listed above.
- Double click on the SAS file.

```
read.seer.research.nov16
filename seer9 './yr1973_2014.seer9/*.TXT';
data in;
  infile seer9 lrecl=362;
  input
    @ 1  PUBCSNUM          $char8. /* Patient ID */
    @ 9  REG               $char10. /* SEER registry */
    @ 19 MAR_STAT         $char1. /* Marital status at diagnosis */
    @ 20 RACE1V           $char2. /* Race/ethnicity */
    @ 23 NHIAD           $char1. /* NHIA Derived Hisp Origin */
    @ 24 SEX              $char1. /* Sex */
    @ 25 AGE_DX           $char3. /* Age at diagnosis */
    @ 28 YR_BRTH         $char4. /* Year of birth */
    @ 35 SEQ_NUM         $char2. /* Sequence number */
    @ 37 MDXRECOMP       $char2. /* Month of diagnosis */
    @ 39 YEAR_DX         $char4. /* Year of diagnosis */
    @ 43 PRIMSITE        $char4. /* Primary site ICD-O-2 (1973+) */
    @ 47 LATERAL         $char1. /* Laterality */
    @ 48 HISTO2V         $char4. /* Histologic Type ICD-O-2 */
    @ 52 BEHO2V         $char1. /* Behavior Code ICD-O-2*/
    @ 53 HISTO3V
    @ 57 BEHO3V
    @ 58 GRADE
    @ 59 DX_CONF
    @ 60 REPT_SRC
    @ 61 EOD10_SZ
    ...
```



- This is what the “read.seer.research.nov16” SAS file contains.
- For this example, we will focus on the “yr2000\_2014.ca\_ky\_lo\_nj\_ga” folder.
- The first thing we need to do is point to the correct data file using the appropriate path.
- We must change the purple text in the top line to indicate where the colorectal cancer text file from the “yr2000\_2014.ca\_ky\_lo\_nj\_ga” folder is located, which we’ll do on the next slide.



The easiest way to do this is to open the “yr2000\_2014.ca\_ky\_lo\_nj\_ga” folder (which is contained in the “incidence” folder of the SEER ASCII data downloaded from the website), and then right click the path name at the top of the folder, and click “Copy address as text”.

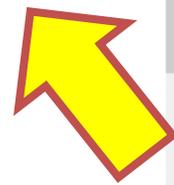
```
read.seer.research.nov16 *
filename seer9 'C:\Documents\SEER_1973_2014_TEXTDATA.d04122017\SEER_1973_2014_TEXTDATA\incidence\yr2000_2014.ca_ky_lo_nj_ga';

data in;
  infile seer9 lrecl=362;
  input
    @ 1 PUBCSNUM          $char8. /* Patient ID */
    @ 9 REG              $char10. /* SEER registry */
    @ 19 MAR_STAT        $char1. /* Marital status at diagnosis */
    @ 20 RACE1V          $char2. /* Race/ethnicity */
    @ 23 NHIAD           $char1. /* NHIA Derived Hisp Origin */
    @ 24 SEX             $char1. /* Sex */
    @ 25 AGE_DX          $char3. /* Age at diagnosis */
    @ 28 YR_BRTH        $char4. /* Year of birth */
    @ 35 SEQ_NUM         $char2. /* Sequence number */
    @ 37 MDXRECMP        $char2. /* Month of diagnosis */
    @ 39 YEAR_DX         $char4. /* Year of diagnosis */
    @ 43 PRIMSITE        $char4. /* Primary site ICD-0-2 (1973+) */
    @ 47 LATERAL         $char1. /* Laterality */
    @ 48 HISTO2V         $char4. /* Histologic Type ICD-0-2 */
    @ 52 BEHO2V          $char1. /* Behavior Code ICD-0-2 */
    @ 53 HISTO3V         $char4. /* Histologic Type ICD-0-3 */
    @ 57 BEHO3V          $char1. /* Behavior code ICD-0-3 */
    @ 58 GRADE           $char1. /* Grade */
    @ 59 DX_CONF         $char1. /* Diagnostic confirmation */
    @ 60 REPT_SRC        $char1. /* Type of reporting source */
    @ 61 EOD10_SZ        $char3. /* EOD 10 - size (1988+) */
```

- This is what the path for my “yr2000\_2014.ca\_ky\_lo\_nj\_ga” folder looks like.
- Yours will likely differ depending on where you saved the SEER data on your computer.
- Next we need to specify the specific .txt file we are interested in as part of the filename.

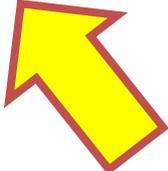
```
read.seer.research.nov16*
filename seer9 'C:\Documents\SEER_1973_2014_TEXTDATA.d04122017\SEER_1973_2014_TEXTDATA\incidence\yr2000_2014.ca_ky_lo_nj_ga\COLRECT.txt';

data in;
  infile seer9 lrecl=362;
  input
    @ 1  PUBCSNUM          $char8.  /* Patient ID */
    @ 9  REG               $char10. /* SEER registry */
    @ 19 MAR_STAT         $char1.   /* Marital status at diagnosis */
    @ 20 RACE1V           $char2.   /* Race/ethnicity */
    @ 23 NHIAD           $char1.   /* NHIA Derived Hisp Origin */
    @ 24 SEX              $char1.   /* Sex */
    @ 25 AGE_DX           $char3.   /* Age at diagnosis */
    @ 28 YR_BRTH          $char4.   /* Year of birth */
    @ 35 SEQ_NUM          $char2.   /* Sequence number */
    @ 37 MDXRECOMP        $char2.   /* Month of diagnosis */
    @ 39 YEAR_DX          $char4.   /* Year of diagnosis */
    @ 43 PRIMSITE         $char4.   /* Primary site ICD-O-2 (1973+) */
    @ 47 LATERAL          $char1.   /* Laterality */
    @ 48 HISTO2V          $char4.   /* Histologic Type ICD-O-2 */
    @ 52 BEHO2V           $char1.   /* Behavior Code ICD-O-2 */
    @ 53 HISTO3V          $char4.   /* Histologic Type ICD-O-3 */
    @ 57 BEHO3V           $char1.   /* Behavior code ICD-O-3 */
    @ 58 GRADE            $char1.   /* Grade */
    @ 59 DX_CONF          $char1.   /* Diagnostic confirmation */
    @ 60 REPT_SRC         $char1.   /* Type of reporting source */
    @ 61 EOD10_SZ         $char3.   /* EOD 10 - size (1988+) */
```



To specify the .txt file with data for colorectal cancers, I added the name of the .txt file with a backslash and the extension, “\COLRECT.txt”, to the end of the address for the “yr2000\_2014.ca\_ky\_lo\_nj\_ga” folder.

```
read.seer.research.nov16 *
@ 330 ADJAJCCSTG      $char2. /* Adjusted AJCC 6th Stage (1988+) */
@ 332 CS7SITE        $char3. /* CS Site-Specific Factor 7 */
@ 335 CS9SITE        $char3. /* CS Site-specific Factor 9 */
@ 338 CS12SITE       $char3. /* CS Site-Specific Factor 12 */
@ 341 HER2           $char1. /* Derived HER2 Recode (2010+) */
@ 342 BRST_SUB       $char1. /* Breast Subtype (2010+) */
@ 348 ANNARBOR       $char1. /* Lymphoma - Ann Arbor Stage (1983+) */
@ 349 CSMETSDXB_PUB  $char1. /* CS mets at DX-bone (2010+) */
@ 350 CSMETSDXBR_PUB $char1. /* CS mets at DX-brain (2010+) */
@ 351 CSMETSDXLIV_PUB $char1. /* CS mets at DX-liver (2010+) */
@ 352 CSMETSDXLUNG_PUB $char1. /* CS mets at DX-lung (2010+) */
@ 353 T_VALUE        $char2. /* T value - based on AJCC 3rd (1988-2003) */
@ 355 N_VALUE        $char2. /* N value - based on AJCC 3rd (1988-2003) */
@ 357 M_VALUE        $char2. /* M value - based on AJCC 3rd (1988-2003) */
@ 359 MALIGCOUNT    $char2. /* Total number of in situ/malignant tumors for patient */
@ 361 BENBORDCOUNT $char2. /* Total number of benign/borderline tumors for patient */
run;
```



- To create a SAS dataset, you will need to add “run;” to the bottom of the code provided by SEER.
- Now you can highlight all of the SAS code in the file and run the program.
- It will read in the colorectal text file and name the new dataset created from that text file “in”.

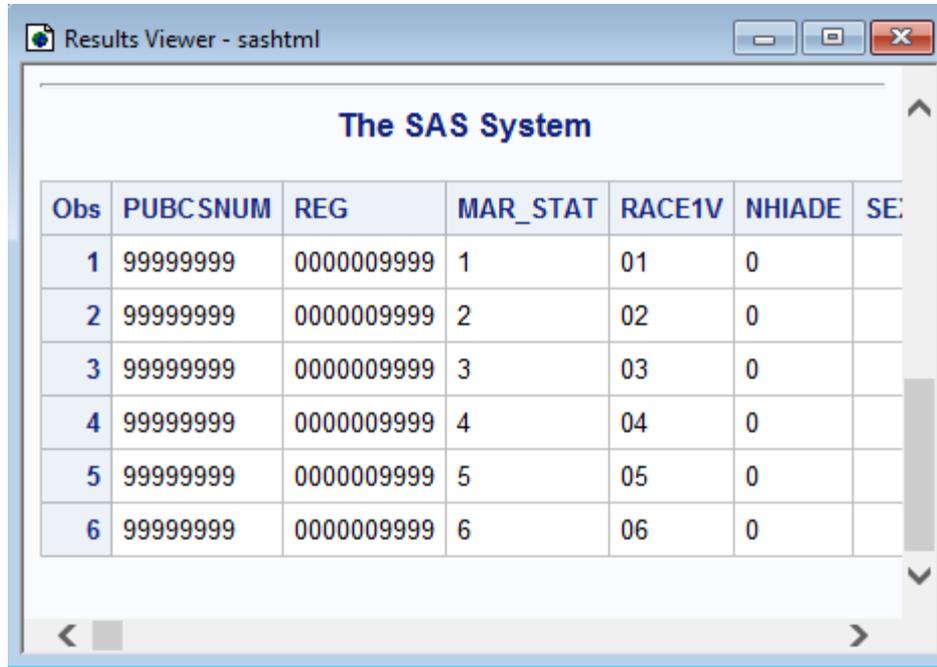
```
read.seer.research.nov16*
@ 330 ADJAJCCSTG      $char2. /* Adjusted AJCC 6th Stage (1988+) */
@ 332 CS7SITE        $char3. /* CS Site-Specific Factor 7 */
@ 335 CS9SITE        $char3. /* CS Site-specific Factor 9 */
@ 338 CS12SITE       $char3. /* CS Site-Specific Factor 12 */
@ 341 HER2           $char1. /* Derived HER2 Recode (2010+) */
@ 342 BRST_SUB       $char1. /* Breast Subtype (2010+) */
@ 348 ANNARBOR       $char1. /* Lymphoma - Ann Arbor Stage (1983+) */
@ 349 CSMETSDXB_PUB  $char1. /* CS mets at DX-bone (2010+) */
@ 350 CSMETSDXBR_PUB $char1. /* CS mets at DX-brain (2010+) */
@ 351 CSMETSDXLIV_PUB $char1. /* CS mets at DX-liver (2010+) */
@ 352 CSMETSDXLUNG_PUB $char1. /* CS mets at DX-lung (2010+) */
@ 353 T_VALUE        $char2. /* T value - based on AJCC 3rd (1988-2003) */
@ 355 N_VALUE        $char2. /* N value - based on AJCC 3rd (1988-2003) */
@ 357 M_VALUE        $char2. /* M value - based on AJCC 3rd (1988-2003) */
@ 359 MALIGCOUNT    $char2. /* Total number of in situ/malignant tumors for patient */
@ 361 BENBORDCOUNT  $char2. /* Total number of benign/borderline tumors for patient */ ;

run;

proc print data=in (obs=10); *display the first 10 observations in your dataset;
run;
|
```

To see the first 10 observations in your dataset, type in the following code:

```
proc print data=in (obs=10); *display the first 10
observations in your dataset;
run;
```



The screenshot shows a window titled "Results Viewer - sashtml". Inside the window, there is a table with the title "The SAS System". The table has 7 columns: "Obs", "PUBCSNUM", "REG", "MAR\_STAT", "RACE1V", "NHIADE", and "SE:". The data is as follows:

Obs	PUBCSNUM	REG	MAR_STAT	RACE1V	NHIADE	SE:
1	99999999	0000009999	1	01	0	
2	99999999	0000009999	2	02	0	
3	99999999	0000009999	3	03	0	
4	99999999	0000009999	4	04	0	
5	99999999	0000009999	5	05	0	
6	99999999	0000009999	6	06	0	

- This is an example of what the data will look like.
- The data has been de-identified, but gives you an idea what the data structure looks like.
- From here, you can do many types of analyses using SAS software procedures.

# Incidence Counts

Example: Breast Cancer

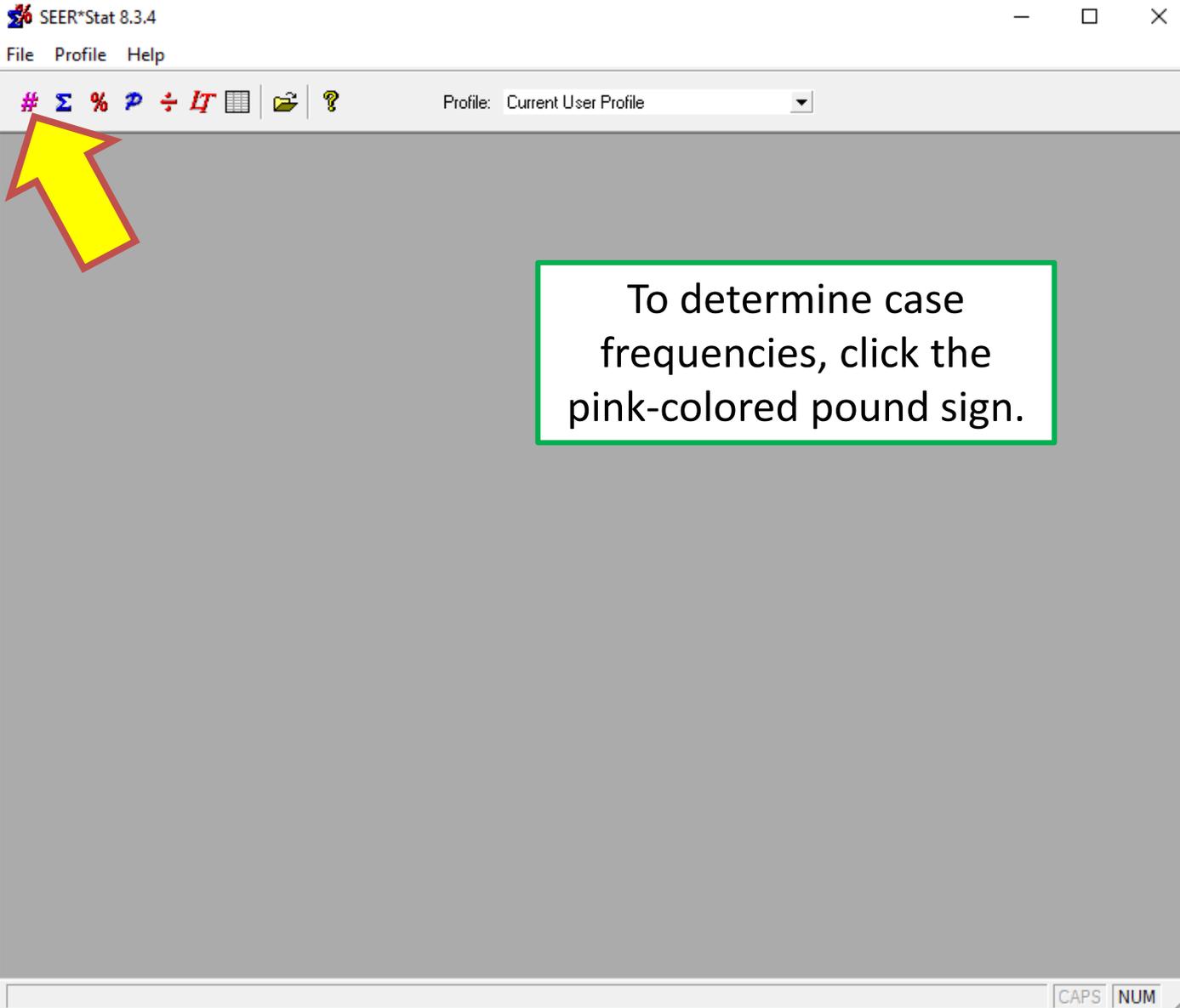


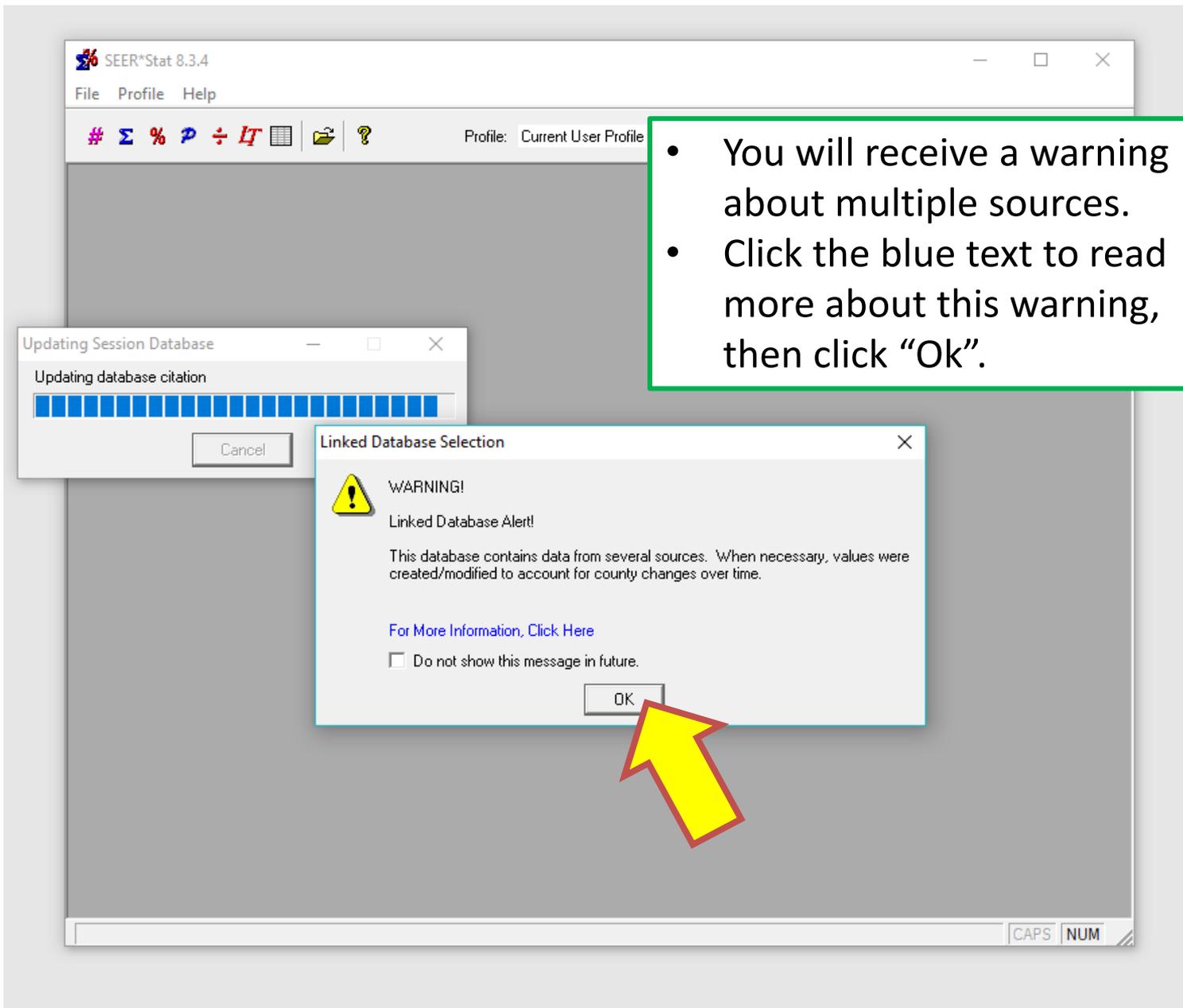
University of Nebraska  
Medical Center

# Incidence Counts

- **This section will instruct you on how to get incidence counts using SEER data**
- **For this example, we will show you to get access to breast cancer incidence counts using the 2016 SEER 18 registries data, with the following criteria:**
  - **Exclude anyone under the age of 15 years old**
  - **Display data by 5-year age categories**
  - **Display data separately by race**
- **This example was done using downloaded binary SEER data for use in SEER\*Stat. If you are using online data instead (i.e. entering your username and password), a couple of your screens and output may be slightly different from the tutorial.**







- You will receive a warning about multiple sources.
- Click the blue text to read more about this warning, then click "Ok".

Frequency Session

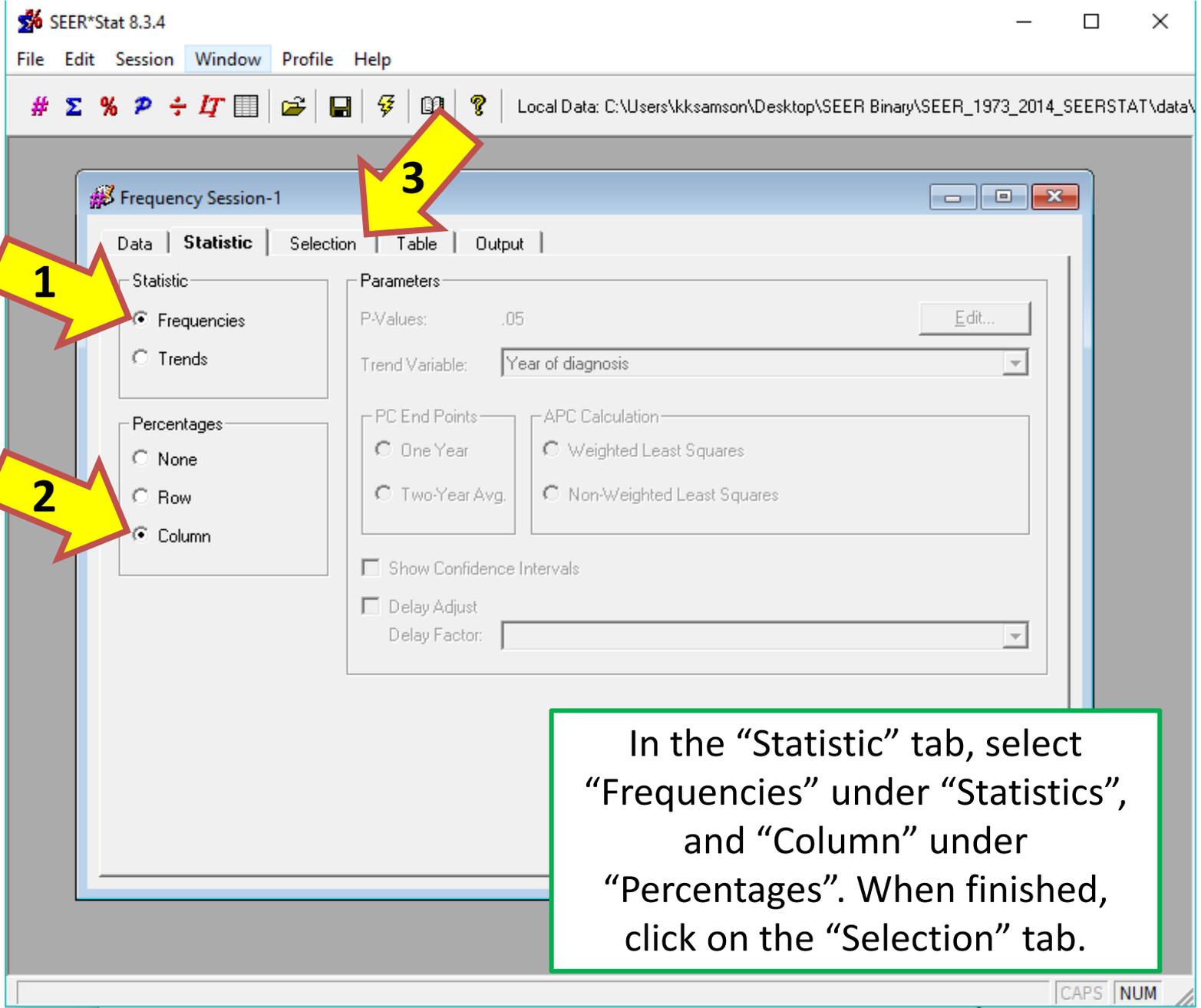
Data | Statistic | Selection | Table | Output

Database Name	Linked
Incidence - SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases, Nov 2016 Sub (1973-2014 varyi... County	County
County Attributes - Total U.S., 1969-2015 Counties	

Find

Suggested citation:  
Surveillance, Epidemiology, and  
Research Data +  
U.S., 1969-2015  
April 2017, based

- For SEER sessions, it's best to work from left to right with the tabs.
- Start by selecting a database from the "Data" tab which will determine what appears in the other tabs.
- Click "Incidence – SEER 18 Regs Research Data...2016" to select it.
  - Note, if you're using online access to SEER\*Data, you will see more options to choose from.
- Next, click on the "Statistic" tab.



In the “Statistic” tab, select “Frequencies” under “Statistics”, and “Column” under “Percentages”. When finished, click on the “Selection” tab.

The screenshot shows the SEER\*Stat 8.3.4 interface. The main window is titled 'Frequency Session-1' and has tabs for 'Data', 'Statistic', 'Selection', 'Table', and 'Output'. The 'Selection' tab is active, showing a 'Select Only' section with four checked options: 'Malignant Behavior', 'Known Age', 'Male or Female Sex', and 'Cases in Research Database'. A yellow arrow points to the 'Edit...' button on the right side of the dialog. Below the 'Select Only' section is a large empty text area. At the bottom of the dialog, there is a checkbox for 'Select Only the First Matching Record for Each Person'. The main window has a menu bar (File, Edit, Session, Window, Profile, Help) and a toolbar with various icons. The status bar at the bottom right shows 'CAPS NUM'.

In the “Selection” tab, click the “Edit...” button to refine which cases you want to include in your query (i.e. inclusion/exclusion criteria).

The screenshot shows the SEER\*Stat 8.3.4 application window. The main window has a menu bar (File, Edit, Session, Window, Profile, Help) and a toolbar with various icons. The local data path is displayed as C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\... A sub-window titled 'Frequency Session-1' is open, showing a 'Selection' tab. In this tab, 'Malignant Behavior' and 'Known Age' are checked. A 'Case Selection' dialog box is overlaid on top. The dialog has three columns: 'Variable', 'Operator', and 'Values'. The 'Variable' list includes folders for 'Age at Diagnosis', 'Race, Sex, Year Dx, Registry, County', 'Site and Morphology', 'Stage - AJCC', 'Stage - TNM', and 'Stage - LRD (Summary and Historic)'. A yellow arrow points to the 'Site and Morphology' folder. The 'Operator' dropdown is set to 'is not = to'. The 'Values' column is empty. At the bottom of the dialog are 'OK', 'Cancel', and 'Help' buttons. A green-bordered text box is overlaid on the bottom left of the screenshot.

- You can use the list of variables to refine your case selection.
- Since we only want breast cancer cases, we will start by clicking the “Site and Morphology” folder.

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

Frequency Session-1

Data | Statistic | **Selection** | Table | Output

Select Only

Malignant Behavior  Known Age [View Statements](#)

Male or Female Sex

**Case Selection**

Variable	Operator	Values
Site and Morphology	is = to	All Sites
Site recode ICD-O-3/WHO 2008	is not = to	Oral Cavity and Pharynx
Behavior recode for analysis		Lip
AYA site recode/WHO 2008		Tongue
Lymphoma subtype recode/WHO 2008		Salivary Gland
ICCC site recode ICD-O-3/WHO 2008		Floor of Mouth
		Gum and Other Mouth
		Nasopharynx
		Tonsil

Modify Conjunction:  [New Line](#)

Subject to change based on evolving ICD-O-3 coding rules. For more information, see <https://seer.cancer.gov/siterecode>.

Selection Statement

(Site and Morphology.Site recode ICD-O-3/WHO 2008) =

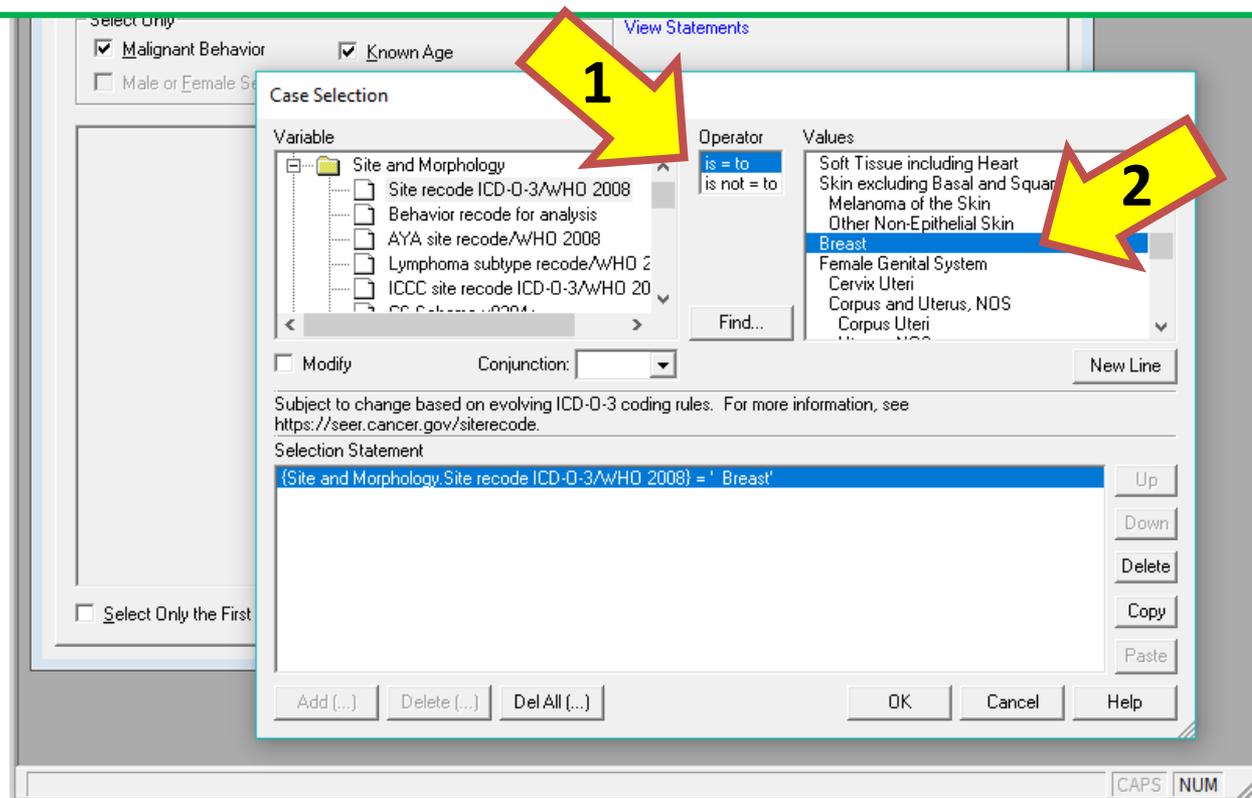
Up Down Delete Copy Paste

OK Cancel Help

CAPS NUM

- Click on the first item listed in the “Site and Morphology” folder called “Site recode ICD-O-3/WHO 2008”.
- When you select this option, values unique to that option show up in the “Values” box.

- Ensure that the “Operator” value “is = to” is selected.
- In the “Values” box, find and select “Breast”.
- In the “Selection Statement” box below, you will see that code was automatically generated to restrict sites to those with values of “Breast”.
- FYI: if you had chosen “is not = to” in the “Operator” box, then your data would have all cancer cases that are not Breast cancer.
- We will now add additional selection criteria by clicking on the “Age at Diagnosis” folder in the “Variable” box.



The screenshot shows the SEER\*Stat 8.3.4 interface. The main window is titled "Case Selection" and contains a tree view of variables on the left, an operator selection area in the middle, and a list of values on the right. The "Selection Statement" field at the bottom contains the text: `{Site and Morphology.Site recode ICD-O-3/WHO 2008} = ' Breast'` and `AND {Age at Diagnosis.Age recode with <1 year olds} = '15-19 years','20-24 years','25-29 years','30-34 years','35-39 years','40-44 years','45-49 years','50-54 years','55-59 years','60-64 years','65-69 years','70-74 years','75-79 years','80-84 years','85+ years','Unknown'`. Four yellow arrows with black outlines point to specific elements: Arrow 1 points to the "Age recode with <1 year olds" folder in the variable list; Arrow 2 points to the "Age recode with <1 year olds" folder in the variable list; Arrow 3 points to the "85+ years" value in the values list; Arrow 4 points to the "OK" button at the bottom of the dialog box.

Variable	Operator	Values
Age at Diagnosis	is = to	50-54 years
Age recode with <1 year olds	is not = to	55-59 years
Race, Sex, Year Dx, Registry, County		60-64 years
Site and Morphology		65-69 years
Stage - AJCC		70-74 years
Stage - TNM		75-79 years
Stage - ICD (Summary and History)		80-84 years
		85+ years
		Unknown

Selection Statement

```
{Site and Morphology.Site recode ICD-O-3/WHO 2008} = ' Breast'
```

```
AND {Age at Diagnosis.Age recode with <1 year olds} = '15-19 years','20-24 years','25-29 years','30-34 years','35-39 years','40-44 years','45-49 years','50-54 years','55-59 years','60-64 years','65-69 years','70-74 years','75-79 years','80-84 years','85+ years','Unknown'
```

Buttons: Up, Down, Delete, Copy, Paste, OK, Cancel, Help

- Click on “Age recode with <1 year olds”, use the operator “is = to” and select all of the age groups between and including 15-19 and 85+ (click “15-19 years”, then hold down shift, then click “85+ years” to select all in the range).
- Now click “Ok”.

