# Exercises on Tolerance Using SPSS and the General Social Survey

# Edward Nelson

# California State University, Fresno

# Table of Contents

 Page

Preface 3

Exercise 1 – Exploring Measurement, Reliability, and Relationships Between Variables 4

Exercise 2 -- Exploring Measurement and Relationships Between Variables with Tolerance
as the Dependent Variable 8

Exercise 3 -- Exploring Relationships Between Variables with Tolerance as the Independent 13
Variable

Appendix – Codebook for data set 17

# Preface

The data set used in this exercise is the 2018 General Social Survey. The name of the data set is GSS18TOLERANCE.SAV which is a subset of the full data set. 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 instructions from the National Opinion Research Center.  These exercises use SPSS to analyze the data but you could also easily use PSPP.  A good reference for using SPSS is [SPSS Statistics (Version 26): A Basic Tutorial](http://www.ssric.org/node/696)by Edward Nelson and John Korey.

The focus of these exercises is on frequency distributions and two-variable crosstabs. The exercises revolve around respondents' willingness to allow others to express opinions that might be very different from their own.

The exercises do not explain how to use SPSS. For that you could use the [SPSS tutorial](http://www.ssric.org/node/696) referred to throughout the exercises. I have added to the data set the recoded variables that students are asked to create. The names of these variables all start with ZZ. That means you could skip some of the SPSS procedures by referring students to these variables in the data set. I have also included in the data set some variables not used in the exercises so you could develop your own exercises around these variables.

The exercises were written so that each exercise is independent of the other exercises. That means that there is some redundancy across the exercises. If you choose to use several exercises you may want to remove some of the redundant material.

You have permission to use these exercises and to revise them to fit your needs which would include adding materials of your own or deleting parts of the exercises that you don't want to use.

# Exercise 1 -- Exploring Measurement, Reliability, and Relationships Between Variables

## **Goal of Exercise**

The goal of this exercise is to create a measure of tolerance where tolerance refers to the willingness to allow others to express opinions that might be very different from their own.  In other words, the willingness to allow people basic civil liberties (i.e., speak in public, teach in college, have books in public libraries) regardless of their own personal opinions.  In subsequent exercises we will focus on discovering which individuals are more or less tolerant and to consider other opinions and behaviors that might be correlated with tolerance.  However, the focus of this exercise will be on creating the tolerance measure.  The exercise also gives you practice in using several SPSS commands – COMPUTE to create new variables, FREQUENCIES to explore how respondents answered the tolerance questions, and CROSSTABS to explore relationships between variables.  You will also be using CRONBACH’S ALPHA to estimate the internal consistency or reliability of the variables that make up the tolerance measure.

## **Part I—Questions Used 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 2018 GSS. Your instructor will tell you how to access this data set which is called GSS18TOLERANCE.SAV. You can also download the data by clicking on this [link](http://ssric.org/node/485).

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 six variables 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 were added 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 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 COLATH, COLCOM, COLHOMO, COLMIL, COLMSLM, COLRAC;
* six questions focusing on having books in a public library – variable names are LIBATH, LIBCOM, LIBHOMO, LIBMIL, LIBMSLM, LIBRAC; and
* six questions focusing on making a public speech in their community – variable names are SPKATH, SPKCOM, SPKHOMO, SPKMIL, SPKMSLM, SPKRAC.

Let’s start by running FREQUENCIES in SPSS for each of these three groups of six variables.  (See Chapter 4, Frequencies in the [online SPSS tutorial](http://ssric.org/node/696).)  The best way to organize your tables is to first request the three tolerance variables for those opposed to religion, then the three tolerance variables for Communists, and so on.  Write several paragraphs describing how respondents in 2018 answered these questions.  Organize your answer by discussing whether more or less tolerance is shown towards some of these six groups.  Think carefully how you are going to present the data to illustrate the patterns you see.

