

Webinar notes for Washington State SNAP-Ed

Nikki Eller, MPH

6/17/2019

SNAPedEvaluation@doh.wa.gov

Downloading data from PEARS

From the PEARS homepage, go to the Analyze dropdown, and under Survey Data click on Responses by Matched ID. You can choose which survey you want to see a summary of, which essentially filters by age group. The “completed” filter will only include activities that have been marked as complete, and the “Active” filter will only show current surveys. It’s a good idea to leave both checked, so that you’re seeing activities that have finished (and will have post-tests entered), and surveys that are currently being used. Date is the only other filtering option you have, PEARS will give you all the data you have access to. So if you want to see a subset for a particular agency or individual, you would have to ask that person to export the data they can see. The “code values” option you can check when you hit download gets you numbers instead of written-out responses in the spreadsheet. Best to leave it unchecked to make it easier to read the output.

Understanding the downloaded data

The first tab is the Codebook, which gives helpful definitions of the fields on the other tabs. The Response Data tab has a row for each individual survey entries, so isn’t all that helpful. The Matching Errors tab lets you know how many surveys were unmatched and why. Often it’s because the participant only took one survey, then you get the error “There is only one question type for this unique id.” Note that you can count the number of unmatched surveys by pre and post-test by counting rows on this page.

The most interesting data is on the Summary Statistics tab. Note the “Pre” and “Post” divisions at the top.

HTML tags sneak into the Question column, making it a little harder to read. Just ignore anything in angle brackets (like), that’s what tells webpages to display a word in **bold**.

You’ll see the same question repeated over several rows, so that there is one row for each response option to each question. The Frequency column gives the number of people who selected a response option, while the Question Total column gives the number of people who answered that question, so if you divide Frequency/Question Total you’ll get Percent of Response, which is the next column. To make it easier to read, you can select the numbers in the Percent of Response column, right-click, choose “Format Cells” and change the format to percentage.

We have all of this information repeated for the post test, so that we can compare it and see how things have changed since the pre-test. Which is really a lot more interesting than looking at them on their own.

Jumping over to the Change in Percent column, this is simply the post Percent of Response column minus the pre Percent of Response column. Scanning down this column, we can look for larger numbers jumping out, which indicate a greater difference between pre and post test. Negative numbers in this column mean a response was more common at pre-test than at post.

The Percent Change column takes the Change in Percent column and divides it by the Percent of Response at pre-test, to give you an idea of how big percent change is. For example, though a change in percent may be only 1.66%, if the pre-test percent of response was only .19%, the Percent Change column tells us that this was a more than 8-fold increase in the number of people choosing that response option. However, it's important to look back at the Frequency column to understand the context: in this case, that 8-fold increase meant 1 person selected a certain response at pre-test, and 6 selected it at post-test, in practical terms, not a very large increase. Also note that if the frequency at pre-test is 0, there will be an N/A in this column because it's impossible to divide by 0.

So now we can look down the Change in Percent and Percent Change columns for big numbers that stand out. It can be helpful to highlight the rows corresponding to a single question and look for patterns, such as decreases in the number of people selecting "no" for a behavior and increases in the number selecting "yes."

Important to Keep in Mind

A problem with this quick analysis tool is that even though it catches matching errors, it still includes unmatched surveys in the pre and post-test statistics, **so they are not necessarily the same people taking the pre and post test.** This means we can't directly interpret these results as individuals reporting changes in health behaviors, because it could be that the people who took the pre-test had different behaviors than the people who took the post-test. As a result, there is more randomness and variation, and more risk that the changes we see from pre to post test are not really happening.

However, this quick analysis can give you a rough idea of which questions you're seeing the most improvement on, and where things are changing with interventions. The evaluation team reports out only matched surveys, so that we can be confident in what the numbers mean. If you are interested in seeing a report of matched surveys only, fill out the "Evaluation Request Form" under Guidance Documents on the Evaluation page of <http://wasnap-ed.org/evaluation/> and the evaluation team can send it to you within a month.