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

# Σ % P ÷ IT Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

Frequency Session-1

Data | Statistic | **Selection** | Table | Output

Select Only

Malignant Behavior  Known Age

Male or Female Sex  Cases in Research Database

[View Statements](#)

{Site and Morphology.Site recode ICD-O-3/WHO 2008} = ' Breast'  
AND {Age at Diagnosis.Age recode with <1 year olds} = '15-19 years','20-24 years','25-29 years','30-34 years','35-'

Edit..

Copy

Clear

- You can see that our new selection criteria has been added to the text box.
- The “Select Only” box has commonly requested selection criteria.
- Let’s keep all of these selections checked.
- Next, move to the “Table” tab.

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

Frequency Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

- Page
- Row
- Column

Available Variables

- Age at Diagnosis
- Race, Sex, Year Dx, Registry, County
- Site and Morphology
- Stage - AJCC
- Stage - TNM
- Stage - LRD (Summary and Historic)
- Therapy
- Extent of Disease - CS
- Extent of Disease - Historic
- Cause of Death (COD) and Follow-up
- Multiple Primary Fields

- While the “Selection” tab selected cases, the “Table” tab organizes how your selected data will be displayed.
- Let’s say we want to see a cross tabulation of age and race for breast cancer cases.
- We’ll have race for columns, and age groups for rows.
- First, lets click on the “Race, Sex, Year Dx, Registry, County” folder to find race.

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

# Σ % P ÷ IT Local Data: C:\Users\mksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

Frequency Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

- Page
- Row
- Column

Available Variables

- Race, Sex, Year Dx, Registry, County
  - Race recode (White, Black, Other)**
  - Sex
  - Year of diagnosis
  - SEER registry
  - Louisiana 2005 - 1st vs 2nd half of year
  - County
  - State-county
  - In research data
  - CHSDA 2012
  - CHSDA Region

1

2

Click on "Race recode (White, Black, Other)" and then click the "Column" button to add this variable as a column variable to our table.

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

# Σ % P ÷ IT Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

Frequency Session-1

Data | Statistic | Selection | Table

Display Variables

- Page
- Row
- Column
  - Race recode (White, Black, Other)

- Notice that “Race” has been added as a column in the “Display Variables” box.
- Next, click “Age of Diagnosis”, then “Age recode with <1 year old” and then “Row”, to add age as a row variable.

Available Variables

- Age at Diagnosis
  - Age recode with <1 year olds
- Race, Sex, Year Dx, Registry, County
- Site and Morphology
- Stage - AJCC
- Stage - TNM
- Stage - LRD (Summary and Historic)
- Therapy
- Extent of Disease - CS
- Extent of Disease - Historic
- Cause of Death (COD) and Follow-up

Remove

Page

Row

Column

Find...

1

2

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

# Σ % P ÷ IT Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

Frequency Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

- Page
- Row
  - Age recode with <1 year olds
- Column
  - Race recode (White, Black, Other)

Available Variables

- Age at Diagnosis
- Age recode with <1 year olds
- Race, Sex, Year Dx, Registry, County
- Site and Morphology
- Stage - AJCC
- Stage - TNM
- Stage - LRD (Summary and Historic)
- Therapy
- Extent of Disease - CS
- Extent of Disease - Historic
- Cause of Death (COD) and Follow-up

Remove

Page

Row

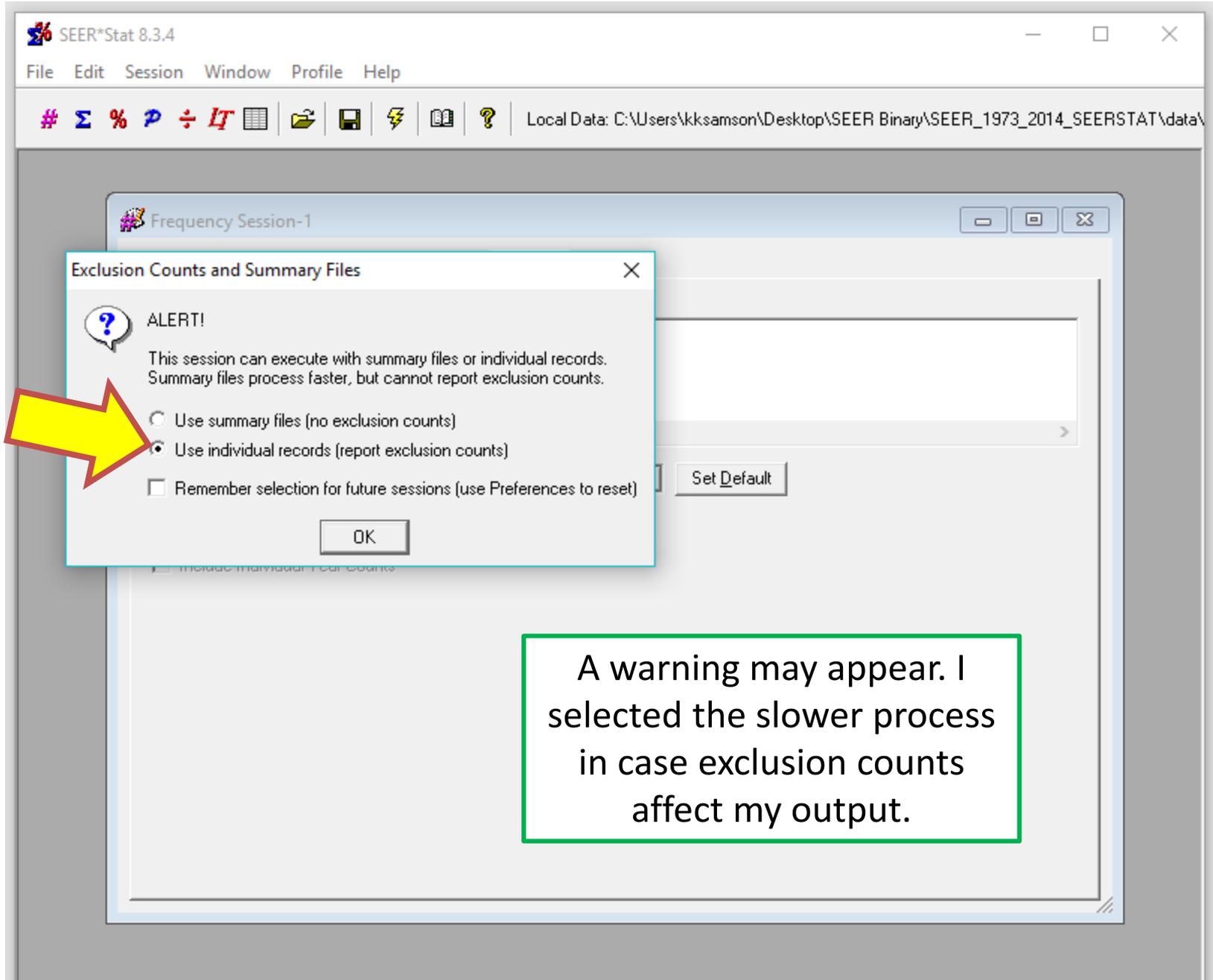
Column

Find...



- Now the age variable has been added as a row in the “Display Variables” box.
- Next, click the “Output” tab.

The image shows a screenshot of the SEER\*Stat 8.3.4 software interface. The main window has a menu bar with 'File', 'Edit', 'Session', 'Window', 'Profile', and 'Help'. Below the menu bar is a toolbar with various icons, including a lightning bolt icon. A yellow arrow points to the lightning bolt icon. The 'Local Data' path is shown as 'C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\'. A dialog box titled 'Frequency Session-1' is open, showing the 'Output' tab. The 'Title (Up to 5 lines)' field contains the text 'Breast Cancer Cross tabulation of Race by Age of Diagnosis (SEER 18)'. Below the title field are three checkboxes: 'Hide Statistics When Fewer Than 25 Cases' and 'Include Individual Year Counts', both of which are unchecked. A green box highlights the text 'Add a meaningful title for your table output, then click the lightning bolt button to generate the table.'.



SEER\*Stat 8.3.4

File Edit **Matrix** Window Profile Help

Retrieve Session  
Options...  
Font...  
Order >  
Filter...  
Lock  
Page >  
Export >  
Properties...

	All races	White	Black	in Indian
01-04 years	0	0	0	0
05-09 years	0	0	0	0
10-14 years	0	0	0	0
15-19 years	104	73	22	
20-24 years	1,019	709	204	
25-29 years	6,188	4,416	1,150	
30-34 years	19,465	14,206	3,169	
35-39 years	43,086	33,003	6,093	
40-44 years	89,000	68,000	13,000	
45-49 years	120,000	90,000	18,000	
50-54 years	140,000	105,000	21,000	
55-59 years	150,000	110,000	22,000	
60-64 years	150,000	110,000	22,000	
65-69 years	150,000	110,000	22,000	
70-74 years	130,000	100,000	20,000	
75-79 years	110,000	85,000	17,000	

- Notice there are a few rows with all zeros.
- There are some cases for ages 10-14, but since we restricted cases to 15 years or older, they are zero in this table.
- To remove rows with all zeros, click “Matrix”, then “Options...”.

SEER\*Stat 8.3.4

File Edit Matrix Window Profile Help

# Σ % P ÷ IT [Grid] [Folder] [Save] [Print] [Help]

Matrix Options

Titles (one per line, up to 5)  
Breast Cancer Cross tabulation of  
Race by Age of Diagnosis (SEER 18)

Delay Statistics

- Trends
- Confidence Intervals
- Counts
- Percentages
- Cumulative Percentages
- Delay Factor

Statistics

- Trends
- Confidence Intervals
- Counts
- Percentages
- Cumulative Percentages

Options

- Display Titles
- Display Flags
- Display Footnotes
- Hide Zero Count Rows
- Hide PC Rows
- Hide APC Rows
- Hide Count Rows

OK Cancel Help

55-59 years	150,526	122,669	15,549
60-64 years	154,638	128,857	14,458
65-69 years	150,348	127,749	12,727
70-74 years			0,432
75-79 years			8,303

Check "Hide Zero Count Rows" and then click "Ok".

SEER\*Stat 8.3.4

File Edit Matrix Window Profile Help

# Σ % P ÷ IT [Grid] [Folder] [Save] [Print] [Back] [Forward] [Print] [Help]

Frequency Session-1

Data | Statistic | Selection | Table | **Output**

Title (Up to 5 lines)

Breast  
Race

Number

Hide

Include

**Frequency Session-1 Matrix-10**

Breast Cancer Cross tabulation of

	All races	White	Black	an Indian/AK
<b>15-19 years</b>	104	73	22	
<b>20-24 years</b>	1,019	709	204	
<b>25-29 years</b>	6,188	4,416	1,150	
<b>30-34 years</b>	19,465	14,206	3,169	
<b>35-39 years</b>	43,986	33,003	6,093	
<b>40-44 years</b>	85,701	65,890	10,629	
<b>45-49 years</b>	126,849	99,809	14,370	
<b>50-54 years</b>	143,013	114,631	15,518	
<b>55-59 years</b>	150,526	122,669	15,549	
<b>60-64 years</b>	154,638	128,857	14,458	
<b>65-69 years</b>	150,348	127,749	12,727	
<b>70-74 years</b>	132,763	114,537	10,432	
<b>75-79 years</b>	114,941	101,150	8,303	
<b>80-84 years</b>	83,638	74,641	5,452	
<b>85+ years</b>	71,421	64,050	4,781	

# Mortality Counts and Rates

Example: Heart Disease



University of Nebraska  
Medical Center

# Mortality Counts and Rates

- This section will instruct you on how to get counts and age-adjusted rates for mortality data.
- Mortality data is only available through SEER\*Stat.
- For this example, we will look at heart diseases mortality counts and age-adjusted rates in three separate states, to be able to compare them:
  - Nebraska
  - North Dakota
  - South Dakota
- These mortality examples were completed using on-line SEER data, so some screens and output may look different if you're using downloaded binary SEER data.





**SEER\*Stat**

Produced by:

- The **Surveillance Research Program** of the Division of Cancer Control and Population Sciences, National Cancer Institute
- **Information Management Services, Inc.**  
3901 Calverton Blvd., Suite 200  
Calverton, MD 20705

Prevalence Session Developed in Collaboration with:

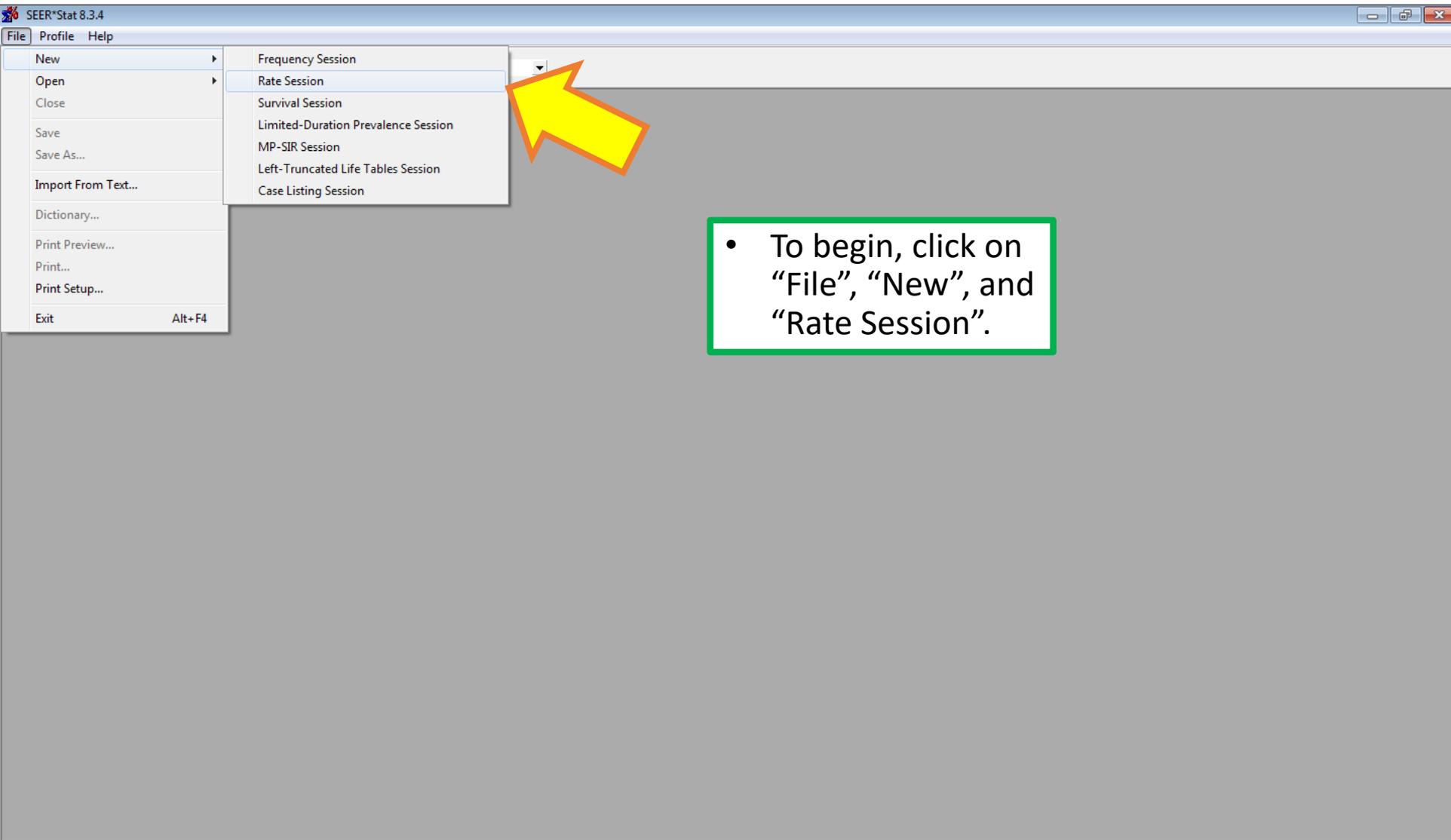
- Unit of Epidemiology, **Istituto Nazionale per lo Studio e la Cura dei Tumori**, Milano, Italia
- Laboratory of Epidemiology and Biostatistics, **Istituto Superiore di Sanita**, Roma, Italia

Multiple Primary - Standardized Incidence Ratio (MP-SIR) Session Developed in Partnership with:

- The **Radiation Epidemiology Branch** of the Division of Cancer Epidemiology and Genetics, National Cancer Institute

**email:** [seerstat@imsweb.com](mailto:seerstat@imsweb.com)  
**web:** [seer.cancer.gov/seerstat](http://seer.cancer.gov/seerstat)

- Open SEER\*Stat
- SEER\*Stat will display a brief informational message.



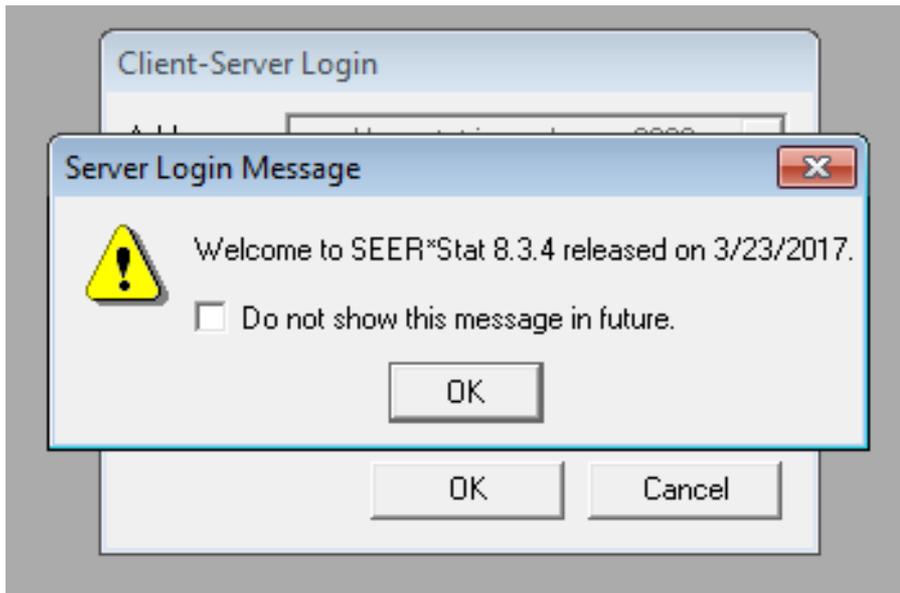
- To begin, click on “File”, “New”, and “Rate Session”.

The screenshot shows the SEER\*Stat 8.3.4 application window. The main window has a menu bar (File, Profile, Help) and a toolbar with various icons. A dropdown menu shows 'Profile: Current User Profile'. Two dialog boxes are overlaid on the main window:

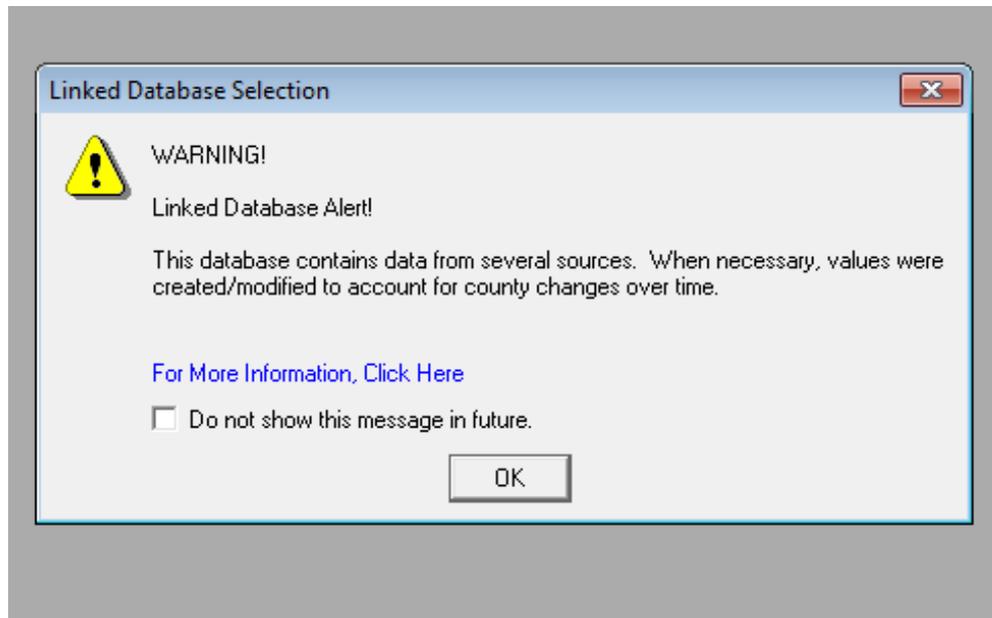
- Verifying Selected Data Locations:** This dialog shows progress bars for 'Current: Establishing Server Connection' and 'Overall: Location #1 of 1'. A 'Cancel' button is at the bottom.
- Client-Server Login:** This dialog has fields for 'Address' (ssp://seerstat.imsweb.com:2038), 'User Name', and 'Password'. It includes a link for forgotten credentials and a 'Remember my password for the future' checkbox. 'OK' and 'Cancel' buttons are at the bottom.

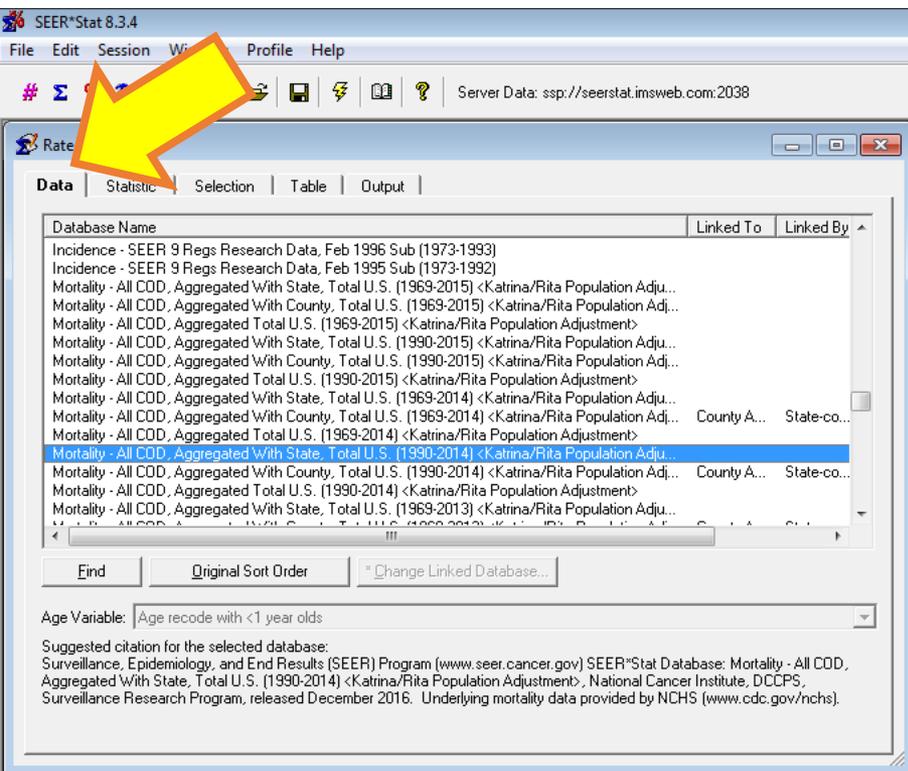
Yellow arrows with numbers 1 and 2 point to the 'User Name' field and the 'OK' button, respectively.

- A login window may appear to connect you to the server.
- Log in using the username and password given to you through email.



Click OK for both the welcome and warning messages to continue.





The rate session begins by displaying a window with various tabs.

- Go to the “Data” tab.
- Scroll down and select the database titled “**Mortality - All COD, Aggregated With County, Total U.S. (1990-2014) <Katrina/Rita Population Adjustment>**”.

Choosing which database will depend on what data you are looking for. Some of the database options are explained at the website:

<https://seer.cancer.gov/data/seerstat/>

### Database Change



#### WARNING!

##### Data Use Restrictions:

The Public Health Service Act (42 U.S.C. 242m(d)) provides that the data collected by the National Center for Health Statistics (NCHS) may be used only for the purpose for which they were obtained; any effort to determine the identity of any reported cases, or to use the information for any purpose other than for health statistical reporting and analysis, is against the law.

Therefore users will:

- Use these data for health statistical reporting and analysis only.
- For sub-national geography, do not present or publish death counts of nine or fewer or death rates based on counts of nine or fewer (in figures, graphs, maps, tables, etc.).
- Make no attempt to learn the identity of any person or establishment included in these data.
- Make no disclosure or other use of the identity of any person or establishment discovered inadvertently and advise SEER of any such discovery.

By clicking the OK button, I agree to these terms of use.

Do not show this message in future.

OK

Cancel

- Clicking on a database will display the warning messages about the data.
- Review and then click “OK” to both of these messages.

### Linked Database Change



#### WARNING!

##### Linked Database Alert!

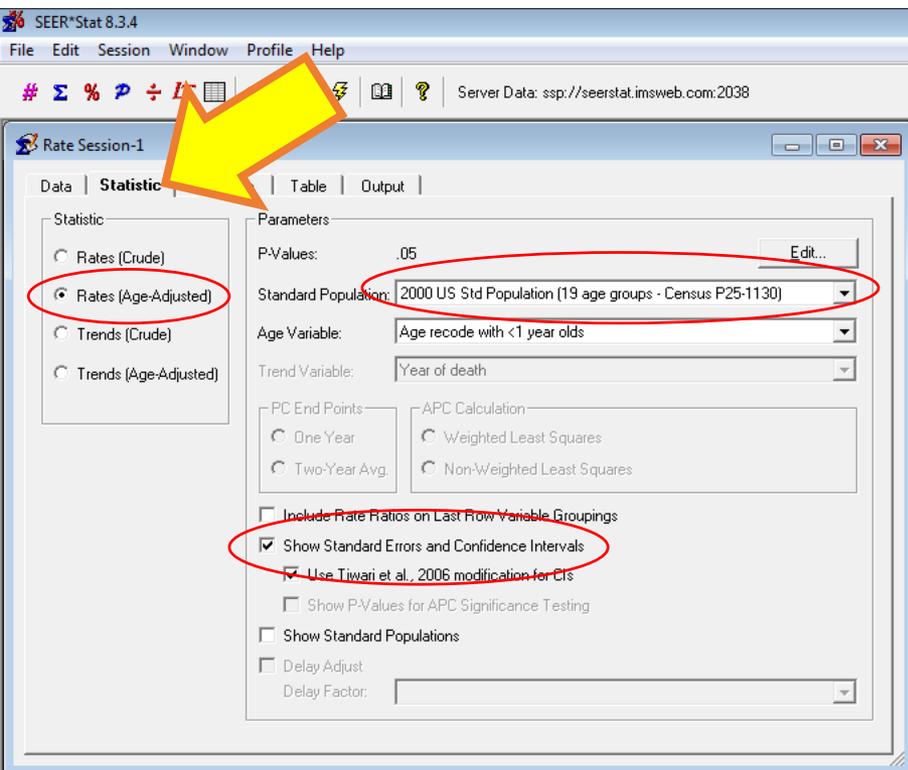
This database contains data from several sources. When necessary, values were created/modified to account for county changes over time.

[For More Information, Click Here](#)

Do not show this message in future.

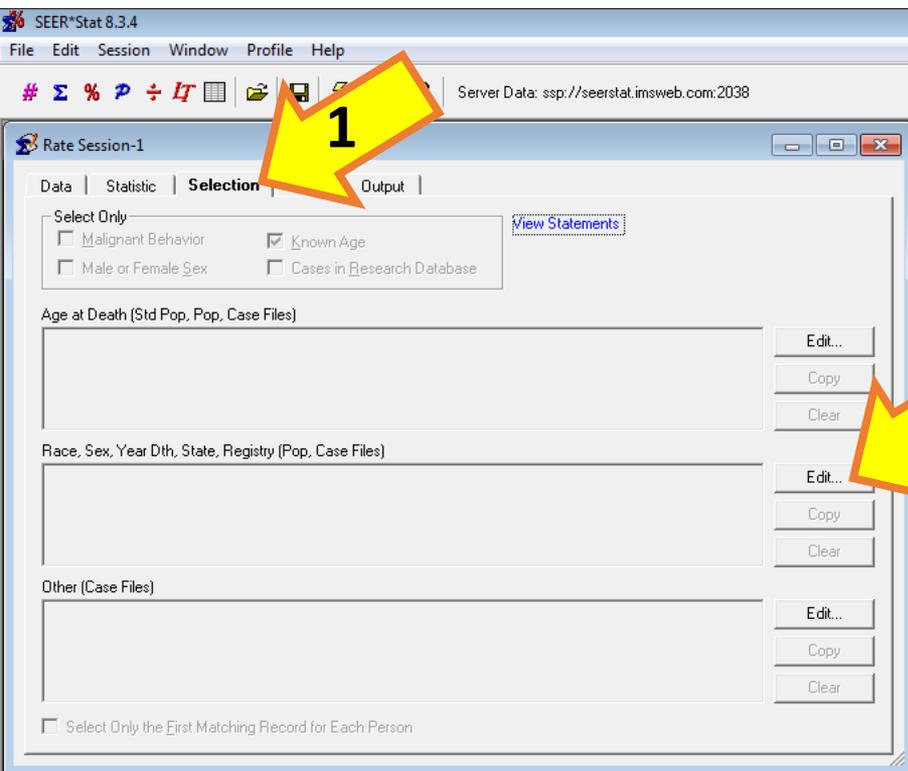
OK

Cancel



- Go to the “Statistic” tab.
- Choose “Rates: Age-Adjusted” under the “Statistic” box.
- Choose “2000 US Std Population” under the “Parameters” box.
- Choose “Show Standard Errors and Confidence Intervals” under the “Parameters” box.

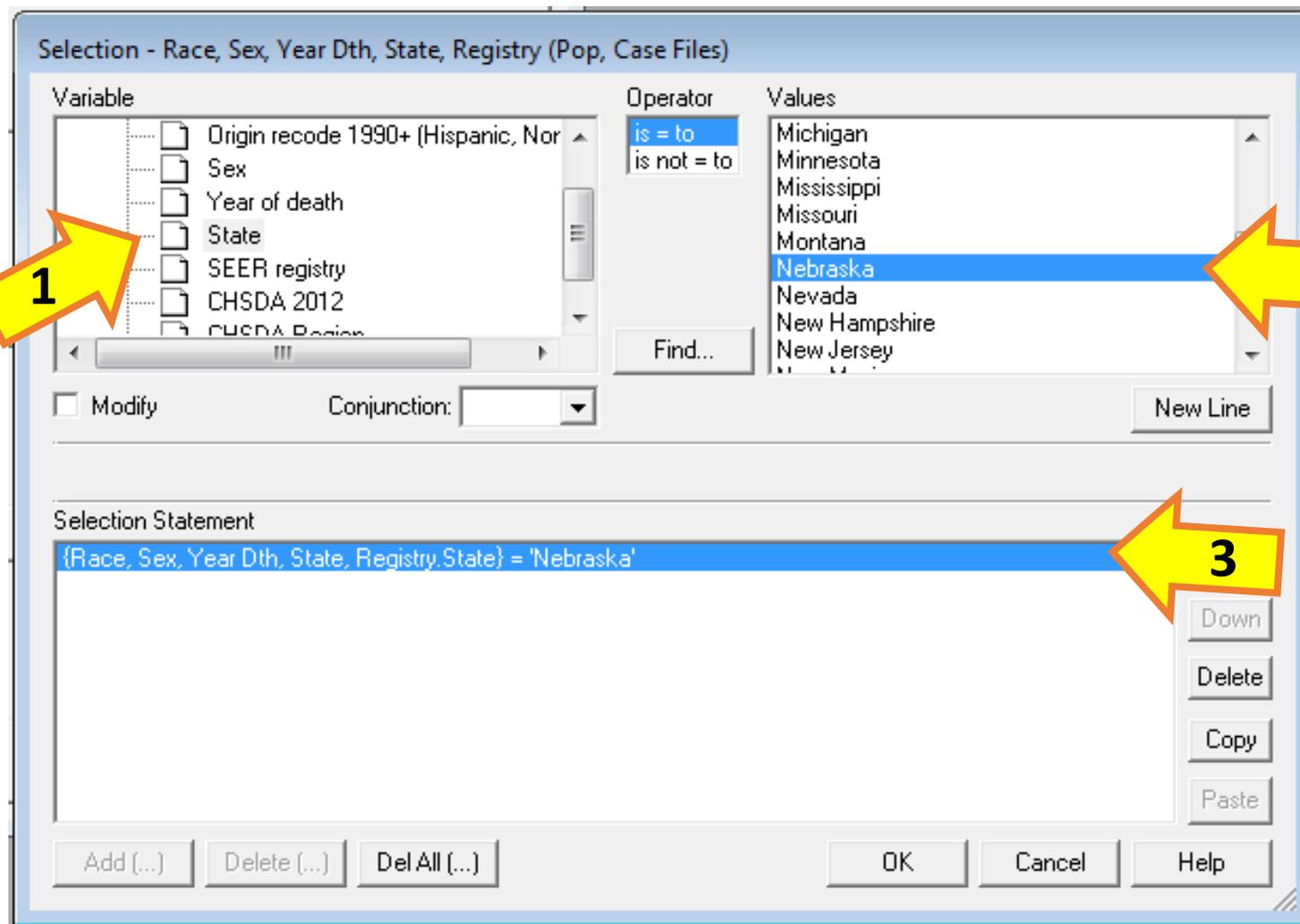
Using an age-adjusted rate adapts the crude rates to standard population age groups. This allows us to accurately compare rates between populations that may not actually have the same age distribution.