## **Part II—Internal Consistency or Reliability of the Variables**

We don’t want to deal with each of these 18 variables separately.  Rather we want to develop composite measures that combine these variables.  Let’s start by developing six composite measures – one for each of the six groups (i.e., those opposed to religion, Communists, homosexuals, militarists, anti-American Muslim clergymen, racists).

Each group has three variables.  We need to determine if these three variables are internally consistent.  This is basically a question of reliability.  Are they consistent in the sense that a person who agrees that someone who is against religion should be allowed to teach in a college also thinks that such a person should also be allowed to make a public speech and have his or her books in public libraries?

A commonly used measure of internal consistency is Cronbach’s Alpha.  This statistic varies from 0 to 1.  Low values indicate low consistency and high values indicate high consistency.  Alpha increases as the number of items increases and as the correlation between items increases.  One rule of thumb that is often used is that an Alpha of .70 or higher is necessary to indicate reliability although some feel that a higher value is required.

Let’s have SPSS compute Cronbach’s Alpha for the three questions about those who are opposed to religion.  Those are COLATH, LIBATH, SPKATH.  Open the data set in SPSS and click on ANALYZE in the menu.  Then point your mouse at SCALE and click on RELIABILITY ANALYSIS.  Click on each of the three variables in the box on the left and move them over to the ITEM box on the right until you have selected all three variables.  Click on the STATISTICS button in the upper right.  This will open the STATISTICS box. Then click on SCALE IF ITEM DELETED in the upper left and CORRELATIONS in the upper right.  Click on CONTINUE and this will take you back to the original box where you selected the variables.  Finally click on OK and SPSS will carry out the calculations you requested.  There is an excellent YouTube video in two parts on Cronbach’s Alpha — ([**Part 1**](https://www.youtube.com/watch?v=2gHvHm2SE5s) and [**Part 2**](https://www.youtube.com/watch?v=9rS49o1rdnk)). My instructions only ask for some of the calculations suggested in this video.

In your output there are three important pieces of information:

* the value of Cronbach’s Alpha,
* the inter-item correlation matrix showing the correlations between each pair of variables, and
* the value of Cronbach’s Alpha if an item (i.e., variable) is deleted.

You want to get a Cronbach’s Alpha of .70 or higher.  The correlations between variables should be consistently high.  Generally the higher the inter-item correlations, the higher the value of Alpha.  It’s also helpful to look at the effect of deleting a variable on Alpha. This will tell you if you ought to consider deleting that variable.

Follow these instructions to compute Alpha for each of the six groups – those who are opposed to religion, Communists, homosexuals, militarists, anti-American Muslim clergymen, and racists.  Based on your analysis, do you think that each of these composite measures is internally consistent or reliable?  Explain why.

## **Part III—Constructing the Six Composite Variables**

Construct six composite variables – one for each of the six groups.  Use COMPUTE in SPSS to create the composite measures.  (See Chapter 3, Creating New Variables Using Compute in the online SPSS tutorial.)  If any cases have missing information for any of the three variables, that composite measure will automatically be given a system missing value for those cases.  Give your measures the following variable names – TOLATH, TOLCOM, TOLHOMO, TOLMIL, TOLMSLM, TOLRAC.  Remember that 1 is always the more tolerant response and 2 is the less tolerant response.  So for your composite measure low values will indicate more tolerance and high values will indicate less tolerance.  Assign value labels to make the output more readable.

It’s easy to make a mistake so I have created six variables that show you what the frequency distributions for your variables ought to look like.  The variables I created are named TOLATHR, TOLCOMR, TOLHOMOR, TOLMILR, TOLMSLMR, TOLRACR.  Notice the R at the end of each variable name.  That distinguishes your variables from my variables.  Run FREQUENCIES for your six variables and the six variables I created.  Your variables and my variables ought to have identical frequency distributions.  If they don’t, you made a mistake and you’ll need to redo your work in SPSS.

## **Part IV—Constructing One Overall Composite Variable**

Now let’s construct one composite variable that will use all 18 variables.  Start by running the Cronbach’s Alpha analysis for all 18 variables.  Look at the Alpha value and the inter-item correlation matrix.  Is your value of Alpha higher or lower than in Part III?  The more items in the measure, the higher the value of Alpha (assuming similar inter-item correlations).  What did you find?

