**TOLERANCE3T: Exercise Using SPSS to Explore Relationships Between Variables with Tolerance as Independent Variable**

Author:   Ed Nelson
Department of Sociology M/S SS97
California State University, Fresno
Fresno, CA 93740
Email:  **ednelson@csufresno.edu**

**Note to the Instructor:** The data set used in this exercise is gss14\_subset\_for\_classes\_TOLERANCE.sav which is a subset of the 2014 General Social Survey. Some of the variables in the GSS have been recoded to make them easier to use and some new variables have been created.  The data have been weighted according to the instructions from the National Opinion Research Center.  The goal of this exercise is explore some possible correlates of tolerance.  This exercise uses RECODE to combine categories of variables, FREQUENCIES to see how respondents answered the questions, and CROSSTABS to explore the relationships between variables.  A good reference on using SPSS is *SPSS for Windows Version 23.0 A Basic Tutorial* by Linda Fiddler, John Korey, Edward Nelson (Editor), and Elizabeth Nelson.  The online version of the book is on the [**Social Science Research and Instructional Council's Website**](http://ssric.org/node/582).  You have permission to use this exercise and to revise it to fit your needs.  Please send a copy of any revision to the author. Included with this exercise (as separate files) are more detailed notes to the instructors, the SPSS syntax necessary to carry out the exercise (SPSS syntax file), and the SPSS output for the exercise (SPSS output file).   Please contact the author for additional information.

I’m attaching the following files.

* [**Data subset**](http://ssric.org/files/gss14_subset_for_classes_TOLERANCE.sav) (.sav format).
* [**Extended notes for instructors**](http://ssric.org/files/Extended_Notes_for_Instructors_for_TOLERANCE3T.docx). MS Word (.docx) format.
* [**SPSS syntax file**](http://ssric.org/files/SPSS_Syntax_for_TOLERANCE3T.sps) (.sps format).
* [**SPSS output file**](http://ssric.org/files/SPSS_Output_for_TOLERANCE3T.spv) (.spv format).
* [**This page**](http://ssric.org/files/TOLERANCE3T.docx) in MS Word (,docx) format.

**Goals of Exercise**

The goal of this exercise is to discover opinions and behaviors that might be correlated with tolerance.  The exercise also gives you practice in using several SPSS commands – RECODE to combine categories of variables, FREQUENCIES to explore how respondents answer various questions, and CROSSTABS to explore relationships between variables.

**Part I—Recoding the Variable We’re Using to Measure Tolerance**

We’re going to use the General Social Survey (GSS) for this exercise.  The GSS is a national probability sample of adults in the United States conducted by the National Opinion Research Center.  The GSS started in 1972 and has been an annual or biannual survey ever since.  For this exercise we’re going to use a subset of the 2014 GSS. Your instructor will tell you how to access this data set which is called gss14\_subset\_for\_classes\_TOLERANCE.sav.

Tolerance refers to the willingness of individuals to allow others to express opinions which might be very different from their own and to exercise their basic civil liberties in the expression of these opinions.  The GSS has a series of 18 variables that we can use to measure tolerance.  These 18 variables are divided into three sets of variables questions each.

One set of variables deals with the willingness of respondents to allow those who might hold these very different opinions to teach in a college.  The questions on which these variables are based start with a general statement that “there are always some people whose ideas are considered bad or dangerous by other people. For instance, somebody who is against all churches and religion.”  This statement is followed by a question – “Should such a person be allowed to teach in a college or university, or not?”  There are six scenarios presented:

* “somebody who is against all churches and religion,”
* “a man who admits he is a communist,”
* “a man who admits he is a homosexual,”
* “a person who advocates doing away with elections and letting the military run the country,”
* “a Muslim clergyman who preaches hatred of the United States,” and
* “a person who believes that Blacks are inferior.”

The second set of questions focus on these same six scenarios but ask whether a book that such a person wrote “should be taken out of your public library.”  The third set asks whether such a person should “be allowed to make a speech in your (city/town/community).”

These questions were originally developed by Samuel Stouffer in his book on *Communism, Conformity, and Civil Liberties* (Doubleday, 1955). He asked about teaching in a college or university, having a book in a public library, and making a public speech for three groups:

* Communists,
* socialists, and
* those against all churches and religions.

These nine questions were included in the first three General Social Surveys in 1972, 1973, and 1974.  The question about socialists was dropped in 1975 and a question about homosexuals was added in 1973, those advocating military control of the country and those who are racists in 1976, and Muslim clergyman who preach hatred of the United States in 2008.  (See Tom W. Smith, “A Review of the Stouffer Civil Liberties Items on the General Social Survey,” GSS Topical Report No. 42, 2009)  The wording of the questions was not changed over time to ensure the comparability of the questions.  While we might prefer to bring the wording of the questions more in line with the way we would ask them today, it’s more important to maintain continuity over time.

So we’re working with 18 variables which are listed below:

* six questions focusing on teaching in a college or university – variable names are T1\_COLATH, T2\_COLCOM, T3\_COLHOMO, T4\_COLMIL, T5\_COLMSLM, T6\_COLRAC;
* six questions focusing on having books in a public library – variable names are T7\_LIBATH, T8\_LIBCOM, T9\_LIBHOMO, T10\_LIBMIL, T11\_LIBMSLM, T12\_LIBRAC; and
* six questions focusing on making a public speech in their community – variable names are T13\_SPKATH, T14\_SPKCOM, T15\_SPKHOMO, T16\_SPKMIL, T17\_SPKMSLM, T18\_SPKRAC.