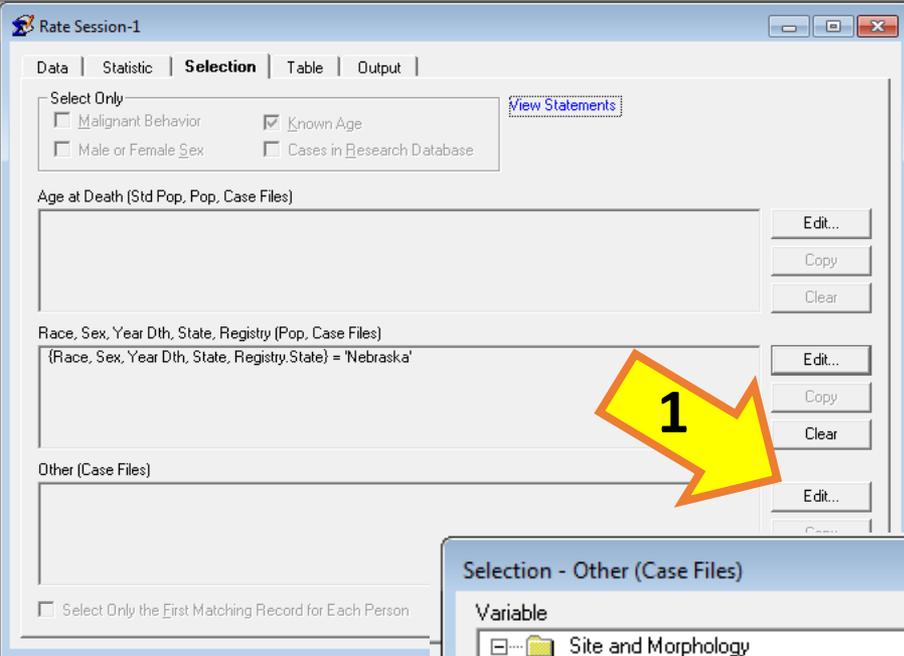
- Go to the “Selection” tab.
- Click on “Edit” for the “Race, Sex, Year Dth, State, Registry.State” category.

A new window will be displayed.

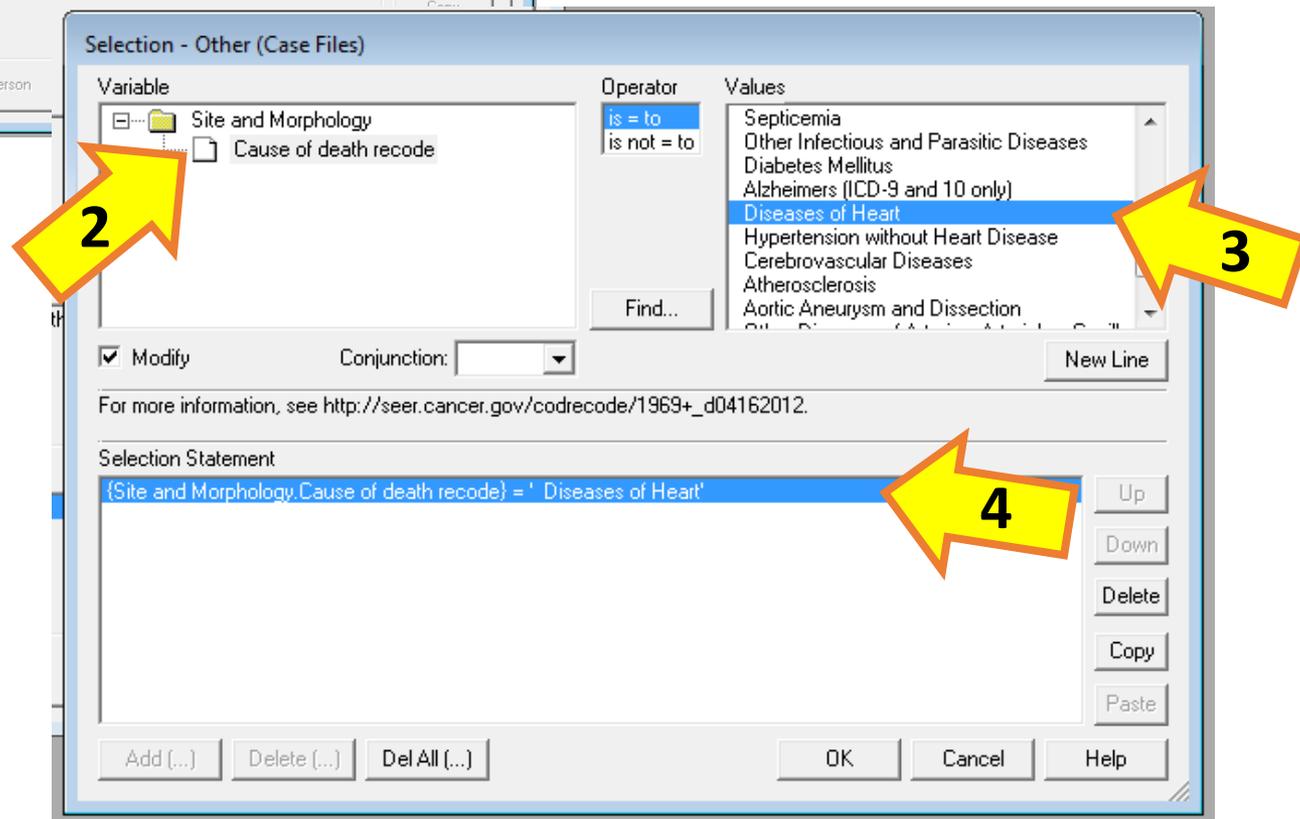
Note: the selection tab determines what data is used for the analysis. This is a good place to specify inclusion/exclusion criteria.



- Click on the “Race, Sex, Year Dth, State, Registry.State” folder in the “Variable” window, and select “State”.
- Select “Nebraska” in the “Values” window. Once selected, the values will display in the “Selection Statement” box.



- Return to the Selection tab.
- Click “Edit” for the “Other” category.
- Open the “Site and Morphology” folder in the “Variable” window.
- Choose “Cause of death recode”, and select “Diseases of Heart” in the “Values” window.
- Click “OK”.



Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

- Page
- Row
- Cause of death recode
- Column

Available Variables

- Age at Death
- Race, Sex, Year Dth, State, Registry
- Site and Morphology
  - Cause of death recode

Move Up

Move Down

Remove

Page

Row

Column

Find...

- Go to the “Table” tab.
- Open the “Site and Morphology” folder.
- Select “Cause of death recode”.
- Click “Row”.

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

Server Data: ssp://seerstat.imswb.com:2038

Rate Session-1

Data | Statistic | Selection | Table | **Output**

Title (Up to 5 lines)

Nebraska Heart Disease Mortality

Display Rates as Cases Per: 100,000 Set Default

Number of Decimal Places for Rates/Trends: 0.1 Set Default

Hide Statistics When Fewer Than 10 Cases

Hide Statistics Based on a Population Count of Less Than 50000

Display All Calculated Statistics In Output Matrix Set Default

- Go to the “Output” tab.
- Type “Nebraska Heart Disease Mortality” in the “Title” box.

CAPS NUM



Rate Session-1 Matrix-1

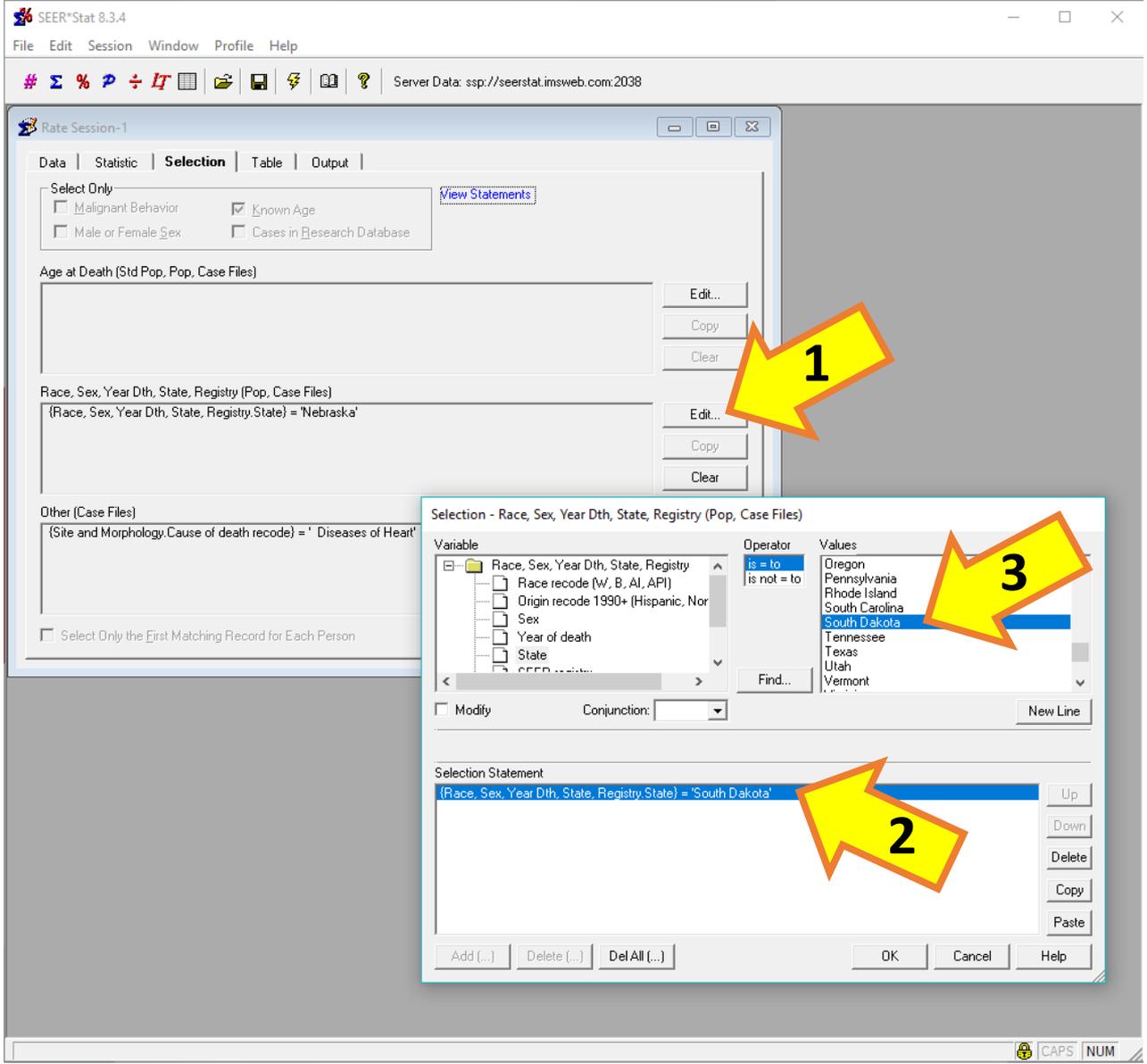
Nebraska Heart Disease Mortality

	Rate	SE	Lower CI	Upper CI	Count	Pop
Other Leukemia	^	^	^	^	^	43,328,940
Other Acute Leukemia	^	^	^	^	^	43,328,940
Aleukemic, Subleukemia	^	^	^	^	^	43,328,940
Miscellaneous Malignant	^	^	^	^	^	43,328,940
In situ, benign or unknown	^	^	^	^	^	43,328,940
Tuberculosis	^	^	^	^	^	43,328,940
Syphilis	^	^	^	^	^	43,328,940
Human Immunodeficiency Virus	^	^	^	^	^	43,328,940
Septicemia	^	^	^	^	^	43,328,940
Other Infectious and Parasitic	^	^	^	^	^	43,328,940
Diabetes Mellitus	^	^	^	^	^	43,328,940
Alzheimers (ICD-9 and 10)	^	^	^	^	^	43,328,940
Diseases of Heart	208.3	0.7	207.0	209.5	104,035	43,328,940
Hypertension without Heart Disease	^	^	^	^	^	43,328,940
Cerebrovascular Disease	^	^	^	^	^	43,328,940
Atherosclerosis	^	^	^	^	^	43,328,940
Aortic Aneurysm and Dissection	^	^	^	^	^	43,328,940
Other Diseases of Arteries	^	^	^	^	^	43,328,940
Pneumonia and Influenza	^	^	^	^	^	43,328,940
Chronic Obstructive Pulmonary Disease	^	^	^	^	^	43,328,940
Stomach and Duodenal Ulcers	^	^	^	^	^	43,328,940
Chronic Liver Disease and Cirrhosis	^	^	^	^	^	43,328,940
Nephritis, Nephrotic Syndrome	^	^	^	^	^	43,328,940

Underlying mortality data provided by NCHS ([www.cdc.gov/nchs](http://www.cdc.gov/nchs)).  
 Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence in  
 ^  
 Statistic not displayed due to fewer than 10 cases.

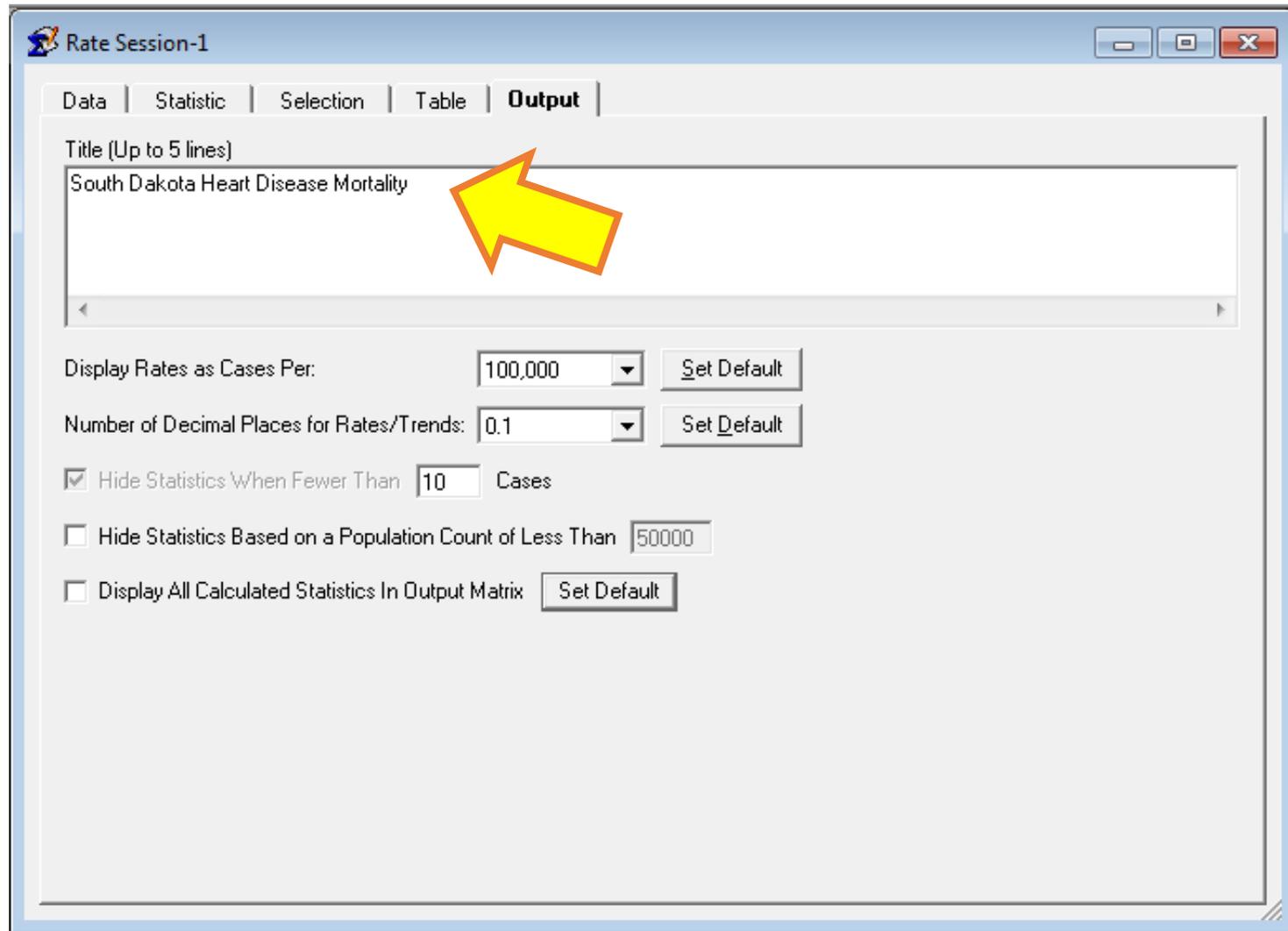
- Click on the lightning bolt icon in the toolbar.

Your table will then be displayed. The table will show all site recode categories for mortality, but will show ^ for all values except those for diseases of heart, since that is the variable we chose in the "Selection" tab.



We have produced the table for Nebraska, but we still need tables for South Dakota and North Dakota for comparison.

- Go back to the same rate session window.
- Go to the “Selection” Tab.
- Click “Edit” for the “Race, Sex, Year Dth, State, Registry” category.
- Click on the text in the “Selection Statement” box.
- Choose “South Dakota” in the “Values” window.
- Click “OK”.



- Go to the “Output” tab and make a new title to avoid confusing your tables!

Rate Session-1 Matrix-3

North Dakota Heart Disease Mortality

	Rate	SE	Lower CI	Upper CI	Count	Pop
In situ, benign or unknow	▲	▲	▲	▲	▲	16,436,834
Tuberculosis	▲	▲	▲	▲	▲	16,436,834
Syphilis	▲	▲	▲	▲	▲	16,436,834
Human Immunodeficiency	▲	▲	▲	▲	▲	16,436,834
Septicemia	▲	▲	▲	▲	▲	16,436,834
Other Infectious and Para	▲	▲	▲	▲	▲	16,436,834
Diabetes Mellitus	▲	▲	▲	▲	▲	16,436,834
Alzheimers (ICD-9 and 10	▲	▲	▲	▲	▲	16,436,834
Diseases of Heart	202.1	1.0	200.1	204.1	40,961	16,436,834
Hypertension without He	▲	▲	▲	▲	▲	16,436,834
Cerebrovascular Disease	▲	▲	▲	▲	▲	16,436,834
Atherosclerosis	▲	▲	▲	▲	▲	16,436,834
Aortic Aneurysm and Dis	▲	▲	▲	▲	▲	16,436,834
Other Diseases of Arterie	▲	▲	▲	▲	▲	16,436,834

Underlying mortality data provided by NCHS (www.cdc.gov/nchs).  
Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence interval not displayed due to fewer than 10 cases.

Rate Session-1 Matrix-2

South Dakota Heart Disease Mortality

	Rate	SE	Lower CI	Upper CI	Count	Pop
In situ, benign or unknow	▲	▲	▲	▲	▲	19,225,44
Tuberculosis	▲	▲	▲	▲	▲	19,225,44
Syphilis	▲	▲	▲	▲	▲	19,225,44
Human Immunodeficiency	▲	▲	▲	▲	▲	19,225,44
Septicemia	▲	▲	▲	▲	▲	19,225,44
Other Infectious and Para	▲	▲	▲	▲	▲	19,225,44
Diabetes Mellitus	▲	▲	▲	▲	▲	19,225,44
Alzheimers (ICD-9 and 10	▲	▲	▲	▲	▲	19,225,44
Diseases of Heart	209.8	1.0	207.9	211.7	48,851	19,225,44
Hypertension without He	▲	▲	▲	▲	▲	19,225,44
Cerebrovascular Disease	▲	▲	▲	▲	▲	19,225,44
Atherosclerosis	▲	▲	▲	▲	▲	19,225,44
Aortic Aneurysm and Dis	▲	▲	▲	▲	▲	19,225,44
Other Diseases of Arterie	▲	▲	▲	▲	▲	19,225,44
Pneumonia and Influenza	▲	▲	▲	▲	▲	19,225,44

Underlying mortality data provided by NCHS (www.cdc.gov/nchs).  
Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence interval not displayed due to fewer than 10 cases.

Rate Session-1 Matrix-1

Nebraska Heart Disease Mortality

	Rate	SE	Lower CI	Upper CI	Count	Pop
In situ, benign or unknow	▲	▲	▲	▲	▲	43,328,940
Tuberculosis	▲	▲	▲	▲	▲	43,328,940
Syphilis	▲	▲	▲	▲	▲	43,328,940
Human Immunodeficiency	▲	▲	▲	▲	▲	43,328,940
Septicemia	▲	▲	▲	▲	▲	43,328,940
Other Infectious and Para	▲	▲	▲	▲	▲	43,328,940
Diabetes Mellitus	▲	▲	▲	▲	▲	43,328,940
Alzheimers (ICD-9 and 10	▲	▲	▲	▲	▲	43,328,940
Diseases of Heart	208.3	0.7	207.0	209.5	104,035	43,328,940
Hypertension without He	▲	▲	▲	▲	▲	43,328,940
Cerebrovascular Disease	▲	▲	▲	▲	▲	43,328,940
Atherosclerosis	▲	▲	▲	▲	▲	43,328,940
Aortic Aneurysm and Dis	▲	▲	▲	▲	▲	43,328,940
Other Diseases of Arterie	▲	▲	▲	▲	▲	43,328,940
Pneumonia and Influenza	▲	▲	▲	▲	▲	43,328,940

Underlying mortality data provided by NCHS (www.cdc.gov/nchs).  
Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence interval not displayed due to fewer than 10 cases.

- Click the lightning bolt icon.
  - Repeat the steps in the last two slides to create your North Dakota table.
- Now we have mortality rates for Nebraska, South Dakota, and North Dakota, and can compare the rates between these states.

Rate Session-1 Matrix-1

Nebraska Heart Disease Mortality

	Rate	SE	Lower CI	Upper CI	Count	Pop
In situ, benign or unknow	▲	▲	▲	▲	▲	43,328,940
Tuberculosis	▲	▲	▲	▲	▲	43,328,940
Syphilis	▲	▲	▲	▲	▲	43,328,940
Human Immunodeficienc	▲	▲	▲	▲	▲	43,328,940
Septicemia	▲	▲	▲	▲	▲	43,328,940
Other Infectious and Para	▲	▲	▲	▲	▲	43,328,940
Diabetes Mellitus	▲	▲	▲	▲	▲	43,328,940
Alzheimers (ICD-9 and 10	▲	▲	▲	▲	▲	43,328,940
Diseases of Heart	208.3	0.7	207.0	209.5	104,035	43,328,940
Hypertension without He	▲	▲	▲	▲	▲	43,328,940
Cerebrovascular Disease	▲	▲	▲	▲	▲	43,328,940
Atherosclerosis	▲	▲	▲	▲	▲	43,328,940
Aortic Aneurysm and Dis	▲	▲	▲	▲	▲	43,328,940
Other Diseases of Arterie	▲	▲	▲	▲	▲	43,328,940
Pneumonia and Influenza	▲	▲	▲	▲	▲	43,328,940

Underlying mortality data provided by NCHS (www.cdc.gov/nchs).  
Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence in  
Statistic not displayed due to fewer than 10 cases.

First, let's compare Nebraska and North Dakota. Nebraska had a count of 104,035 deaths from heart disease between 1990-2014; North Dakota had 40,961 deaths. This seems like a large difference until we notice the discrepancy in population size.

Rate Session-1 Matrix-3

North Dakota Heart Disease Mortality

	Rate	SE	Lower CI	Upper CI	Count	Pop
In situ, benign or unknow	▲	▲	▲	▲	▲	16,436,834
Tuberculosis	▲	▲	▲	▲	▲	16,436,834
Syphilis	▲	▲	▲	▲	▲	16,436,834
Human Immunodeficienc	▲	▲	▲	▲	▲	16,436,834
Septicemia	▲	▲	▲	▲	▲	16,436,834
Other Infectious and Para	▲	▲	▲	▲	▲	16,436,834
Diabetes Mellitus	▲	▲	▲	▲	▲	16,436,834
Alzheimers (ICD-9 and 10	▲	▲	▲	▲	▲	16,436,834
Diseases of Heart	202.1	1.0	200.1	204.1	40,961	16,436,834
Hypertension without He	▲	▲	▲	▲	▲	16,436,834
Cerebrovascular Disease	▲	▲	▲	▲	▲	16,436,834
Atherosclerosis	▲	▲	▲	▲	▲	16,436,834
Aortic Aneurysm and Dis	▲	▲	▲	▲	▲	16,436,834
Other Diseases of Arterie	▲	▲	▲	▲	▲	16,436,834

Underlying mortality data provided by NCHS (www.cdc.gov/nchs).  
Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence interva  
Statistic not displayed due to fewer than 10 cases.

Nebraska had a rate of 208.3 deaths per 100,000. North Dakota had a rate of 202.1 deaths per 100,000. Nebraska had a slightly higher mortality rate due to heart disease between 1990-2014.

Rate Session-1 Matrix-1

Nebraska Heart Disease Mortality

	Rate	SE	Lower CI	Upper CI	Count	Pop
In situ, benign or unknow	^	^	^	^	^	43,328,940
Tuberculosis	^	^	^	^	^	43,328,940
Syphilis	^	^	^	^	^	43,328,940
Human Immunodeficienc	^	^	^	^	^	43,328,940
Septicemia	^	^	^	^	^	43,328,940
Other Infectious and Para	^	^	^	^	^	43,328,940
Diabetes Mellitus	^	^	^	^	^	43,328,940
Alzheimers (ICD-9 and 10	^	^	^	^	^	43,328,940
Diseases of Heart	208.3	0.7	207.0	209.5	104,035	43,328,940
Hypertension without He	^	^	^	^	^	43,328,940
Cerebrovascular Disease	^	^	^	^	^	43,328,940
Atherosclerosis	^	^	^	^	^	43,328,940
Aortic Aneurysm and Dis	^	^	^	^	^	43,328,940
Other Diseases of Arteri	^	^	^	^	^	43,328,940
Pneumonia and Influenza	^	^	^	^	^	43,328,940

Underlying mortality data provided by NCHS (www.cdc.gov/nchs).  
 Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence in  
 Statistic not displayed due to fewer than 10 cases.

Nebraska has higher mortality counts than South Dakota. South Dakota has 209.8 deaths per 100,000 and Nebraska has 208.3 deaths per 100,000. South Dakota has a slightly higher mortality rate due to heart disease than Nebraska.

Rate Session-1 Matrix-2

South Dakota Heart Disease Mortality

	Rate	SE	Lower CI	Upper CI	Count	Pop
In situ, benign or unknow	^	^	^	^	^	19,225,44
Tuberculosis	^	^	^	^	^	19,225,44
Syphilis	^	^	^	^	^	19,225,44
Human Immunodeficienc	^	^	^	^	^	19,225,44
Septicemia	^	^	^	^	^	19,225,44
Other Infectious and Para	^	^	^	^	^	19,225,44
Diabetes Mellitus	^	^	^	^	^	19,225,44
Alzheimers (ICD-9 and 10	^	^	^	^	^	19,225,44
Diseases of Heart	209.8	1.0	207.9	211.7	48,851	19,225,44
Hypertension without He	^	^	^	^	^	19,225,44
Cerebrovascular Disease	^	^	^	^	^	19,225,44
Atherosclerosis	^	^	^	^	^	19,225,44
Aortic Aneurysm and Dis	^	^	^	^	^	19,225,44
Other Diseases of Arteri	^	^	^	^	^	19,225,44
Pneumonia and Influenza	^	^	^	^	^	19,225,44

Underlying mortality data provided by NCHS (www.cdc.gov/nchs).  
 Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence  
 Statistic not displayed due to fewer than 10 cases.

Rate Session-1 Matrix-3

North Dakota Heart Disease Mortality

	Rate	SE	Lower CI	Upper CI	Count	Pop
In situ, benign or unknow	▲	▲	▲	▲	▲	16,436,834
Tuberculosis	▲	▲	▲	▲	▲	16,436,834
Syphilis	▲	▲	▲	▲	▲	16,436,834
Human Immunodeficienc	▲	▲	▲	▲	▲	16,436,834
Septicemia	▲	▲	▲	▲	▲	16,436,834
Other Infectious and Para	▲	▲	▲	▲	▲	16,436,834
Diabetes Mellitus	▲	▲	▲	▲	▲	16,436,834
Alzheimers (ICD-9 and 10	▲	▲	▲	▲	▲	16,436,834
Diseases of Heart	202.1	1.0	200.1	204.1	40,961	16,436,834
Hypertension without He	▲	▲	▲	▲	▲	16,436,834
Cerebrovascular Disease	▲	▲	▲	▲	▲	16,436,834
Atherosclerosis	▲	▲	▲	▲	▲	16,436,834
Aortic Aneurysm and Dis	▲	▲	▲	▲	▲	16,436,834
Other Diseases of Arterie	▲	▲	▲	▲	▲	16,436,834

Underlying mortality data provided by NCHS (www.cdc.gov/nchs).  
Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence interval not displayed due to fewer than 10 cases.

North Dakota has a lower count of heart disease deaths and a lower mortality rate than South Dakota for heart disease mortality.

Rate Session-1 Matrix-2

South Dakota Heart Disease Mortality

	Rate	SE	Lower CI	Upper CI	Count	Pop
In situ, benign or unknow	▲	▲	▲	▲	▲	19,225,44
Tuberculosis	▲	▲	▲	▲	▲	19,225,44
Syphilis	▲	▲	▲	▲	▲	19,225,44
Human Immunodeficienc	▲	▲	▲	▲	▲	19,225,44
Septicemia	▲	▲	▲	▲	▲	19,225,44
Other Infectious and Para	▲	▲	▲	▲	▲	19,225,44
Diabetes Mellitus	▲	▲	▲	▲	▲	19,225,44
Alzheimers (ICD-9 and 10	▲	▲	▲	▲	▲	19,225,44
Diseases of Heart	209.8	1.0	207.9	211.7	48,851	19,225,44
Hypertension without He	▲	▲	▲	▲	▲	19,225,44
Cerebrovascular Disease	▲	▲	▲	▲	▲	19,225,44
Atherosclerosis	▲	▲	▲	▲	▲	19,225,44
Aortic Aneurysm and Dis	▲	▲	▲	▲	▲	19,225,44
Other Diseases of Arterie	▲	▲	▲	▲	▲	19,225,44
Pneumonia and Influenza	▲	▲	▲	▲	▲	19,225,44

Underlying mortality data provided by NCHS (www.cdc.gov/nchs).  
Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence interval not displayed due to fewer than 10 cases.

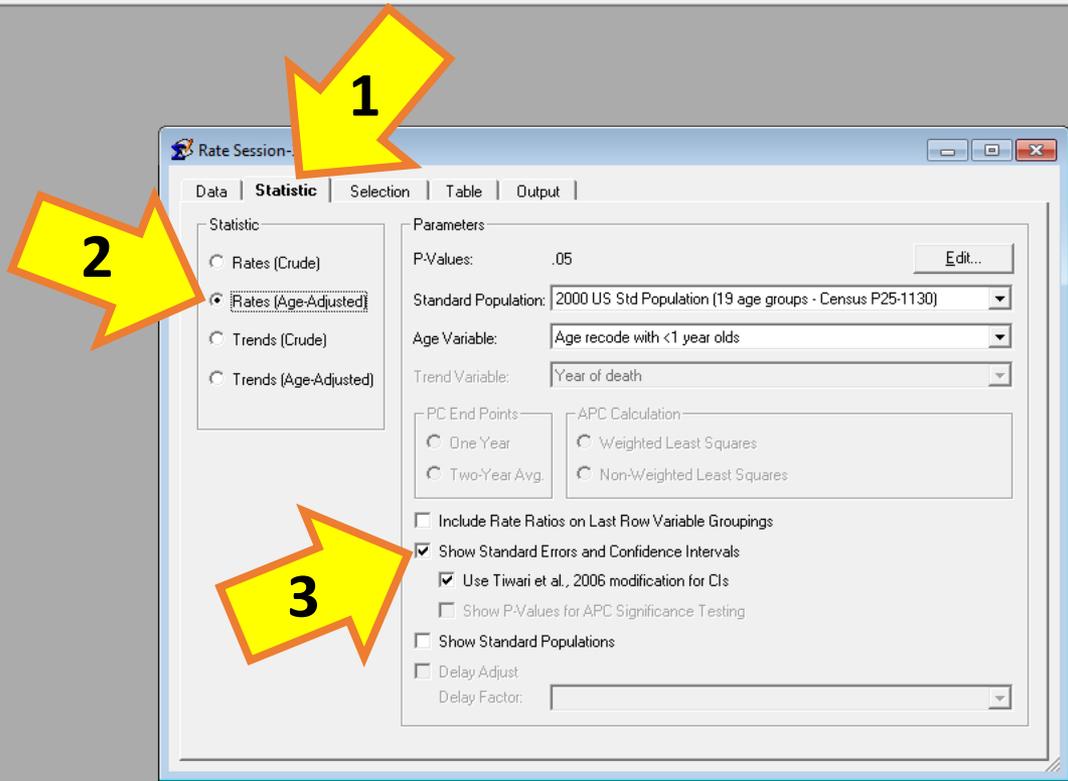
In our comparisons, South Dakota had the highest mortality rate, followed by Nebraska. North Dakota had the lowest mortality rate of these states from 1990-2014.

# Mortality Counts and Rates (cont.)

- To provide another example, we will look at heart diseases mortality counts and age-adjusted rates in different races within the same state, to be able to compare them:
  - American Indians
  - Whites
- Again, this mortality example was completed using on-line SEER data, so some screens and output may look different if you're using downloaded binary SEER data.







- In the “Statistic” tab, choose an age-adjusted rate, standardized to the 2000 US Population, with standard errors and confidence intervals.

Using an age-adjusted rate adapts the crude rates to standard population age groups. This allows us to accurately compare rates between populations that may not actually have the same age distribution.