Once you have convinced yourself that the variables are internally consistent or reliable, use COMPUTE to create the 18-variable composite measure.  Name your variable TOL.  Add value labels to make it easier to read your output.  Run FREQUENCIES for your variable and for the variable I created which is named TOLR.  Compare the two frequency distributions.  They should be identical.  If they aren’t, you made a mistake and will have to redo your work in SPSS.

## **Part V—Analysis**

How are our six composite measures – TOLATH, TOLCOM, TOLHOMO, TOLMIL, TOLMSLM, TOLRAC – related to each other?  To answer this question would require running 15 tables in CROSSTABS.  (See Chapter 5, Crosstabulations in the online SPSS tutorial.)  That’s a lot of tables.  So let’s focus on how TOLRAC is related to each of the other variables.  Use CROSSTABS in SPSS to see how TOLRAC is related to TOLATH, TOLCOM, TOLHOMO, TOLMIL, and TOLMSLM.  That means you will need to run five tables in CROSSTABS. Put TOLRAC in the column and ask for the column percents.  Be sure to also get Chi Square and an appropriate measure of association.  Write a paragraph describing how TOLRAC is related to each of the other variables.

## **Part VI – Summary and Next Exercise**

In this exercise we created several different measures of tolerance. In the next exercise we're going to determine which individuals are more or less tolerant of those who express opinions that might be very different from their own.

# Exercise 2 -- Exploring Measurement and Relationships Between Variables with Tolerance as the Dependent Variable

## **Goal of Exercise**

The goal of this exercise is to discover which individuals are more or less tolerant of those who express opinions which might be very different from their own.  We will consider such variables as age, sex, education, family income, and the region of the country in which the respondents live to see if these variables are related to tolerance.  In a subsequent exercise we will consider other 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 2018 GSS. Your instructor will tell you how to access this data set which is called GSS18TOLERANCE.SAV. You can also download the data by clicking on this [link](http://ssric.org/node/485).

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 six variables 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 were added 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 COLATH, COLCOM, COLHOMO, COLMIL, COLMSLM, COLRAC;
* six questions focusing on having books in a public library – variable names are LIBATH, LIBCOM, LIBHOMO, LIBMIL, LIBMSLM, LIBRAC; and
* six questions focusing on making a public speech in their community – variable names are SPKATH, SPKCOM, SPKHOMO, SPKMIL, SPKMSLM, SPKRAC.

In Exercise 1, 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 tutorial](http://ssric.org/node/696).)

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 tutorial.)

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 often 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]](#footnote-1)  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 TOL1R.  They should be identical.  If you made a mistake, redo this part of the exercise.

## **Part II – More Recoding**

Now that we have our tolerance measure in a form that we can use in our analysis, let’s think about variables that might help us discover which individuals are more or less tolerant.  Let’s use the following variables:

* education – DEGREE is the respondent’s highest educational degree,
* family income – INCOME16 is the family’s yearly income for the year preceding the survey (2017),[[2]](#footnote-2)
* sex – SEX,
* age – AGE, and
* the region in which the respondent lives – REGION.

Let’s start by running FREQUENCIES for each of these variables. Some of these variables have too many categories.  So let’s recode three of them into fewer categories.

* age,
* family income, and
* region.

Look at your frequency distributions and decide on how you could recode age and family income into four categories. Try to get categories that have about the same number of cases.  You won’t be able to do this exactly, but you can approximate it.

Region uses the Census classification of states into nine Census regions.  We’re going to recode these regions into the four Census divisions using the following Census coding scheme:

* West combines the Pacific and Mountain regions,
* Midwest combines the West North Central and the East North Central regions,
* Northeast combines the Middle Atlantic and New England regions, and
* South combines the West South Central, East South Central, and South Atlantic regions.