In a previous exercise (TOLERANCE1T) we created an overall measure of tolerance based on these 18 variables.  The measure students created was named TOL.  I created the same measure and named it TOLR in the data set.  This is to avoid confusion between the measure students created and the measure I created.  In each variable that makes up this composite measure the value 1 refers to the tolerant answer and the value 2 refers to the non-tolerant answer.  So if we sum these 18 variables we’ll get a new variable with 18 being the lowest possible value and 36 being the highest possible value. Low values indicate greater tolerance and high values indicate less tolerance.  Let’s start by running FREQUENCIES in SPSS for this variable.  (See Chapter Three, Frequencies in the online SPSS book.)

There are 19 different categories in our overall measure of tolerance.  That’s too many to work with.  So we’re going to recode this variable into fewer categories.  When you use RECODE in SPSS, you can recode in two different ways—into the same variable or into different variables.  If you recode into the same variable, be careful.  It’s easier, but if you make a mistake, you will not be able to go back and recode it again.  You will have to close SPSS without saving the data set and then reopen the data set to get a fresh, clean copy of the data. So for this exercise recode into different variables.  (See Chapter 3, Recoding into Different Variables in the online SPSS book.)

There are two guidelines to follow when recoding.

* Try not to have so few categories that you lose too much information.  Recoding into two categories almost always results in too much loss of information.
* Try not to have too many categories.  You’ll find that too many categories make it hard to interpret the data and are confusing to the reader of your report.

A good rule of thumb is to recode into three to five categories.

We’re going to recode into four categories but what should those categories be?  It’s a good idea to avoid a category (when possible) that has a very large percent of the cases or a very small percent.  What we can do is try to construct categories that have about 25 percent of all the cases in each category.**[[1]](http://ssric.org/node/490/edit%22%20%5Cl%20%22_ftn1%22%20%5Co%20%22)**  You won’t be able to have exactly 25 percent in each category but you can approximate it.  We can accomplish this by creating the following four categories:

* category 1 will be 18 through 19,
* category 2 will be 20 through 23,
* category 3 will be 24 through 27, and
* category 4 will be 28 through 36.

You’ll have to give your recoded variable a new name.  Call it TOL1.  To make your output more readable, assign value labels to these categories.

To make sure you didn’t make a mistake, run FREQUENCIES for your recoded variable (TOL1) and compare it to the frequency distribution for the variable I created which is named TOLR1.  They should be identical.  If you made a mistake, redo this part of the exercise.  If you recoded into the same variable, you will have to exit SPSS (or close your file) being sure **NOT** to save it.  Then get back into SPSS and open the gss14\_subset\_for\_classes\_TOLERANCE.sav file again.  The reason for this is that you have altered the coding of these this variable and will have to get another copy of the data file to start over.  If you saved the data file, then you would have written over the original copy.  So be careful.  That’s why we said to recode into different variables in this exercise.

**Part II—Trust as a Possible Correlate of Tolerance**

Now that we have our tolerance measure in a form that we can use in our analysis, let’s think about variables that might be correlated with tolerance.  There are several variables in the GSS that deal with the trust that the respondent has for others.

* tf1\_trust – The question is “generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?” The response categories are can trust, depends, and cannot trust.
* tf2\_fair – The question is “do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?  The response categories are take advantage, depends, and fair.
* tf5\_helpful – The question is “Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?”  The response categories are helpful, depends, and lookout for selves.

Let’s start by running FREQUENCIES for each of these variables.  Write a couple of paragraphs describing the trust that respondents have in others.

Choose one of the trust variables for your analysis.  Select a variable which you think will be related to tolerance.

1. Write a hypothesis stating how you expect tolerance to be related to your trust variable.
2. Write a paragraph or two that indicates why you think tolerance will be related to trust.  In other words, write an argument to support your hypothesis.
3. Use SPSS to run the crosstabulation of TOLR1 and your variable.  Think about which is the independent and dependent variable.  Remember to get the correct percentages.  Use Chi Square and an appropriate measure of the strength of the relationship.
4. Write a paragraph interpreting the table that SPSS gave you and indicate whether the data support your hypothesis.  Use the percents, Chi Square, and the measure of association to help you interpret the table.  Why do you think you found this relationship?

**Part III – Voting as a Possible Correlate of Tolerance**

This time let’s think about political correlates of tolerance.  There are four variables in the GSS that tell us how respondents said they voted in the 2008 and 2012 presidential elections and whether they said they voted.

* p5\_pres08 – Did they say they voted for Obama or McCain?
* p6\_pres12 – Did they say they voted for Obama or Romney?
* p7\_vote08 – Did they say they voted or didn’t vote in the 2008 presidential election or were they ineligible?
* p8\_vote12 – Did they say they voted or didn’t vote in the 2012 presidential election or were they ineligible?

Choose one of the years (2008 or 2012) for your analysis.

1.         Write a hypothesis stating how you expect tolerance to be related who they said they voted for and whether they voted.

2.         Write a paragraph or two that indicates why you think tolerance will be related to voting.  In other words, write an argument to support your hypotheses.

3.         Use SPSS to run the crosstabulation of TOLR1 and your two variables.  Think about which is the independent and dependent variable.  Remember to get the correct percentages.  Use Chi Square and an appropriate measure of the strength of the relationship.

4.         Write a paragraph interpreting the tables that SPSS gave you and indicate whether the data support your hypotheses.  Use the percents, Chi Square, and the measure of association to help you interpret the table.  Was the relationship between tolerance and your dependent variable stronger for whether they voted or for whom they voted?  How do you know?

**Part IV – Conclusions**

What did you learn about the relationship of tolerance to trust and voting?  Try to explain why tolerance is related to trust and/or voting.  In other words, why did you find these relationships?

**[[1]](http://ssric.org/node/490/edit%22%20%5Cl%20%22_ftnref1%22%20%5Co%20%22)** What you are doing is dividing the data into four quartiles.