Rate Session-1

Data | Statistic | **Selection** | Table | Output

Select Only

Malignant Behavior  Known Age [View Statements](#)

Male or Female Sex  Cases in Research Database

Age at Death (Std Pop, Pop, Case Files)

Edit...  
Copy  
Clear

Race, Sex, Year Dth, State, Registry (Pop, Case Files)

Edit... **1**  
Copy

Other (Case Files)

Select Only the First Matching Record for E

**2**

Selection - Race, Sex, Year Dth, State, Registry (Pop, Case Files)

Variable	Operator	Values
<input type="checkbox"/> Sex	is = to	Oklahoma
<input type="checkbox"/> Year of death	is not = to	Oregon
<input type="checkbox"/> State		Pennsylvania
<input type="checkbox"/> SEER registry		Rhode Island
<input type="checkbox"/> CHSDA 2012		South Carolina
<input type="checkbox"/> CHSDA Region		<b>South Dakota</b> <b>3</b>
		Tennessee
		Texas
		Utah

Modify Conjunction:

Find... **4**

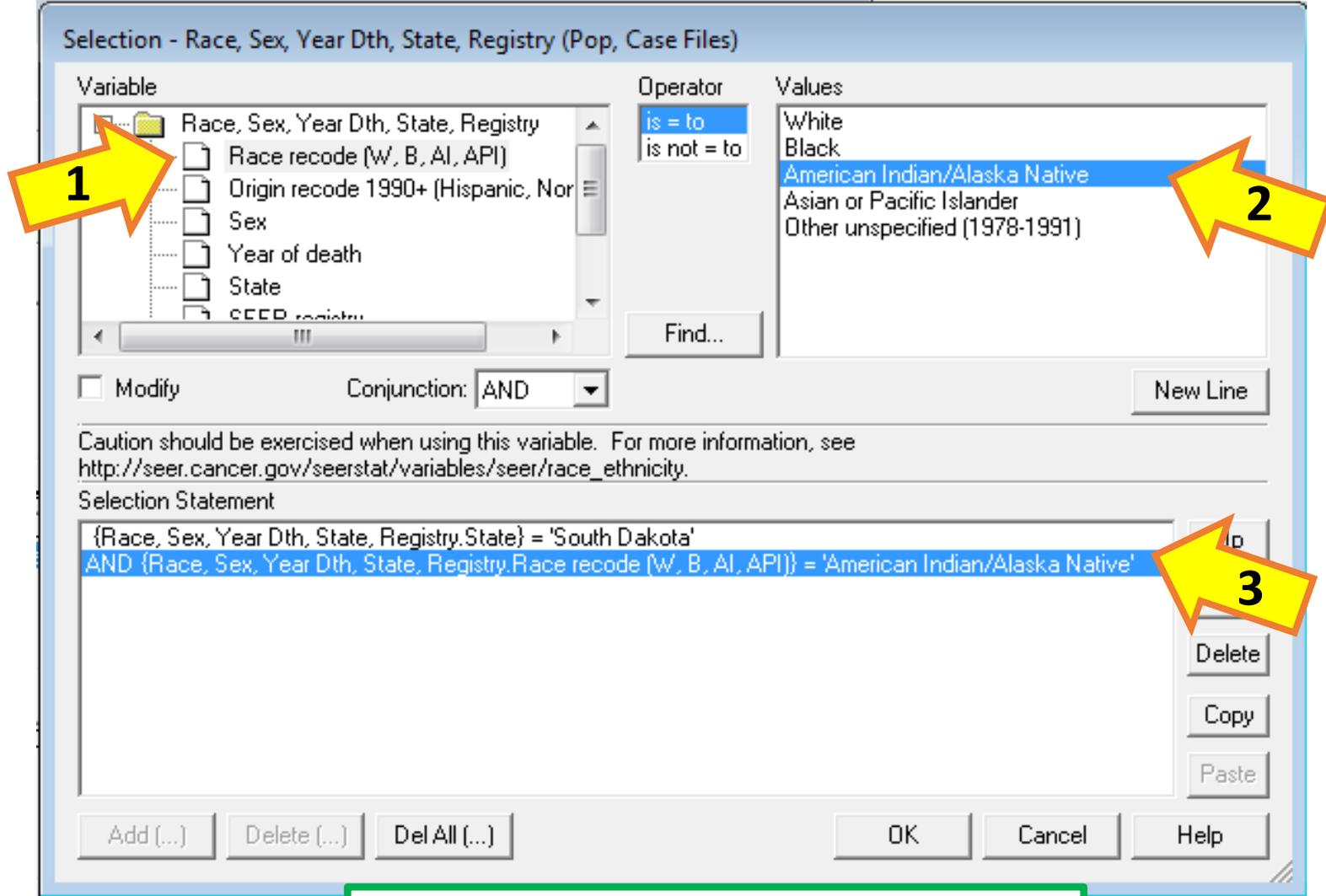
Selection Statement

{Race, Sex, Year Dth, State, Registry.State} = 'South Dakota'

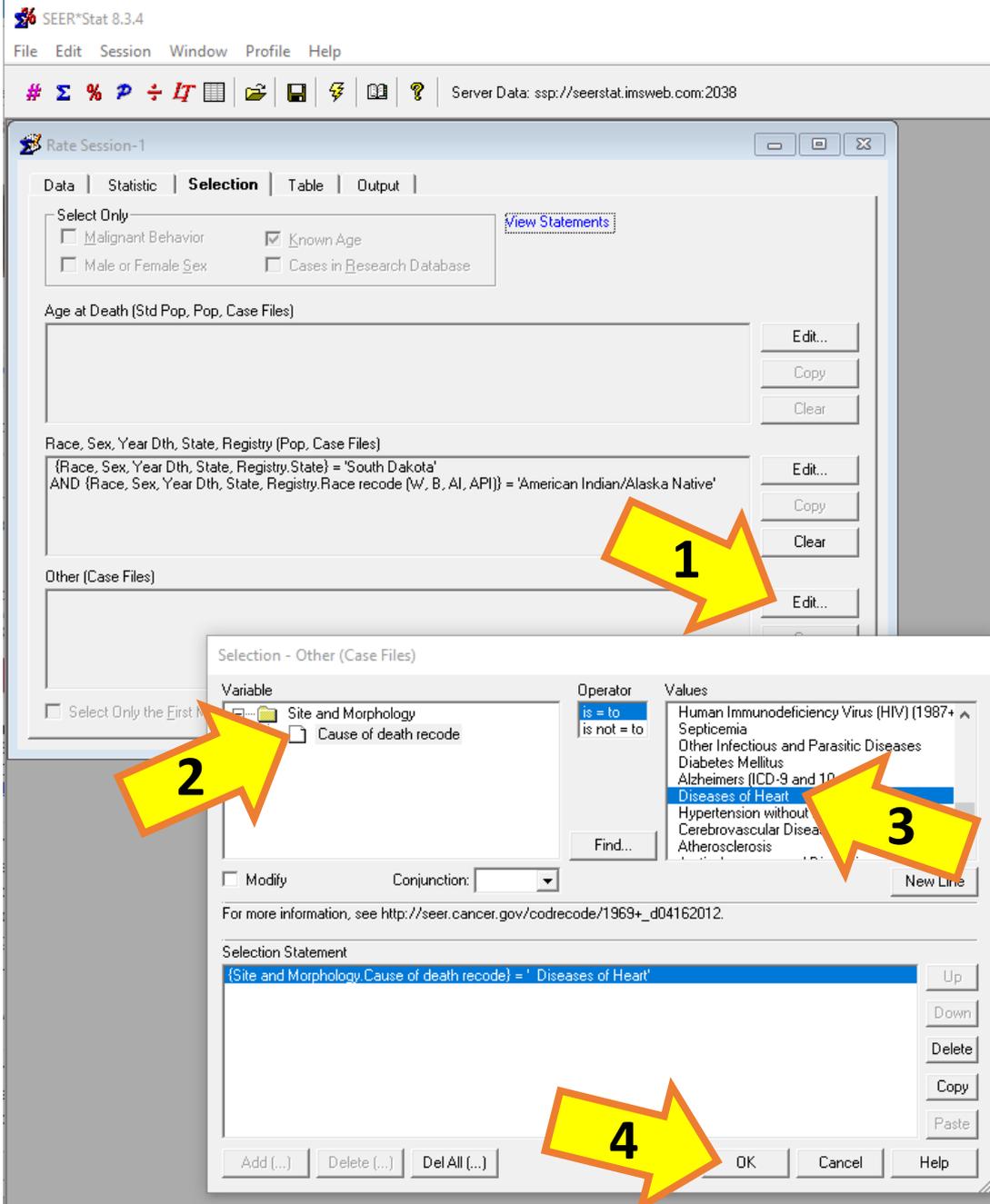
Up  
Down  
Delete  
Copy  
Paste

Add (...) Delete (...) Del All (...) OK Cancel Help

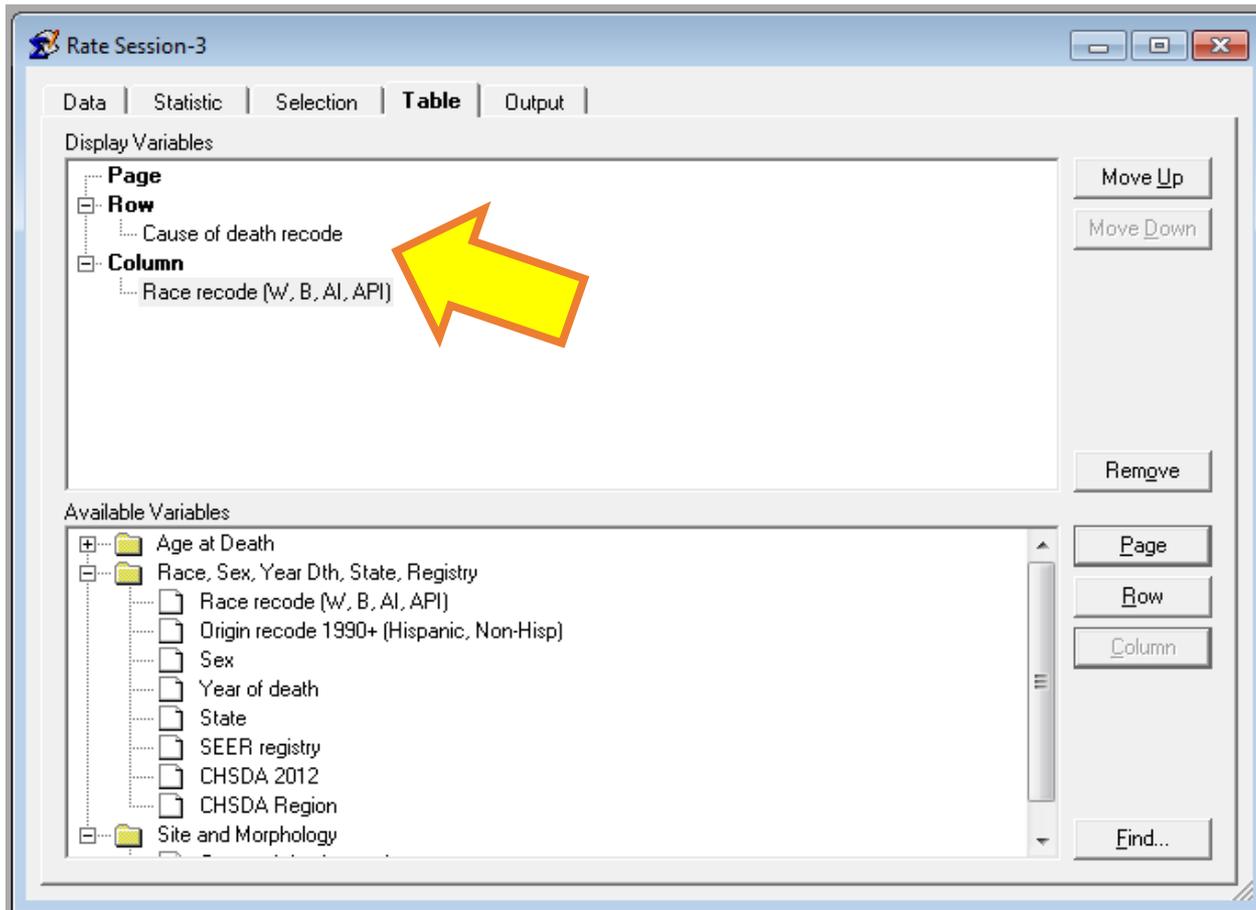
- In the “Selection” tab, click “Edit” on the “Race, Sex, Year Dth, State, Registry” category.
- Open the folder titled “Race, Sex, Year Dth, State, Registry” in the “Variables” window.
- Select “State” in this folder.
- Select “South Dakota” in the “Values” window.



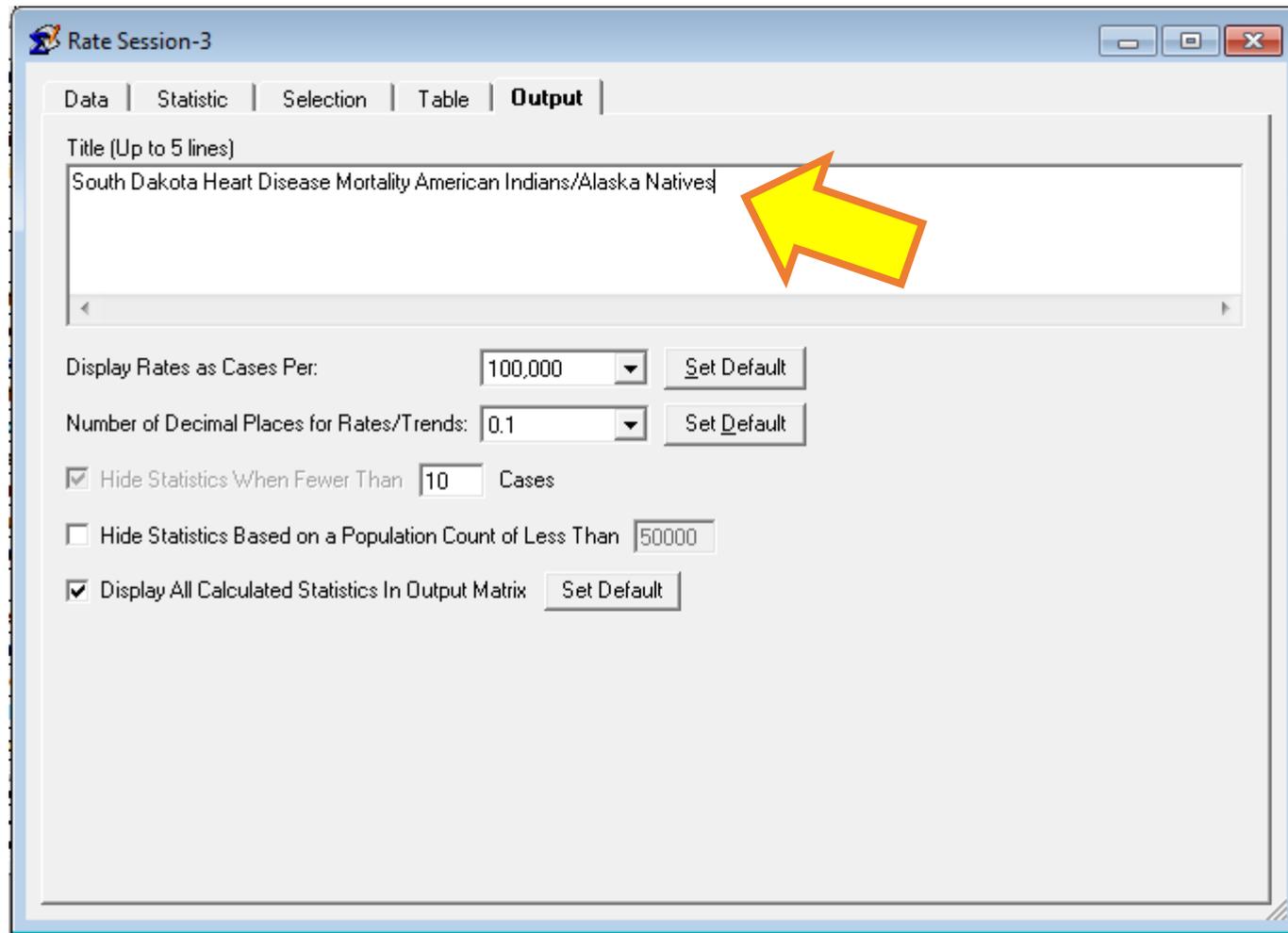
- Within the same “Variable” folder, click on “Race recode” .
- Choose “American Indian/Alaska Native”.
- Click “OK”.



- Click “Edit” in the “Other” category.
- Select “Cause of death recode” in the “Site and Morphology” folder.
- Select “Diseases of Heart” in the “Values” window.
- Click “OK”.



- Go to the Table tab.
- Choose "Cause of death recode" as the row variable.
- Choose "Race recode" as the column variable.



- Go to the “Output” tab.
- Give the table the title “South Dakota Heart Disease Mortality American Indians/Alaska Natives”.

Rate Session-3 Matrix-3

South Dakota Heart Disease Mortality American Indians/Alaska Natives

	Rate	SE	Lower CI	Upper CI	Count	Pop
Aleukemic, Subleuker	^	^	^	^	^	1,663,665
Miscellaneous Malignan	^	^	^	^	^	1,663,665
In situ, benign or unknow	^	^	^	^	^	1,663,665
Tuberculosis	^	^	^	^	^	1,663,665
Syphilis	^	^	^	^	^	1,663,665
Human Immunodeficienc	^	^	^	^	^	1,663,665
Septicemia	^	^	^	^	^	1,663,665
Other Infectious and Para	^	^	^	^	^	1,663,665
Diabetes Mellitus	^	^	^	^	^	1,663,665
Alzheimers (ICD-9 and 10	^	^	^	^	^	1,663,665
<b>Diseases of Heart</b>	324.3	7.5	309.8	339.2	2,225	1,663,665
Hypertension without He	^	^	^	^	^	1,663,665
Cerebrovascular Disease	^	^	^	^	^	1,663,665
Atherosclerosis	^	^	^	^	^	1,663,665
Aortic Aneurysm and Dis	^	^	^	^	^	1,663,665
Other Diseases of Arterie	^	^	^	^	^	1,663,665
Pneumonia and Influenza	^	^	^	^	^	1,663,665
Chronic Obstructive Pulm	^	^	^	^	^	1,663,665
Stomach and Duodenal Ul	^	^	^	^	^	1,663,665
Chronic Liver Disease an	^	^	^	^	^	1,663,665
Nephritis, Nephrotic Synd	^	^	^	^	^	1,663,665
Complications of Pregna	^	^	^	^	^	1,663,665

Underlying mortality data provided by NCHS ([www.cdc.gov/nchs](http://www.cdc.gov/nchs)).  
 Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence in  
 ^ Statistic not displayed due to fewer than 10 cases.

- Click the lightning bolt icon to produce the table.

Server Data: ssp://seerstat.imsweb.com:2038

Rate Session-1

Data | Statistic | **Selection** | Table | Output

Select Only

Malignant Behavior  Known Age [View Statements](#)

Male or Female Sex  Cases in Research Database

Age at Death (Std Pop, Pop, Case Files)

Race, Sex, Year Dth, State, Registry (Pop, Case Files)

{Race, Sex, Year Dth, State, Registry, State} = 'South Dakota'  
AND {Race, Sex, Year Dth, State, Registry, Race recode (W, B, AI, API)} = 'American Indian/Ala'

Other (Case Files)

{Site and Morphology, Cause of death recode} = ' Diseases of Heart'

Select Only the First Matching P

**1**

Selection - Race, Sex, Year Dth, State, Registry (Pop, Case Files)

Variable	Operator	Values
Race, Sex, Year Dth, State, Registry	is = to	White
Race recode (W, B, AI, API)	is not = to	Black American I Asian or Pac Other unsped (1970

Modify Conjunction: AND New Line

Caution should be exercised when using this variable. For more information, see [http://seer.cancer.gov/seerstat/variables/seer/race\\_ethnicity](http://seer.cancer.gov/seerstat/variables/seer/race_ethnicity).

Selection Statement

{Race, Sex, Year Dth, State, Registry, State} = 'South Dakota'  
AND {Race, Sex, Year Dth, State, Registry, Race recode (W, B, AI, API)} = 'White'

**2**

**3**

**4**

Add (...) Delete (...) Del All (...) OK Cancel Help

- Return to the “Selection” tab in the same rate session window.
- Click “Edit” for the “Race, Sex, Year Dth, State, Registry” category.
- Click the “AND {Race, Sex,...” option in the Selection Statement box
- Select “White” in the “Values” window.
- Rename the table in the “Output” tab.
- Click the lightning bolt icon to produce the second table.

Rate Session-3 Matrix-4

South Dakota Heart Disease Mortality whites

	Rate	SE	Lower CI	Upper CI	Count	Pop
Aleukemic, Subleuker						17,187,759
Miscellaneous Malignan						17,187,759
In situ, benign or unknow						17,187,759
Tuberculosis						17,187,759
Syphilis						17,187,759
Human Immunodeficienc						17,187,759
Septicemia						17,187,759
Other Infectious and Para						17,187,759
Diabetes Mellitus						17,187,759
Alzheimers (ICD-9 and 10						17,187,759
Diseases of Heart	205.0	1.0	203.1	206.9	46,485	17,187,759
Hypertension without He						17,187,759
Cerebrovascular Disease						17,187,759
Atherosclerosis						17,187,759
Aortic Aneurysm and Dis						17,187,759
Other Diseases of Arterie						17,187,759
Pneumonia and Influenza						17,187,759
Chronic Obstructive Pulm						17,187,759
Stomach and Duodenal Ul						17,187,759
Chronic Liver Disease an						17,187,759
Nephritis, Nephrotic Synd						17,187,759
Complications of Pregnai						17,187,759

Underlying mortality data provided by NCHS ([www.cdc.gov/nchs](http://www.cdc.gov/nchs)). Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence interval not displayed due to fewer than 10 cases.

Rate Session-3 Matrix-3

South Dakota Heart Disease Mortality American Indians/Alaska Natives

	Rate	SE	Lower CI	Upper CI	Count	Pop
Aleukemic, Subleuker						1,663,665
Miscellaneous Malignan						1,663,665
In situ, benign or unknow						1,663,665
Tuberculosis						1,663,665
Syphilis						1,663,665
Human Immunodeficienc						1,663,665
Septicemia						1,663,665
Other Infectious and Para						1,663,665
Diabetes Mellitus						1,663,665
Alzheimers (ICD-9 and 10						1,663,665
Diseases of Heart	324.3	7.5	309.8	339.2	2,225	1,663,665
Hypertension without He						1,663,665
Cerebrovascular Disease						1,663,665
Atherosclerosis						1,663,665
Aortic Aneurysm and Dis						1,663,665
Other Diseases of Arterie						1,663,665
Pneumonia and Influenza						1,663,665
Chronic Obstructive Pulm						1,663,665
Stomach and Duodenal Ul						1,663,665
Chronic Liver Disease an						1,663,665
Nephritis, Nephrotic Synd						1,663,665
Complications of Pregnai						1,663,665

Underlying mortality data provided by NCHS ([www.cdc.gov/nchs](http://www.cdc.gov/nchs)). Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence interval not displayed due to fewer than 10 cases.

With both tables produced, we can now compare heart disease mortality between American Indians/Alaska Natives and whites in South Dakota.

American Indians/Alaska Natives have a lower count of heart disease compared to whites in South Dakota. However, when we look at the mortality rate per 100,000 population, we see that the heart disease mortality rate for American Indians/Alaska Natives is 324.3 deaths per 100,000, compared to the lower 205.0 deaths per 100,000 for whites for the period of 1990-2014.

# What if we want to compare age-adjusted mortality for two races in the same table?

We will start by selecting the same population in the “Data” tab, and following the same steps as before for the “Statistic” tab.

- Start a new rate session.
- Select “Mortality - All COD, Aggregated With County, Total U.S. (1990-2014) <Katrina/Rita Population Adjustment>”.

- In the “Statistic” tab, choose an age-adjusted rate, standardized to the 2000 US Population, with standard errors and confidence intervals.

Server Data: ssp://seerstat.imsweb.com:2038

Rate Session-1

Data | Statistic | **Selection** | Table | Output

Select Only

Malignant Behavior  Known Age [View Statements](#)

Male or Female Sex  Cases in Research Database

Age at Death (Std Pop, Pop, Case Files)

Race, Sex, Year Dth, State, Registry (Pop, Case Files)

Other (Case Files)

Select Only

Selection - Race, Sex, Year Dth, State, Registry (Pop, Case Files)

Variable	Operator	Values
<ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Race, Sex, Year Dth, State, Registry</li> <li><input type="checkbox"/> Race recode (W, B, AI, API)</li> <li><input type="checkbox"/> Origin recode 1990+ (Hispanic, Nor)</li> <li><input type="checkbox"/> Sex</li> <li><input type="checkbox"/> Year of death</li> <li><input type="checkbox"/> State</li> <li><input type="checkbox"/> SEER registry</li> </ul>	<ul style="list-style-type: none"> <li><input checked="" type="radio"/> is = to</li> <li><input type="radio"/> is not = to</li> </ul>	<ul style="list-style-type: none"> <li>Oregon</li> <li>Pennsylvania</li> <li>Rhode Island</li> <li>South Carolina</li> <li><b>South Dakota</b></li> <li>Tennessee</li> <li>Texas</li> <li>Utah</li> <li>Vermont</li> </ul>

Modify Conjunction:

Selection Statement

{Race, Sex, Year Dth, State, Registry.State} = 'South Dakota'

- In the “Selection” tab, click “Edit” on the “Race, Sex, Year Dth, State, Registry” category.
- Open the folder titled “Race, Sex, Year Dth, State, Registry” in the “Variables” window.
- Select “State” in this folder.
- Select “South Dakota” in the “Values” window and click “OK”.

Server Data: ssp://seerstat.imsweb.com:2038

Rate Session-1

Data | Statistic | **Selection** | Table | Output

Select Only

Malignant Behavior  Known Age [View Statements](#)

Male or Female Sex  Cases in Research Database

Age at Death (Std Pop, Pop, Case Files)

Edit...  
Copy  
Clear

Race, Sex, Year Dth, State, Registry (Pop, Case Files)

{Race, Sex, Year Dth, State, Registry.State} = 'South Dakota'

Edit...  
Copy  
Clear

Other (Case Files)

Edit...  
Copy

Select Only the First

Selection - Other (Case Files)

Variable

Site and Morphology  
 Cause of death recode

Operator

is = to  
is not = to

Values

Human Immunodeficiency Virus (HIV) (1987+)  
Septicemia  
Other Infectious and Parasitic Diseases  
Diabetes Mellitus  
Alzheimers (ICD-9 and 10)  
**Diseases of Heart**  
Hypertension without  
Cerebrovascular Disease  
Atherosclerosis

Find...  
New Case

Modify Conjunction:

For more information, see [http://seer.cancer.gov/codrecode/1969+\\_d04162012](http://seer.cancer.gov/codrecode/1969+_d04162012).

Selection Statement

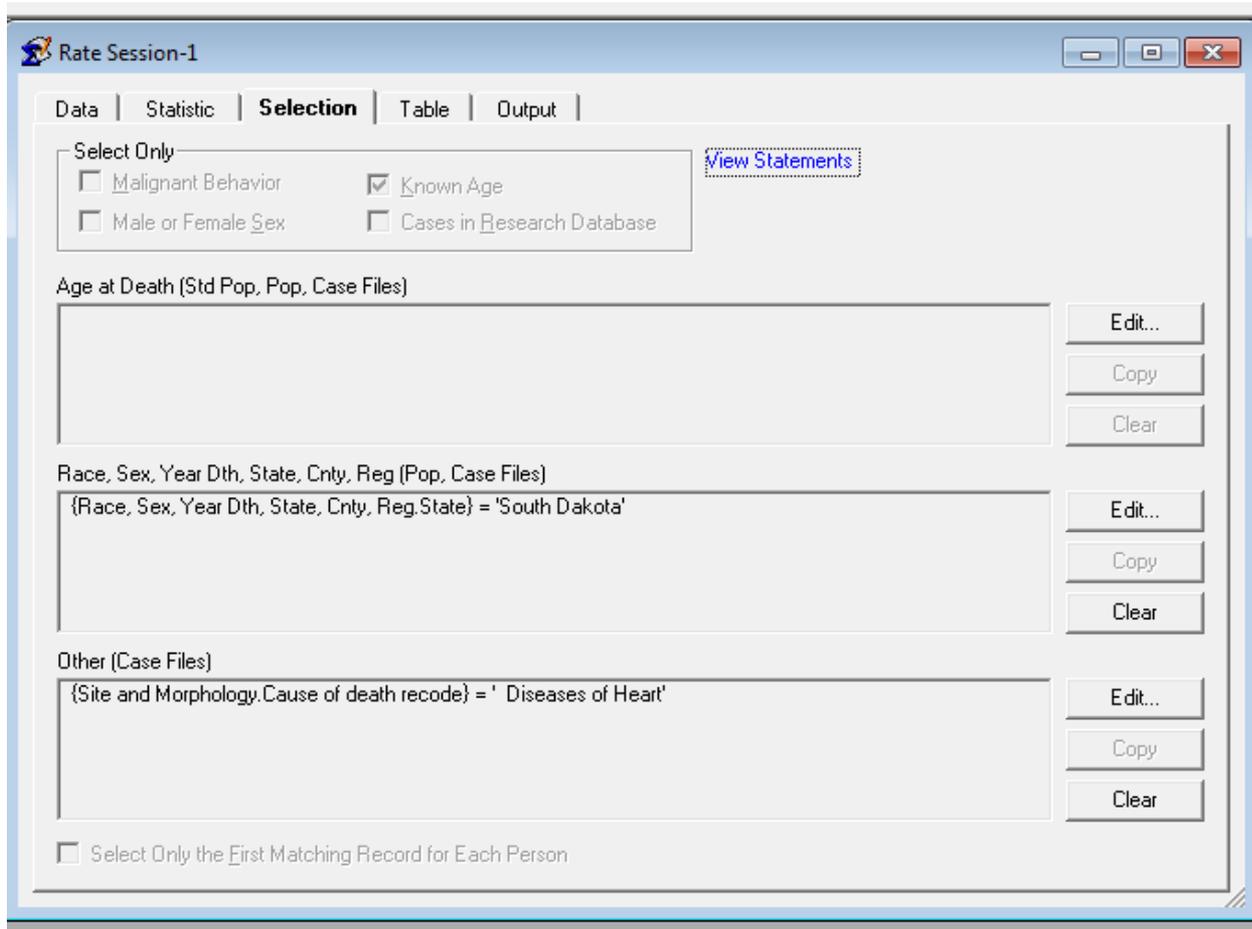
{Site and Morphology.Cause of death recode} = 'Diseases of Heart'

Up  
Down  
Delete  
Copy  
Paste

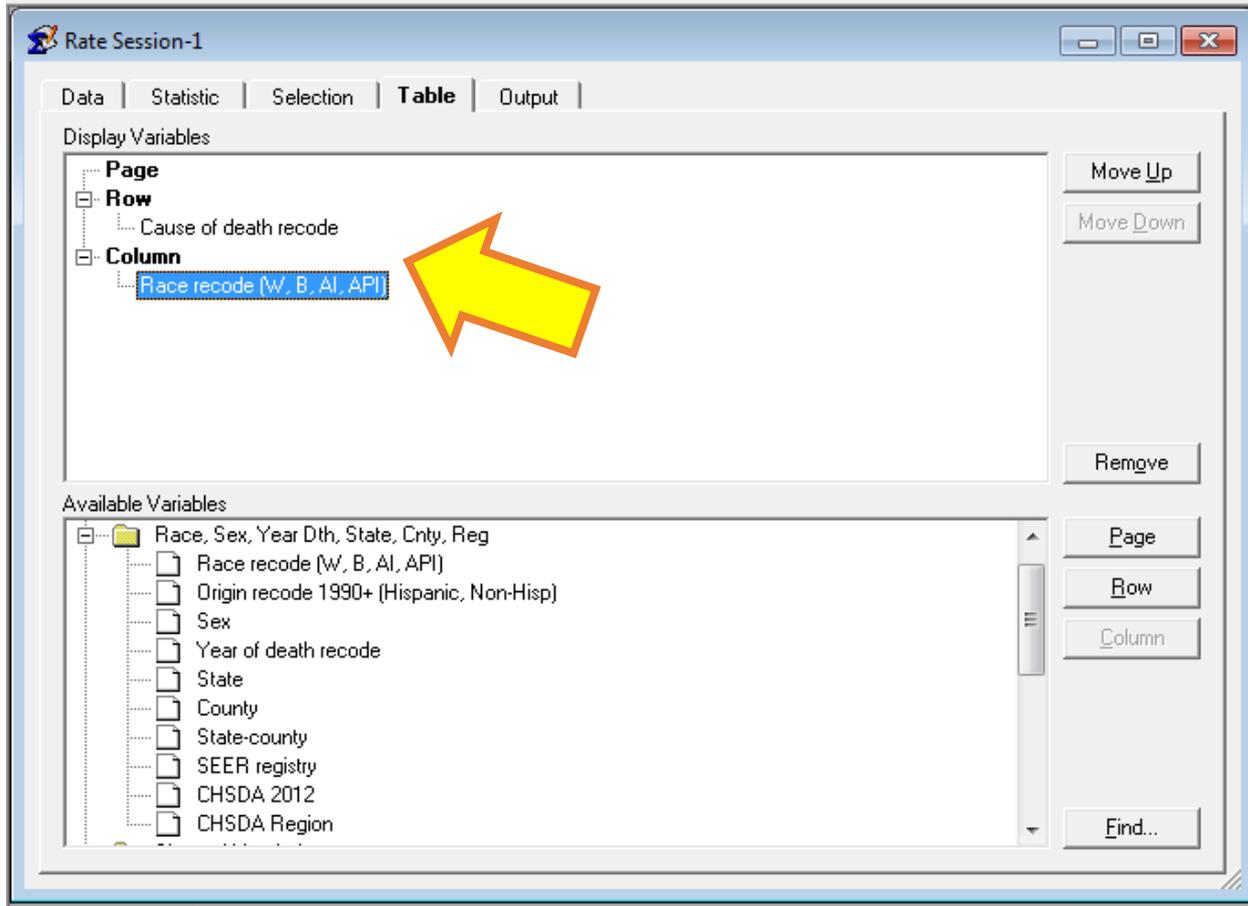
Add (...) Delete (...) Del All (...)