Name your recoded values INCOME161, AGE1, and REGION1.  Add value labels to make your output more readable.  Remember to recode into different variables.

After you are done recoding, run FREQUENCIES for both the unrecoded and the recoded variables to make sure you didn’t make a mistake.  If you make a mistake, you’ll need to redo the recoding of that variable.

## **Part III—Analysis**

Now we’re ready to analyze the data.  All research starts with a question.  Your question is why some people are more tolerant than others.  So TOL1 will be your dependent variable.  The dependent variable is always what you are trying to explain.  Your independent variables are the variables that you think will help you explain the variation in your dependent variable. Put the dependent variables in the row and the independent variables in the column.   When you set your tables up this way, you’ll always want the column percents.  In addition to requesting the column percents you’ll want to get Chi Square and an appropriate measure of association.

Your independent variables are:

* education – DEGREE is the respondent’s highest educational degree,
* family income – INCOME161 is the family’s yearly income for the year preceding the survey (2017),[[3]](#footnote-3)
* sex – SEX,
* age – AGE1, and
* region in which the respondent lives – REGION1.

Run the tables in CROSSTABS and interpret them using the percents, Chi Square, and your measure of association to help you describe the relationship between the two variables in your crosstabulations.

## **Part IV.  Summary and Next Exercise**

What have you learned in this exercise?  Which variables are statistically related to tolerance?  Which of these variables are more strongly related to tolerance?  How do you know?  What did you discover about the relationship of each independent variable to tolerance?

In the next exercise we'll explore opinions and behavior that might be correlated with tolerance.

# Exercise 3 -- Exploring Relationships Between Variables with Tolerance as the Independent Variable

## **Goal 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 2018 GSS. Your instructor will tell you how to access this data set which is called GSS18TOLERANCE.SAV. You can also download the data by clicking on this [link](http://ssric.org/node/485).

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 six variables 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 were added 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 COLATH, COLCOM, COLHOMO, COLMIL, COLMSLM, COLRAC;
* six questions focusing on having books in a public library – variable names are LIBATH, LIBCOM, LIBHOMO, LIBMIL, LIBMSLM, LIBRAC; and
* six questions focusing on making a public speech in their community – variable names are SPKATH, SPKCOM, SPKHOMO, SPKMIL, SPKMSLM, SPKRAC.

In Exercise 1, 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 tutorial](http://ssric.org/node/696).)

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 tutorial.)

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 often 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.[[4]](#footnote-4)  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.

## **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 two variables in the GSS that deal with the trust that the respondent has for others.

* 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.
* 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.

Both of these variables seem to be variations on the theme of trust. Let's start by running FREQUENCIES for all three 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 the trust variable.  In other words, write an argument to support your hypothesis.
3. Use SPSS to run the crosstabulation of TOL1 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 2012 and 2016 presidential elections and whether they said they voted.

* PRES12 – Did they say they voted for Obama or Romney?
* PRES16 – Did they say they voted for Trump or Clinton?
* VOTE12 – Did they say they voted or didn’t vote in the 2012 presidential election or were they ineligible?
* VOTE16 – Did they say they voted or didn’t vote in the 2016 presidential election or were they ineligible?

Choose one of the years (2012 or 2016) 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 TOL1 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 – Summary**

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?

# **Appendix**

# **Codebook for the Subset of the 2018 General Social Survey**

The General Social Survey (GSS) is a large, national probability sample of adults in the United States. It began in 1972 and continued on an almost yearly basis until 1996. In 1996, the GSS became a biannual survey and the sample size increased. Many questions are asked on each survey, while other questions are rotated from survey to survey. This subset from the 2018 GSS includes all the cases (2,348) and 78 variables. This data set has already been weighted using the weight variable supplied by the GSS (WTSS). Some of the original GSS variables were recoded and a few new variables created. Some of the new variables have names similar to those in the original GSS data set. The data set was created to accompany these exercises and is named GSS18TOLERANCE.SAV.