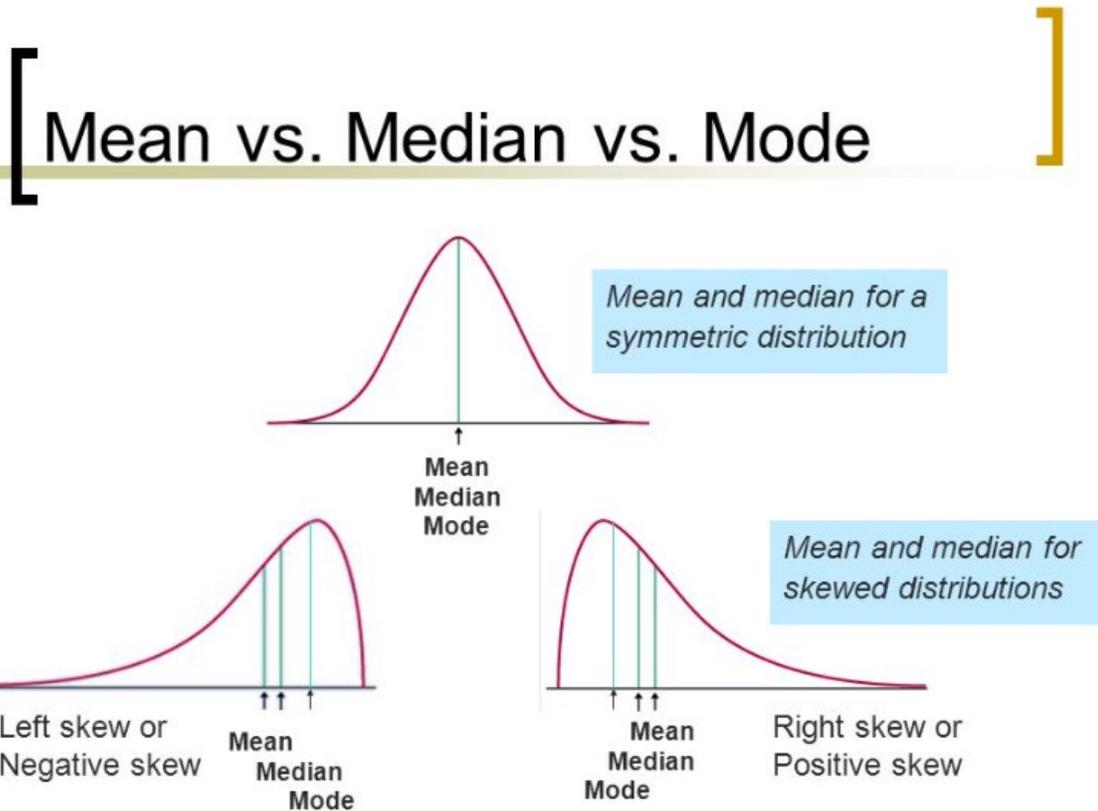
Extra Credit

The three columns, mean, median, and mode, are what we call "measures of central tendency," because they all tell us about where the middle of the data is, and combined, can help us figure out the shape of our data.

The Mean column gives the average response for each question, and the number given corresponds to the numbered response options. So a mean of 4.92 for the question "How old are you?" actually means that the average age was between response options 4 (age 11) and option 5 (age 12), which makes sense when you look at the frequency of the responses.

The median is the value that sits 50% of the way through the data, so if you sorted your data from smallest to largest, counted through half of it, the value that sits there is the median. Mode simply means the most common response to a question.

Often, responses tend to follow what's called a bell curve (top image, below), where most responses cluster around the average in the middle, and the frequency of answers goes down the farther they are from the average. Data that look like the lower two images are what we call skewed and look more like this:



Author: Jason Turner. Downloaded from: <http://slideplayer.com/slide/6992598/>

The lower left picture shows data with a left skew, because though the bulk of responses are clustered together, there's a long tail of responses going off to the left.

We can see from this picture that when the data fit a bell curve, mean, median, and mode are all the same, but when those three values are different the data are skewed.

Moving on to the next column, the Standard Deviation, has a complicated formula I won't get into, but is meant to give you an idea of the spread of the data. 68% of data falls within 1 standard deviation of the mean, so larger standard deviations means data is more spread out. See the diagram, below:

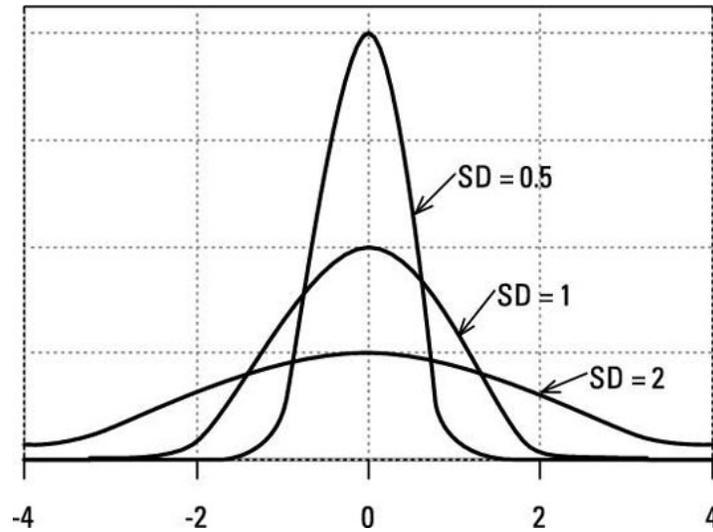


Image from article by Patrick Weaver. Downloaded from: <http://projectmanager.com.au/can-you-use-standard-deviation-in-project-management/>

All of these curves have an average of 0, but as you can see they flatten out as the standard deviation gets larger, because responses are less clustered around the average.

EARS Data

PEARS also stores all of the demographic information collected from SNAP-Ed participants. From the Analyze drop-down, select EARS Data under Reports to get to a form where you can select a timeframe, county, or region to view demographic information. Fill out the form with the area and time period you're interested in, click Run Report and PEARS will open up a new tab in your browser (you may need to allow pop-ups).

Keep in mind that the data PEARS can report is only as good as the data that was entered into it, so, if age, sex, race and ethnicity were not collected for participants, they will not show up in the EARS report. While EARS tries to produce an unduplicated count of participants, the way program activities are entered may affect its ability to do so. Contact SNAPedEvaluation@doh.wa.gov for any questions specific to your data.

Item 1 shows the parameters you selected when you ran the EARS report, Items 2 and 3 show participants by age group, sex, race, and ethnicity. The Actual Count column displays a count of participants who self-reported their demographic information, and the Estimated Count displays the number of participants where SNAP-Ed staff used some other data source to estimate the number of participants who would have chosen a particular category.

Item 4 gives a breakdown of the number and length of sessions used in direct education. Item 5 looks complicated, but column F lets you see how many participants were reached with each intervention listed in column A. The letter codes can be deciphered by following the link in the header to "SNAP-Ed EARS Form" and scrolling down to page 6.