OK Cancel Help

- Click “Edit” in the “Other” category.
- Select “Cause of death recode” in the “Site and Morphology” folder.
- Select “Diseases of Heart” in the “Values” window and click “OK”.



You may notice that we did not select the races that we want to compare to our “Race, Sex, Year Dth, State, Cnty, Reg” category. This will allow us to keep all races in the same table.



- Go to the Table tab.
- Choose “Cause of death recode” as the row variable.
- Choose “Race recode” as the column variable.

Rate Session-1

Data | Statistic | Selection | Table | **Output**

Title (Up to 5 lines)  
South Dakota Heart Disease Mortality American Indians/Alaska Natives Table 2

Display Rates as Cases Per: 100,000 Set Default

Number of Decimal Places for Rates/Trends: 0.1 Set Default

Hide Statistics When Fewer Than 10 Cases

Hide Statistics Based on a Population Count of Less Than 50000

Display All Calculated Statistics In Output Matrix Set Default

- Go to the “Output” tab.
- Give the table the title “South Dakota Heart Disease Mortality American Indians/Alaska Natives Table 2”.

Rate Session-1 Matrix-2

South Dakota Heart Disease Mortality American Indians/Alaska Natives Table 2

	White						Black						American Indian/Alaska Native		
	Rate	SE	Lower CI	Upper CI	Count	Pop	Rate	SE	Lower CI	Upper CI	Count	Pop	Rate	SE	Lower CI
Miscellaneous Malignan	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
In situ, benign or unknow	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Tuberculosis	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Syphilis	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Human Immunodeficienc	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Septicemia	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Other Infectious and Para	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Diabetes Mellitus	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Alzheimers (ICD-9 and 10	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
<b>Diseases of Heart</b>	<b>205.0</b>	<b>1.0</b>	<b>203.1</b>	<b>206.9</b>	<b>46,485</b>	<b>17,187,759</b>	<b>132.4</b>	<b>15.9</b>	<b>103.1</b>	<b>166.4</b>	<b>86</b>	<b>215,965</b>	<b>324.3</b>	<b>7.5</b>	<b>301.1</b>
Hypertension without He	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Cerebrovascular Disease	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Atherosclerosis	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Aortic Aneurysm and Dis	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Other Diseases of Arterie	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Pneumonia and Influenza	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Chronic Obstructive Pulm	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Stomach and Duodenal Ul	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Chronic Liver Disease an	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	
Nephritis, Nephrotic Synd	▲	▲	▲	▲	▲	17,187,759	▲	▲	▲	▲	▲	215,965	▲	▲	

Underlying mortality data provided by NCHS ([www.cdc.gov/nchs](http://www.cdc.gov/nchs)).  
Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence intervals (Tiwarí mod) are 95% for rates.  
Statistic not displayed due to fewer than 10 cases.

- Click the lightning bolt icon to produce the table.

All races with data for heart disease mortality in South Dakota are now displayed in the same table.

# Finding Age-Adjusted Incidence Rates

Example: Breast Cancer



University of Nebraska  
Medical Center

# Age-Adjusted Incidence Rates

- This section will instruct you on how to get age-adjusted incidence rates using SEER data
- For this example, we will show you to get access to breast cancer incidence counts using the SEER 9 registries data, with the following criteria:
  - Only include females
  - Exclude 'In Situ' cases
  - Display data by custom age categories
    - < 50 years or 50+ years
  - Display data separately by race
  - Show data by year (2000 – 2014)
- This example was done using downloaded binary SEER data for use in SEER\*Stat. If you are using online data instead (i.e. entering your username and password), a couple of your screens and output may be slightly different from the tutorial.



Browse the SEER Cancer Statistics Review 1975-2014

NATIONAL CANCER INSTITUTE  
Surveillance, Epidemiology, and End Results Program

Cancer Statistics  
Statistical Summaries | Interactive Tools | Publications

For Researchers  
Datasets and Software

Home > Statistical Summaries > Cancer Statistics Review, 1975-2014 > Browse the Tables

Browse the SEER Cancer Statistics Review 1975-2014

The navigation below allows you to jump to any table or figure within the SEER Cancer Statistics Review 1975-2014.

1. First select the CSR Section, then a Table/Figure from that section.
2. Use the Submit button to view the table or figure. You will have options to download or print the table or figure.

To learn more about the Cancer Statistics Review and the statistics presented in this report, click on the link below.

Section:

Table/Figure:

Download and Print: [Download Printer-friendly PDF](#) [Download data](#)

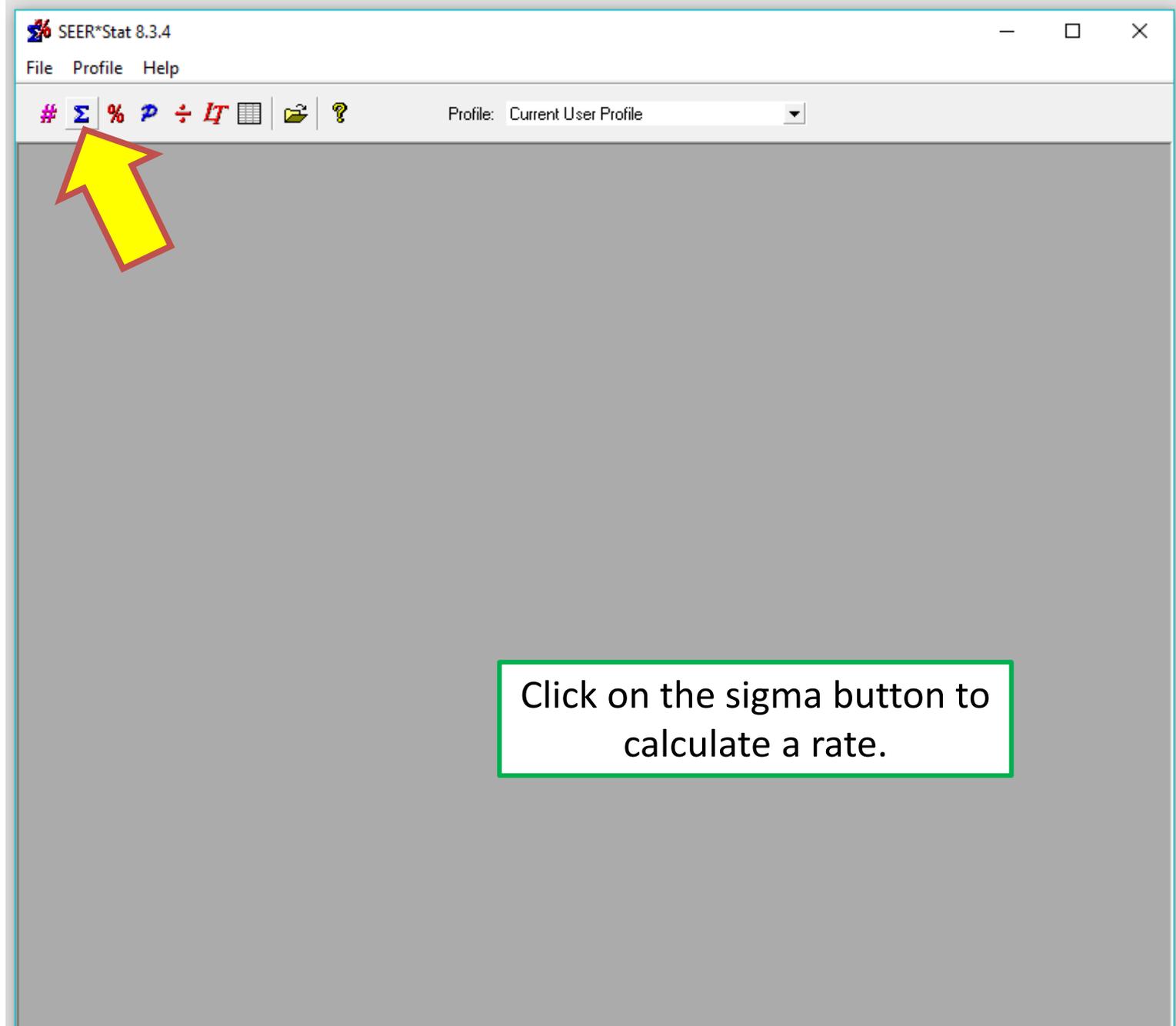
For this example, the table of age-adjusted incident rates we will generate is very similar to this table from the following SEER website:

[https://seer.cancer.gov/csr/1975\\_2014/browse\\_csr.php?sectionSEL=4&pageSEL=sect\\_0\\_4\\_table.08.html#](https://seer.cancer.gov/csr/1975_2014/browse_csr.php?sectionSEL=4&pageSEL=sect_0_4_table.08.html#)

**Table 4.8**  
**Cancer of the Female Breast (Invasive)**

Age-adjusted SEER Incidence<sup>a</sup> Rates by Year, Race and Age

Year of Diagnosis	All Races, Females			White Females			Black Females		
	All Ages	Ages <50	Ages 50+	All Ages	Ages <50	Ages 50+	All Ages	Ages <50	Ages 50+
1975-2014	126.21	43.60	342.52	129.82	44.00	354.55	118.14	44.33	311.40
1975	105.08	40.64	273.82	107.39	40.72	281.95	93.57	42.97	226.09
1976	101.94	40.00	264.15	104.78	40.77	272.40	85.72	38.05	210.55
1977	100.80	39.09	262.40	103.34	39.28	271.08	87.14	37.87	216.14
1978	100.62	38.87	262.30	103.59	39.32	271.90	86.29	40.20	206.95
1979	102.09	38.01	269.90	104.70	38.61	277.77	86.94	37.76	215.71
1980	102.24	37.77	271.06	105.10	38.28	280.09	89.48	37.69	225.12
1981	106.36	39.21	282.19	109.91	40.10	292.72	94.35	38.86	239.64
1982	106.49	40.55	279.14	109.92	41.20	289.85	93.58	38.02	239.06



SEER\*Stat 8.3.4

File Edit Session Window Profile Help

# Σ % P ÷ *IT* [Grid] [Folder] [Save] [Lightning] [Book] [Help] Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

Rate Session-1

Data | Statistic | Selection | Table | Output

Database Name	Linked To	Linked By
Incidence - SEER 9 Regs Research Data, Nov 2016 Sub (1973-2014) <Katrina/Rita Population ...	County A...	State-co...
Incidence - SEER 13 Regs Research Data, Nov 2016 Sub (1992-2014) <Katrina/Rita Population...	County A...	State-co...
Incidence - SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases, Nov 2...	County A...	State-co...

Find Original

Age Variable: Age recode with

Suggested citation for the selected data: Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER\*Stat Database: Incidence - SEER 9 Regs Research Data, Nov 2016 Sub (1973-2014) <Katrina/Rita Population Adjustment> - Linked To County Attributes - Total U.S., 1969-2015 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released April 2017, based on the November 2016 submission.

- Choose the appropriate dataset.
- To match the output on the website, we'll choose the SEER 9 dataset, and verify that the "Age Variable" is set to "Age recode with < 1 year olds".
- Next, click the "Statistic" tab.

1

2

3

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

Rate Session-1

Data | **Statistic** | Selection | Table | Output

Statistic

- Rates (Crude)
- Rates (Age-Adjusted)
- Trends (Crude)
- Trends (Age-Adjusted)

Parameters

P-Values: .05

Standard Population: 2000 US Std Population (19 age groups - Census P25-1130)

Age Variable: Age recode with <1 year olds

Trend Variable: Year of diagnosis

PC End Points

- One Year
- Two-Year Avg.

- Include Rate Ratios or
- Show Standard Errors
  - Use Tiwari et al., 2
  - Show P-Values for
- Show Standard Popula
- Delay Adjust
  - Delay Factor:

- For age-adjusted incidence rates, make sure “Rates (Age-Adjusted)” is selected.
- Also, we’ll use the “2000 US standard population”, and “Age recode with < 1 year olds”.



SEER\*Stat 8.3.4

File Edit Session Window Profile Help

# Σ % P ÷  $\int$  Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

Rate Session-1

Data | Statistic | **Selection** | Table | Output

Select Only

Malignant Behavior  Known Age [View Statements](#)

Male or Female Sex  Cases in Research Database

Age at Diagnosis (Std Pop, Pop, Case Files)

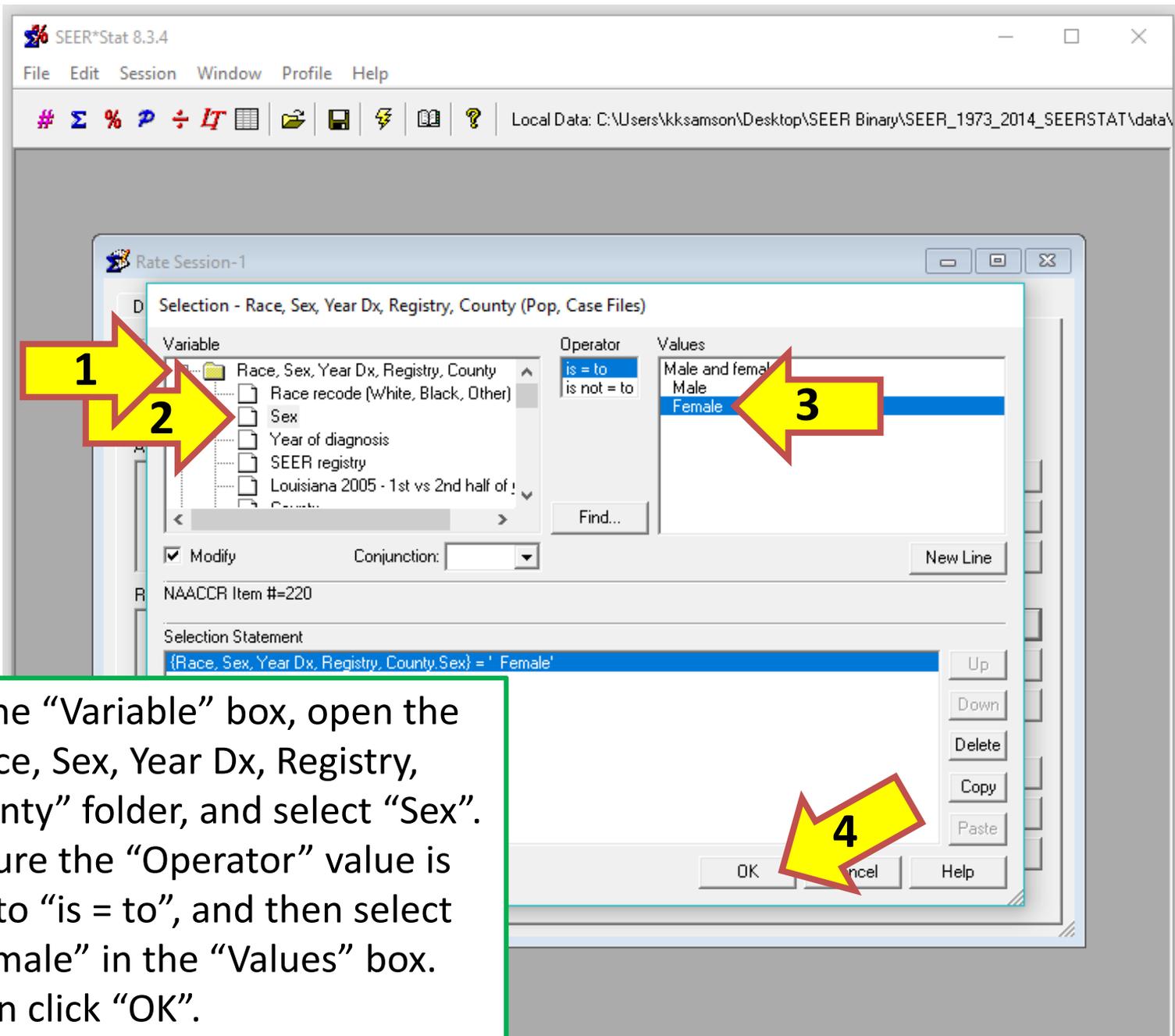
Edit...  
Copy  
Clear

Race, Sex, Year Dx, Registry, County (Pop, Case Files)

Edit...  
Copy  
Clear

Edit...  
Copy  
Clear

- For selection criteria, we can restrict which cases we want to use for our calculations.
- Start from the top box and move down.
- We plan to use all ages, so we won't use the first box.
- For the second box, we can restrict the calculations to only include females.
- To do this, click the "Edit..." button for the second box.



- In the “Variable” box, open the “Race, Sex, Year Dx, Registry, County” folder, and select “Sex”.
- Ensure the “Operator” value is set to “is = to”, and then select “Female” in the “Values” box.
- Then click “OK”.

The screenshot shows the SEER\*Stat 8.3.4 application window. The 'Rate Session-1' dialog box is open, with the 'Selection' tab selected. In the 'Select Only' section, the 'Malignant Behavior' checkbox is checked, and the 'Known Age' checkbox is also checked. The 'Age at Diagnosis (Std Pop, Pop, Case Files)' section is empty. The 'Race, Sex, Year Dx, Registry, County (Pop, Case Files)' section contains the text '{Race, Sex, Year Dx, Registry, County.Sex} = ' Female''. The 'Other (Case Files)' section is empty. A yellow arrow points to the 'Edit...' button for the 'Other (Case Files)' section.

- Our selection criteria for sex has been set.
- Now, let's limit cases to breast cancer cases only, and only those that are malignant (i.e. exclude 'in situ' cases).
- Click the "Edit..." button for the bottom box.

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

Rate Session-1

Selection - Other (Case Files)

Variable	Operator	Values
Site and Morphology	is = to	Soft Tissue including Heart
Site recode ICD-O-3/WHO 2008	is not = to	Skin excluding Basal and Squamous
Behavior recode for analysis		Melanoma of the Skin
AYA site recode/WHO 2008		Other Non
Lymphoma subtype recode/WHO 2		Breast
ICCC site recode ICD-O-3/WHO 20		Female Gen
CC Site recode ICD-O-3/WHO 2004		Cervix Uteri
		Corpus and Uterus, NOS
		Corpus Uteri

Modify Conjunction: [v] New Line

Subject to change based on evolving ICD-O-3 coding rules. For more information, see <https://seer.cancer.gov/siterecode>.

Selection Statement

```
{Site and Morphology.Site recode ICD-O-3/WHO 2008} = ' Breast'
```

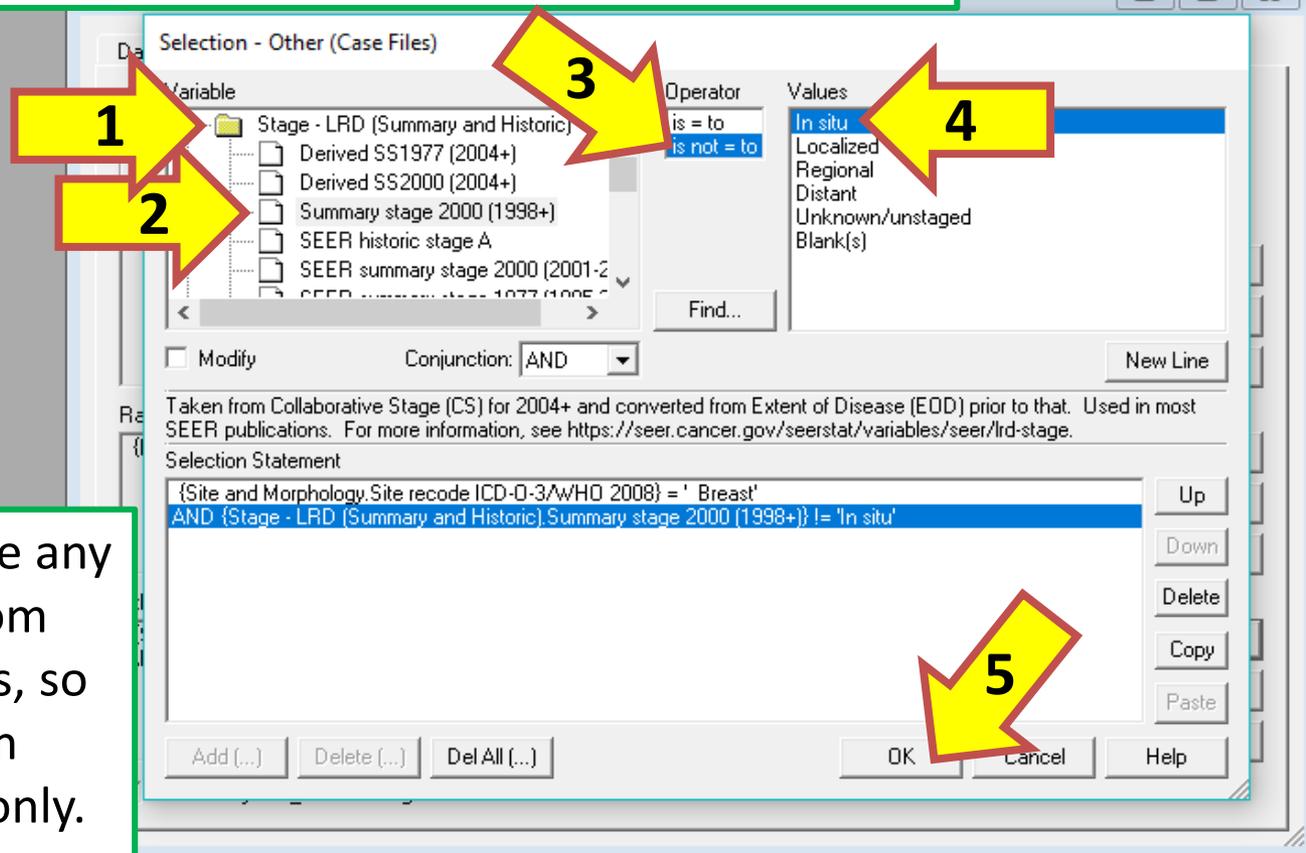
Up Down Delete Copy Paste Cancel Help

1

2

In the “Variable” box, open the “Site and Morphology” folder, and select “Site recode ICD-O-3/WHO 2008”. Ensure the “Operator” value is set to “is = to”, and then select “Breast” in the “Values” box. Don’t hit “OK” quite yet – we have another criteria to enter.

- In the “Variable” box, open the “Stage – LRD (Summary and Historic)” folder, and select “Summary stage 2000 (1998+)”.
- Ensure the “Operator” value is set to “is not = to”, and then select “In situ” in the “Values” box.



- This will exclude any in situ cases from our calculations, so we can focus on invasive cases only.
- Click “OK”.

The screenshot shows the SEER\*Stat 8.3.4 application window. The main menu includes File, Edit, Session, Window, Profile, and Help. The status bar indicates the local data path: C:\Users\vkksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\.

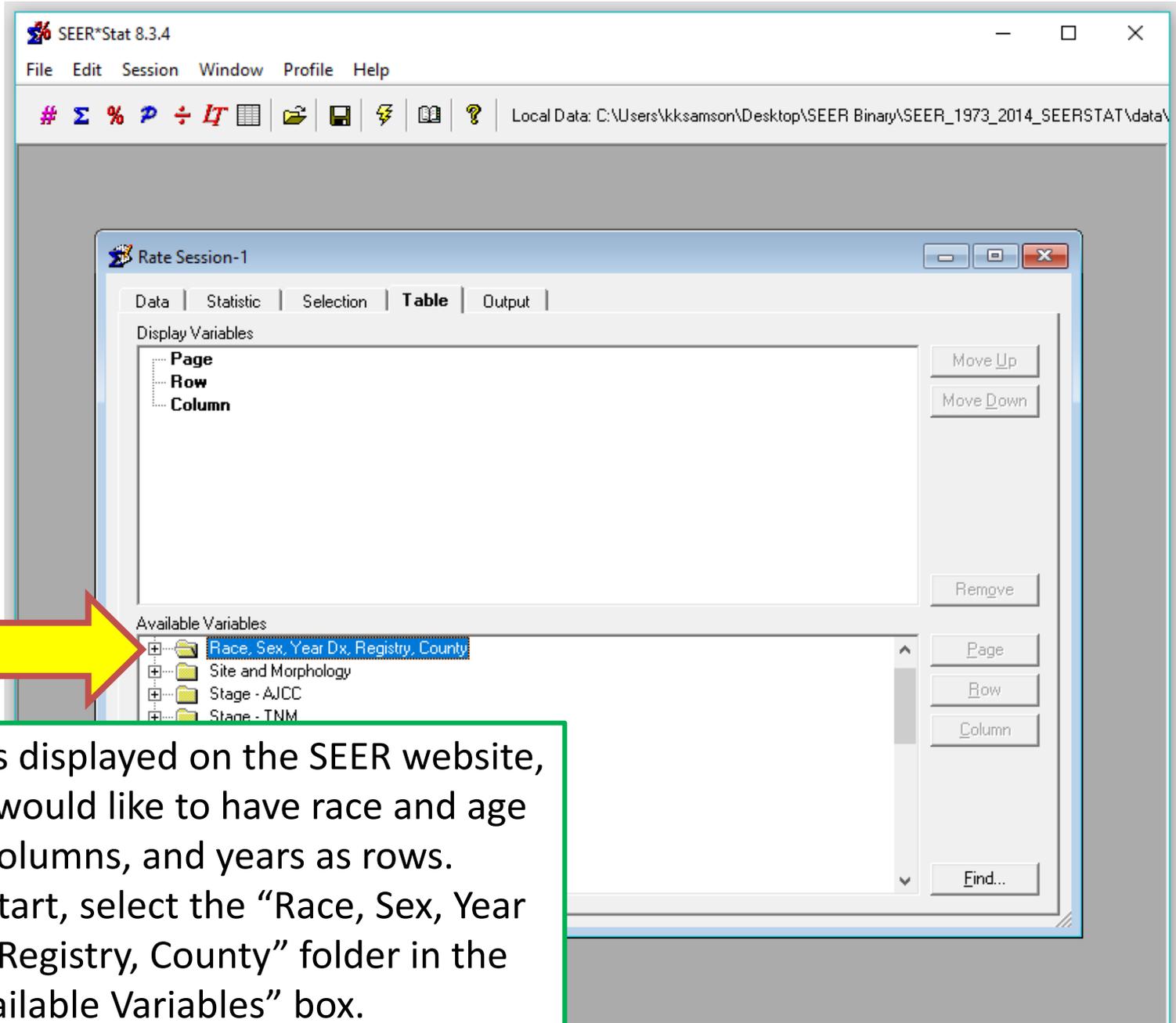
The 'Rate Session-1' dialog box is open, with the 'Selection' tab selected. A yellow arrow points to the 'Selection' tab. The dialog box contains the following sections:

- Select Only:**
  - Malignant Behavior
  - Known Age
  - Male or Female Sex
  - Cases in Research Database
- Age at Diagnosis (Std Pop, Pop, Case Files):** (Empty text box)
- Race, Sex, Year Dx, Registry, County (Pop, Case Files):**  
{Race, Sex, Year Dx, Registry, County.Sex} = ' Female'
- Other (Case Files):**  
{Site and Morphology.Site recode ICD-0-3/WHO 2008} = ' Breast'  
AND {Stage - LRD (Summary and Historic).Summary stage 2000 (1998+)} != 'In situ'

At the bottom of the dialog box, there is a checkbox:  Select Only the First Matching Record for Each Person.

A green-bordered text box at the bottom of the screenshot contains the following text:

All of our selection criteria has been set so we can move to the "Table" tab.



- As is displayed on the SEER website, we would like to have race and age as columns, and years as rows.
- To start, select the “Race, Sex, Year Dx, Registry, County” folder in the “Available Variables” box.

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

# Σ % P ÷ IT Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

- Page
- Row
- Column

Available Variables

- Race, Sex, Year Dx, Registry, County
- Race recode (White, Black, Other)**
- Sex
- Year of diagnosis
- SEER registry
- Louisiana 2005 - 1st vs 2nd half of year
- County
- State-county
- In research data
- CHSDA 2012
- CHSDA Region

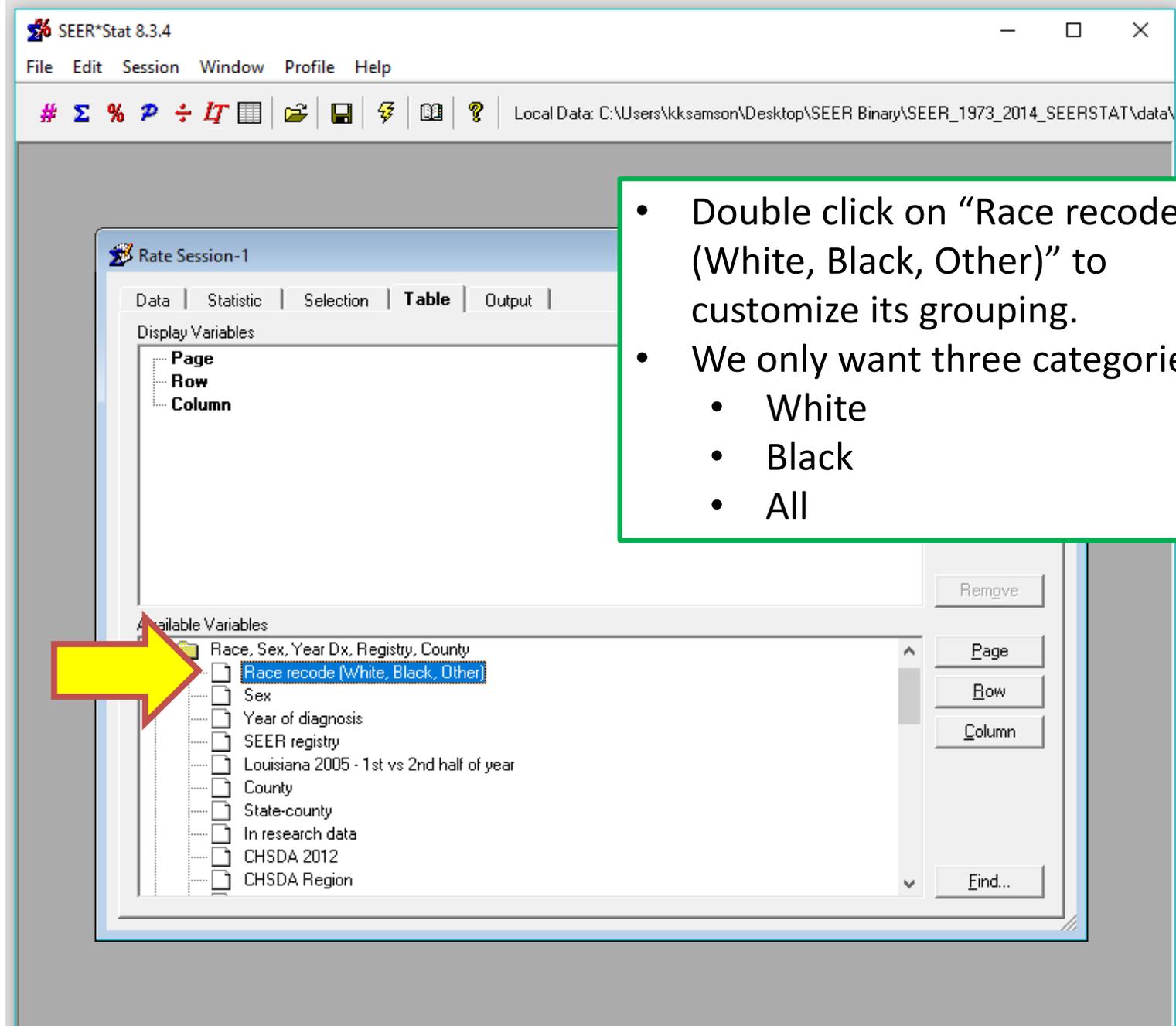
Remove

Page

Row

Column

Find...



- Double click on “Race recode (White, Black, Other)” to customize its grouping.
- We only want three categories:
  - White
  - Black
  - All

Rate Session-1

Data | Statistic

Display Variables

- Page
- Row
- Column

Available Variables

- Race, Sex, Year of Dx, Registry, County
- Race recode (White, Black, Other)
- Sex
- Year of diagnosis
- SEER registry
- Louisiana 2005 - 1st vs 2nd half of year
- County
- State-county
- In research data
- CHSDA 2012
- CHSDA Region
- State
- Site and Morphology
- Stage - AJCC
- Stage - TNM
- Stage - LRD (Summary and Historic)
- Therapy

Dictionary

Database: Incidence - SEER 9 Regs Research Data, Nov 2016 Sub (1973-2014) <Katrina/Rita Population Adjustment>

Close

Help

Create...

Edit...

Delete...

Merge...

Import...

Export...

Categories...

Find...

Move Up

Move Down

Remove

Page

Row

Column

Find...

CAPS NUM

Click on "Race recode (White, Black, Other)" and then click the "Create..." button.

Edit Variable - Race recode (White, Black, O)

Name:  Category:

Description: Caution should be exercised when using this variable. For more information, see [https://seer.cancer.gov/seerstat/variables/seer/race\\_ethnicity](https://seer.cancer.gov/seerstat/variables/seer/race_ethnicity).