**Variable** **Description of Variable**

ABANY Abortion if woman wants for any reason

ABDEFECT Abortion if strong chance of serious defect

ABHLTH Abortion if woman's health seriously endangered

ABNOMORE Abortion if married and wants no more children

ABPOOR Abortion if low income and can't afford more children

ABRAPE Abortion if pregnant as result of rape

ABSINGLE Abortion if not married

AGE Age of respondent

ATTEND How often respondent attends religious services

CAPPUN Favor or oppose death penalty for murder

CHILDS Number of children

CLASS Subjective class identification

COLATH Allow anti‑religionist to teach

COLCOM Allow communist to teach

COLHOMO Allow homosexual to teach

COLMIL Allow militarist to teach

COLMSLM Allow anti-American Muslim Clergyman to teach in college

COLRAC Allow racist to teach

DEGREE Respondent's highest degree

EDUC Highest year of school completed

FAIR People fair or try to take advantage

FEAR Afraid to walk at night in neighborhood

FINRELA Opinion of family income

FUND Fundamentalism of respondent’s religion

GRASS Should marijuana be made legal?

GUNLAW Favor or oppose gun permits

ID Respondent’s identification (id) number

INCOME16 Total family income (2017)

LIBATH Allow anti‑religious book in library

LIBCOM Allow communist's book in library

LIBHOMO Allow homosexual's book in library

LIBMIL Allow militarist's book in library

LIBMSLM Allow anti-American Muslim clergyman's book in library

LIBRAC Allow racist's book in library

MADEG Mother's highest degree

MAEDUC Highest year school completed, mother

MARITAL Marital status

PADEG Father's highest degree

PAEDUC Highest year school completed, father

PARTYID Political Party Affiliation

POLVIEWS Think of self as liberal or conservative

PORNLAW Feelings about pornography laws

PRAY How often does respondent pray?

PRAYER Support Supreme Court Decision on prayer in public schools

PRES12 Vote for Romney or Obama in 2012

PRES16 Vote for Clinton or Trump in 2016

RACE Race of respondent

REGION Region of interview

RELIG Respondent's religious preference

RELITEN Strength of religious affiliation

RELPERSN Respondent considers self a religious person

SATFIN Satisfaction with financial situation

SEX Respondent's sex

SIZE Size of place respondent lives in thousands

SPDEG Spouse's highest degree

SPEDUC Highest year school completed, spouse

SPKATH Allow anti‑religionist to speak

SPKCOM Allow communist to speak

SPKHOMO Allow homosexual to speak

SPKMIL Allow militarist to speak

SPKMSLM Allow anti-American Muslim clergymen to speak

SPKRAC Allow racist to speak

TOLATHR Composite measure – tolerance for anti-religionists

TOLCOMR Composite measure – tolerance for communists

TOLHOMOR Composite measure – tolerance for homosexuals

TOLMILR Composite measure – tolerance for militarists

TOLMSLMR Composite measure – tolerance for anti-American Muslim Clergymen

TOLRACR Composite measure – tolerance for racists

TOLR Composite measure – overall tolerance measure

TOL1R Composite measure – recoded overall tolerance measure

TRUST Can people be trusted?

VOTE12 Did respondent vote in 2012 election?

VOTE16 Did respondent vote in 2016 election?

WTSS Weight variable for GSS18 (data subset already weighted by the variable WTSS)

YEAR Year of survey (2018 for all respondents)

ZZINCOME161 Recoded family income
ZZAGE1 Recoded age
ZZREGION1 Recoded region
ZZTRUST1 Recoded can people be trusted
ZZFAIR1 Recoded people are fair or take advantage

1. What you are doing is dividing the data into four quartiles. [↑](#footnote-ref-1)
2. Note that the 16 in the variable name does not mean this was the family income in 2016.   Rather it means that this variable uses a set of categories that was developed in 2016.  [↑](#footnote-ref-2)
3. Education and family income are often used as indicants of socioeconomic status. [↑](#footnote-ref-3)
4. What you are doing is dividing the data into four quartiles. [↑](#footnote-ref-4)