Groupings

- All races
- White
- Black
- Other (American Indian/AK Native, Asian/Pacific)
- Unknown

Values (0 Selected)

- White
- Black
- Other (American Indian/AK Native, Asian/Pacific)
- Unknown

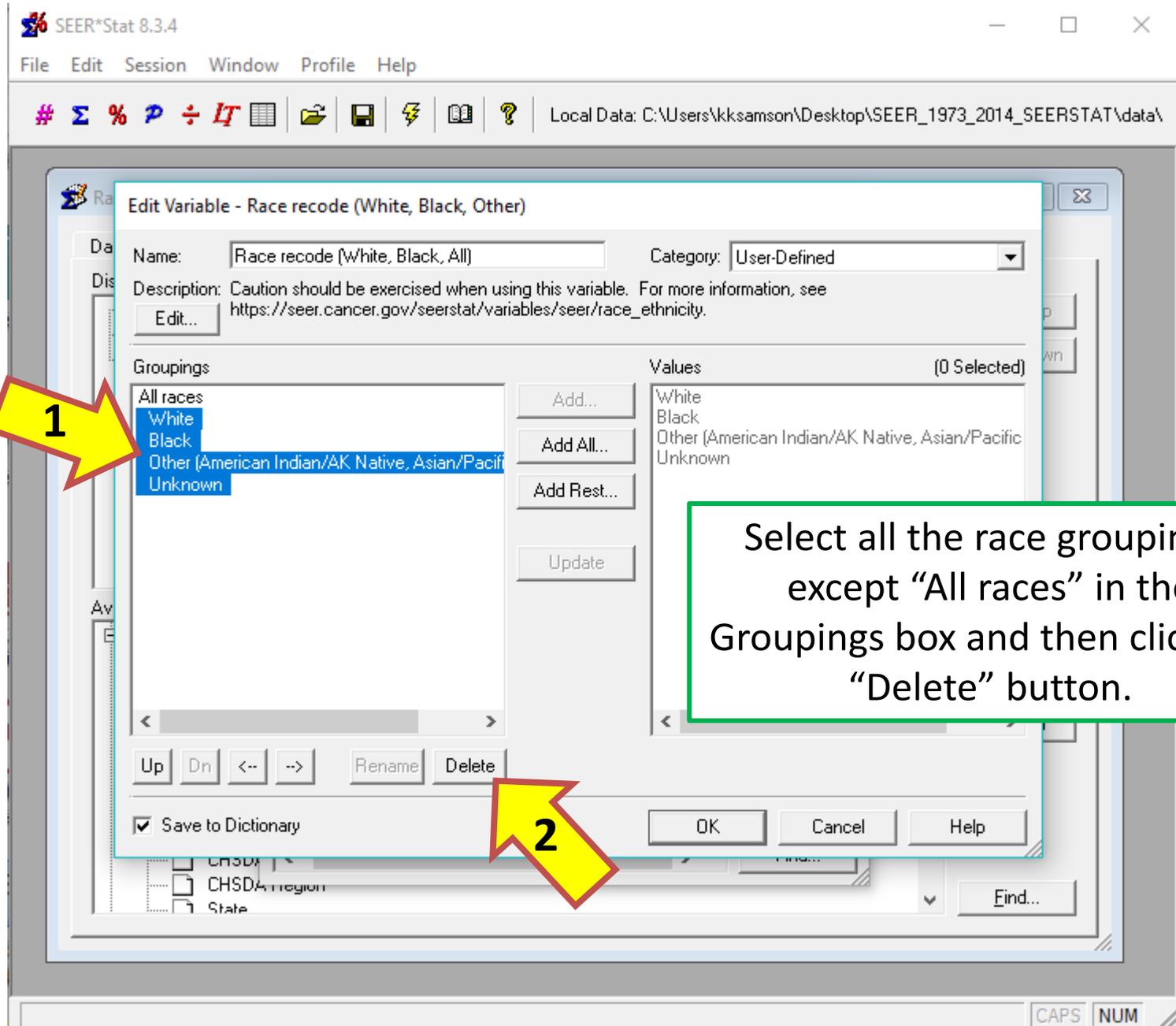
Buttons: Add..., Add All..., Add Rest..., Update, OK, Cancel, Help

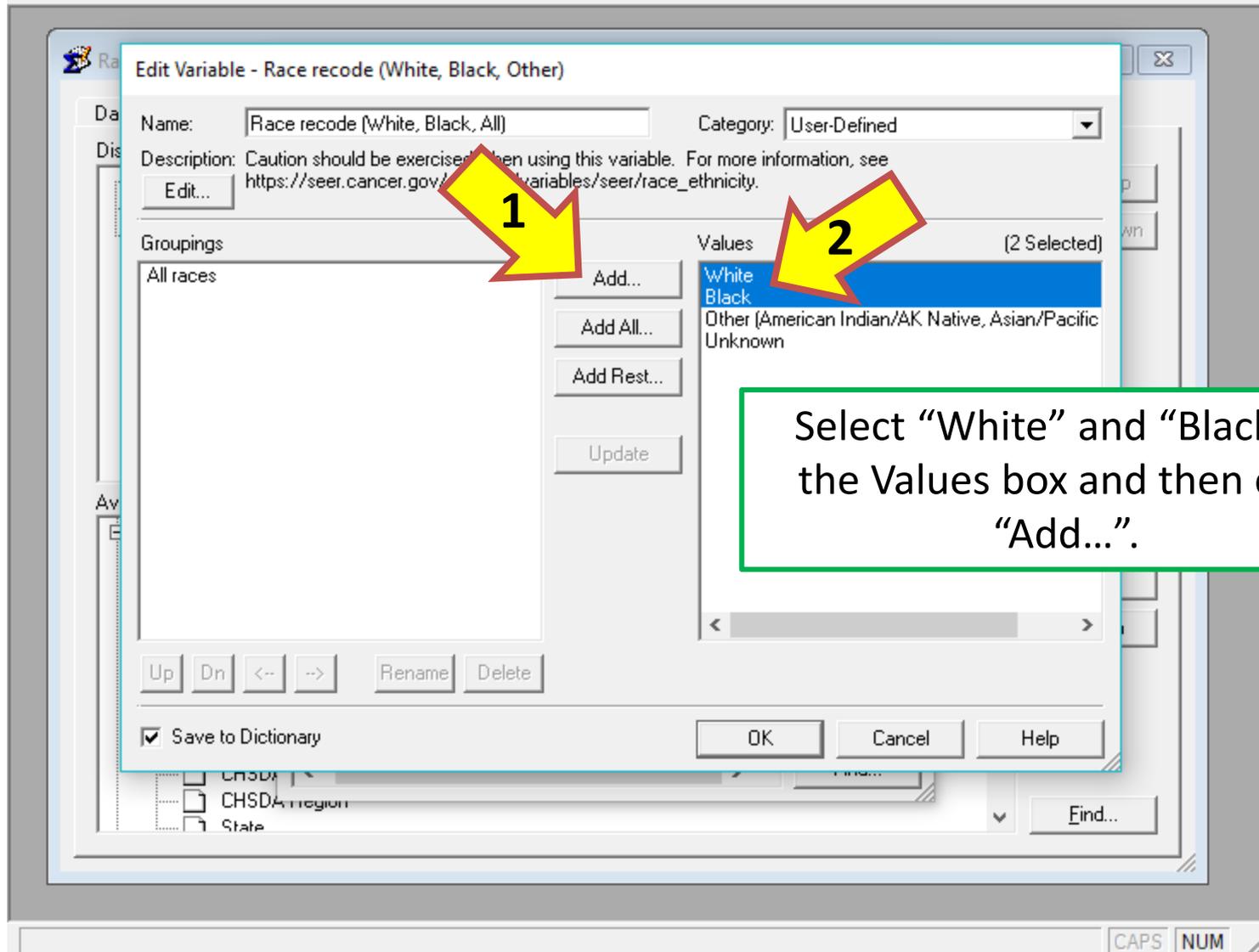
Find...

CHSDA region  
State

CAPS NUM

Rename this variable "Race recode (White, Black, All)" in the Name box.





Select "Added as separate groupings (one for each value)" and then click "OK".

**Edit Variable - Race recode (White, Black, All)**

Name: Race recode (White, Black, All) Category: User-Defined

Description: Caution should be exercised when using this variable. For more information, see [https://seer.cancer.gov/seerstat/variables/seer/race\\_ethnicity](https://seer.cancer.gov/seerstat/variables/seer/race_ethnicity).

Groupings	Values
All races	White
	Black
	Other (American Indian/AK Native, Asian/Pacific Islander)

**Add Selected Values**

The selected values will be:

- Added as one grouping (all values combined)
- Added as separate groupings (one for each value)
- Use Underlying Data Values Instead of Labels For Grouping Names

Buttons: OK, Cancel, Help

**Edit Variable - Race recode (White, Black, Other)**

Name: Race recode (White, Black, All) Category: User-Defined

Description: Caution should be exercised when using this variable. For more information, see [https://seer.cancer.gov/seerstat/variables/seer/race\\_ethnicity](https://seer.cancer.gov/seerstat/variables/seer/race_ethnicity).

Groupings

All races	Add...	White
White	Add All...	Black
Black	Add Rest...	Other (American Indian/AK Native, Asian/Pacific Unknown)
	Update	

Values (2 Selected)

Up Dn <- -> Rename Delete

Save to Dictionary

OK Cancel Help

Find...

Rate Session-1

Data | Statistic

Display Variables

Page  
Row  
Column

Available Variables

Race, Sex, ...  
Race r  
Sex  
Year o  
SEER  
Louisia  
County  
State-d  
In  
CA region  
State

Dictionary

Database: Incidence - SEER 9 Regs Research Data, Nov 2016 Sub (1973-2014) <Katrina/Rita Population Adjustment>

- Site and Morphology - Historic
- Race and Age (case data only)
- Dates
- Site Specific Sequence Numbers
- Other
- County attributes ACS - 2011-2015
- County attributes ACS - 2010-2014
- County attributes ACS - 2009-2013
- County attributes ACS - 2008-2012
- County attributes ACS - 2007-2011
- County attributes 2010s
- County attributes 2000s
- County attributes - modeled
- County attributes 1990s
- County attributes 1980s
- County attributes 1970s
- County attributes General
- System-Supplied
- Merged System-Supplied
- User-Defined
- Race recode (White, B

Close  
Help  
Create...  
Edit...  
Delete...  
Merge

Move Up  
Move Down  
Remove

Find...  
Find...

CAPS NUM

- Note that our new variable definition appears in a folder called "User-Defined".
- Click "Close".

Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

- Page
- Row
- Column

Available Variables

- County attributes 2010s
- County attributes 2000s
- County attributes - modeled small area estimates
- County attributes 1990s
- County attributes 1980s
- County attributes 1970s
- County attributes General
- System-Supplied
- Merged System-Supplied
- User-Defined
- Race recode (White, Black, All)**

Page

Row

Column

Find...

CAPS NUM

- Find your new race variable in the Available Variables box.
- It will be in the “User-Defined” folder.
- Click on the “Race recode (White, Black, All)” and then click “Column” to add this variable as a new column variable for our output table.

Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

Page

Row

Column

Race recode (White, Black, All)

Available Variables

Age at Diagnosis

Age recode with &lt;1 year olds

Race, Sex, Year Dx, Registry, County

Site and Morphology

Stage - AJCC

Stage - TNM

Stage - LRD (Summary and Historic)

Therapy

Extent of Disease - CS

Extent of Disease - Historic

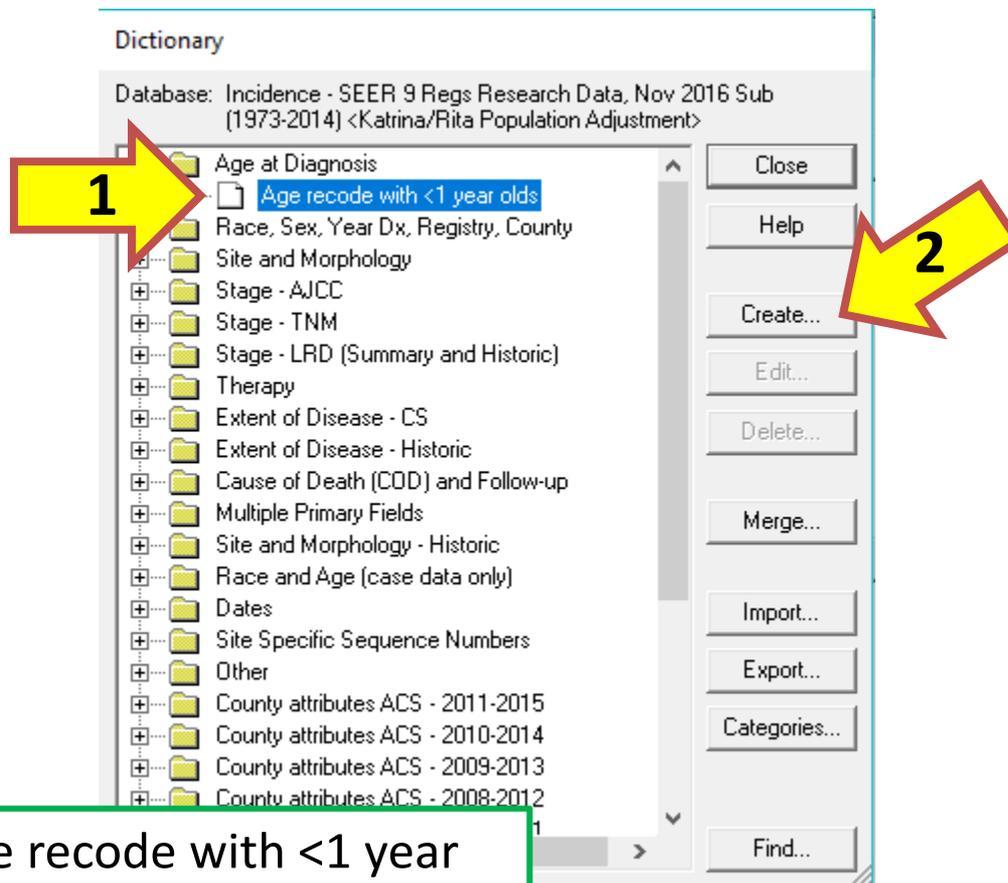
Cause of Death (COD) and Follow-up

Multiple Primary Fields

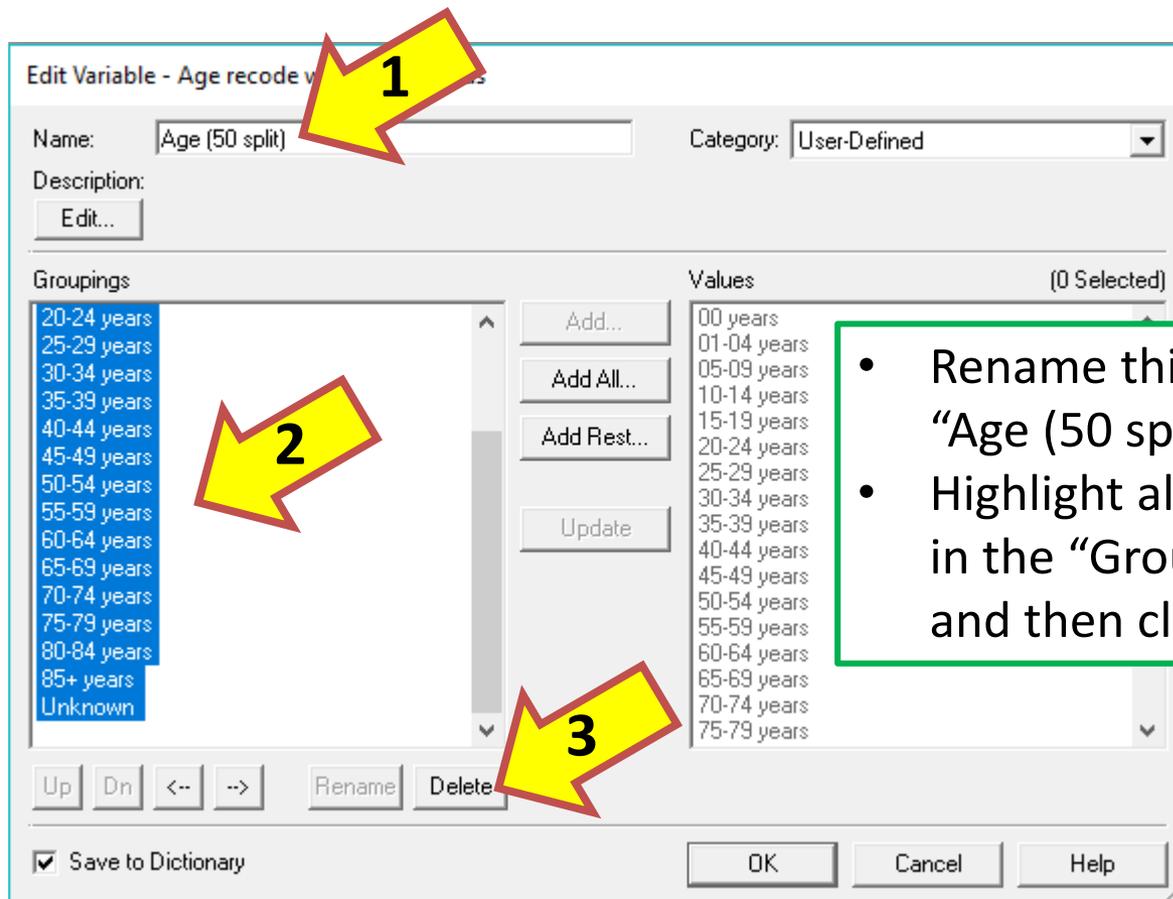
1

2

- Notice that Race was added as a column variable in the Display Variables box.
- Next, we'll want to add our age variable.
- As on the SEER website table we're modeling this exercise off of, we want to display age in three groups:
  - All ages
  - < 50 years
  - 50+ years
- To do this, we'll have to create a new variable grouping.
- Double click on the "Age recode with <1 year olds" in the "Age at Diagnosis" folder.



Make sure “Age recode with <1 year olds” is selected and then click “Create...”.



**Edit Variable - Age recode with <1 year olds**

Name: Age (50 split) Category: User-Defined

Description: Edit...

Groupings

Values (19 Selected)

- 10-14 years
- 15-19 years
- 20-24 years
- 25-29 years
- 30-34 years
- 35-39 years
- 40-44 years
- 45-49 years
- 50-54 years
- 55-59 years
- 60-64 years
- 65-69 years
- 70-74 years
- 75-79 years
- 80-84 years
- 85+ years
- Unknown

Buttons: Add... Add All... Add Rest...

Buttons: Up Dn <- -> Rename Delete

Save to Dictionary

Buttons: OK Cancel Help

Find...

- Next, highlight all of the values in the “Values” box, except for the “Unknown” value.
- Then click the “Add...” button.

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

# Σ % P ÷ IT [Grid] [Folder] [Save] [Lightning] [Book] [Question] Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERS

Edit Variable - Age recode with <1 year olds

Name: Age (50 split) Category: User-Defined

Description: Edit...

Groupings Values (19 Selected)

10-14 years  
15-19 years  
20-24 years  
25-29 years  
30-34 years  
35-39 years  
40-44 years  
45-49 years  
50-54 years  
55-59 years  
60-64 years  
65-69 years  
70-74 years  
75-79 years  
80-84 years  
85+ years  
Unknown

Add Selected Values

? The selected values will be:

Added as one grouping (all values combined)

Added as separate groupings (one for each value)

Use Underlying Values Instead of Labels For Grouping Names

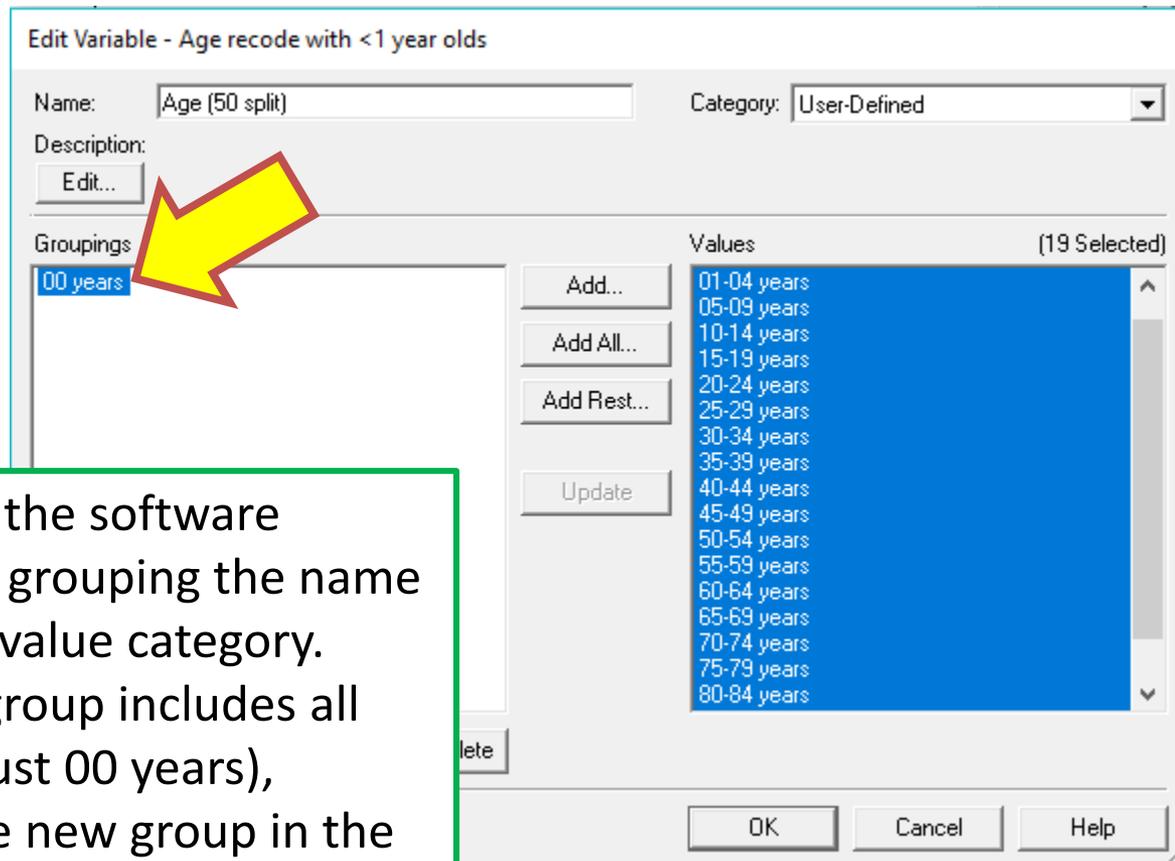
OK Cancel Help

1

2

OK Cancel Help Find... Find...

- When prompted, select “Added as one grouping”, since we want all ages to be a single group.
- Click “OK”.



- By default, the software names this grouping the name of the first value category.
- Since this group includes all ages (not just 00 years), rename the new group in the Groupings box “All ages”.
- If the name is not editable, click it once to highlight it, then click again to type.

Edit Variable - Age recode with <1 year olds

Name: Age (50 split) Category: User-Defined

Description:  
Edit...

Groupings  
All Ages

Add...  
Add All...  
Add Rest...  
Update

Values (11 Selected)

00 years
01-04 years
05-09 years
10-14 years
15-19 years
20-24 years
25-29 years
30-34 years
35-39 years
40-44 years
45-49 years
50-54 years
55-59 years
60-64 years
65-69 years
70-74 years
75-79 years

OK Cancel Help



- Next, we want to add a new grouping that contains all ages less than 50 years old.
- Select all values between and including “00 years” and “45-49 years” in the Values box.
- Click “Add”.

As with the previous group creation, select “Added as one grouping” and click “Ok”.

Edit Variable - Age recode with <1 year olds

Name: Age (50 split) Category: User-Defined

Description: Edit...

Groupings: All Ages Add... Values: (11 Selected)

00 years  
01-04 years

70-74 years  
75-79 years

**1** Add Selected Values

The selected values will be:

- Added as one grouping (all values combined)
- Added as separate groupings (one for each value)

Use Underlying Values Instead of Labels For Grouping Names

OK Cancel Help

Up Dn <- -> Rename Delete

Save to Dictionary OK Cancel Help

Edit Variable - Age recode with < 1 year olds

Name: Age (50 split) Category: User-Defined

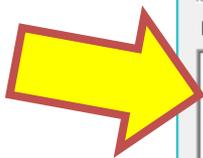
Description:  
Edit...

Groupings	Values (11 Selected)
All ages	00 years
<50 Years	01-04 years
	05-09 years
	10-14 years
	15-19 years
	20-24 years
	25-29 years
	30-34 years
	35-39 years
	40-44 years
	45-49 years
	50-54 years
	55-59 years
	60-64 years
	65-69 years
	70-74 years
	75-79 years

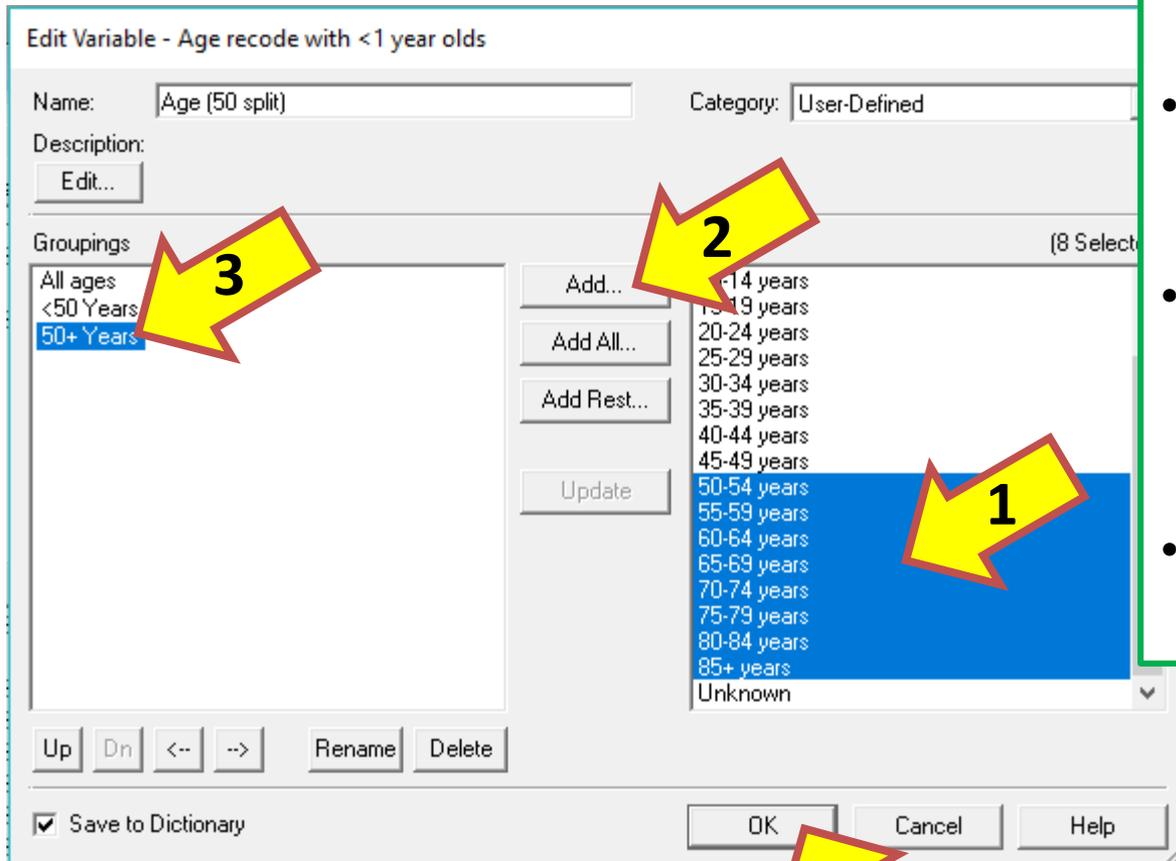
Buttons: Add..., Add All..., Add Rest..., Update

Save to Dictionary

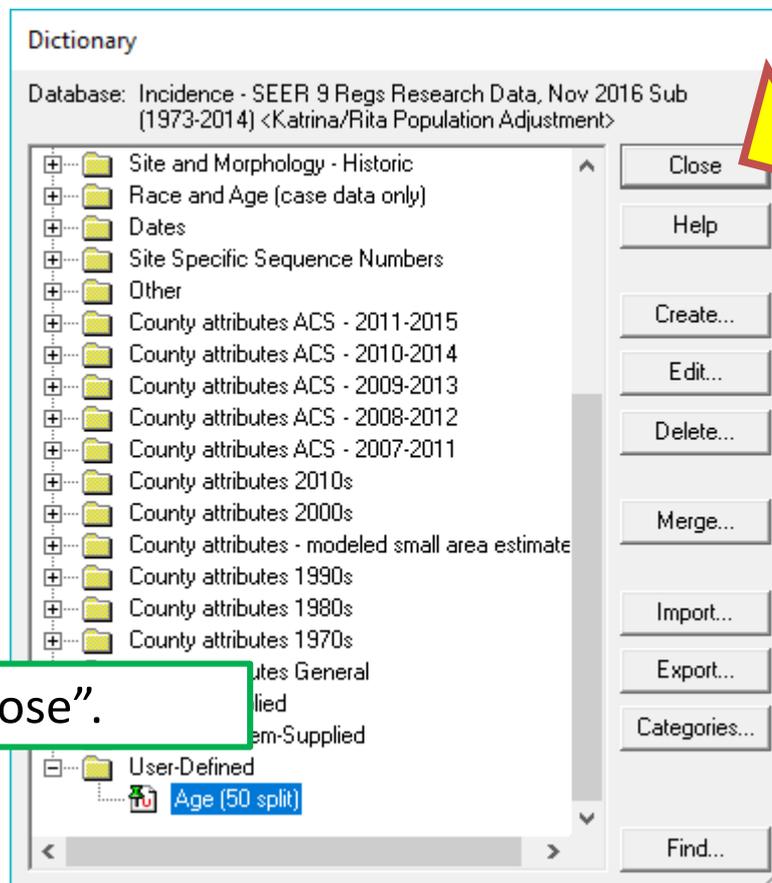
Buttons: OK, Cancel, Help



Change the new grouping name from "00 years" (default) to "< 50 years".



- Follow the previous steps once more, selecting “50-54 years” to “85+ years” (exclude the Unknown group).
- Click “Add...”, select “Added as one grouping”, click “Ok”
- Change the name for this grouping from “50-54 years” (default) to “50+ years”.
- Once you’re finished, click “OK”.



Click "Close".

Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

- Page
- Row
- Column
  - Race recode (White, Black, All)

Available Variables

- County attributes 2000s
- County attributes - modeled s
- County attributes 1990s
- County attributes 1980s
- County attributes 1970s
- County attributes General
- System-Supplied
- Merged System-Supplied
- User-Defined
  - Age (50 split)
  - Race recode (White, Black, All)

- Find...

Move Up

Move Down

Remove

Page

Row

Column

1

2

3

- Find the “User-Defined” folder in the “Available Variables” box.
- Find and select your newly defined variable grouping “Age (50 split)” and click the “Column” button.

## Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

Page

Row

Column

Race recode (White, Black, All)

Age (50 split)

Available Variables

1

Race, Sex, Year Dx, Registry, County  
Race recode (White, Black, Other)  
Sex  
Year of diagnosis  
SEER registry  
Louisiana 2005 - 1st  
County  
State-county  
In research data  
CHSDA 2012  
CHSDA Region  
State

2

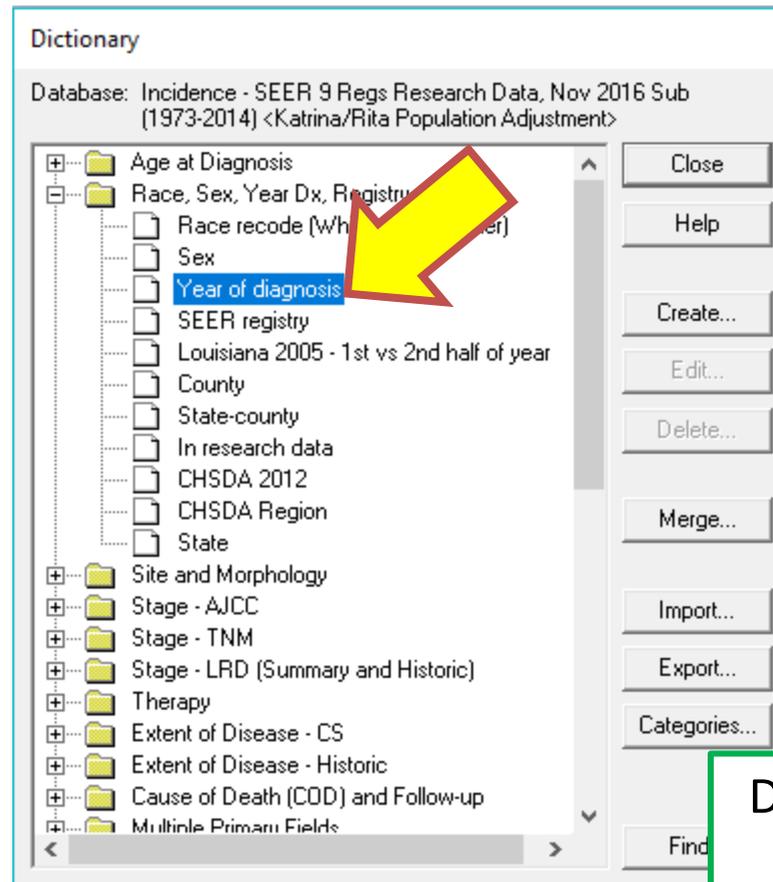
Page

Row

Column

Find...

- Notice that our new, user-defined age variable was added as a column variable in the top box.
- Next we'll add year of diagnosis as a row variable, but we want to redefine it so it only includes years of interest (i.e. 2000 – 2014).
- Within the “Race, Sex, Year Dx, Registry, County” folder, double click “Year of diagnosis”.



Double click on “Year of diagnosis” in the Dictionary pop-up.

**Edit Variable - Year of diagnosis**

Name: Year of diagnosis Category: User-Defined

Description: NAACCR Name=Date of Diagnosis--Year, Item #=390

Groupings

1988	1997
1989	1998
1990	1999
1991	2000
1992	2001
1993	2002
1994	2003
1995	2004
1996	2005
1997	2006
1998	2007
1999	2008
2000	2009
2001	2010
2002	2011
2003	2012
	2013

Values (0 Selected)

1

2

- Select all years prior to 1999 in the Groupings box, including the first category listed (i.e. “1973-2014”).
- Then click “Delete” to exclude these earlier years from the output table.

Up Dn <- -> Rename Delete

Save to Dictionary

OK Cancel Help

Age (in years) Find...

The screenshot shows the 'Edit Variable - Year of diagnosis' dialog box in SEER\*Stat 8.3.4. The 'Name' field is set to 'Year of diagnosis (2000+)' and the 'Category' is 'User-Defined'. The 'Description' is 'NAACCR Name=Date of Diagnosis--Year, Item #=390'. The 'Groupings' list contains years from 2000 to 2014. The 'Values' list contains years from 1997 to 2013. A yellow arrow labeled '1' points to the 'Name' field. A yellow arrow labeled '2' points to the 'OK' button. A green box contains instructions for renaming the variable and verifying the groupings.

**1**

Name: Year of diagnosis (2000+) Category: User-Defined

Description: NAACCR Name=Date of Diagnosis--Year, Item #=390

Groupings

Groupings	Values
2000	1997
2001	1998
2002	1999
2003	2000
2004	2001
2005	2002
2006	2003
2007	2004
2008	2005
2009	2006
2010	2007
2011	2008
2012	2009
2013	2010
2014	2011

Up Dn << >> Rename Delete

Save to Dictionary

OK Cancel Help

Age (5u spirs)

**2**

- Rename the “Name” of the variable to “Year of diagnosis (2000+)”.
- Verify that the only years you have left in the “Groupings” box are 2000 to 2014.
- Then click “OK”.

Rate Session-1

Data | Statistic

Display Variables

Page  
Row  
Column

Race recod  
Age (50 spli

Available Variables

Race, Sex  
Race  
Sex  
Year  
SEER  
Louis  
Coun

CHST  
CHST  
State

Dictionary

Database: Incidence - SEER 9 Regs Research Data, Nov 2016 Sub (1973-2014) <Katrina/Rita Population Adjustment>

Close

Help

Create...

Edit...

Delete...

Merge...

Import...

Export...

Categories...

Find...

Move Up

Move Down

Row

Column

Find...

1

2

- Notice that your new variable definition has been added to the dictionary.
- Then click “Close”.

CAPS NUM

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

# Σ % P ÷ IT [Grid] [Folder] [Save] [Lightning Bolt] [Book] [Question Mark] Local Data: C:\Users\... \Desktop\SEER\_1973\_2014\_SEERSTAT\...

Rate Session-1

Data | Statistic | Selection | **Table** | Output

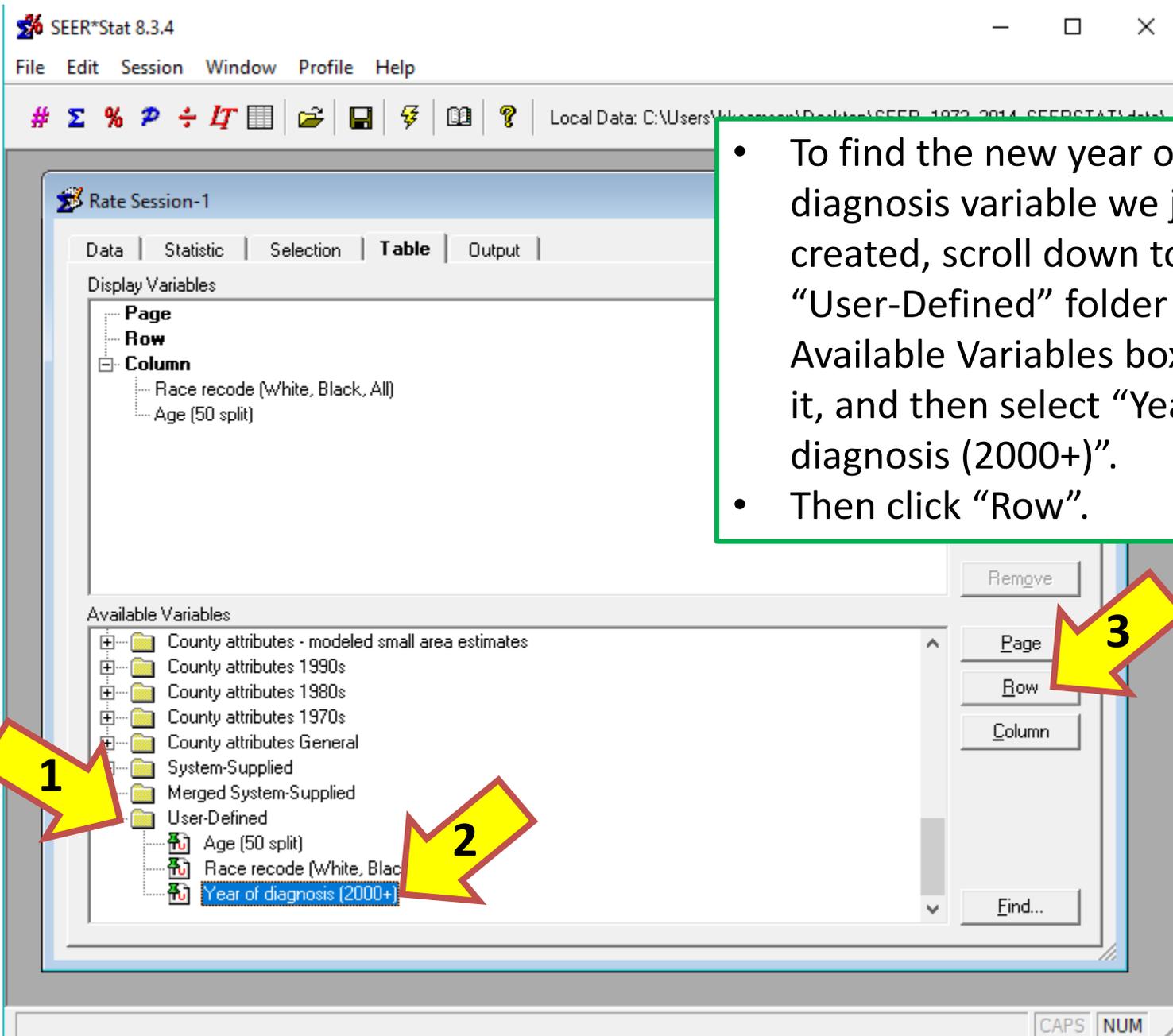
Display Variables

- Page
- Row
- Column
  - Race recode (White, Black, All)
  - Age (50 split)

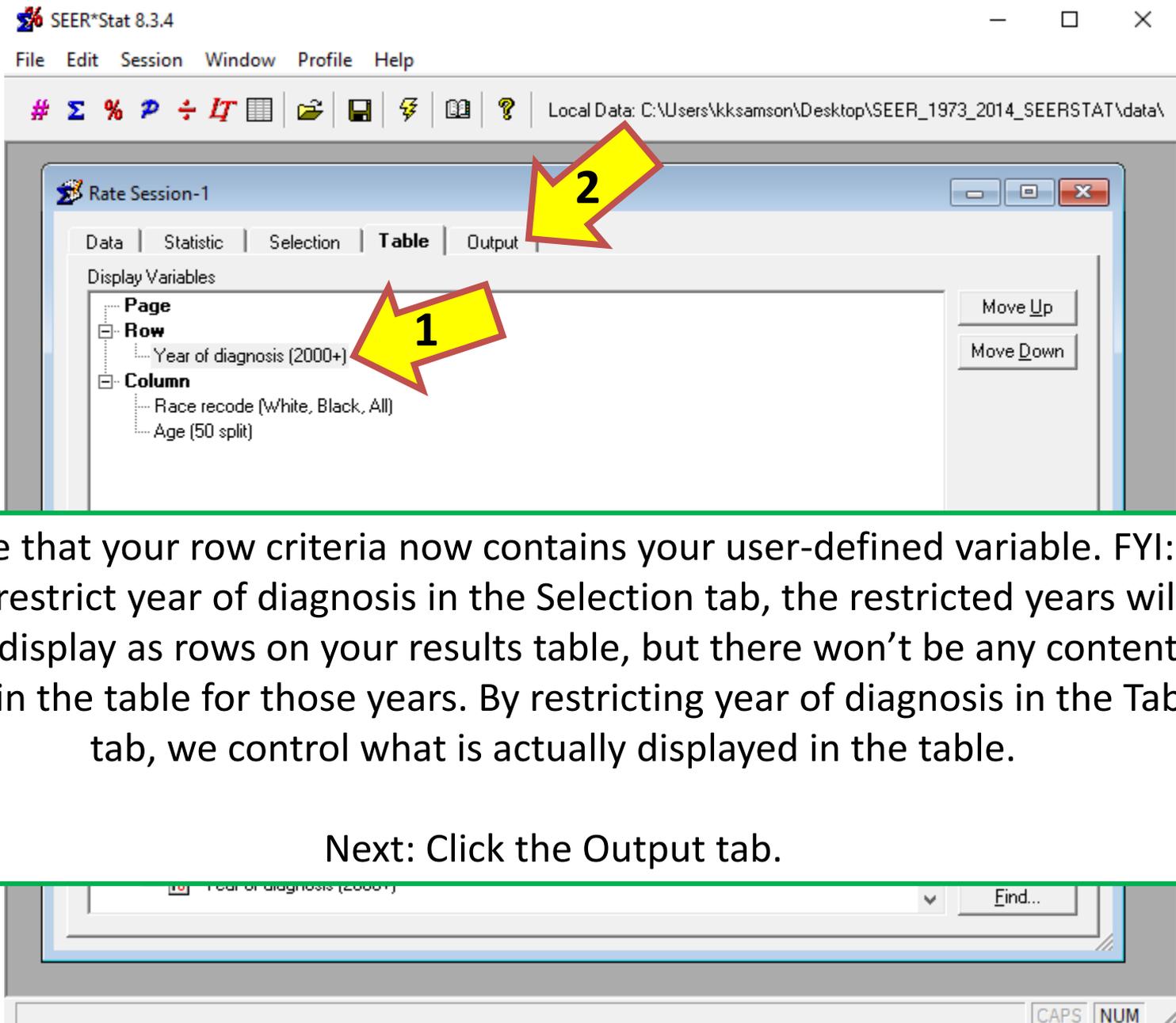
Available Variables

- County attributes - modeled small area estimates
- County attributes 1990s
- County attributes 1980s
- County attributes 1970s
- County attributes General
- System-Supplied
- Merged System-Supplied
- User-Defined
  - Age (50 split)
  - Race recode (White, Black)
  - Year of diagnosis (2000+)**

Remove Page Row Column Find...

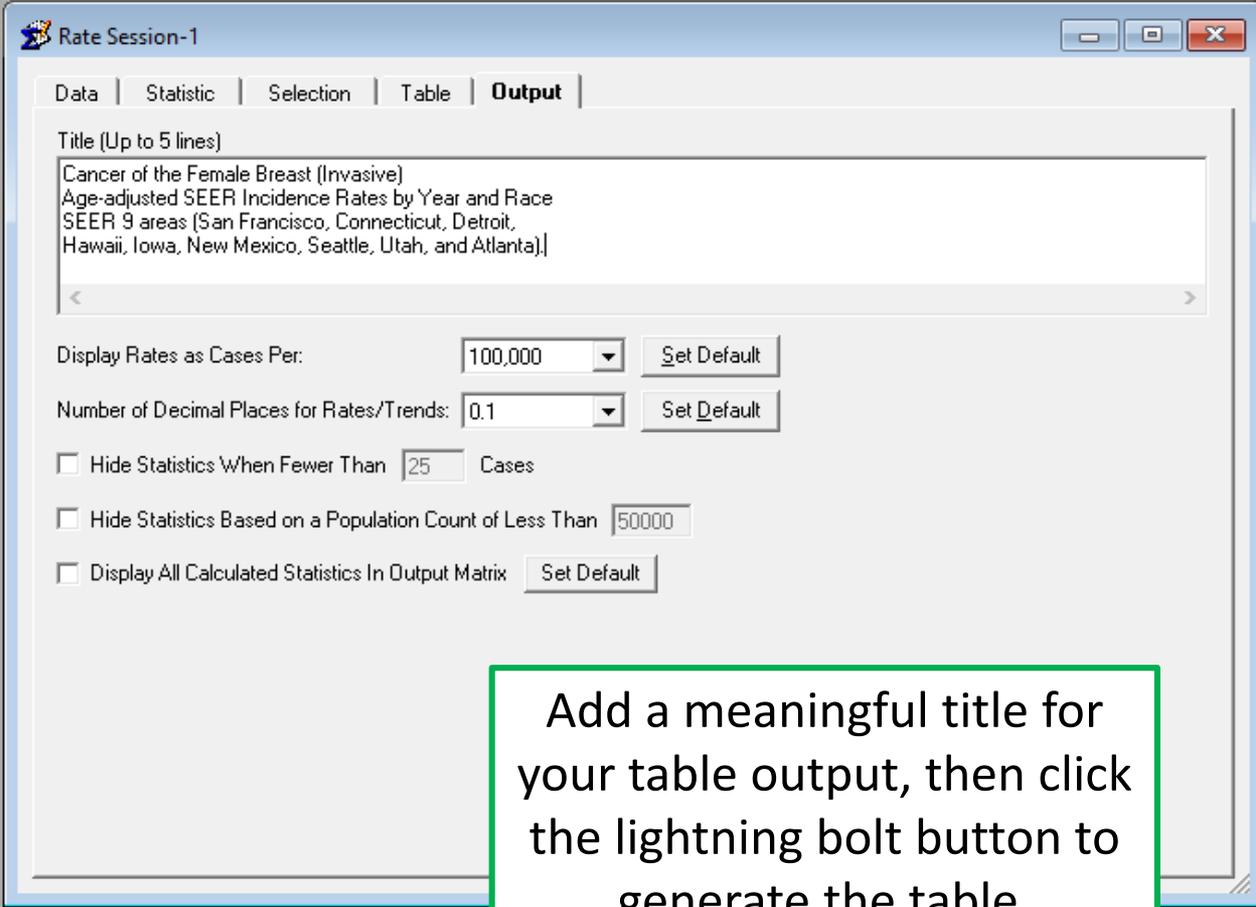
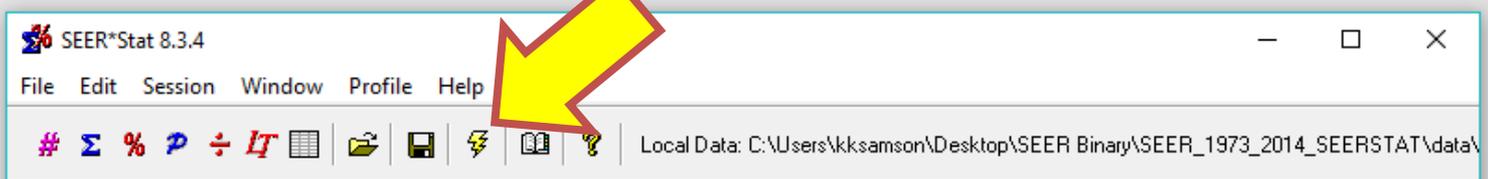


- To find the new year of diagnosis variable we just created, scroll down to the “User-Defined” folder in the Available Variables box, click it, and then select “Year of diagnosis (2000+)”.
- Then click “Row”.



Notice that your row criteria now contains your user-defined variable. FYI: If you restrict year of diagnosis in the Selection tab, the restricted years will still display as rows on your results table, but there won't be any content listed in the table for those years. By restricting year of diagnosis in the Table tab, we control what is actually displayed in the table.

Next: Click the Output tab.



Add a meaningful title for your table output, then click the lightning bolt button to generate the table.

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

# Σ % P ÷ IT [Grid] [Folder] [Save] [Lightning] [Book] [Question] Local Data: C:\Users\kksamson\Desktop\SEER Binary\SEER\_1973\_2014\_SEERSTAT\data\

**Variable Warning** [X]

 **WARNING!**  
Summary stage 2000 (1998+) <Selection>  
Caution should be exercised when using this variable.  
Press OK to continue.  
[For More Information, Click Here](#)

OK Cancel

Rate Selection

Data

Title (Up

Cancer

Age-ad

SEER

Hawai

Display Rate: 100,000 [Set Default]

Number of Decimal Places for Rates/Trends: 0.1 [Set Default]

Hide Statistics When Fewer Than 25 Cases

Hide Statistics Based on a Population Count of Less Than 50000

Display All Calculated Statistics In Output Matrix [Set Default]

- A warning comes up regarding the stage variable we used to exclude in situ cases.
- Click the blue text to read what this warning refers to.
- Then click “OK”.



Rate Session-1

Rate Session-1 Matrix-5

Cancer of the Female Breast (Invasive)

	All races						
	All ages					<50 Year:	
	Rate	SE	Lower CI	Upper CI	Count	Pop	Rate
2000	136.6	1.0	134.6	138.5			
2001	138.8	1.0	136.9	140.8			
2002	135.9	1.0	134.0	137.8			
2003	127.1	0.9	125.3	129.0			
2004	128.2	0.9	126.4	130.0			
2005	126.7	0.9	124.9	128.5			
2006	126.4	0.9	124.6	128.2			
2007	128.3	0.9	126.5	130.1			
2008	128.5	0.9	126.7	130.3			
2009	130.9	0.9	129.1	132.7			
2010	127.0	0.9	125.2	128.7			
2011	130.4	0.9	128.6	132.2			
2012	130.1	0.9	128.3	131.8			
2013	130.8	0.9	129.1	132.6			
2014	132.8	0.9	131.0	134.6			

Check out the similarities of the rates in this table to the one found here (note that the online table doesn't have confidence intervals or standard errors):  
[https://seer.cancer.gov/csr/1975\\_2014/browse\\_csr.php?sectionSEL=4&pageSEL=sect\\_04\\_table\\_08.html#a](https://seer.cancer.gov/csr/1975_2014/browse_csr.php?sectionSEL=4&pageSEL=sect_04_table_08.html#a)

Keep this session open if you would like to practice formatting and exporting tables for use with Joinpoint software.

# Reformatting Incidence Output and Exporting Files for Use in Joinpoint

- **To run an example in Joinpoint, we will use the data from the previous age-adjusted incidence rate session from this tutorial.**
- **While the table in the previous example is straightforward and easy to read, it cannot be used in Joinpoint software. Therefore, we will restructure the table so that we can use it.**





Rate Session-1

Rate Session-1 Matrix-5

Cancer of the Female Breast (Invasive)

	All races						
	All ages						<50 Year:
	Rate	SE	Lower CI	Upper CI	Count	Pop	Rate
2000	136.6	1.0	134.6	138.5	18,965	13,595,589	43.5
2001	138.8	1.0	136.8	140.8	19,574	13,705,794	44.0
				137.8	19,451	13,781,716	42.9
				129.0	18,484	13,856,312	43.5
				130.0	18,937	13,942,122	45.0
				128.5	19,037	14,045,528	44.1
				128.2	19,355	14,190,920	44.2
				130.1	19,954	14,325,687	45.4
				130.3	20,307	14,465,662	45.6
				132.7	21,118	14,606,190	45.5
				128.7	20,829	14,727,666	44.2
				132.2	21,728	14,851,045	45.5
2012	130.1	0.9	128.3	131.8			
2013	130.8	0.9	129.1	132.6			
2014	130.6	0.9	128.8	132.3			

To be able to utilize this output in Joinpoint software to be able to assess trends over time, we need to slightly restructure the table.

Close out of the Matrix table.

< >  
 ~ Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence Interval could not be calculated.

Rate Session-1

Data | **Statistic** | Selection | Table | **Print**

Title (Up to 5 lines)

Cancer of the Female Breast (Invasive) 2000 - 2014  
Age-Adjusted SEER Incidence Rates by Year and Race  
SEER 9 areas (San Francisco, Connecticut, Detroit,  
Hawaii, Iowa, New Mexico, Seattle, Utah, and Atlanta).

Display Rates as Cases Per:

100,000

Set Default

Number of Decimal Places for Rates/Trends:

0.1

Set Default

 Hide Statistics When Fewer Than 25 Cases Hide Statistics Based on a Population Count of Less Than 50000 Display All Calculated Statistics In Output Matrix Set Default

Click on the Table tab.

Local Data: C:\Users\kksamson\Desktop\SEER\_1973\_2014\_SEERSTAT\data\

Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

- Page
- Row
  - Year of diagnosis (2000+)
- Column
  - Race recode (White, Black, All)
  - Age (50 split)

Remove Up

Remove

Page

Row

Available Variables

- County attributes - modeled small area estimates
- County attributes 1990s
- County attributes 1980s
- County attributes 1970s

In general, we will need all variables listed as Row variables, but the year of diagnosis variable should be listed last. For details on table construction for use in Joinpoint, check out this website:  
<https://surveillance.cancer.gov/help/joinpoint/setting-parameters/input-file-tab/input-data-file/using-seer-stat-to-create-joinpoint-input-data-files>

Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

- Page
- Row
- Column
  - Race recode (White, Black, All)**
  - Age (50 split)



Click on the “Race recode (White, Black, All)” in the Display Variables box and then click the “Remove” button.

Available Variables

- County attributes - modeled small area estimates
- County attributes 1990s
- County attributes 1980s
- County attributes 1970s
- County attributes General
- System-Supplied
- Merged System-Supplied
- User-Defined
  - Age (50 split)
  - Race recode (White, Black, All)
  - Year of diagnosis (2000+)

Remove

Page

Row

Column

Find...



Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

Page

Row

Column

Age (50 split)

1

Click on the "Age (50 split)" in the Display Variables box and then click the "Remove" button.

Remove

2

Available Variables

- County attributes - modeled small area estimates
- County attributes 1990s
- County attributes 1980s
- County attributes 1970s
- County attributes General
- System-Supplied
- Merged System-Supplied
- User-Defined
  - Age (50 split)
  - Race recode (White, Black, All)
  - Year of diagnosis (2000+)

Page

Row

Column

Find...

Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

Page

Row

Column

Move Up

Click on the "Race recode (White, Black, All)" in the Available Variables box and then click the "Row" button.

Available Variables

- ⊕ County attributes - modeled small area estimates
- ⊕ County attributes 1990s
- ⊕ County attributes 1980s
- ⊕ County attributes 1970s
- ⊕ County attributes General
- ⊕ System-Supplied
- ⊕ Merged System-Supplied
- ⊖ User-Defined
  - Age (50 split)
  - Race recode (White, Black, All)**
  - Year of diagnosis (2000+)

Remove

Page

Row

Column

Find...



Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

- Page
- Row
  - Race recode (White, Black, All)
- Column

Available Variables

- County attributes - modeled small area estimates
- County attributes 1990s
- County attributes 1980s
- County attributes 1970s
- County attributes General
- System-Supplied
- Merged System-Supplied
- User-Defined
  - Age (50 split)**
  - Race recode (White, Black, All)
  - Year of diagnosis (2000+)

Remove

Page

Row

Column

Find...

Click on the "Age (50 split)" in the Available Variables box and then click the "Row" button.



Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

Page

Move Up

Row

Race recode (White, Black, All)

Age (50 split)

Column

Finally, click on the “Year of diagnosis (2000+)” in the Available Variables box and then click the “Row” button.

Available Variables

- [-] County attributes - modeled small area estimates
- [-] County attributes 1990s
- [-] County attributes 1980s
- [-] County attributes 1970s
- [-] County attributes General
- [-] System-Supplied
- [-] Merged System-Supplied
- [-] User-Defined
  - [-] Age (50 split)
  - [-] Race recode (White, Black)
  - [-] **Year of diagnosis (2000+)**

Remove

Page

Row

Column

Find...

1

2

SEER\*Stat 8.3.4

File Edit Session Window Profile Help

# Σ % P ÷ IT Local Data: C:\Users\kksamson\Desktop\SEER\_1973\_2014\_SEERSTAT\data\

Rate Session-1

Data | Statistic | Selection | **Table** | Output

Display Variables

- Page
- Row
  - Race recode (White, Black, All)
  - Age (50 split)
  - Year of diagnosis (2000+)
- Column

Available Variables

- County attributes - modeled small area estimates
- County attributes 1990s
- County attributes 1980s
- County attributes 1970s
- County attributes General
- System-Supplied
- Merged System-Supplied
- User-Defined
  - Age (50 split)
  - Race recode (White, Black, All)
  - Year of diagnosis (2000+)

Move Up

Move Down

Remove

Page

Row

Find...

CAPS NUM

Rate Session-1

Data | Statistic | Selection | Table | **Output**

Title (Up to 5 lines)

Cancer of the Female Breast (Invasive) 2000 - 2014 (Joinpoint Table Structure)  
Age-Adjusted SEER Incidence Rates by Year and Race  
SEER 9 areas (San Francisco, Connecticut, Detroit,  
Hawaii, Iowa, New Mexico, Seattle, Utah, and Atlanta).

Display Rates as Cases Per:

100,000

Set Default

Number of Decimal Places for Rates/Trends:

0.1

Set Default

 Hide Statistics When Fewer Than 25 Cases Hide Statistics Based on a Population Count of Less Than 50000 Display All Calculated Statistics In Output Matrix Set Default

Modify the table title if you would like, then hit the lightning bolt button and click "OK" on the warning pop-up.

SEER\*Stat 8.3.4

File Edit Matrix Window Profile Help

# % P IT

Session Statistics

(...ive) 2000 - 2014 (Joinpoint Table Structure)

			Rate	SE	Lower CI	Upper CI	Count	Pop
All races	All ages	2000	136.6	1.0	134.6	138.5	18,965	13,595,589
		2001	138.8	1.0	136.9	140.8	19,574	13,705,794
		2002	135.9	1.0	134.0	137.8	19,451	13,781,716
		2003	127.1	0.9	125.3	129.0	18,484	13,856,312
		2004	128.2	0.9	126.4	130.0	18,937	13,942,122
		2005	126.7	0.9	124.9	128.5	19,037	14,045,528
		2006	126.4	0.9	124.6	128.2	19,355	14,190,920
		2007	128.3	0.9	126.5	130.1	19,954	14,325,687
		2008	128.5	0.9	126.7	130.3	20,307	14,465,662
		2009	130.9	0.9	129.1	132.7	21,118	14,606,190
		2010	127.0	0.9				
		2011	130.4	0.9				
		2012	130.1	0.9				
		2013	130.8	0.9				
2014	130.6	0.9						
	< 50 Years	2000	43.5	0.7				
		2001	44.0	0.7				
		2002	42.9	0.7				
		2003	43.5	0.7				
		2004	45.0	0.7				
		2005	44.1	0.7	42.8	45.5	4,447	9,744,772
		2006	44.2	0.7	42.9	45.5	4,459	9,778,410
		2007	45.4	0.7	44.0	46.7	4,544	9,797,625
		2008	45.6	0.7	44.2	46.9	4,535	9,822,750

Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence interv:

CAPS NUM

Notice the new layout of the table. All of our variables of interest are in the Rows. This is the structure your data needs to be in to use it for Joinpoint analysis.

SEER\*Stat 8.3.4

File Edit **Matrix** Window Profile Help

Retrieve Session  
Options...  
Font...  
Order >  
Filter...  
Lock  
Page >  
**Export** >  
Properties...

Results as Text File...  
MP-SIR Rate...

2014 (Joinpoint Table Structure)

	SE	Lower CI	Upper CI	Count	Pop
36.6	1.0	134.6	138.5	18,965	13,595,589
38.8	1.0	136.9	140.6		
35.9	1.0	134.0	137.8		
129.0					
130.0					
128.5					
128.2					
130.1					
130.3					
126.4	0.9	126.4	126.4		
2007	128.3	0.9	128.3		
2008	128.5	0.9	128.5		
2009	130.9	0.9	129.1		
2010	127.0	0.9	125.2		
2011	130.4	0.9	128.6		
2012	130.1	0.9	128.3		
2013	130.8	0.9	129.1	132.6	22,522
2014	130.6	0.9	128.8	132.3	22,909
< 50 Years					
2000	43.5	0.7	42.2	44.8	4,321
2001	44.0	0.7	42.7	45.3	4,406
2002	42.9	0.7	41.6	44.2	4,316
2003	43.5	0.7	42.2	44.8	4,371
2004	45.0	0.7	43.7	46.3	4,528
2005	44.1	0.7	42.8	45.5	4,447
2006	44.2	0.7	42.9	45.5	4,459
2007	45.4	0.7	44.0	46.7	4,544
2008	45.6	0.7	44.2	46.9	4,535

Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence interv:

Export to a text file

To export this table to be able to use in Joinpoint software, go to “Matrix” at the top of the page, then “Export”, and then “Results as Text File...”.

**Export**

Data File (\*.txt) or Compressed Data File (\*.gz)

Export Dictionary File (\*.dic)

Generate SAS Code to Read Data (SAS File: \*.sas)

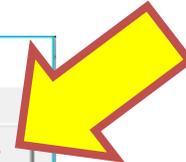
Output Variables as  
 Numeric Representation  
 Labels Enclosed in Quotes  
 Labels Without Quotes

Line Delimiter  
 DOS/Windows (CR/LF)  
 UNIX (LF)

Missing Character  
 Space  
 Period  
 "NA"

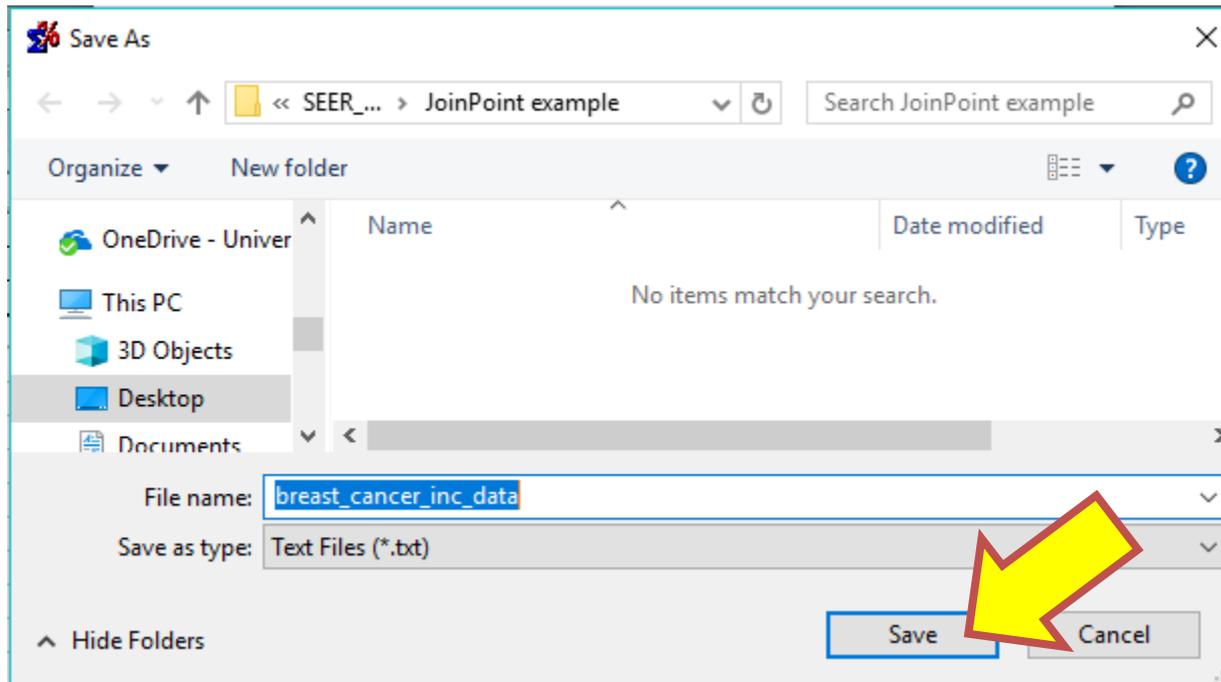
Field Delimiter  
 Tab  
 Space  
 Comma  
 Semi-Colon

Enclose Fields Containing Delimiter in Quotes  
 Remove All Thousands Separators (Commas)  
 Remove Flags (Footnote), Prefix and Suffix  
 Output Variable Names Before Data  
 Preserve Matrix Columns and Rename Fields  
 Include Full Path for Input File in SAS Code



- You will need to save two separate files: a data file and a dictionary file.
- Click the top “Browse...” button to locate a folder on your computer where you will save the data file.

Give the file a descriptive name and then click “Save”.



**Export**

Data File (\*.txt) or Compressed Data File (\*.gz)  
C:\Users\kksamson\Desktop\SEER\_1973\_2014\_SEERSTAT\JoinPoint example\br Browse...

Export Dictionary File (\*.dic)  
C:\Users\kksamson\Documents\export.dic Browse...

Generate SAS Code to Read Data (SAS File: \*.sas)  
Browse...

Output Variables as  
 Numeric Representation  
 Labels Enclosed in Quotes  
 Labels Without Quotes

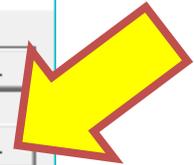
Line Delimiter  
 DOS/Windows (CR/LF)  
 UNIX (LF)

Missing Character  
 Space  
 Period  
 "NA"

Field Delimiter  
 Tab  
 Space  
 Comma  
 Semi-Colon

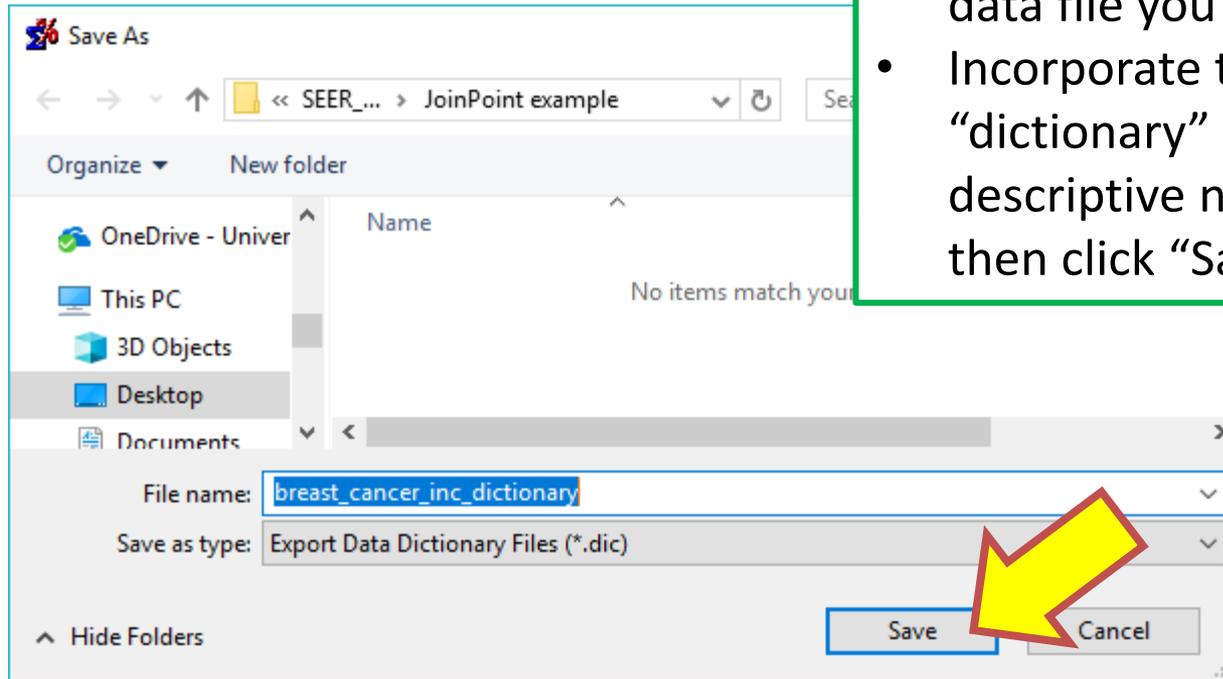
Enclose Fields Containing Delimiter in Quotes  
 Remove All Thousands Separators (Commas)  
 Remove Flags (Footnote), Prefix and Suffix Ch  
 Output Variable Names Before Data  
 Preserve Matrix Columns and Rename Fields  
 Include Full Path for Input File in SAS Code

Defaults CSV Defaults Set Default OK Cancel Help



Next, you'll save the dictionary file. Click the second "Browse..." button.

- Save the dictionary file in the same folder as the data file you just saved.
- Incorporate the word “dictionary” in the descriptive name and then click “Save”.



Export

Data File (\*.txt) or Compressed Data File (\*.gz)  
 \sktop\SEER\_1973\_2014\_SEERSTAT\JoinPoint example\breast\_cancer\_inc

Export Dictionary File (\*.dic)  
 \SEER\_1973\_2014\_SEERSTAT\JoinPoint example\breast\_cancer\_inc\_dic

Generate SAS Code to Read Data (SAS File: \*.sas)

Output Variables as

Numeric Representation  
 Labels Enclosed in Quotes  
 Labels Without Quotes

Line Delimiter

DOS/Windows (CR/LF)  
 UNIX (LF)

Field Delimiter

Tab  
 Space  
 Semi-Colon

Enclose Fields Containing Delimiter in Quotes  
 Remove All Thousands Separators (Commas)  
 Remove Flags (Footnote), Prefix and Suffix Characters  
 Output Variable Names Before Data  
 Preserve Matrix Columns and Rename Fields  
 Include Full Path for Input File in SAS

Defaults CSV Defaults Set Default OK Cancel Help

2007	45.4	0.7	44.0	46.7	4,544	9,797,625
2008	45.6	0.7	44.2	46.9	4,535	9,822,750

Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence interv:

- Be sure to check “Remove Flags (Footnote), Prefix and Suffix Characters”.
- If you don’t select this option, your file won’t work in Joinpoint software.
- The click “Ok”.

# Downloading Joinpoint Software



University of Nebraska  
Medical Center

Download Joinpoint Desktop Software

https://surveillance.cancer.gov/joinpoint/download

NIH NATIONAL CANCER INSTITUTE  
Division of Cancer Control & Population Sciences

Print Email Glossary

Search SRP

## Surveillance Research Program

HOME RESEARCH AREAS **METHODS & TOOLS** CANCER STATISTICS PUBLICATIONS FUNDING & GRANTS ABOUT

### Joinpoint:

- Joinpoint Home
- Download Joinpoint Desktop Software**
- Application for Command-Line Version
- Joinpoint Revision History
- Sample Joinpoint Analyses
- Frequently Asked Questions
- Conference Presentations
- Getting Help
- Jump and Comparability Ratio Models
  - Overview
  - Melanoma Example
  - CSR Alternative Tables

### Related Materials:

- Methods & Tools for Population-based Cancer Statistics

## Download Joinpoint Desktop Software

Thank you for your interest in using the Joinpoint statistical software for the analysis of continuous linear trends with change points, i.e. joinpoints. This version of the software is supplied with a graphical user interface for use interactively. Joinpoint is a Microsoft Windows application and can only be used on a Mac in conjunction with a Windows emulator.

### Version 4.5.0.1 (released June 12, 2017)

#### Bug Fixes

- ▶ As of version 4.5.0.0, users can set the Alpha levels for the Joinpoint Local Comparability Ratio. Upon installation of that version, the default values for the Alpha levels should have been 0.05. **Installing the new version of Joinpoint requires setting to 0.05.**
- ▶ The significance level footnote on the Model Selection tab referenced the

#### New Features

- ▶ The value 100 has been added to the list of possible “Rates per” (population Adjusted Rates).

Before downloading **Joinpoint version 4.5.0.1**, you must read and sign the Terms of Use Agreement below and complete the registration form. After completing the form you will receive an e-mail with a link to the software download. The download link may be used once and then it will expire. If your link has expired, request a new one using the [Already Registered? form](#).

### Terms of Use Agreement for Joinpoint Statistical Software

Please review the following terms carefully.

Go to the following website to download Joinpoint software from the National Cancer Institute:  
<https://surveillance.cancer.gov/joinpoint/download>

Download Joinpoint Desktop Sc X

https://surveillance.cancer.gov/joinpoint/download

### Register for Software

The information below is needed to record acceptance of the Terms of Use agreement, to evaluate the types of software users in order to be responsive to user needs, and to notify users of updates. If you would like to be notified of updates, please indicate so below. **All fields are required.**

I agree to the Terms of Use Agreement

Email

Name

Organization

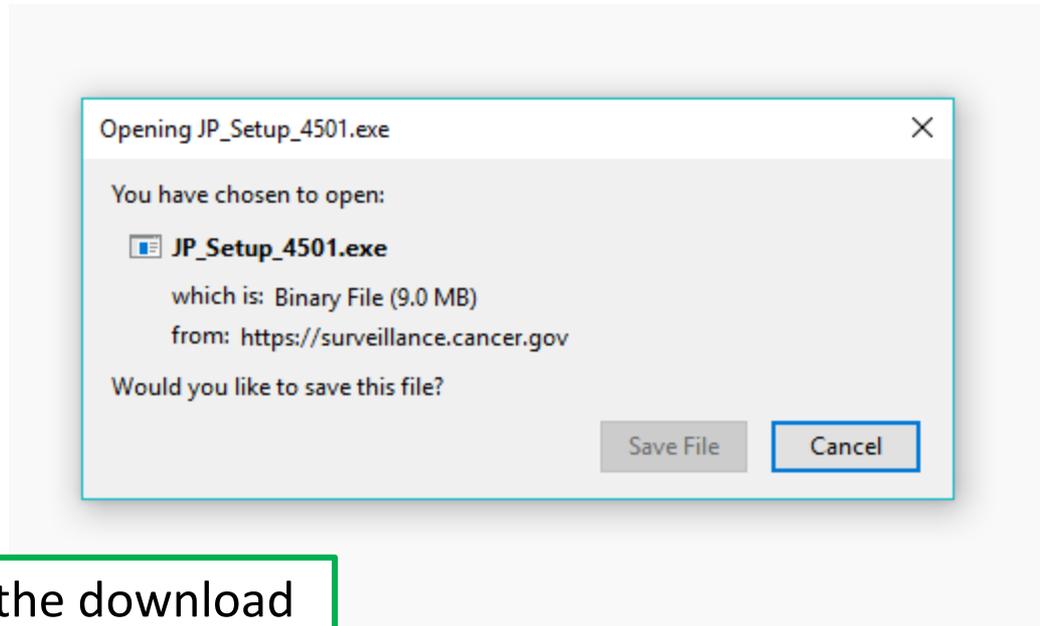
Type of Organization  
--- select ---

*IF OTHER, PLEASE SPECIFY (REQUIRED IF "OTHER" IS SELECTED ABOVE)*

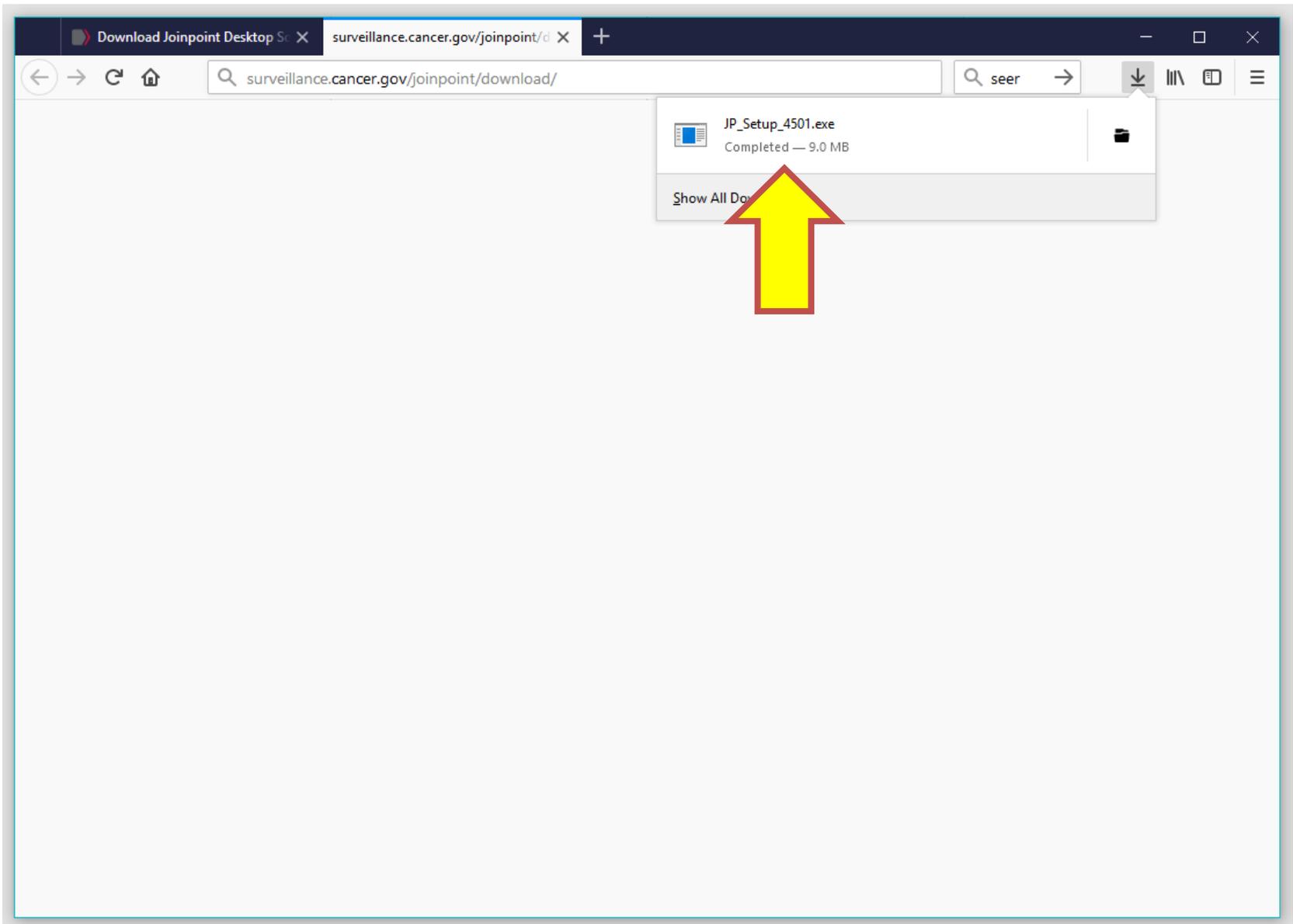
State/Province

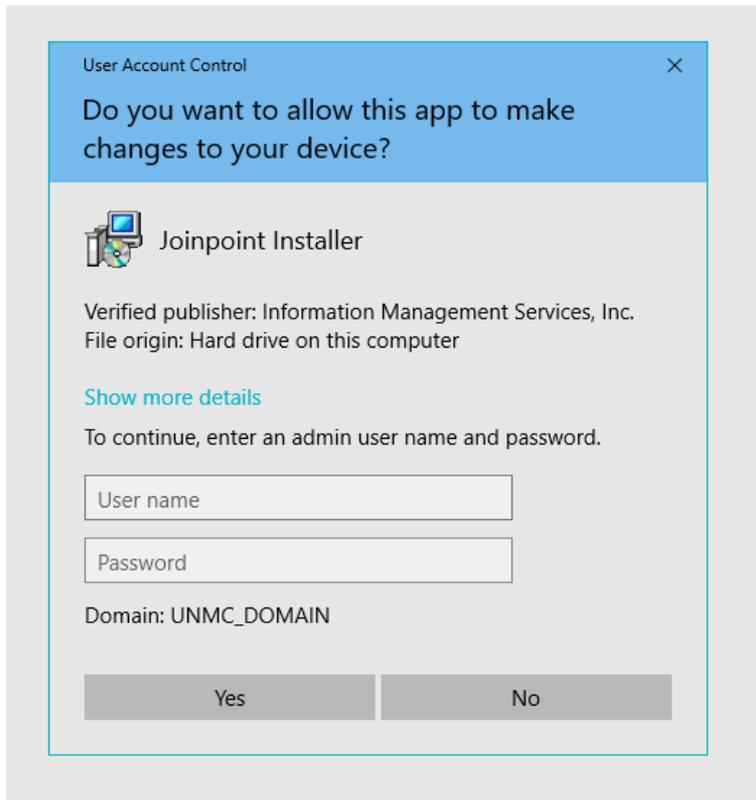
Country

Fill out the form on that same screen and wait for an email confirmation with a link to download the software.



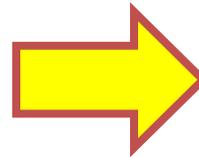
Click on the download link in your email.





- After clicking on the executable file, follow the prompts to download the software.
- You may need to contact your IT department to get the administrative rights to download executable files.

When the download has completed, you should have a Joinpoint shortcut on your desktop.



# Using Joinpoint software to detect significant changes in temporal trends

Example: Breast Cancer Age-Adjusted Incidence Rates



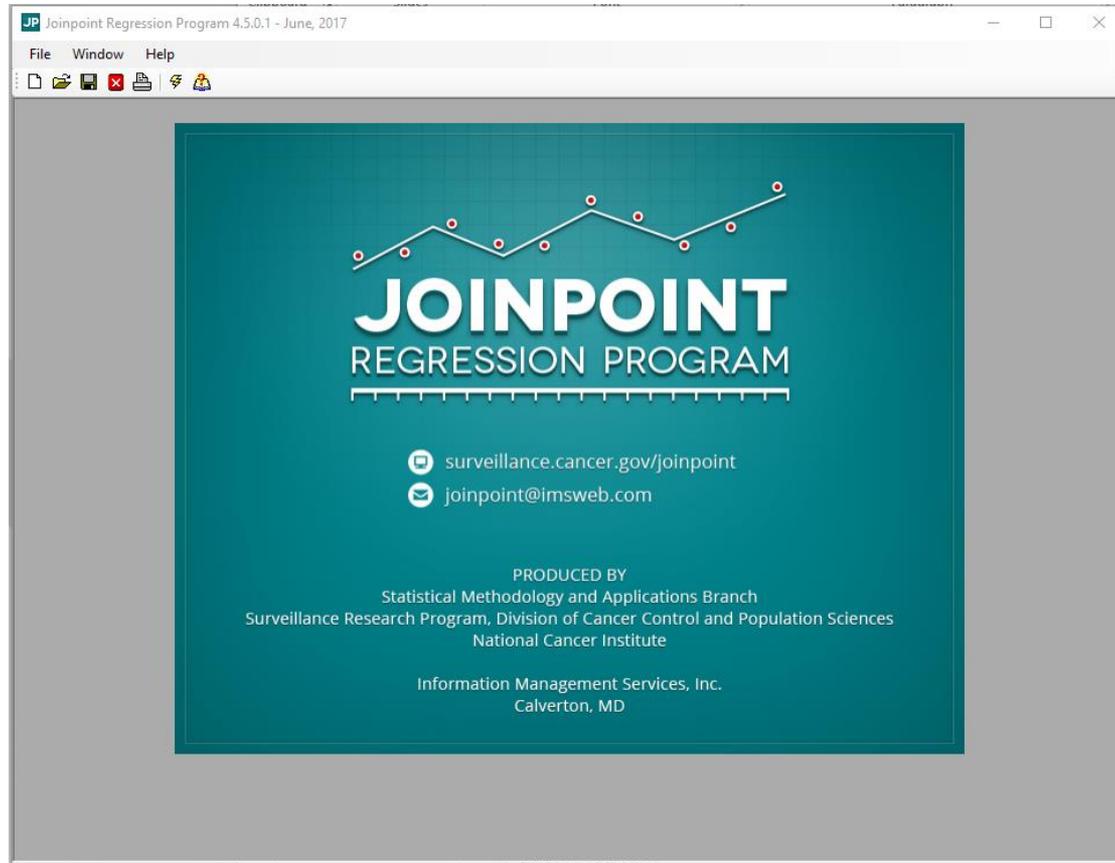
University of Nebraska  
Medical Center

# Using Joinpoint

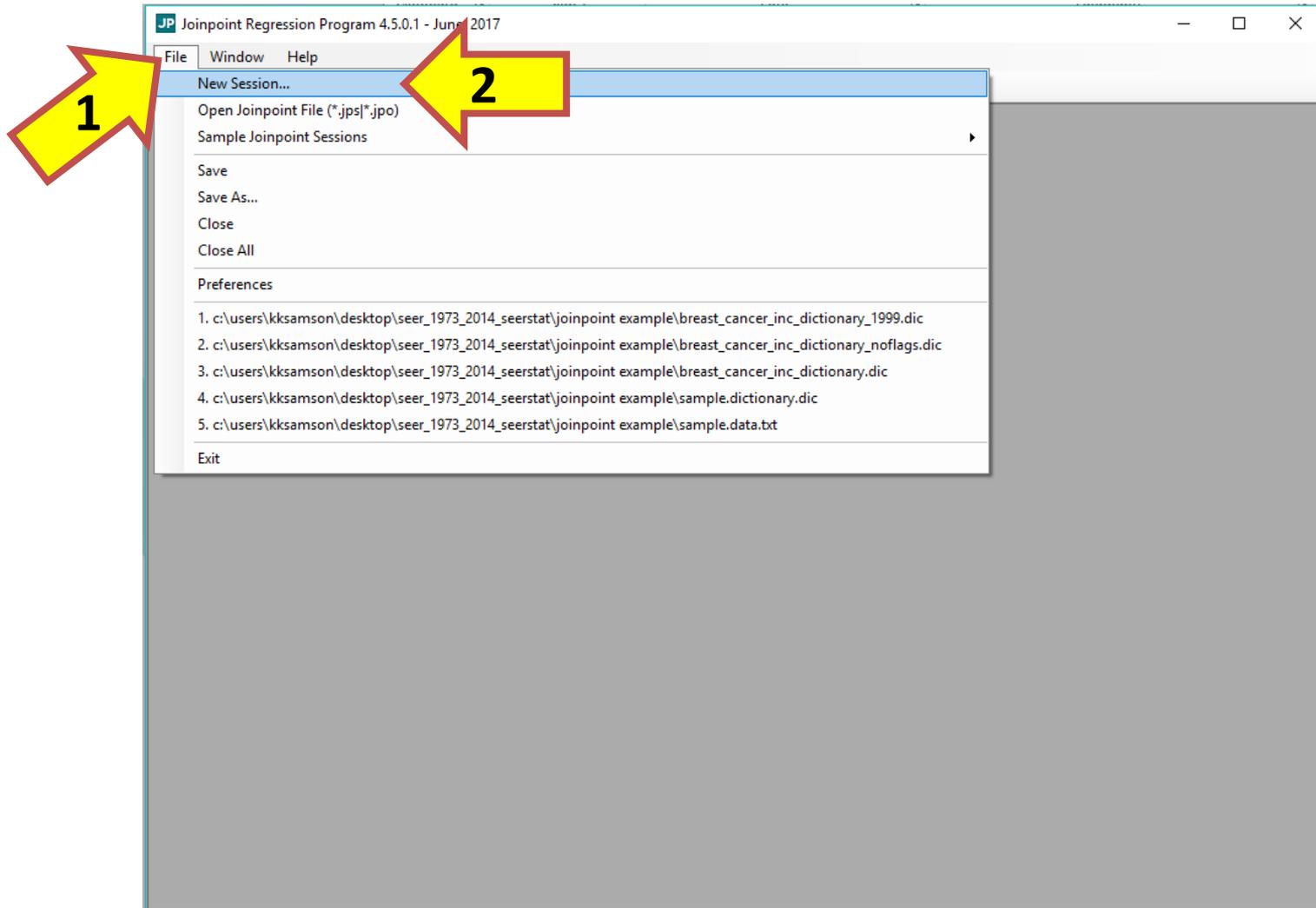
- The primary objective of this tutorial is showing you how to import SEER\*Stat results into Joinpoint and then run a session.
- There are many nuances and options within this software, and we encourage you to check out more information about Joinpoint here:
  - <https://surveillance.cancer.gov/joinpoint/example.html>
  - <https://surveillance.cancer.gov/joinpoint/faq/>
- For this example, we will be importing the breast cancer age-standardized rates from the previous tutorial sections.

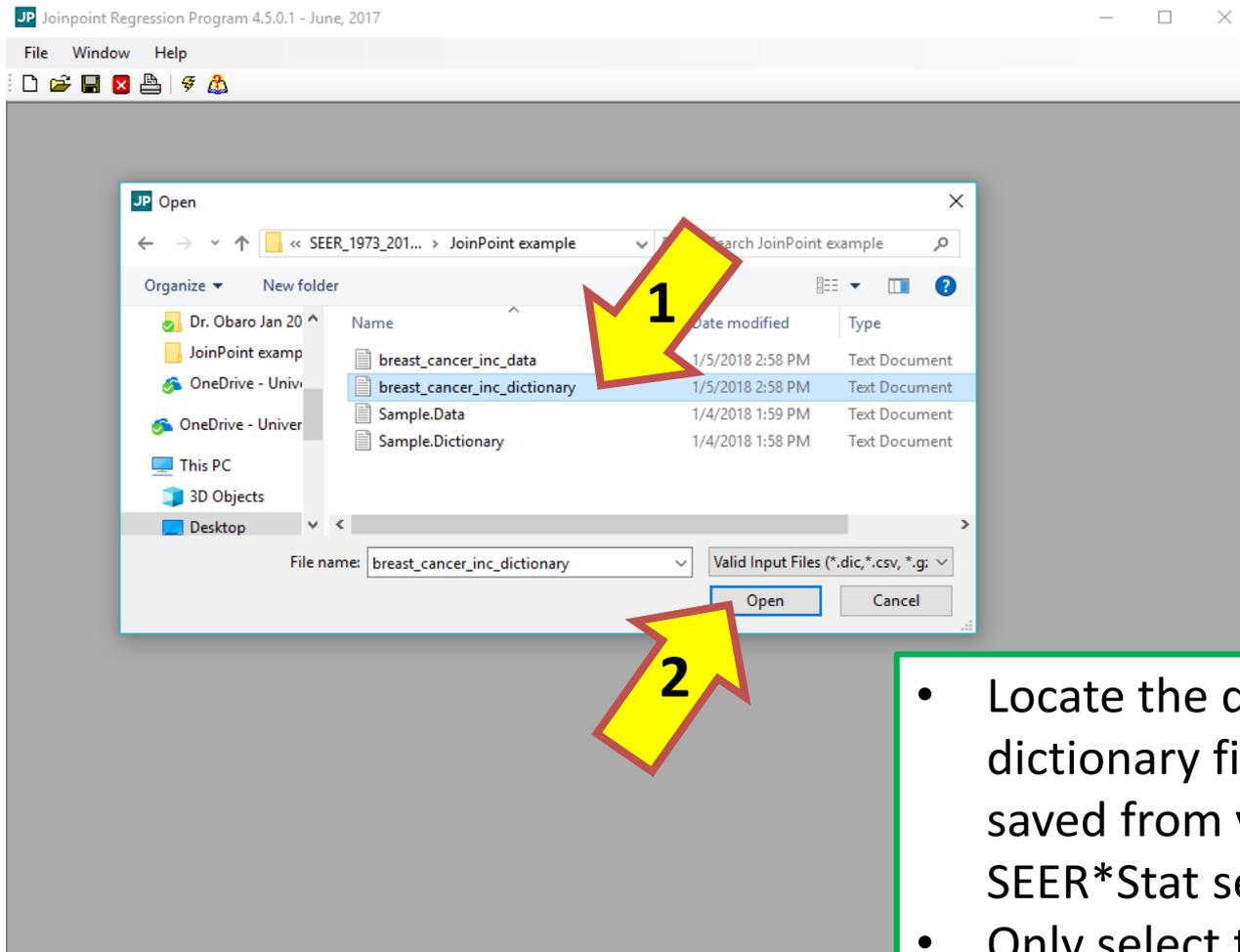


Double click the Joinpoint icon on your desktop to open the software.

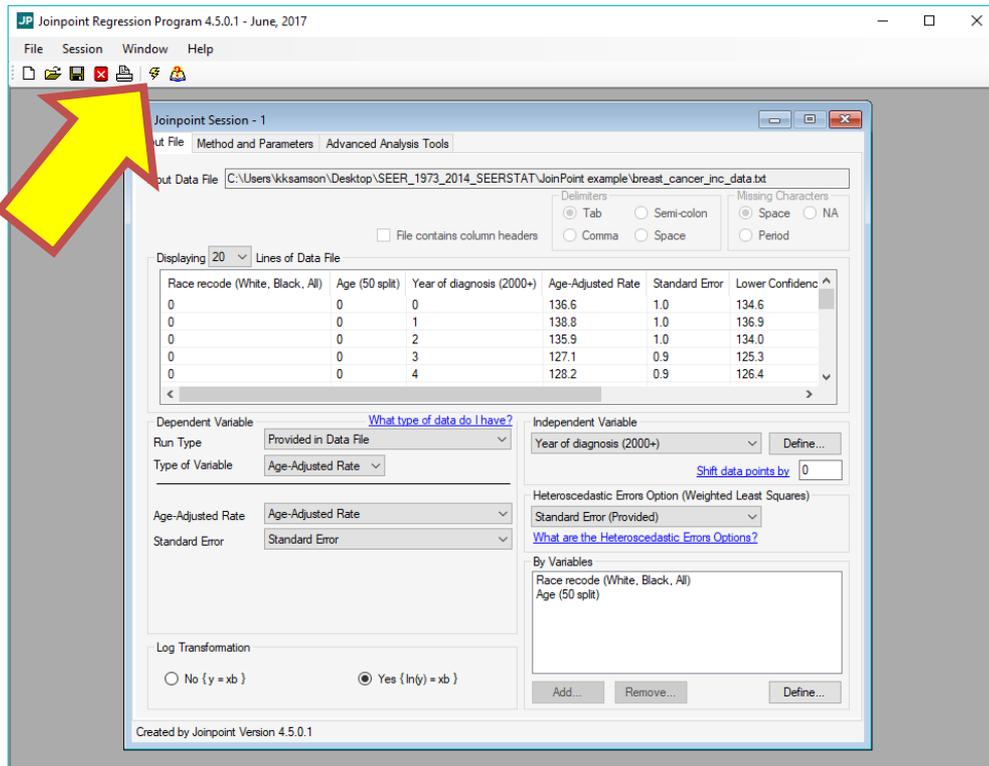


To start a new session, go to File and then “New Session...”





- Locate the data and dictionary files you saved from your SEER\*Stat session.
- Only select the **dictionary file** and then click “Open”.

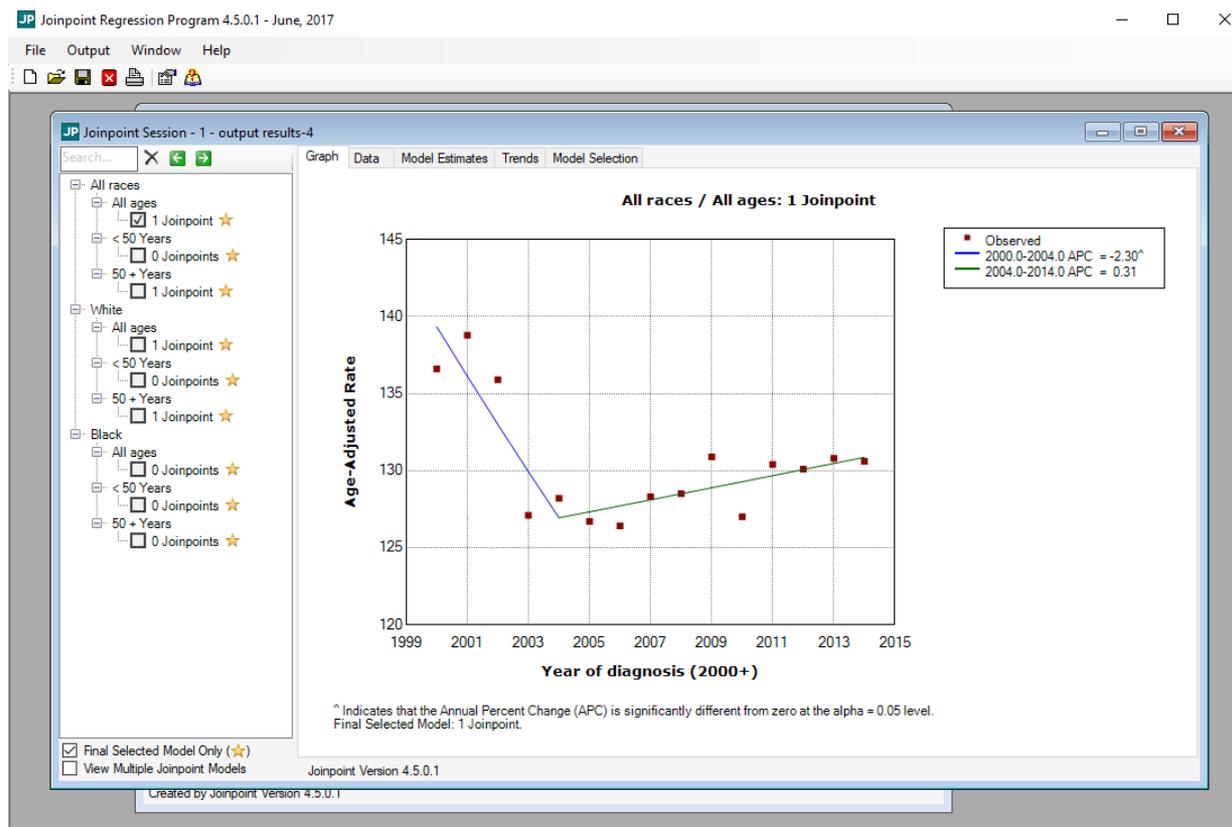


For the purposes of this exercise, we will use the default settings. Note that because we imported the data and dictionary from SEER\*Stat, the Joinpoint software automatically detected Run Type, Type of Variable, Age-Adjusted Rate, Standard Error, and Independent/By Variables. If you imported your own data, you would have to manually specify these.

Click the lightning bolt button to run the session.

The software generated a Joinpoint regression for each combination of the By variables. Here, we are seeing the trend for all races and ages. You may click through the other combinations on the side panel to see other Joinpoint regressions. For additional resources on what Joinpoint software does, please see this website:

<https://surveillance.cancer.gov/help/joinpoint/tech-help/frequently-asked-questions>



# Contact Information

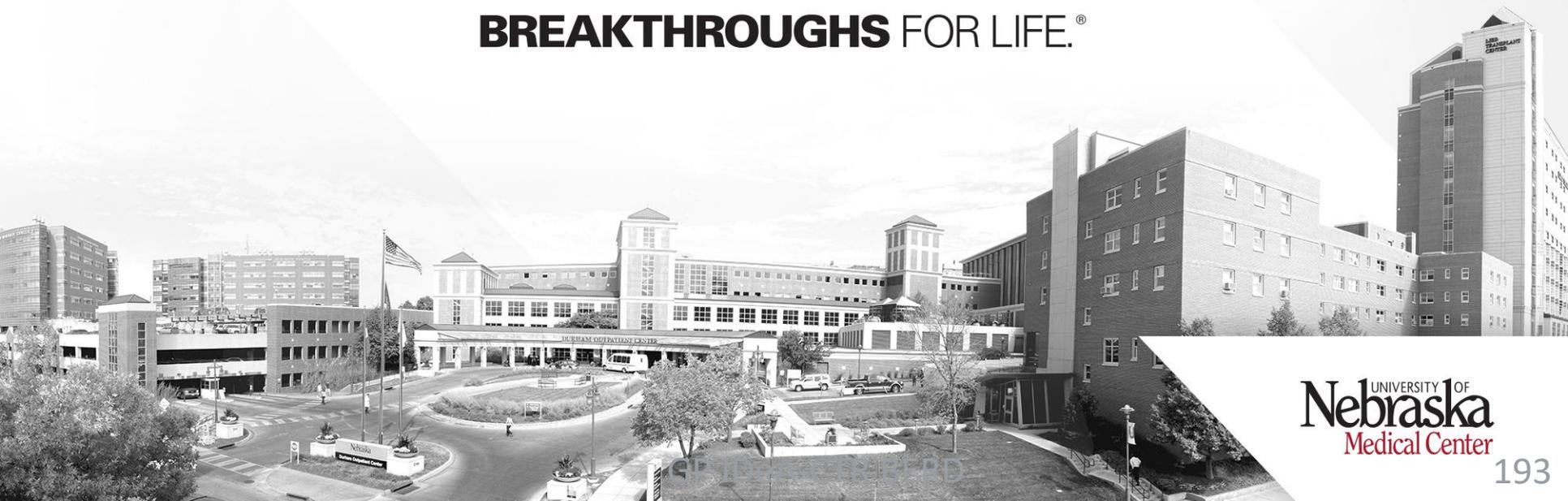
If you have any questions related to the content of this presentation, or have noticed that this presentation needs to be updated due to changes with SEER data or software, please contact Kaeli Samson at [kksamson@unmc.edu](mailto:kksamson@unmc.edu).





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

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



UNIVERSITY OF  
**Nebraska**  
Medical Center